The Analog Readout of the LHCb Vertex Detector and Study of the Measurement of the Bs Oscillation Frequency

THÈSE N^O 4122 (2008)

PAR

Jérémie BOREL

ingénieur physicien diplômé EPF
de nationalité suisse et originaire de Vevey (VD) et Neuchâtel (NE)

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2008

Abstract

𝕿HE LHCb detector is one of the four experimental setups built to detect high-energy proton collisions to be produced by the Large Hadron Collider (LHC). Located at CERN (Geneva, Switzerland), the LHC machine and the LHCb experiment are expected to start in 2008, and will then operate for several years.

Being the largest collider of its kind, the LHC will open the way to new investigations, in the very-high energies, but also in terms of statistics for the study of rare-phenomena and flavor physics. In this framework LHCb is dedicated to precise measurements of CP-violating and rare decays of beauty hadrons, in order to test (or over-constrain) the Standard Model of particle physics.

From the hardware point of view, the construction of such detectors represents several challenges; one of them is the routing at a very high frequency of many signals in a harsh radiation environment. We designed to this purpose a hardware setup and a software filter which together reduce the cross-talk present in the readout of the LHCb vertex detector to a level of $(1 \pm 2)\%$, leading to an improved signal quality in the acquisition chain.

From the physics point of view, many of the CP-violation measurements performed at LHCb using B_s decays, for example using $B_s \to D_s^{\mp}K^{\pm}$ decays, will require as input the B_s–$\overline{B}_s$ oscillation frequency Δm_s. Thus the measurement of B_s–$\overline{B}_s$ oscillations, which are best observed using the flavor-specific $B_s \to D_s^{-}\pi^{+}$ decays, will play an important role. We have developed a complete selection of $B_s \to D_s^{-}\pi^{+}$ and $B_s \to D_s^{\mp}K^{\pm}$ events, based on Monte Carlo simulations. Assuming 2 fb^{-1} of data, we expect a $B_s \to D_s^{-}\pi^{+}$ signal yield, after the first level of trigger, of 155 k events over a background between 0.6 k and 7.8 k events at 90% confidence level. Moreover, we assess with fast Monte Carlo studies the corresponding statistical sensitivity on the B_s oscillation frequency, $\sigma(\Delta m_s) = 0.008$ ps^{-1}, on the wrong tag fraction, $\sigma(\omega) = 0.003$, as well as on other parameters related to the B_s-meson system.

We also addressed an important aspect of the systematics associated with the Δm_s measurement, and developed a method to calibrate and assess the length scale. This calibration is performed through the reconstruction of secondary interactions occurring in the material of the vertex detector. We show that the statistical relative precision of this approach quickly matches 6×10^{-5}, obtained from the survey measurements of the detector.

Keywords: CERN, LHC, LHCb, Standard Model, CP violation, B_s–$\overline{B}_s$ oscillations, FIR, Cross-talk, Secondary interactions, length calibration

Resumé

$\mathcal{L}$E détecteur LHCb est l'un des quatre dispositifs expérimentaux construits pour détecter des collisions de protons à haute énergie produites par le Large Hadron Collider (LHC). Situés au CERN (Genève, Suisse), le LHC et l'expérience LHCb commenceront la prise de données en 2008 et fonctionneront ensuite pendant plusieurs années.

Etant le plus grand collisionneur en son genre, le LHC permettra de nouvelles recherches, tant dans le domaine des hautes énergies que dans celui des phénomènes rares ou de la physique des saveurs. En effet la quantité de données accumulées sera rapidement plus grande que celle accessible à ce jour. LHCb a été conçu pour des mesures de précision dans des désintégrations de hadrons beaux rares violant la symétrie CP, ceci afin de tester le Modèle Standard de la physique des particules.

Du point de vue technologique, la construction de tels détecteurs représente de nombreux défis. L'un deux est la transmission de signaux, à des fréquences très élevées et dans un environnement soumis à de fortes irradiations. Nous avons conçu à cette fin un circuit analogique ainsi qu'un programme de filtrage qui réduisent les couplages diaphoniques dans la chaîne d'acquisition du détecteur de vertex, ceci à un niveau de $(1 \pm 2)\%$, améliorant ainsi la qualité globale du signal.

Du point de vue de la physique, la plupart des mesures effectuées sur la violation de CP, par exemple dans les désintégrations $B_s \rightarrow D_s^{\mp} K^{\pm}$, nécessitent la valeur de la fréquence d'oscillation des mésons B_s (Δm_s), d'où l'importance de cette mesure dans le cadre de l'expérience LHCb. Ces oscillations peuvent être mesurées en utilisant les désintégrations à saveur spécifique $B_s \rightarrow D_s^{-} \pi^{+}$. Nous avons développé une sélection complète des évènements $B_s \rightarrow D_s^{-} \pi^{+}$ et $B_s \rightarrow D_s^{\mp} K^{\pm}$, basée sur des simulations Monte Carlo. En considérant une statistique de 2 fb^{-1}, nous montrons que 155 k événements $B_s \rightarrow D_s^{-} \pi^{+}$ sont sélectionnés après le premier niveau de trigger. Nous montrons également que le niveau de bruit combinatoire, pour la même statistique, se trouve entre 0.6 k et 7.8 k avec un niveau de confiance de 90%. La sensibilité de l'expérience LHCb aux oscillations des mésons B_s est extraite au moyen d'études Monte Carlo simplifiées. Nous obtenons une sensibilité à la fréquence Δm_s de $\sigma(\Delta m_s) = 0.008$ ps^{-1}, et une sensibilité à l'amplitude de l'oscillation de $\sigma(\omega) = 0.003$.

Afin d'étudier l'erreur systématique sur la mesure de Δm_s, nous étudions le problème de la calibration de l'échelle des longueurs. En effet, celle-ci représente la principale contribution à l'erreur systématique. Cette étude utilise la reconstruction des interactions secondaires qui ont lieu dans le détecteur de vertex. Elle montre que la précision statistique sur la calibration rejoint rapidement la limite de 6×10^{-5}, une limite absolue, due aux erreurs de mesure de la métrologie.

Mots-clés : CERN, LHC, LHCb, Modèle Standard, violation de CP, oscillations des mésons B_s, FIR, couplage diaphonique, interactions secondaires, calibration des longueurs

Remerciements

Contents

Introduction

AT the end of the nineteenth century, the physicists believed that they understood the most fundamental principles of nature. Atoms were solid building blocks, people trusted Newtonian laws of motion and no more real surprises were expected. Three decades later, quantum mechanics and the theory of relativity were well established, and the scientific vision of the world had undergone major changes.

Today, the lesson seems to be learned: the Standard Model of elementary particles is describing the known particle physics phenomenology with great accuracy and has so far successfully passed all the experimental tests. Still, because some aspects are unexplained, particle physicists are convinced that there must be something beyond, new phenomena called New Physics (NP), a very handy term which includes everything that is not known today. To investigate it, physicists have done what they usually do: theorists have developed new ideas, among them the Supersymmetry, the Technicolor and Superstring theory, to mention only the most famous ones; and experimentalists have designed new, larger and more precise particle detectors. For these, the Large Hadron Collider (LHC) has been built at CERN, and it will start operation in 2008.

There are already hints where NP could be found, and one of these is the study of the CP symmetry, which is the symmetry relating matter and antimatter (in fact, a combination of the so-called charge-conjugation C and the space inversion P). Indeed, although the observed CP violation can be accommodated within the Standard Model, its amplitude is too small to explain the observed matter-antimatter asymmetry in the universe. However, some predicted CP violation effects are still unmeasured or only poorly constrained by experiment, for instance those induced by particle-antiparticle mixing in the B_s system. What the LHC will discover at the high-energy frontier or in the precise study of CP violation is of course still unknown, but history shows that, most probably, it will come as a surprise.

This document summarizes the studies conducted by the author during the past four years. Although covering very different aspects and based on different approaches, those studies can be centralized under a unique theme: the precise measurement of the B_s–$\overline{B}_s$ oscillation frequency (Δm_s) at the Large Hadron Collider beauty experiment (LHCb).

The first chapter is a short introduction to the Standard Model, presenting the modern formalism of particle physics as well as the theoretical basis implied by the B_s–$\overline{B}_s$ oscillation mechanism. Chapter 2 gives a description of the LHCb experiment, which is one of the experiments located at CERN on the Large Hadron Collider (LHC), the largest proton synchrotron ever built.

In Chapter 3, we enter in the very details of the LHCb vertex detector and its analog acquisition chain. We present our own study to enhance the signal transmission quality, both from software and hardware point of views. Although this study is not directly related to the Δm_s measurement and of wider scope, we show that it has an impact on the tracking residuals, and thus indirectly on the vertexing, which in turn will influence the quality of the Δm_s measurement.

The second analysis is more directly related to particle physics. We first present in Chapter 4 the general software program involved in the simulation, the reconstruction and the analysis of the LHCb data. We then describe in Chapter 5 our specific selection of $B_s \rightarrow D_s^- \pi^+$ decays, which LHCb will use to perform the Δm_s measurement. We assess in this analysis the selection efficiency and background level expected at LHCb. We then present in Chapter 6 a fast Monte Carlo model that we have used to simulate a year of data and extract a Δm_s value. Repeating this procedure many times allows to assess the statistical error on the Δm_s measurement.

Chapter 7 addresses one of the systematics which will affect all the proper-time related measurements: the length scale calibration. In this analysis, we reconstruct the secondary interactions occurring in the vertex-detector material and compare the obtained spatial distribution with the detector geometry known from survey measurements.

Chapter 1

CP violation and the beauty system
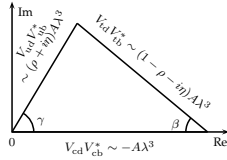
We introduce in this chapter the basic concept of the Standard Model and describe the CKM matrix and the unitary triangles, outlining their role in the framework of CP violation studies. We then describe the mixing occurring in neutral B meson systems.

THE Standard Model (SM) is a quantum field theory which describes the elementary particles and their interactions through the electromagnetic, the weak and the strong forces. Despite its numerous successful predictions, it is not yet a complete theory, for all attempts to describe gravity at the quantum level have failed up to now, and the existence of the predicted Higgs boson, which would generate the particle masses, is still not confirmed experimentally. Moreover the lack of explanation for the observed matter–anti-matter asymmetry or the numerous free parameters that the SM contains and and the actual value of some of these (e.g. leading to the mass hierarchy problem) also denotes an incomplete theory. All these arguments show that the SM has to be extended and will probably become a low energy description of some more fundamental theory.

In this context, the study of CP violation in the weak interaction provides an excellent probe to test the SM. The decay amplitudes of weak processes can be calculated and, in some cases, offer a clean theoretical prediction when the effect of the strong interaction is small or cancels out in the computation. Furthermore the study of the B mesons provides insight in a region of the CKM matrix which is still not known very precisely and where CP violating effects are expected. We use throughout this document the notoation used in [1].

1.1 Standard Model summary

The Standard Model of particle physics was developed in the seventies. It is based on quantum field theory and accommodates both quantum mechanics and special relativity. To date, almost all experimental tests have agreed with its predictions. The SM is extensively discussed in textbooks on particle physics and we shortly recall here its main elements.

Mathematically, the SM is based on an unification of the electroweak theory [2, 3] and quantum chromodynamics (QCD) into a structure denoted by the gauge groups $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$. The latter is spontaneously broken into $SU(3)_C \otimes U(1)_Q$ by the

introduction of a non-gauge boson, the Higgs.

The SM defines different gauge fields. Eight of them ($G^k_{\nu\mu}$, $k = 1 \ldots 8$) belong to the $SU(3)_C$ group, where the subscript C stands for color[1]. They are called gluons and carry a charge which can be either red, green or blue. They are the mediators of the strong interaction. Three boson fields (W^a_ν, $a = 1 \ldots 3$) belong to the $SU(2)_L$ group, where L stands for left handed. Finally, one boson field B_μ belongs to the $U(1)_Y$ group, where the subscript Y stands for hypercharge.

The SM also includes non-gauge fields. There are three generations of down-type quarks, which we note $D' \in \{d', s', b'\}$ and their associated up-type quarks $U' \in \{u', c', t'\}$. The former have an electric charge $Q = -\frac{1}{3}$, the latter have an electric charge $Q = +\frac{2}{3}$. All of them are spin-½ fermions and are eigenstates of the weak interaction (mass eigenstates do not have the $'$). Moreover, each quark is associated with the existence of an anti-quark that has the same mass but opposite charges. The quark fields are grouped in two components, characterized by their helicity, and written

- left-handed doublets of $SU(2)_L$: $\begin{pmatrix} U' \\ D' \end{pmatrix}$,

- right-handed singlets of $SU(2)_L$: $(U')_R$ and $(D')_R$.

There exist also three generations of leptons and their associated anti-leptons. Each generation is composed of a neutrino $\nu_\ell \in \{\nu_e, \nu_\mu, \nu_\tau\}$ which has no electric charge and is assumed to be massless, and a lepton $\ell^- \in \{e^-, \mu^-, \tau^-\}$ which has a minus one charge. They are represented as

- doublets $\begin{pmatrix} \nu_\ell \\ \ell^- \end{pmatrix}$,

- singlets $(\ell^-)_R$,

with no right-handed neutrinos. Still, recent measurements have proved a non-zero mass difference between two of the neutrinos and the SM theory will have to accommodate this result.

The existence of many more particles is known experimentally. Yet they are all bound states of quarks. The two main groups are the mesons, composed of one quark and one anti-quark (e.g. the pions and the kaons, or the B_s which is the bound state of an s quark and a $\bar{b}$ quark) and the baryons, which have three quarks. The most famous are the proton (uud) and the neutron (udd).

Up to now, none of the quantum fields involved is associated with a charge which would account for the mass. This is expected as mass terms are not possible in a framework where gauge symmetry is conserved. Yet the SM assumes the existence of an additional non-gauge, scalar field Φ, called the Higgs field. The latter introduces a potential term in the Lagrangian which happens to be minimized for a non-zero value of the Higgs field. This phenomenon is called the Spontaneous Symmetry Breaking (SSB) of the $SU(2)_L$ group. It has the remarkable feature to introduce couplings with the other fields which can be associated to mass terms: this is called the Higgs mechanism. It gives masses to the gauge fields, that is the three W^a_ν and the B_μ fields. The latter can be associated, after an appropriate change of basis, with the $W^\pm$ bosons, the Z^0 boson and the photon γ. Note that the masses of the Z^0 and the $W^\pm$ bosons are not independent and the prediction of this dependence with a great accuracy is one of the SM successes.

---

[1]Greek letters denote Lorentz indices.

1.1.1 The CKM matrix and the unitary triangles

For the quark masses, the same mechanism can be applied. However, it has been measured that the weak eigenstates are not equal to the mass eigenstates (i.e. lepton couplings have not the same amplitude as quark coupling in the weak interaction). This can be accounted for with a change of basis, which is parametrized with the so-called Cabibbo–Kobayashi–Maskawa (CKM) unitary matrix V_{CKM}. It connects the weak-interaction eigenstates to the mass eigenstates through the relation:

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

where the matrix coefficients are complex numbers.

In perturbation theory, the quantum mechanical transition amplitude for a given process can be written in terms of currents which describe the exchange of a boson. After the SSB, the charged current of a W^- exchange can be written as

$$J^{\mu-} = (\bar{\mathrm{u}}_{\mathrm{L}}, \bar{\mathrm{c}}_{\mathrm{L}}, \bar{\mathrm{t}}_{\mathrm{L}}) \, \gamma^\mu V_{\mathrm{CKM}} \begin{pmatrix} \mathrm{d}_{\mathrm{L}} \\ \mathrm{s}_{\mathrm{L}} \\ \mathrm{b}_{\mathrm{L}} \end{pmatrix} ,$$

where γ^μ are the Dirac matrices. Similarly, the exchange of a W^+ is obtained from the hermitian conjugate. Hence the CKM elements are involved in any weak process calculation.

Because the CKM matrix is unitary and because one is free to re-phase the quark fields, the CKM matrix actually has only 4 independent parameters, 3 Euler angles and one complex phase which is the only possible source of CP in the SM.

A very useful parametrization of the CKM matrix has been developed by Wolfenstein [4] and is an expansion in terms of the small quantity $\sin\theta_{\mathrm{C}} \equiv \lambda$, where θ_{C} is the so-called Cabibbo angle. In this parametrization, the CKM matrix reads:

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

where A is or order 1, and ρ and η are real. The correction up to $\mathcal{O}(\lambda^5)$ reads:

$$\begin{pmatrix} -\frac{1}{8}\lambda^4 & 0 & 0 \\ A^2\lambda^5(\frac{1}{2} - \rho - i\eta) & -\frac{1}{8}\lambda^4(1 + 4A^2) & 0 \\ \frac{1}{2}A\lambda^5(\rho + i\eta) & A\lambda^4(\frac{1}{2} - \rho - i\eta) & -\frac{1}{2}A^2\lambda^4 \end{pmatrix} + \mathcal{O}(\lambda^6) .$$

The unitary of the CKM matrix, $V_{\mathrm{CKM}}^\dagger V_{\mathrm{CKM}} = \mathbb{1}$, yields twelve constraints, among which six are the following orthogonality relations:

$$\text{(ds)} \quad V_{\mathrm{ud}}V_{\mathrm{us}}^* + V_{\mathrm{cd}}V_{\mathrm{cs}}^* + V_{\mathrm{td}}V_{\mathrm{ts}}^* = 0 , \tag{1.1}$$

$$\text{(sb)} \quad V_{\mathrm{us}}V_{\mathrm{ub}}^* + V_{\mathrm{cs}}V_{\mathrm{cb}}^* + V_{\mathrm{ts}}V_{\mathrm{tb}}^* = 0 , \tag{1.2}$$

$$\text{(db)} \quad V_{\mathrm{ud}}V_{\mathrm{ub}}^* + V_{\mathrm{cd}}V_{\mathrm{cb}}^* + V_{\mathrm{td}}V_{\mathrm{tb}}^* = 0 , \tag{1.3}$$

$$\text{(cu)} \quad V_{\mathrm{ud}}V_{\mathrm{cd}}^* + V_{\mathrm{us}}V_{\mathrm{cs}}^* + V_{\mathrm{ub}}V_{\mathrm{cb}}^* = 0 , \tag{1.4}$$

$$\text{(tc)} \quad V_{\mathrm{cd}}V_{\mathrm{td}}^* + V_{\mathrm{cs}}V_{\mathrm{ts}}^* + V_{\mathrm{cb}}V_{\mathrm{td}}^* = 0 , \tag{1.5}$$

$$\text{(tu)} \quad V_{\mathrm{ud}}V_{\mathrm{td}}^* + V_{\mathrm{us}}V_{\mathrm{ts}}^* + V_{\mathrm{ub}}V_{\mathrm{tb}}^* = 0 . \tag{1.6}$$

These can be represented as triangles in the complex plane, all with the same surface. The CKM matrix can also be parametrized using four independent angles. In terms of CKM coefficients, they are:

$$\gamma \quad \equiv \quad \phi_3 \equiv \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right) \; , \tag{1.7}$$

$$\beta \quad \equiv \quad \phi_1 \equiv \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right) \; , \tag{1.8}$$

$$\beta_s \quad \equiv \quad \chi \equiv \arg\left(-\frac{V_{cb}V_{cs}^*}{V_{tb}V_{ts}^*}\right) \; , \tag{1.9}$$

$$\beta_K \quad \equiv \quad \chi' \equiv \arg\left(-\frac{V_{us}V_{ud}^*}{V_{cs}V_{cd}^*}\right) \; . \tag{1.10}$$

Two of them, β and γ are angles of the (db) triangle (Eq. 1.3). One, β_K, is visible in the (ds) triangle (Eq. 1.1). The fourth one belongs to the (sb) triangle (Eq. 1.2). Figure 1.1 sketches the (db), the (ds) and the (sb) triangles in the complex plane.
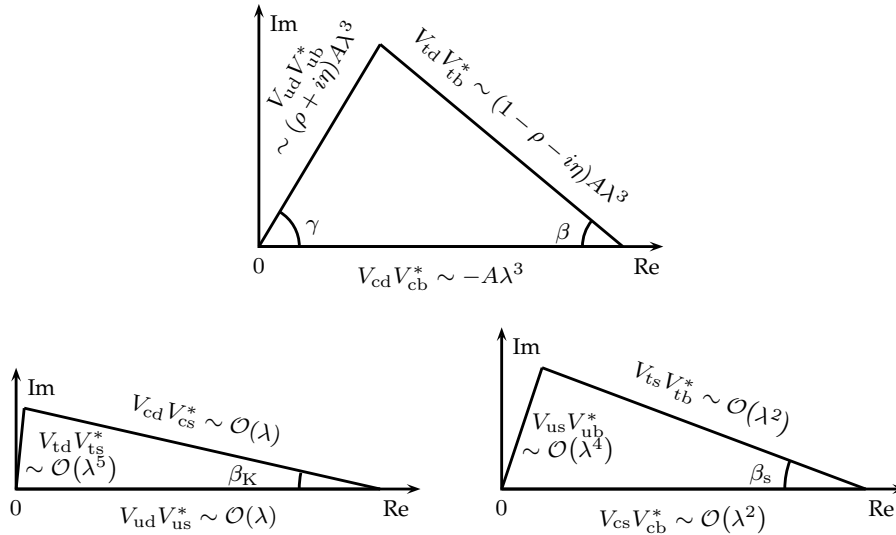
Figure 1.1: Illustration of three of the unitary triangles in the complex plane (not to scale). The (db) triangle on the top, the (ds) triangle on the left and the (sb) triangle on the right. The side lengths are shown in term of V_{CKM} coefficients as well as in term of the Wolfenstein parametrization. We also show the angles defined in Eqs. 1.7–1.10.

Hence measuring the unitary triangle parameters (angles and sides) allows to put constraints on the CKM coefficients. These in turn have only four degrees of freedom with a single value which can create CP violation. Therefore, CP violation measurements, as any CKM related measurements are all tests of the SM validity.

1.2 The Neutral B–$\overline{\text{B}}$ System

There are two systems of neutral B mesons. We denote their flavor eigenstates with

$$|B_q\rangle = |\bar{b}q\rangle \; , \quad |\overline{B}_q\rangle = |b\bar{q}\rangle \; , \tag{1.11}$$

where q is one of the d or s quark, forming the B_d system or the B_s system respectively.

1.2.1 B -meson mixing

Assuming only the existence of the strong and the electromagnetic interactions, the states $|B_q\rangle$ and $|\overline{B}_q\rangle$ are stable and we have $\langle \overline{B}_q|B_q\rangle = 0$. As soon as the weak Lagrangian is switched on, the neutral mesons starts to mix and decay, i.e. a B_q meson can become a $\overline{B}_q$ meson and vice versa before it decays. This mixing, or oscillation, happens through the box diagrams shown in Figure 1.2.
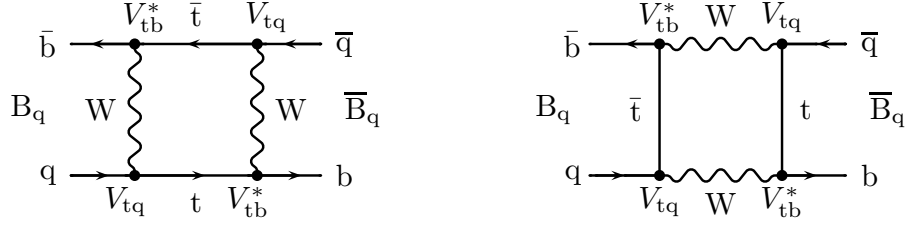
Figure 1.2: Feynman box diagrams for B mixing. The q quark can be either a d or a s quark. The t quarks can be replaced by a c or u quarks, but the dominant contribution comes from the heaviest quark. The charge-conjugated process for $\overline{B}_q \rightarrow B_q$ is obtained by replacing all quarks by anti-quarks (and vice versa) and taking the complex-conjugates of the CKM elements.

The general state of a B-meson system (not considering its possible decays for the moment) is described as a quantum superposition of $|B_q\rangle$ and $|\overline{B}_q\rangle$:

$$a|B_q\rangle + b|\overline{B}_q\rangle \ .$$

Its time evolution is governed by the Schrödinger equation:

$$i\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = H \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = (M - \frac{i}{2}\Gamma) \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} \ , \tag{1.12}$$

where M and Γ are two hermitian matrices. Assuming that CPT is conserved, B_q and $\overline{B}_q$ have equal masses ($M_{11} = M_{22}$) and equal decay widths ($\Gamma_{11} = \Gamma_{22}$). The eigenstates of the above Schrödinger equation are the mass eigenstates, we define them as

$$|B_{H,L}\rangle \equiv p|B_q\rangle \mp q|\overline{B}_q\rangle \ ,$$

where H and L stand for heavy and light respectively. We stress that the signs in front of q is only a convention. Moreover, using the same coefficients p and q for the heavy and light state (i.e., implicitly $q_H = q_L = q$, $p_H = p_L = p$) implies a re-phasing of the kets and is possible only under the assumption that CPT is conserved.

The states $|B_H\rangle$ and $|B_L\rangle$ have well-defined masses and decay widths. Thus the evolution, starting from a state $|B_H\rangle$ at $t = 0$ is governed by

$$|B_H(t)\rangle = \mathrm{e}^{-(im_H+\Gamma_H/2)t}|B_H\rangle \ ,$$

and similarly for the $|B_L\rangle$ state. The mass and decay width differences are defined as

$$\begin{aligned} \Delta m &\equiv m_H - m_L \ , \\ \Delta\Gamma &\equiv \Gamma_H - \Gamma_L \ , \end{aligned}$$

where the sign of Δm is positive by definition and the sign of $\Delta\Gamma$ is determined experimentally. The average mass and decay time are:

$$m = \frac{1}{2}(m_H + m_L)\,,$$

$$\Gamma = \frac{1}{2}(\Gamma_H + \Gamma_L)\,.$$

Solving the Schrödinger equation leads to the expression of q/p in terms of the matrix elements of M and Γ

$$\frac{q}{p} = -\sqrt{\frac{M_{12}^* - i\Gamma_{12}^*/2}{M_{12} - i\Gamma_{12}/2}}\,. \tag{1.13}$$

When produced through the strong interaction, B mesons are created in one of the flavor eigenstates (Eq. 1.11). The time evolution of a B meson initially produced as a $|B_q\rangle$ or a $|\overline{B}_q\rangle$ can be expressed as

$$
\begin{aligned}
|B_{\mathrm{phys}}(t)\rangle &= g_+(t)|B_q\rangle + \tfrac{q}{p}g_-(t)|\overline{B}_q\rangle\,, \\
|\overline{B}_{\mathrm{phys}}(t)\rangle &= g_+(t)|\overline{B}_q\rangle + \tfrac{p}{q}g_-(t)|B_q\rangle\,,
\end{aligned}
\tag{1.14}
$$

where the $g_\pm$ reads:

$$g_\pm(t) = \frac{1}{2}\left(\mathrm{e}^{-(im_L+\Gamma_L/2)t} \pm \mathrm{e}^{-(im_H+\Gamma_H/2)t}\right)\,.$$

Thus, the probability to observe a B_q meson or a $\overline{B}_q$ meson that is originally produced as either $|B_q\rangle$ or $|\overline{B}_q\rangle$ are

$$
\begin{aligned}
|\langle B_q|B_{\mathrm{phys}}(t)\rangle|^2 &= |g_+(t)|^2\,, \\
|\langle \overline{B}_q|B_{\mathrm{phys}}(t)\rangle|^2 &= \left|\frac{q}{p}\right|^2 |g_-(t)|^2\,, \\
|\langle B_q|\overline{B}_{\mathrm{phys}}(t)\rangle|^2 &= \left|\frac{p}{q}\right|^2 |g_-(t)|^2\,, \\
|\langle \overline{B}_q|\overline{B}_{\mathrm{phys}}(t)\rangle|^2 &= |g_+(t)|^2\,,
\end{aligned}
\tag{1.15}
$$

with

$$|g_\pm(t)|^2 = \frac{1}{2}\mathrm{e}^{-\Gamma t}\left\{\cosh\frac{\Delta\Gamma}{2}t \pm \cos\Delta mt\right\}\,.$$

We plot in Fig. 1.3, the probability for a meson produced as B_q to decay as B_q or $\overline{B}_q$ (Eq. 1.15) for both the B_d and B_s systems.

Looking at Eqs 1.15, we see that if $|q/p| \neq 1$, the processes $B \to \overline{B}$ and $\overline{B} \to B$ are asymmetric: this is called CP violation in the mixing. In the limit $|\Gamma_{12}| \ll |M_{12}|$, Eq. 1.13 becomes $q/p = -\sqrt{M_{12}^*/M_{12}}$ and can be written in terms of CKM matrix elements:

$$\left(\frac{q}{p}\right)_{B_d} = \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} = \mathrm{e}^{-2i\beta}\,, \qquad \left(\frac{q}{p}\right)_{B_s} = \frac{V_{tb}^* V_{ts}}{V_{tb} V_{ts}^*} = \mathrm{e}^{-2i\beta_s}\,,$$

where β and β_s are the unitarity triangles defined in Eqs. 1.7 to 1.10.
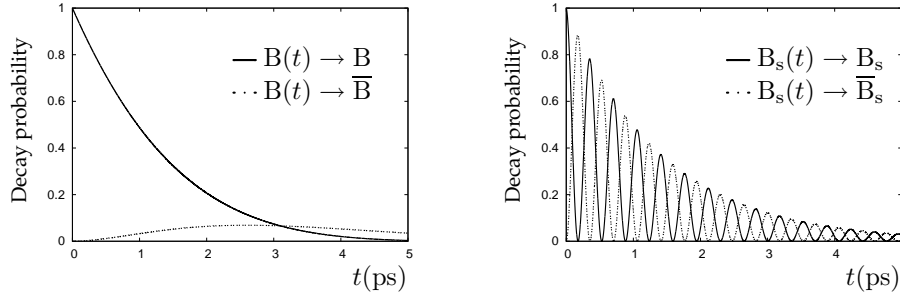
Figure 1.3: Probability for a bottom meson produced as B to decay as B or $\overline{\text{B}}$ for the B_d system (left) and the B_s system (right). In this example, $\Delta m_\text{d} = 0.507 \text{ ps}^{-1}$, $\Delta\Gamma_\text{d}/\Gamma_\text{d} = 0$, $\Delta m_\text{s} = 17.77 \text{ ps}^{-1}$ and $\Delta\Gamma_\text{s}/\Gamma_\text{s} = 0.1$.

1.2.2   Neutral B-meson decays

Considering now that the B mesons will decay into various channels, we define the decay amplitudes of a B_q or a $\overline{\text{B}}_\text{q}$ to a final state f as:

$$A_f = \langle f|T|\text{B}_\text{q}\rangle , \qquad \overline{A}_f = \langle f|T|\overline{\text{B}}_\text{q}\rangle , \tag{1.16}$$

where T is the transition matrix. Similarly we define the amplitudes $A_{\overline{f}}$ and $\overline{A}_{\overline{f}}$, which are the transition amplitudes to the charge conjugate state $\overline{f}$. Clearly, if $|A_f| \neq |\overline{A}_{\overline{f}}|$ we have CP violation. However, this kind of CP symmetry-breaking can occur only when at least two phases contribute to the same final state. This can be seen by defining the following parametrization for A_f and $\overline{A}_{\overline{f}}$:

$$A_f = \sum_k r_k \text{e}^{i\delta_k} \text{e}^{i\phi_k} , \qquad \overline{A}_f = \sum_k r_k \text{e}^{i\delta_k} \text{e}^{-i\phi_k} ,$$

where r_k are real numbers, equal for the B_q and the $\overline{\text{B}}_\text{q}$, δ_k are strong phases, which are invariant under CP, and ϕ_k are weak phases, which change sign under the CP transformation. If only one term is present in the sum, then $|A_f| = |\overline{A}_{\overline{f}}|$ and CP is conserved. If there is a second term in the sum, then the two may interfere and result in $|A_f| \neq |\overline{A}_{\overline{f}}|$ i.e. CP violation. This kind of CP violation is called CP violation in the decay amplitudes.

Still, if only one decay amplitude is possible, there is actually another phase which is present and may produce an interference leading to the so-called mixing-induced CP violation. Indeed, the equations show that one is sensitive to the relative phase between q/p and $\overline{A}_f/A_f$, it is usually written with the parameter λ_f

$$\lambda_f \equiv \frac{q}{p}\frac{\overline{A}_f}{A_f} , \quad \lambda_{\overline{f}} \equiv \frac{q}{p}\frac{\overline{A}_{\overline{f}}}{A_{\overline{f}}} , \quad \overline{\lambda}_f \equiv \frac{1}{\lambda_f} , \quad \overline{\lambda}_{\overline{f}} \equiv \frac{1}{\lambda_{\overline{f}}} , \tag{1.17}$$

and the conservation of CP implies $\arg(\lambda_f) + \arg(\lambda_{\overline{f}}) = 0$.

Thus it is possible to compute the decay rate of a neutral B meson to a final state f:

$$\Gamma_{\text{B}\to f}(t) = |\langle f|T|\text{B}_\text{phys}(t)\rangle|^2 ,$$

with equivalent expressions for $\overline{\text{B}}$ and $\overline{f}$. Using Equations 1.14, 1.16 and 1.17, the decay rates eventually become

$$\Gamma_{B\to f}(t) = |A_f|^2(1+|\lambda_f|^2)\frac{e^{-\Gamma t}}{2}\left\{\cosh\frac{\Delta\Gamma}{2}t+\right.$$
$$\left. D_f\sinh\frac{\Delta\Gamma}{2}t + C_f\cos\Delta mt - S_f\sin\Delta mt\right\}, \tag{1.18}$$

$$\Gamma_{\overline{B}\to f}(t) = |A_f|^2\left|\frac{p}{q}\right|^2(1+|\lambda_f|^2)\frac{e^{-\Gamma t}}{2}\left\{\cosh\frac{\Delta\Gamma}{2}t+\right.$$
$$\left. D_f\sinh\frac{\Delta\Gamma}{2}t - C_f\cos\Delta mt + S_f\sin\Delta mt\right\}, \tag{1.19}$$

$$\Gamma_{\overline{B}\to\overline{f}}(t) = |\overline{A}_{\overline{f}}|^2\left|\frac{p}{q}\right|^2(1+|\overline{\lambda}_{\overline{f}}|^2)\frac{e^{-\Gamma t}}{2}\left\{\cosh\frac{\Delta\Gamma}{2}t+\right.$$
$$\left. D_{\overline{f}}\sinh\frac{\Delta\Gamma}{2}t + C_{\overline{f}}\cos\Delta mt - S_{\overline{f}}\sin\Delta mt\right\}, \tag{1.20}$$

$$\Gamma_{B\to\overline{f}}(t) = |\overline{A}_{\overline{f}}|^2\left|\frac{p}{q}\right|^2(1+|\overline{\lambda}_{\overline{f}}|^2)\frac{e^{-\Gamma t}}{2}\left\{\cosh\frac{\Delta\Gamma}{2}t+\right.$$
$$\left. D_{\overline{f}}\sinh\frac{\Delta\Gamma}{2}t - C_{\overline{f}}\cos\Delta mt + S_{\overline{f}}\sin\Delta mt\right\}, \tag{1.21}$$

where

$$D_f = \frac{2\mathrm{Re}\lambda_f}{1+|\lambda_f|^2}, \quad C_f = \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2}, \quad S_f = \frac{2\mathrm{Im}\lambda_f}{1+|\lambda|^2},$$
$$D_{\overline{f}} = \frac{2\mathrm{Re}\overline{\lambda}_{\overline{f}}}{1+|\overline{\lambda}_{\overline{f}}|^2}, \quad C_{\overline{f}} = \frac{1-|\overline{\lambda}_{\overline{f}}|^2}{1+|\overline{\lambda}_{\overline{f}}|^2}, \quad S_{\overline{f}} = \frac{2\mathrm{Im}\overline{\lambda}_{\overline{f}}}{1+|\overline{\lambda}_{\overline{f}}|^2}. \tag{1.22}$$

The factors D_f ($D_{\overline{f}}$) and S_f ($S_{\overline{f}}$) are sensitive to the phase of λ_f ($\overline{\lambda}_{\overline{f}}$), and thus to CP violation.

1.2.3 Accessible observables with $B_s \to D_s^-\pi^+$ and $B_s \to D_s^\mp K^\pm$ decays

From an experimental point of view, the observation of B–$\overline{B}$ oscillations is easiest with the knowledge of the flavor of the B both at production and at decay time. This happens to be possible with $B_s \to D_s^-\pi^+$ decays, because this decay chain is flavor specific, which means that the B meson flavor at its decay time unambiguously determines the charges of the final-state particles. This can be seen in the quark diagrams of $B_s \to D_s^-\pi^+$ and $\overline{B}_s \to D_s^+\pi^-$ shown in Fig. 1.4. With $B_s \to D_s^-\pi^+$ decays, only one decay amplitude is present and no interference between the mixing and the decay can occur. Thus the time evolution of $B_s \to D_s^-\pi^+$ decays is governed by Eqs. 1.15. Furthermore, the term $|p/q|$, which quantifies CP violation in the mixing, is expected to be very close to 1. The determination of the B flavor at creation time is performed with a technique called tagging (see Sec. 1.2.1); the outcome of the flavor tagging is sometimes wrong and this dilutes the observed asymmetry, such that Eqs. 1.15 can be written as:

$$\Gamma_{\mathrm{unmixed}}(t) = \frac{e^{-\Gamma_s t}}{2}\left\{\cosh\frac{\Delta\Gamma_s}{2}t + D\cos\Delta m_s t\right\}, \tag{1.23}$$

$$\Gamma_{\mathrm{mixed}}(t) = \frac{e^{-\Gamma_s t}}{2}\left\{\cosh\frac{\Delta\Gamma_s}{2}t - D\cos\Delta m_s t\right\}, \tag{1.24}$$

where 'mixed' refers to $B_s \to D_s^+\pi^-$ and $\overline{B}_s \to D_s^-\pi^+$ decays together and 'unmixed' refers to $\overline{B}_s \to D_s^+\pi^-$ and $B_s \to D_s^-\pi^+$ decays together. The amplitude of the oscillations
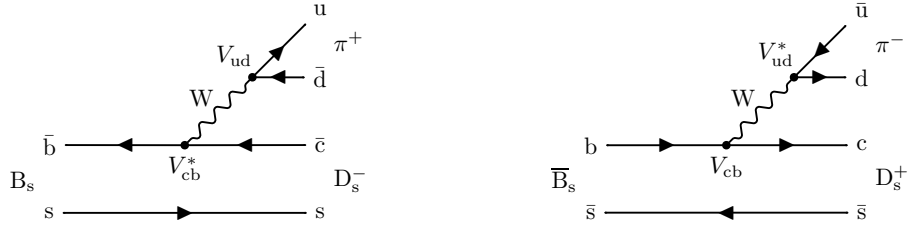
Figure 1.4: Feynman diagrams for the decays $B_s \to D_s^- \pi^+$ (left) and $\overline{B}_s \to D_s^+ \pi^-$ (right).

is reduced through the dilution factor D. This effect is due to the imperfect tagging; the dilution amplitude being $D = 1 - 2\omega$, where ω is the probability for the outcome of the flavor tagging to be wrong.

In addition to giving access to the Δm_s measurement, we see that tagged $B_s \to D_s^- \pi^+$ decays offer sensitivity to the dilution amplitude D and also to Γ_s and $\Delta\Gamma_s$.

The $B_s \to D_s^\mp K^\pm$ Cabibbo-suppressed decays are non flavour-specific, as can be seen from the Feynman diagrams of Fig. 1.5. Contrary to the case where only a B_s can decay to the $D_s^- \pi^+$ state, both a B_s and a $\overline{B}_s$ can decay to the $D_s^- K^+$ final state. The four decay rates from Eqs. 1.18 and 1.21 can be used to extract λ_f and $\overline{\lambda}_{\overline{f}}$:

$$
\begin{aligned}
\lambda_{D_s^- K^+} &= \frac{q}{p} \frac{\overline{A}_{D_s^- K^+}}{A_{D_s^- K^+}} = \left(\frac{V_{tb}^* V_{ts}}{V_{tb} V_{ts}^*} \right) \left(\frac{V_{ub} V_{cs}^*}{V_{cb}^* V_{us}} \right) \left| \frac{A_2}{A_1} \right| e^{i\Delta_s} \\
&= |\lambda_{D_s^- K^+}| e^{i(\delta_{DK} - (\gamma - 2\beta_s))} , \qquad\qquad (1.25) \\
\overline{\lambda}_{D_s^+ K^-} &= \frac{p}{q} \frac{A_{D_s^+ K^-}}{\overline{A}_{D_s^+ K^-}} = \left(\frac{V_{tb} V_{ts}^*}{V_{tb}^* V_{ts}} \right) \left(\frac{V_{ub}^* V_{cs}}{V_{cb} V_{us}^*} \right) \left| \frac{A_2}{A_1} \right| e^{i\Delta_s} \\
&= |\lambda_{D_s^- K^+}| e^{i(\delta_{DK} + (\gamma - 2\beta_s))} , \qquad\qquad (1.26)
\end{aligned}
$$

where γ and β_s are the angles defined in Eqs. 1.7–1.10, $|A_2/A_1|$ is the ratio of the hadronic amplitudes and δ_{DK} is the strong phase difference between the top-left and bottom-right diagram of Fig 1.5.

By measuring the phase of $\lambda_{D_s^- K^+}$ and $\overline{\lambda}_{D_s^+ K^-}$, both the strong phase δ_{DK}, and the weak phase $\gamma - 2\beta_s$ can be extracted using the relations:

$$
\begin{aligned}
\gamma - 2\beta_s &= \frac{1}{2} \left[\arg(\overline{\lambda}_{\overline{f}}) - \arg(\lambda_f) \right] , \\
\delta_{DK} &= \frac{1}{2} \left[\arg(\overline{\lambda}_{\overline{f}}) + \arg(\lambda_f) \right] .
\end{aligned}
$$

Hence this decay channel provides an important method to probe the CKM angle γ as first pointed out in Ref. [5] and recently studied in Ref. [6].
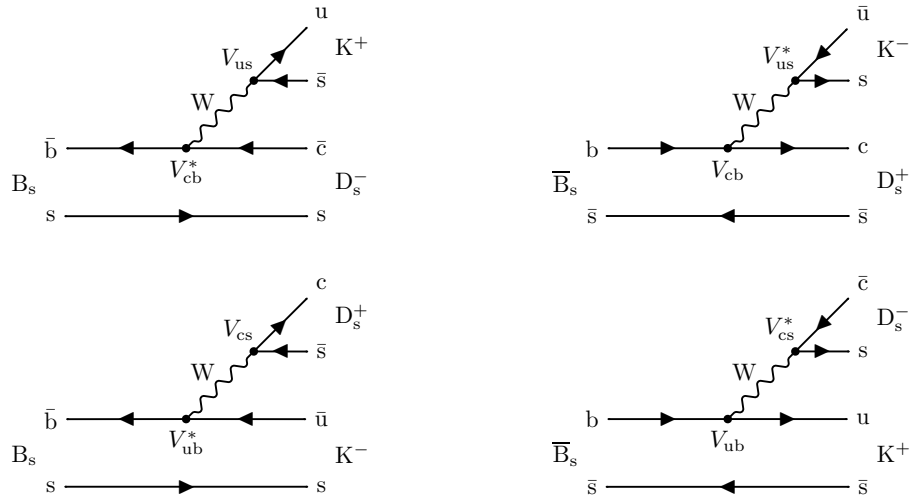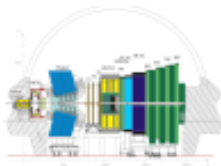
Figure 1.5: Feynman diagrams for the B_s (left) and for the $\overline{B}_s$ (right) decays into the $D_s^{\pm}K^{\mp}$ final states.

Chapter 2

The LHCb detector

We introduce the Large Hadron Collider beauty experiment, describing its main elements and their performances. In particular, we focus on the vertex detector and the way its data are transmitted through the acquisition chain.

T HE European Organization for Nuclear Research (CERN) [7] was founded in 1954 after Louis de Broglie's suggestion to have an international science laboratory. It quickly became one of the leading laboratories for particle physics and its growing success motivated physicists to build larger installations exploring higher energies and therefore going deeper into the structure of matter. Some of its key achievements are the discovery of the Weak Neutral Current in the seventies with the Gargamelle bubble chamber; the famous Super Proton Synchrotron (SPS) opened in 1976, which leads in 1983 to the discovery of the Z and W bosons, unifying the weak and the electromagnetic forces and the Large Electron Position Collider (LEP) which, among numbers of other measurements, determined the number of lepton families.

The next milestone will be the inauguration of the Large Hadron Collider (LHC) [8], in 2008, with its four associated particle detectors, called ATLAS, CMS, LHCb and ALICE.

2.1 The Large Hadron Collider

LHC's first goal is to test the Standard Model (SM) validity. One of its predictions, the existence of the yet unobserved Higgs boson, has driven the design of the ATLAS and CMS experiments. LHCb is primarily meant for precise CP-violation measurements in the b-physics sector. ALICE will be used during dedicated heavy-ion runs (e.g. Pb-Pb, Ca-Ca) to study the behavior of nuclear matter in extreme conditions and the formation of quark-gluon plasma. Yet as the SM is known to be an incomplete theory, the LHC experiments also aims to reveal what is called New Physics, e.g. SUSY or Technicolor.

The LHC will provide proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 14$ TeV. It is installed in the original LEP tunnel, which has a circumference of about 27 km. The Super Proton Synchrotron, a 7 km circular accelerator, will be used as injector for the LHC. Figure 2.1 shows the positions of the LHC main detectors.

The LHC accelerates protons, and to ensure the highest possible luminosity, it collides two proton beams running in opposite directions. This is more complicated from the
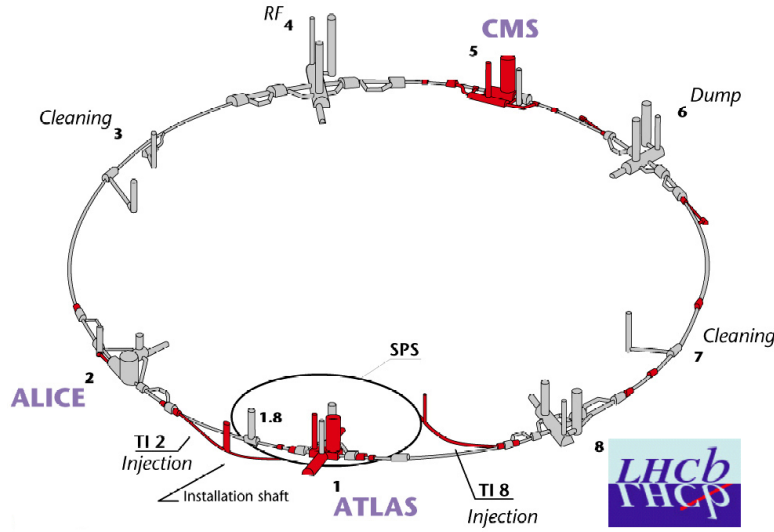
Figure 2.1: The LHC complex with its four main experiments.

engineering point of view (especially for the magnets), but avoids the use of anti-protons which are difficult to produce in large amounts. During nominal operation, each beam will contain 2808 bunches of 10^{11} protons each. The time between two consecutive bunch crossings is 25 ns. This fixes all the hardware clocks of the machine and detectors to a frequency of 40 MHz. However, due to the complicated filling pattern, some of the RF buckets are empty. Moreover, the nominal interaction point of some of the detectors (in particular LHCb) could not be perfectly laid at a crossing point. Thus, the actual average bunch crossing frequency depends on each detector (e.g. for LHCb, a frequency of ~ 30 MHz is foreseen).

The luminosity is a value to characterize the collision rate. It depends on the transverse size of the beams at the interaction points, the number of particles in a bunch, and the bunch crossing frequency. The LHC design luminosity obtained with focusing quadrupole magnets is 10^{34} cm^{-2}s^{-1}, however such a value will most probably require some training before it can be reached, and the luminosity at startup will be much lower.

In this document, we assume, based on a center-of-mass energy of 14 TeV, a total pp cross section of $\sigma_{\mathrm{pp}} = 100$ mb, with an inelastic cross-section of $\sigma_{inel} = 80$ mb [9] and a b$\bar{\mathrm{b}}$ cross section[1] of $\sigma_{\mathrm{b\bar{b}}} = 500$ μb. The number of inelastic pp collisions per bunch crossing follows a Poisson distribution, with a mean value depending on the luminosity and the pp cross section. At the design luminosity, the average number of collisions per bunch crossing is 20. This represents a tremendous track multiplicity to cope with. For this reason, and to protect from radiation damage its hardware which stands very close to the interaction region, LHCb will run at an average luminosity of 2×10^{32} cm^{-2}s^{-1}, which will lower the average number of collisions per bunch crossing to around 0.5. Still, studies are currently done to see how LHCb could cope with higher luminosities in order to increase its available statistics.

---

[1]The b$\bar{\mathrm{b}}$ cross-section at an energy of 14 TeV, has large uncertainties (see Ref. [10]).

2.2   The Large Hadron Collider beauty experiment

The Large Hadron Collider beauty experiment (LHCb) [11, 12] is a single arm spectrometer (see Fig. 2.2). Indeed, to avoid significant civil engineering in the cavern, the detector
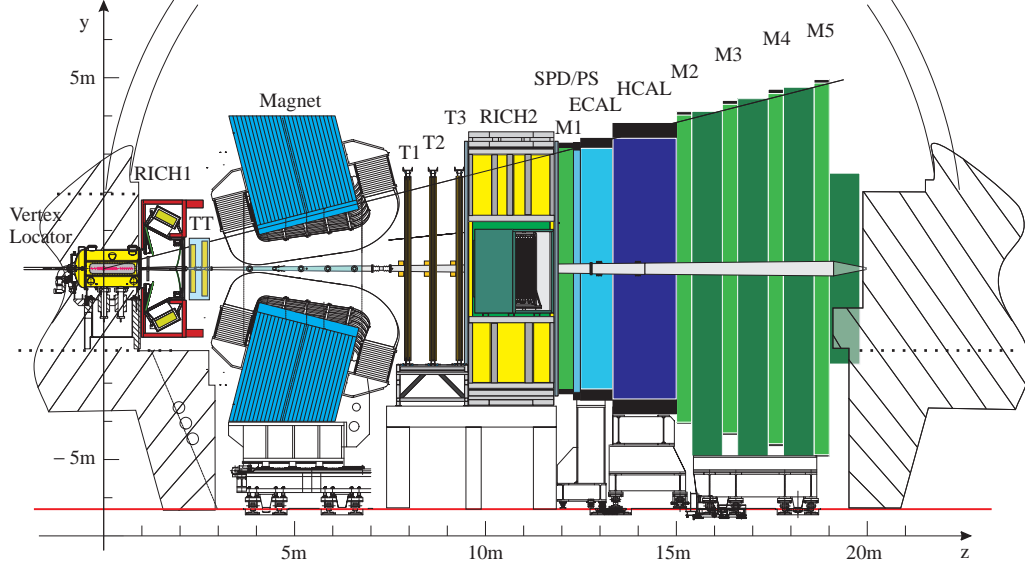
Figure 2.2: Side view of the LHCb detector.

had to fit in the present cavern and this constrained its total length and precluded a second arm. This particular layout is suitable because of the combined effect of the boost, due to the high beam energy, and the quark fragmentation, which tends to produce all the outgoing particles in a rather small cone.

The LHCb coordinate system is centered at the interaction points, which is in the middle of the Vertex Locator. The z axis is parallel to the beam and points downstream, i.e. in the direction of the detector, also called the forward direction. The y axis is vertical and points to the top, the x axis is horizontal, it points outside the LHC ring and defines what is called the left side ($x > 0$) and the right side ($x < 0$) of the detector[2].

The LHCb detector is essentially composed of a tracking system, made of the VeLo, the TT stations, the magnet, the T stations and finally the muon stations. The other components are the charged hadron identification system, consisting of the two Ring Imaging Cherenkov detectors (RICH) and the calorimeter system (pre-shower, electronic and hadronic calorimeters). We summarize below the sub-detector main characteristics. The Vertex Locator is more extensively depicted as its information will be used in Chapter 7.

2.2.1   Vertex Locator

The Vertex Locator (VeLo) [13] aims at providing precise measurements of the charged tracks close to the interaction region. They are used to reconstruct the primary vertices (PV) as well as the displaced secondary vertices, the latter being a characteristic of b-hadron decays. The VeLo being the only part of the LHCb detector which surrounds the interaction point, it provides some information about the backward side of the event, which helps disentangling multiple PVs.

---

[2]The left and right denomination are mostly used for the Vertex Locator, which, as we will see, is split in a left and a right part.

The VeLo is built out of a succession of 300 μm thick silicon half disks arranged along the beam line (Fig. 2.3). Each half-disk provides a measurement of either the r or the ϕ
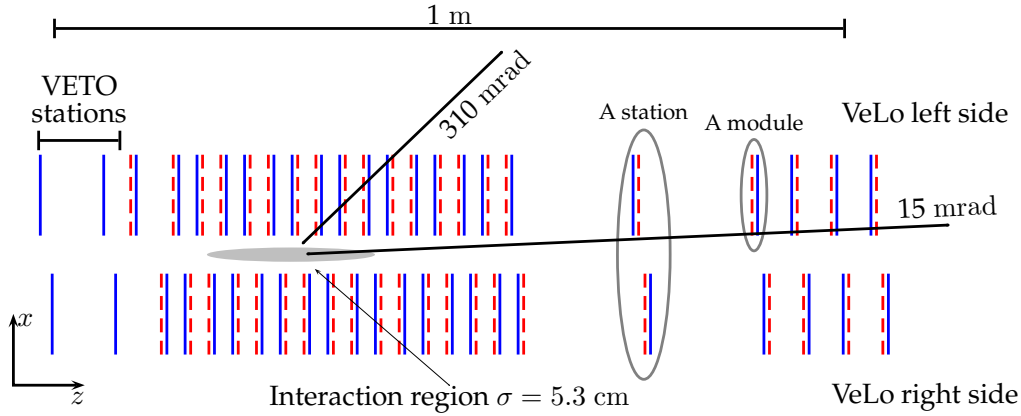
Figure 2.3: Top view of the VeLo sensors region adapted from Ref. [13]. The horizontal (xz plan) acceptance and the interaction region are also shown. On this sketch, the z scale is three times smaller than the x scale, and the space between each sensor pair is enlarged. Figure adapted from Ref. [13].

coordinate and is called r sensor or ϕ sensor. A *module* is made of an r and a ϕ sensor held together –with a gap of about 2 mm– by a carbon fiber structure. Two modules at about the same z position on each side of the VeLo (left and right) define a *station*, which can be seen as a plane able to measure a 2D point, the third dimension being known from the plane position. The VeLo has 21 stations and four additional r sensors in the backward region called the VETO stations and used by the hardware trigger. The VeLo will allow a resolution on primary vertices of 40 μm in z and 8 μm in x and y, and for secondary vertices between 150 and 300 μm in z depending on the number of tracks.

Figure 2.4 shows a front view of both sensor types in the closed detector position (see below). Both have 2048 strips (n-on-n silicon junctions). The routing is done via a double metal layer on top of the silicon. The strip topology and readout order differ for the two types of sensors. As these are important for the cross-talk issues discussed in Chapter 3, a detailed description is given here.

r sensors

An r sensor is divided along the ϕ dimension in four sectors of $45°$ each. Each sector has 512 strips. The distance between the strips (the pitch) varies from 40 μm at the innermost radius to 100 μm at the outermost radius. The strips are numbered from the innermost to the outermost, with the sector increasing anti-clockwise. Due to mechanical constraints, the routing lines have a readout order which starts at strip 127 down to 0 and then from 128 up to 512. This pattern is reversed in adjacent sector. A full description of the sensor dimensions and its readout order is given in Ref. [14].

ϕ sensors

A ϕ sensor boundary along the r dimension –at a radius of 17.2 mm– splits the 683 inner from the 1365 outer strips (2048 in total). The strips have a stereo angle to reduce pattern
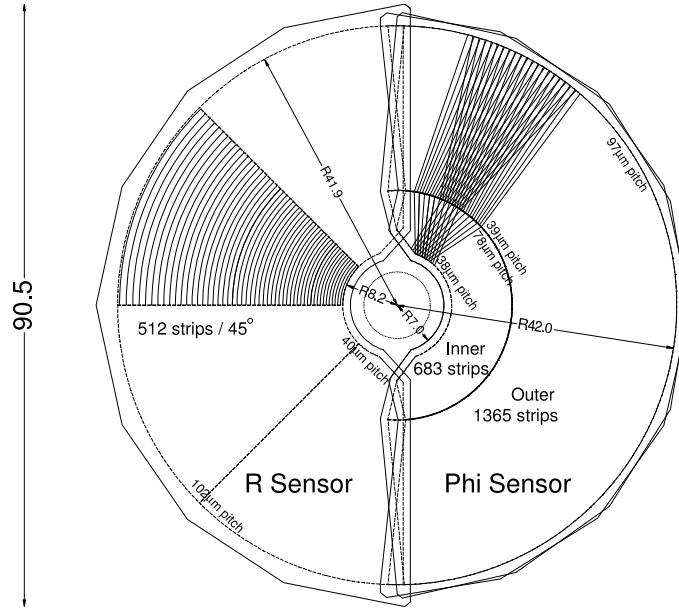
Figure 2.4: Front view of an r and a ϕ sensor in the closed detector position. The different strip regions are shown. For clarity, the strips are drawn only in one area. The sensitive region is also shown, the diameter of the sensor is 90 mm.

recognition ambiguities. They are straight lines and their pitch varies between 35 μm and 96 μm.

The readout topology of a ϕ sensor, sketched in Fig. 2.5, is complex. The silicon strips have a numbering scheme which goes counter-clockwise, whereas the acquisition chip channels run clockwise. Then there are mechanical constraints: the routing lines have to accommodate the fact that there are twice as many outer strips than inner strips. This is achieved by picking one inner then two outer strips and repeating this scheme over the whole sensor. This creates a modulo-three pattern in the strip readout. The sensor pads are laid around the sensor on two rows. Associating them to the pitch adaptors creates a modulo-six pattern which is called here the *strip scrambling*. Eventually, the pitch adaptors re-arrange the strips to allow the bonding to the readout chips on four rows. This again creates a modulo-four pattern.

Altogether, the channel order in the acquisition chips is scrambled with respect to the strip ordering through a modulo-12 pattern (the smallest common multiple of 3, 4 and 6). A full description of the ϕ sensor dimensions and readout scheme is given in Ref. [15].

Mechanical aspects

The VeLo is split in two halves along the yz plane. They can be retracted such that the aperture between them becomes 6 cm, allowing the LHC beam injection. While the detector is in the fully closed position, the two halves overlap and the detection region starts at a radius of 8 mm. This mechanical feature requires that no physical pipe is present. Therefore, the vacuum is ensured by the VeLo vacuum vessel [16, 17] which surrounds the detector (Fig. 2.6).

We show in Fig. 2.7 a representation of one module with its mechanical structure. The hybrid is the part actually holding the r and the ϕ silicon-sensors. The hybrid itself is

Figure 2.5: ϕ-sensor readout scheme. All the permutations occurring in the readout are shown. The same global pattern is repeated for each group of 12 channels. As the strip numbering goes counter-clockwise and the chip channel numbering clockwise, the first chip channels corresponds to the last strips of both the inner and outer rows.
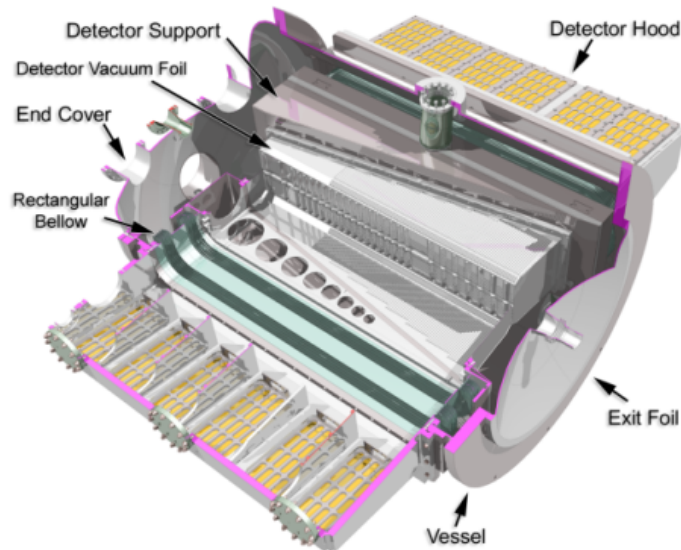
Figure 2.6: Open view of the VeLo tank.

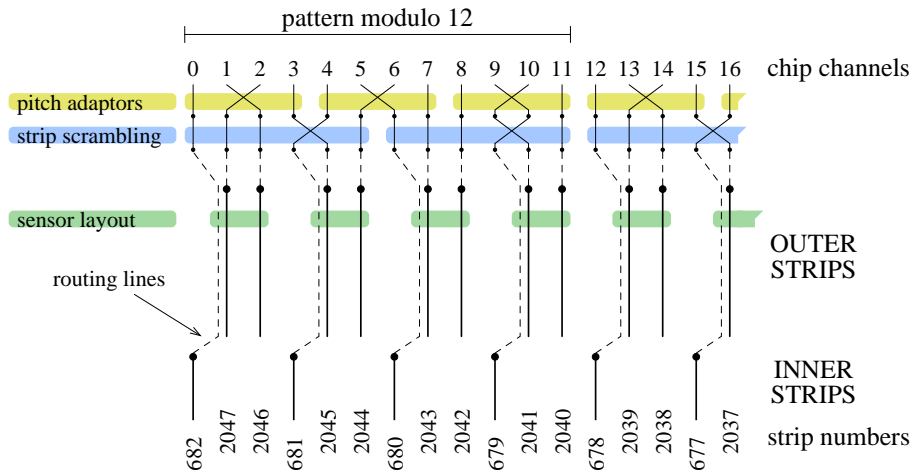fixed to the module support, or paddle, which holds the cables and the cooling tube for this module. All the paddles are screwed on the module bases (one on each VeLo side, not shown in Fig. 2.7), which are 1 m long aluminum structures holding the cooling system, the modules and the RF foils (see below). Because the metrology measurements indicated that the modules were mechanically too unstable, a carbon-fiber (CF) structure, running on both hybrid sides, was built to avoid relative movements between the modules.
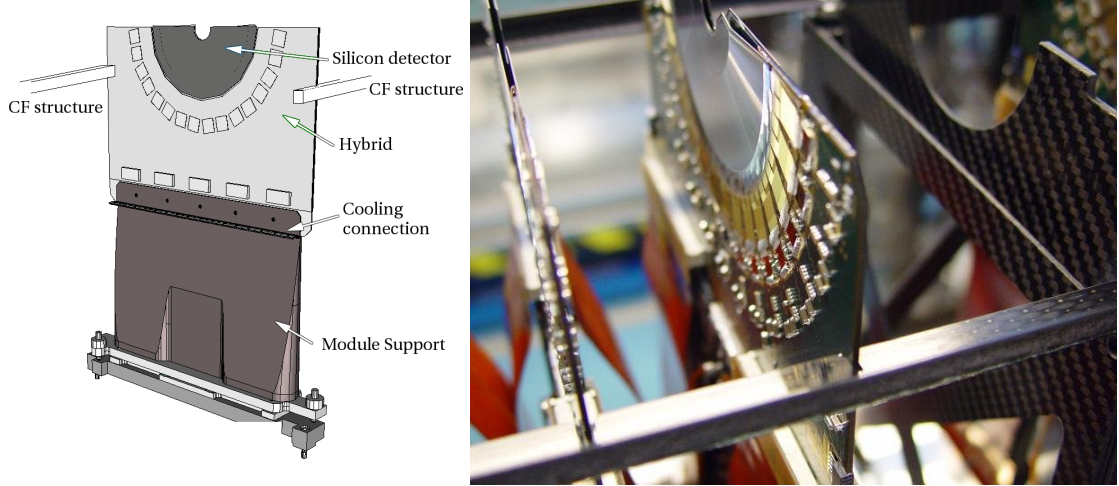
Figure 2.7: Left: one VeLo module with its mechanical structure. Right: photograph of a module, showing the silicon sensors, the acquisition chips and the carbon-fiber structure on the side of the hybrid.

The silicon disks and the front-end electronics are protected from the beam radio-frequency pickup by a 300 μm thick aluminum (AlMg$_3$) box called the RF foil. To allow the overlapping of the two sensitive regions, the foil has a complex shape in the yz plane. Figure 2.8 shows a close-up of some of the VeLo sensors with the vertical part of the RF foil.

2.2.2 The Dipole magnet

The dipole magnet [18] consists of two trapezoidal coils bent at $45°$ on the two transverse sides, arranged inside an iron yoke. It is located close to the interaction point to keep it relatively small, but after the VeLo, RICH1 and TT detectors. The magnetic field is vertical and bends the track in the horizontal xz plane. Its maximal intensity is 1.1 T and its average integrated intensity is 4 Tm. The magnet is made of aluminum conducting wires (9 km in total). It is warm rather than super-conducting, mainly for cost reason. To compensate for a possible detector left-right asymmetry, one can flip the magnetic field polarity.

2.2.3 Trigger Tracker and Tracking stations

The VeLo, the Trigger Tracker (TT) and the T stations (T1,T2,T3) are the main components of the tracking system.

The Trigger Tracker is located downstream of the first RICH detector and just in front of the magnet. It consists of two stations separated by a distance of 27 cm. Each station consists of two layers of silicon detectors, for a total of four layers which have a small stereo angle between them to resolve tracking ambiguities (see Fig. 2.9). The Trigger
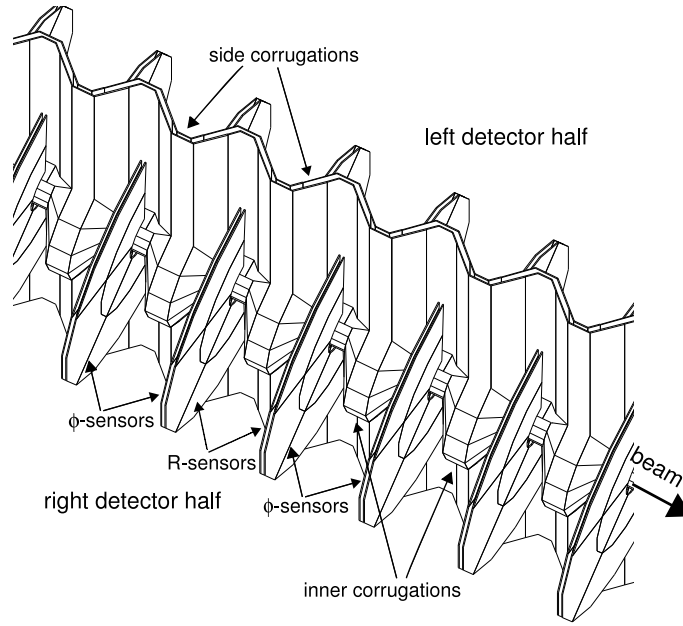
Figure 2.8: The silicon sensors (without the hybrids) and the RF foil, in the closed position [13].

Tracker main function is to provide momentum information to the trigger system; it is therefore mostly used for the so-called upstream tracks and to get a first estimate of the momentum, as it captures the fringe field of the magnet.

The T stations, located after the magnet, cover an area of about 6×5 m^2. In order to cope with a large particle flux gradient, the innermost part, called the Inner Tracker (IT) [19] is based on silicon-strip technology, while the outer part, called the Outer Tracker (OT) [20] is equipped with straw tubes (see text below). The IT covers approximately 2% of the surface but detects about 20% of the track passing through the T stations. An IT station consists of four boxes with four layer silicon sensors, placed around the beam pipe in a cross shape. It spans about 125 cm in width and 40 cm in height (Fig. 2.9).

The Outer Tracker stations are made of drift cells called straw tubes. These have a 5 mm diameter and 75 μm thick walls. The gas filling the cells is an Ar/CF$_4$/CO$_2$ mixture, which has a drift and collection time of less than 50 ns. The spatial resolution obtained in a test beam experiment with this gas is 200 μm.

2.2.4   Particle identification system

The particle identification relies on RICH detectors and the calorimeters.

The Ring Imaging Cherenkov system (RICH) [21] comprises: RICH1, located between the VeLo and the TT, and RICH2 located between the last T stations and the calorimeters. This combined system achieves a K–π separation of 3σ in the momentum range 3–90 GeV. This separation is required for example to allow a distinction of the B$_s \rightarrow$ D$_s^{\mp}$K$^{\pm}$ decays from the B$_s \rightarrow$ D$_s^{-}\pi^{+}$ decays.

The calorimeter system [22] identifies photons, electrons and hadrons and measures their position and energy. Its is made of three elements.

- The preshower consists of 12 mm of lead followed by 15 mm of scintillators and allows the separation of photons and electrons by the different shapes of the electromagnetic showers induced in the Electromagnetic Calorimeter.
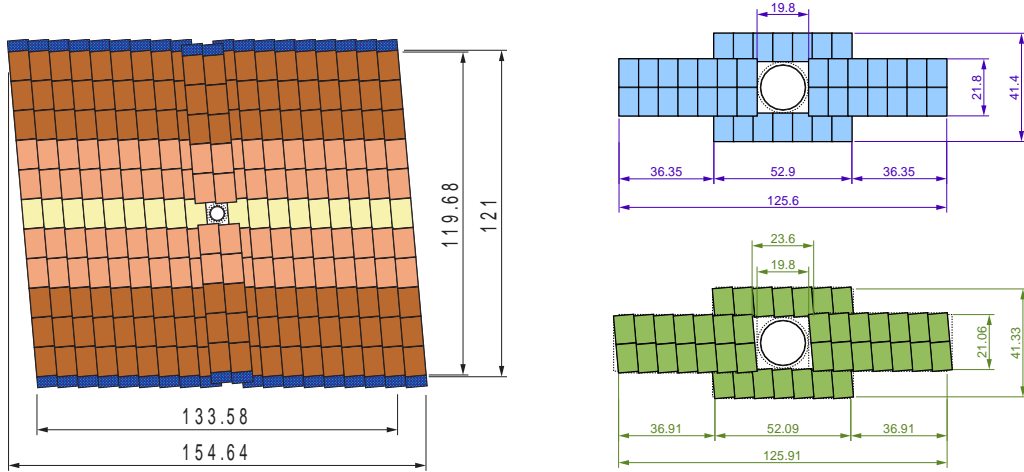
Figure 2.9: Left: front view of one of the stereo Trigger Tracker stations [11], each small square is a silicon sensor. Right: two stations of the Inner Tracker [19]. The dimensions are in centimeters.

- The Electromagnetic Calorimeter (ECAL) uses lead as absorber material. It is segmented in three resolution zones in order to optimize the π^0 reconstruction.

- The Hadronic Calorimeter (HCAL) identifies hadrons via inelastic interactions in a dense material. The charged products of these interactions, mainly pions, are detected in a scintillating medium.

  The HCAL is made of a succession of 16 mm iron and 4 mm thick scintillating tiles which are parallel to the beam. Ionizing particles crossing the scintillator produce light in the UV range which is converted to visible light by scintillating dopants. The light is collected at the end of the tile by wavelength shifting fibers.

2.2.5   Muon system

Muons are the only charged particles which go through the calorimeters. They are thus identified by a dedicated muon system. There are five muon stations, the first one located just in front of the calorimeters, the other four stations end the detector and are separated by an 80 cm thick iron plate which acts as filter for the hadronic background. The muon stations are for the greatest part equipped with multi-wire proportional chambers (MWPC) which allow a full collection of the signal within 20 ns.

The muon system plays an important role in the trigger, hence its fast readout requirement, and is used in the offline tracking to confirm the muon hypothesis.

2.2.6   VeLo data acquisition system

The VeLo Data Acquisition system (DAQ) is highly constrained by the fact that any equipment sitting in the detector area must be tolerant to high radiation doses. The damage to silicon at the most irradiated area for one nominal year of running, i.e. an accumulated luminosity of 2 fb^{-1}, is expected to be equivalent to that of 1 MeV neutrons with a flux of $1.3 \times 10^{14} \text{ particles/cm}^2$, whereas the irradiation in the outer regions should be at least 25 times less severe.

In order to use standard electronics and computers, all the signals are routed behind a concrete wall, in the barracks located about 30 m away from the detector. For this purpose, the VeLo detector uses analog copper cables. A sketch of the full acquisition chain is represented in Fig. 2.10. Its main elements are the Beetle chips [23], a dedicated
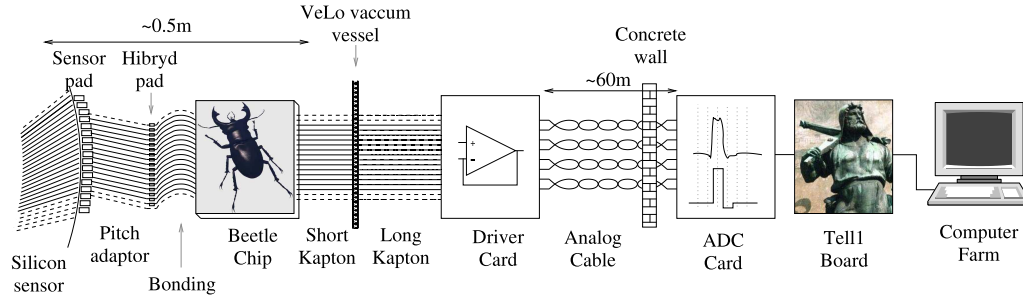
Figure 2.10: VeLo acquisition chain schematic. The numbers of wires are arbitrary.

chip for the signal amplification, the Trigger ELectronic of Level 1 boards (TELL1) [24] in the barracks, for sampling and first processing of the data, and the PC farm which runs the High Level Trigger (HLT).

The acquisition starts in the silicon where the signals are routed via the pitch adaptors glued to the sensors and bonded to the Beetle chips. Each Beetle has 128 inputs. It amplifies the incoming signal and stores it every 25 ns in a FIFO (First In First Out) buffer for at most 800 ns while waiting for a possible positive trigger decision. If this is the case, the signals are multiplexed into four streams of 32 channels serialized in time. These are sent to the so called analog links which comprises a driver card, the copper cable itself, and a receiver board.

In the Driver Card, the signal undergoes a two-stage correction, which amplifies the high frequencies in a way that will be canceled by the copper line transfer function, giving an output signal as close as possible to the original input. The cables are similar to the standard CAT6 model with differential pairs individually shielded. More details on the driver card can be found in Ref. [25].

Each Receiver card can handle 4 input lines. It samples the incoming signals on ten bits. The clock –25 ns frequency– is given by the TELL1 board[3] but the phase is adjusted to maximize the signal amplitude while minimizing the pick-up (cross-talk) from the neighboring channels. For example if one samples the signal too early, one might sample the next channel on the falling edge of the pulse and therefore have a big cross-talk in the output.

The signal is then stored in the TELL1 board, where different algorithms are applied:

- First, header effects are corrected. Indeed, the two header bits, which carry various event-related information [26], are known to influence the very first channel coming after them.

- Then a pedestals subtraction is performed: for each channel, one computes a running average (over time) of the pulse height. The result, called the *pedestal* typically converges to a stable value after a few thousand events. It can evolve to accommodate drifts occurring in the signal baseline. The pedestal value is then subtracted from the current channel-amplitude (we then speak about *pedestal subtracted data*).

---

[3]All the TELL1 are synchronized by a single control board, which distributes a 40 MHz clock.

- A Finite Impulse Response Filter correction is applied[4], as described in Sec. 3.5.

- The Linear Common Mode Subtraction (LCMS) algorithm [27] is applied to remove global fluctuations, e.g. coming from the beam pick-up or from the Beetle baseline that fluctuates. The LCMS acts on each group of 32 channels, i.e. all the channels relayed by the same physical cable. Essentially, it fits a linear function through this 32 channels and subtracts the fit result from the channel amplitudes.

- The channels are reordered to match the sensor layout.

- The data is clusterized. Any channel above a given 'high' threshold is used as seed and neighboring channels are added to from a *cluster* if their amplitudes are above a 'low' threshold.

- Channels not belonging to a cluster are ignored and the remaining data, called zero-suppressed data, are then sent through a standard giga-bit Ethernet network to the PC farm.

---

[4]Note that the exact sequence of the algorithms is not yet fully fixed with regard to the FIR filter. It may also be applied after the LCMS correction. The position shown here reflects the current implementation.

Chapter 3

Cross-talk in the VeLo acquisition chain
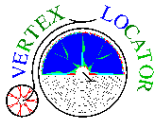
We introduce the Vertex Locator acquisition chain and the electronic cross-talk. We then present the analysis of test-beam data, disentangle and quantify the different cross-talk sources. Finally, we describe an algorithm that has been implemented to correct these effects.

THE Vertex Locator (VeLo) acquisition chain is a succession of electronic devices and cables leading the signal from the VeLo silicon sensors to the PC farm. Through all of them, the signal is spoiled by noise sources but also by capacitive and inductive coupling effects. Moreover, the routing of the signals depends on a quite complicated topology which means that two neighboring channels in a VeLo sensor may not be next to each other in the analog transmission.

This chapter describes analyzes done on the VeLo analog readout chain. It aims at measuring the different sources of cross-talk occurring in the full VeLo setup. We then describe a digital filter which can correct cross-talk effects down to a level of $\sim 2\%$. We conclude by showing that there is, before the correction, a visible effect on the cluster position and that this effect is removed after the cross-talk correction is applied.

3.1 Cross-talk definition

In electronics, the term cross-talk refers to any phenomenon by which a signal transmitted on one circuit or channel of a transmission system creates an undesired effect in another circuit or channel. Cross-talk is usually caused by undesired capacitive, inductive, or conductive coupling from one circuit, part of a circuit, or channel, to another. A well known case of cross-talk in the old analog telecommunication domain were these distinguishable pieces of speech leaking from other people's connections.

The cross-talk value (simply named the cross-talk hereafter) is defined as the fraction of one channel signal amplitude spreading into another channel. The two channels are called the *aggressor channel* and the *aggressed channel* respectively. The rest of this document will use the notation A_s and A_x for their associated
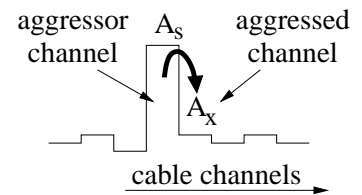
amplitudes (the subscript s and x stand for *signal* and *x-talk* respectively).

In the VeLo system, cross-talk happens in various places, however we define four main categories:

Cable cross-talk, which occurs in the analog 60 m copper cable itself. As the signals are serialized in time before being sent through the cable, this cross-talk is time related and therefore purely causal, i.e. the data signal transmitted at time t_0 can only disturb subsequent ($t>t_0$) signals, and cannot influence already transmitted information ($t<t_0$).

Analog-line cross-talk, which is the cross-talk occurring in the setup made of the Driver card, the analog copper cable and the Receiver card. Note that this cross-talk includes the cable cross-talk defined above, it is also purely causal, and as we will see later on, the Driver card actually aims to cancel the cable cross-talk.

Sensor cross-talk, which originates from the silicon strip and the sensor routing line interactions. This can be sub-divided, e.g. for a ϕ sensor, cross-talk between the inner strips, between the outer strips or between an outer strip and the inner-strip routing line running over it.

Readout cross-talk, which includes everything. It is dominated by contributions from the analog-line cross-talk (time related), and from the Beetle chip cross-talk, that is, inside the chip itself. The latter is space related and as such, can spread in many neighbors (e.g. the one on the left and the one on the right of the channel). Hence, when sent through the cable, this kind of cross-talk mimics the behavior of a non-causal propagation. Finally, the sensor cross-talk also contributes.

In a first study, we have measured in the laboratory the actual analog-line cross-talk contribution. The results of our measurements have been presented in Ref. [25]. For completeness, we have summarized in the methodology and some of our results in Sec. 3.2.

Then in the rest of this chapter, we adopt another approach which consists in analyzing test beam data coming out of a full VeLo acquisition chain. These data contain a superposition of all the possible noises and couplings that will be seen in the real data. We will show that their analysis allows to partially disentangle the different cross-talk sources and that the influence of some of these sources can be removed from the data with a digital filter.

3.2   Cable cross-talk studies

We summarize here the cross-talk analyzes that were conducted in laboratory in order to select the best cable type and to design the Driver Card. The goal was to minimize the amount of cross-talk induced by the system, with a setup reproducing the situation in LHCb. A pulse generator was used to send signals through the Driver Card prototype and then through 60 m of cable. The cable was terminated by the Receiver Card prototype, which was connected to an oscilloscope with high impedance active probes, at the level of the ADC input. We recall that the main idea of the system is to compensate the bandwidth loss in the cable by a pre-emphasis of the high frequencies in the Driver Card. We expect a better signal-over-noise ratio, compared to a 'post correction' (in the Receiver) of the frequency spectrum.

In parallel, we developed a simulation based on the SPICE software [28]. The copper cable transfer function was emulated by a chain of three low-pass filters. The Driver and

Receiver Card electronics were simulated as well. The implemented electronic scheme is shown in Fig. 3.1. All the transfer functions were compared to real-data measurements and found in excellent agreement. Figure 3.2 gives an example of measured and simulated response to a 25 ns wide pulse, with an indication of the cross-talk amplitude on the next channel.
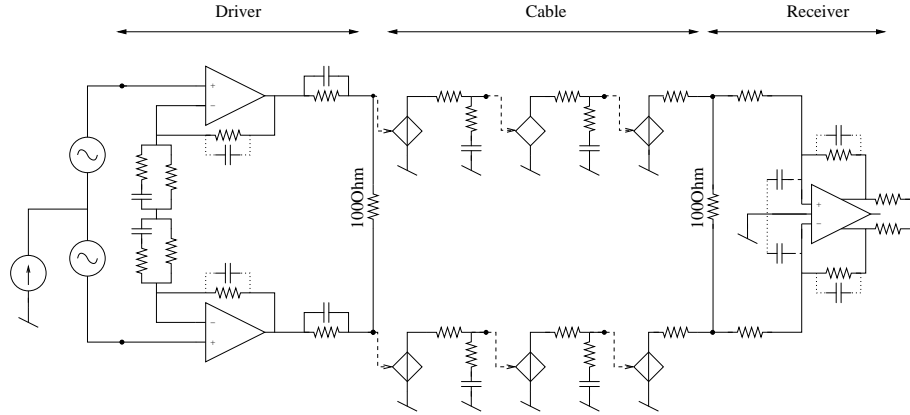
Figure 3.1: SPICE setup of the VeLo analog transmission line.
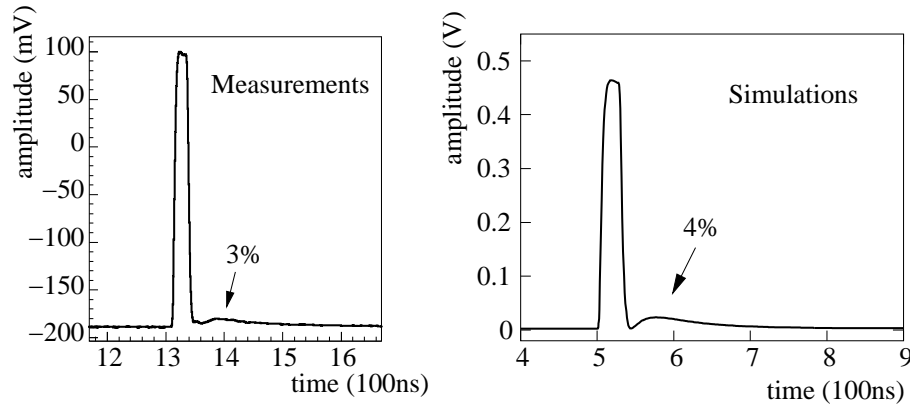
Figure 3.2: Measured and simulated response of the VeLo analog transmission line to a 25 ns wide pulse. The cross-talk level sampled on the next pulse is shown.

We concluded from this study that the maximum cross-talk contribution from the transmission in the cable could be easily kept under the 5% level, by a careful tuning of the compensation network in the driver card. We must notice that the amplitudes can vary depending on the sampling point chosen, e.g. sampling the pulse at the very beginning of its plateau will force the next channel to be sampled right in the small depletion visible after the pulse. Hence the cross-talk in the next channel will be almost zero (on the other hand, this will affect the value of the next-to-next channel). More details can be found in Ref. [25].

3.3   Cross-talk estimation techniques

3.3.1   Cross-talk model

A simple toy model has been used to simulate the experimental system, which includes a simulation of the signal amplitude distribution, the cross-talk, the presence of random noise, and also a possible 'common mode' (CM) background. The model assumes pedestal subtracted data (see Sec. 2.2.6). Note that a CM correction procedure is already present in the VeLo readout-scheme (implemented in the TELL1 firmware) [27]. Here we should consider cases where the correction procedure is not yet applied or where residual effects after CM correction exist.

We use the following expressions for the aggressor and aggressed channel amplitudes (A_s and A_x respectively):

$$
\begin{aligned}
A_s^i &= \mathrm{L}(23, 2) + A_{\mathrm{cm}}^i \,, \\
A_x^i &= \mathrm{G}(0, 2) + A_s^i \cdot \chi + A_{\mathrm{cm}}^i \,, \\
A_{\mathrm{cm}}^i &= \mathrm{G}(0, \sigma_{\mathrm{cm}}) \,,
\end{aligned}
$$

where $\mathrm{G}(\mu, \sigma)$ (resp. $\mathrm{L}(a, b)$) are random variables following a Gaussian (resp. a Landau[1]) distribution and χ is the cross-talk fraction. A Landau distribution is used for the aggressor channel amplitude. Gaussian distributions are used to simulate the noise and, if wanted, a non-zero CM value A_{cm} that can be added to both A_x and A_s. The fixed parameters of G and L are extracted from a fit of the distribution of Fig. 3.3, which has been measured for a real sensor during a test beam. We see that the 10–20 ADC region,
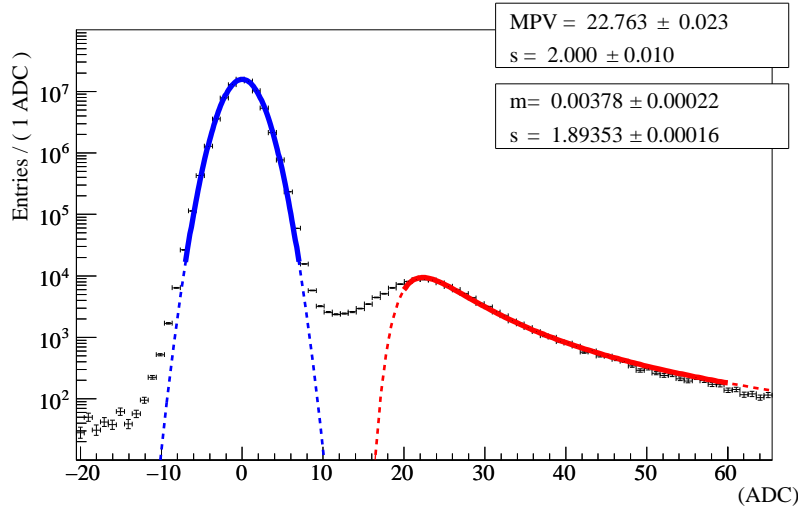
Figure 3.3: Distribution of the ADC values for all channels of one real sensor. The noise and the signal peaks are fitted with a Gaussian and a Landau function respectively, shown as a solid (dashed) curve inside (outside) the fitting regions.

which is due to high noise fluctuations and perhaps also to charge sharing effects, is poorly modeled. However, due to the cut on the aggressor and the aggressed channel amplitudes, this region is in any case not considered in our analysis.

---

[1]The Landau distribution follows the complex integral $p(x) = \frac{1}{\pi} \int_0^\infty \exp[-t \log(t) - xt] \sin(\pi t) \mathrm{d}t$ and is generated with `ranlan(G110)` from CERNLIB.

3.3.2 Cross-talk estimators

We consider here three different statistic that can estimate the amount of cross-talk in a set of amplitudes $\{(A_s^i, A_x^i)\}$ for a given aggressor channel and its aggressed channel.

Slope of linear fit: Fitting a straight line to the scatter plot of A_x vs A_s allows to extract a slope describing the amount of correlation between the aggressor and the aggressed channel.

Mean ratio: The mean value $\frac{1}{N}\sum_i A_x^i/A_s^i$ is an estimator for the cross-talk.

Ratio of the means: A slightly different estimator for the cross-talk is $(\sum_i A_x^i)/(\sum_k A_s^k)$. This is actually equivalent to the slope of a linear fit constrained to go through the origin of the scatter plot.

Figure 3.4 illustrates the three different estimators in the case of a fast Monte Carlo sample, generated using the model described above with $\chi = 0.07$ and $A_{\mathrm{cm}} = 0$.

(a) Linear fit of A_x vs A_s.

(b) Distribution of A_x/A_s. The different peaks are due to the fact that A_x and A_s are integer variables with relatively small values.

(c) Ratio of the means. Only the center of the scatter plot and the origin are used to define the slope.
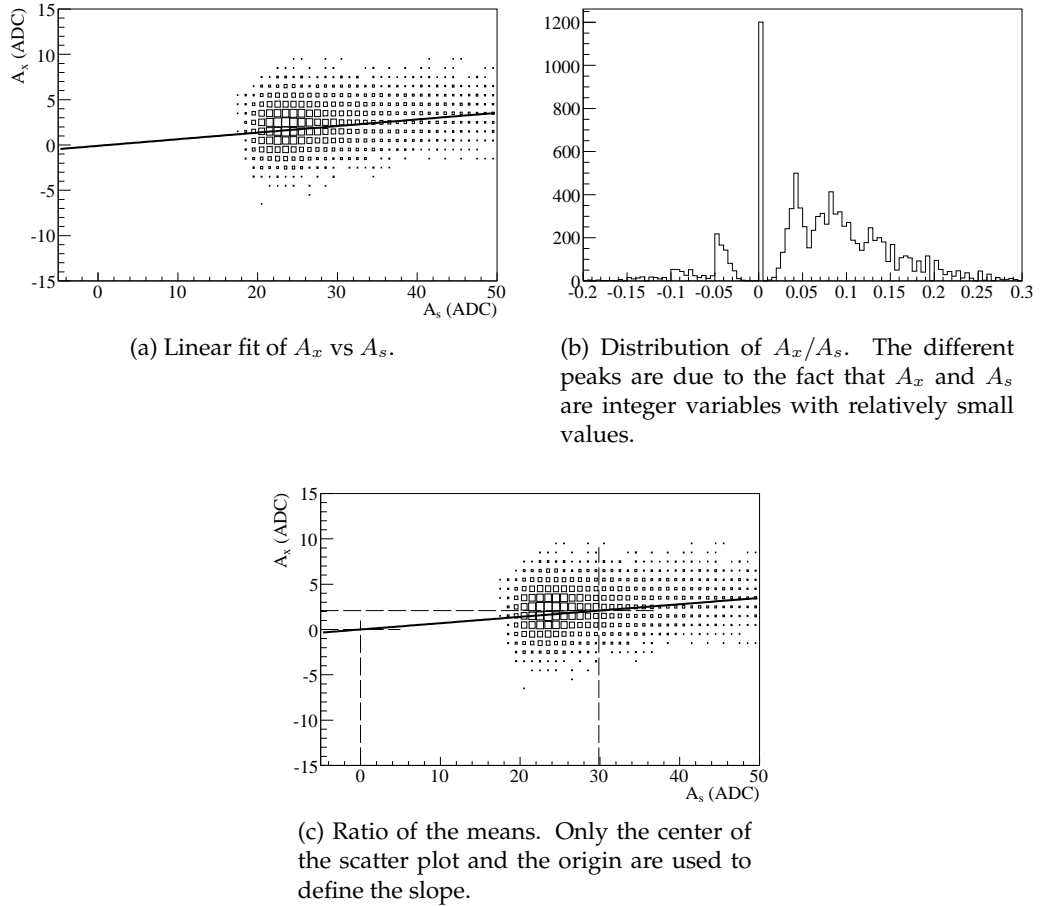
Figure 3.4: Fast Monte Carlo simulation of an aggressor and an aggressed channel with 7% of cross-talk and no common mode noise.

Estimators bias

In order to compare the estimators, we generate 1000 pairs of (A_s^i, A_x^i) values with a known cross-talk χ and compute the values of the three estimators. We repeat this 1000

times in order to have an estimate of the fluctuations. We then repeat the above simulation, scanning cross-talk values between -30% and $+30\%$. Figure 3.5 shows the deviation from the true value of χ. The standard deviations of the estimators are also shown. The most precise and stable estimator is the linear fit, still, the two other statistics have satisfactory accuracies.
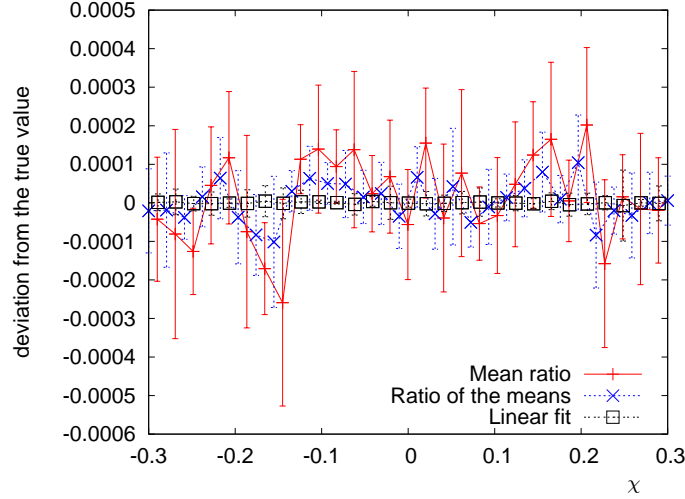
Figure 3.5: Deviation of the estimated cross-talk from the true cross-talk as a function of the true cross-talk for the three different cross-talk estimators. Each point is based on 1000 fast Monte Carlo events and is repeated 1000 times to estimate the error.

Common mode effect

We recall here that we assume pedestal subtracted data (see Sec. 2.2.6). It means that the amplitude distribution of any channel is centered around zero. Both the uncorrelated noise and the common-mode noise, if any, will smear this distribution but will not change the mean value.

We examine the robustness of the three estimators when some common-mode noise is present in the analyzed data, either because it has not yet been removed or because some residual effect is still present. We show in Fig. 3.6 the deviation from the true cross-talk as a function of the common-mode amplitude, i.e. a scan of the σ_{cm} parameter. The 'mean ratio' is quickly biased, whereas the linear fit and the ratio of the means successfully extract the cross-talk value.

From the two Fig. 3.4b and 3.6, we conclude that the 'mean ratio' is not a good cross-talk estimator, we will henceforth not consider it anymore.

Robustness of the estimators

Up to now, the 'ratio of the means' and the linear fit estimators have shown similar performances. However we expect (and have observed) that in real life other additional effects contribute, e.g. noisy strips that have not been removed from the readout, pedestal values that have not yet converged or common-mode noise correlated over a small number of
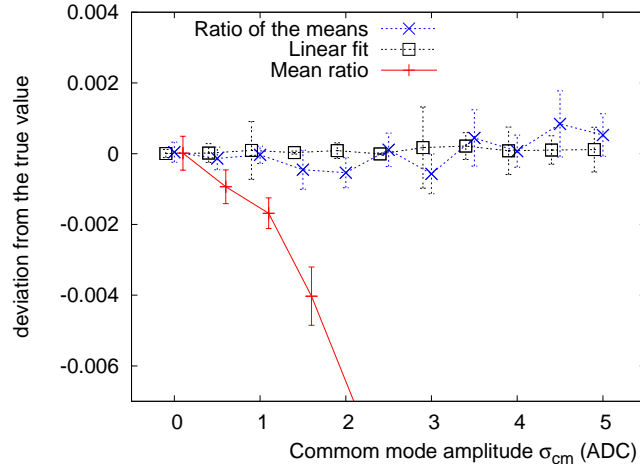
Figure 3.6: Deviation of the estimated cross-talk from the true cross-talk as a function of the common-mode amplitude σ_{cm}.

channels[2]. All these effects contribute to create correlations that cannot be handled with a first-order correction[3]. To illustrate this point, we artificially create an example shown in Fig. 3.7. The scatter plot is generated with a negative correlation, but it has been shifted in y toward positive values. In such a situation, the linear fit and the 'ratio of the means' estimators extract cross-talk values of opposite signs.
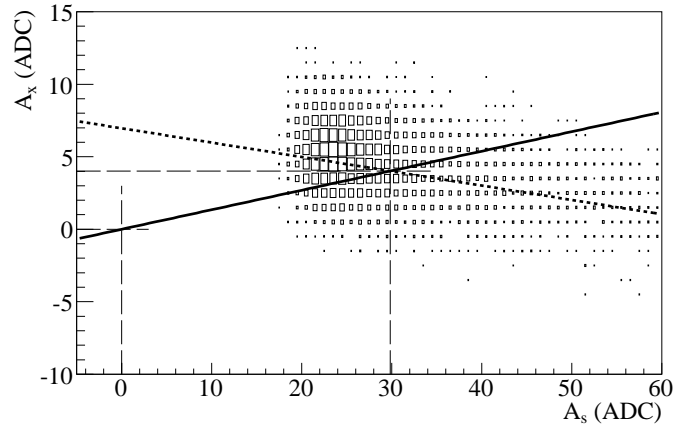
Figure 3.7: Artificial example illustrating a phenomenon observed in real data (although with a smaller amplitude): the scatter plot has a negative correlation, to which the linear fit is sensitive (dashed line) but the mean aggressed channel amplitude is positive, to which the ratio of the means is sensitive (solid line).

Clearly, to remove the correlation between A_s and A_x, one must alter both the slope and the intercept of the distribution. However the algorithm which correct the data will be eventually implemented in the TELL1, and a second-order correction is too heavy in

---

[2]As the common-mode is computed over 32 channels, fluctuations that occur over less than ~ 10 channels are not corrected for.

[3]By first-order correction, we mean a correction which is defined with a single parameter.

terms of TELL1 resources. Therefore, we consider here only first-order corrections.

After having removed the measured cross-talk as extracted by the linear fit, we get the distribution shown in Fig. 3.8a. The slope is indeed removed, but the mean value of the A_x is left almost unchanged (it actually slightly increases). On the other hand, we get, after correction of the cross-talk measured by the 'ratio of the means' estimator, the distribution shown in Fig. 3.8b. There, A_x is centered around zero, but the slope of the distribution is increased.

(a) Correction for the slope (linear fit esti-mator). The mean value of the aggressed channel is still above zero.

(b) Correction for the mean value (ratio of the means estimator).
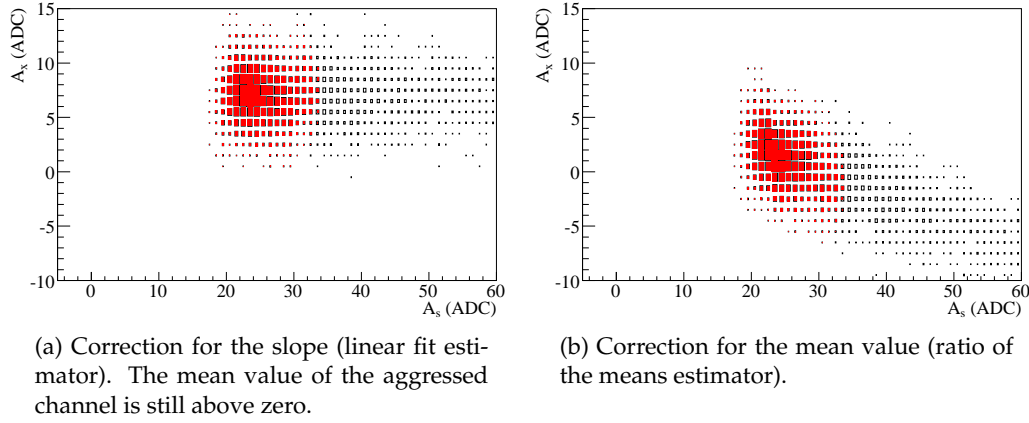
Figure 3.8: Same as in Fig. 3.7, but after the cross-talk correction with the 'linear fit' model (a) and the 'ratio of the means' model (b). The filled boxes, in red, represents, 75% of the statistics, centered on the maximum of the distribution.

Between these two estimators, we have chosen the latter for pragmatic reasons. Indeed, although we are not able to tell why the data sometimes exhibit such second order correlations, we see that the 'ratio of the means' will always tend to shift the aggressed channel distribution towards zero. This is especially visible when looking at the filled red part of the histograms (which represents 75% of the statistics) from Fig. 3.8. In the linear fit case, this fraction of the statistics is left almost uncorrected whereas the correction based on the 'ratio of the means' shifts it to zero. This is consistent with the idea that a signal with zero amplitude should not create a cross-talk contribution.

Summary

The cross-talk seen in a dataset as well as the correction one can hope to obtain from a linear filter will depend on the choice of the cross-talk estimator. All of them require pedestal subtracted data. The linear fit and the 'ratio of the means' estimators exhibit similar performances, however, in some limit cases, the 'ratio of the means' is more robust by construction, as it always tends to shift the mean value of the aggressed channel toward zero, regardless to the sign of the correlation. Therefore, we will use this estimator throughout this analysis. Explicitly, we will estimate the cross-talk as:

$$\hat{\chi} = \frac{\sum_i A_s^i}{\sum_k A_x^k} \ . \tag{3.1}$$

From a statistical point of view, one can easily compute the standard deviation of this estimator. However, in the case of the VeLo, the cross-talk amplitude will be estimated

by averaging different strip pairs that are assumed to have the same cross-talk, e.g. we average on the 32 channels of an analog cable to measure one cable readout cross-talk. This approximation happens to be the dominant factor in the cross-talk error and is rather hard to quantify. The author estimates the cross-talk precision to be $\sim 1\%$. Unless specified, we will assume this value throughout this document.

3.3.3 ϕ-sensor cross-talk

Because of the topology of the routing lines, the cross-talk in the ϕ sensors can be studied with data that have been sent through the analog acquisition chain while the cross-talk in the r sensors can only be studied in laboratory conditions. We explain why below.

As described in Sec. 2.2.1, the readout scheme of a ϕ sensor is non trivial. To better show the topologies which can provoke sensor cross-talk, we re-arrange the readout scheme from Fig. 2.5 into the one shown in Fig. 3.9. Each small square in the figure can be seen as a silicon strip or its associated routing line. If two squares are close to each other, sensor cross-talk is possible between them. Moreover, the chip channel number are also given, allowing to see which pairs will be affected by readout cross-talk. For
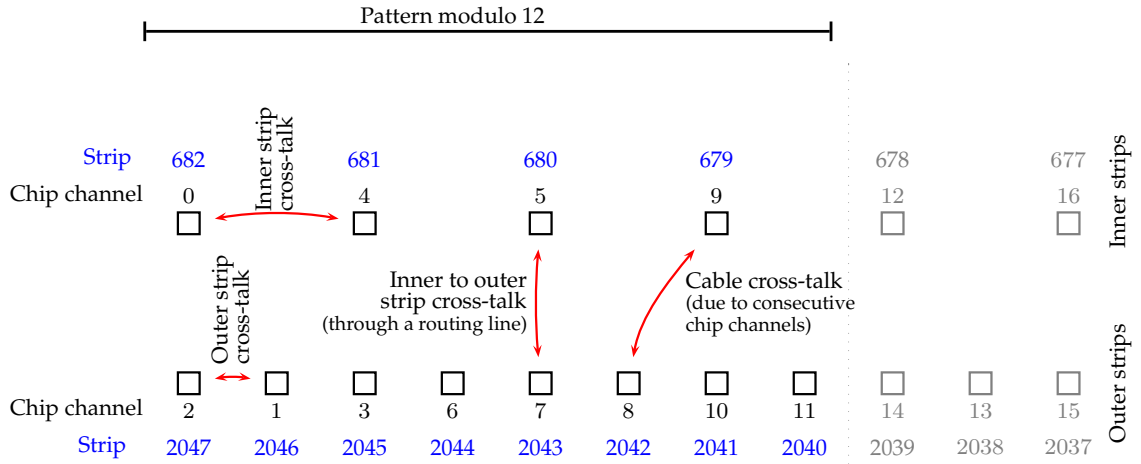
Figure 3.9: Chip channel and strip numbering in a ϕ sensor. The small squares on the top row represent the inner strip routing lines that run over the outer strips. The squares on the bottom row are the outer strips. Some examples of cross-talk are shown.

example chip channel pairs like $(0,1)$ or $(5,6)$ are not neighbors in the sensor but are next to each other in the Beetle chips and the analog cables. Therefore, analyzing the cross-talk only on these pairs (and all their modulo 12 equivalents) will give a measurement of the cross-talk coming from the readout with a negligible contribution from the sensor. In contrast, chip channel pairs such as $(1, 3)$, $(3, 6)$ and $(8, 10)$ are not neighbors in the readout (their chip channel numbers are not consecutive) but are neighboring sensor outer strips. Hence the dominant contribution will be the cross-talk from the sensor outer strips, but obviously these pairs will also be sensitive to cross-talk between second- and third-neighbor chip channels. Therefore, by carefully selecting the strip pairs, one can, to some extent study different cross-talk contributions. This method will of course be less sensitive than laboratory studies, but it has the advantage to be usable with real data.

3.3.4 Measuring the cross-talk with test pulses

For testing purposes, the Beetle chip has an internal mechanism allowing the generation of a test pulse pattern. These pulses have alternating polarities for neighboring chip channels and for consecutive events. This scheme raises different problems.

The pedestal algorithm implemented in the TELL1 will require —when applied on a test pulse— a very long time to converge because is assigns a smaller weight to amplitudes that differ from the current mean value, which is always the case for test pulses. Moreover, the pedestal will be wrongly estimated as it will compute the mean value of the negative and positive pulses which are known to have different amplitudes. This explains why, as shown in Fig. 3.10, the cross-talk computed separately on the two pulse polarities are different. Although the relative dependencies in the neighbor offset is quite
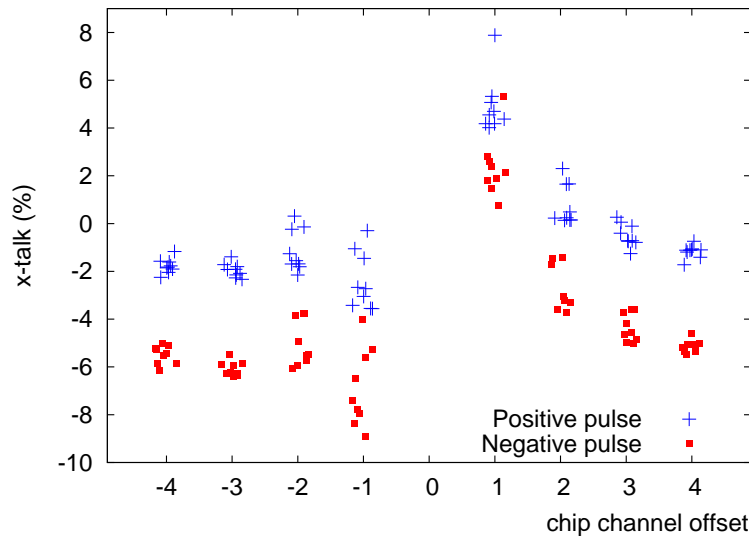
Figure 3.10: Readout cross-talk computed on positive (+ sign) and negative (squares) test pulses for each analog cable. The shift toward the negative values is probably due to a pedestal which has not yet converged (due to lack of statistics). For readability, the data points have been slightly spread along the x axis.

similar, a $\sim 4\%$ global shift of the cross-talk is visible. Moreover, the presence of a long range cross-talk is rather non-physical. Still, it is not possible to tell whether this is a real physical effect of a fake shift created by the way we compute the cross-talk, for example if the pedestal has not converged due to lack of statistics. A way to avoid this problem is to do a dedicated run with no pulse, to compute an unbiased pedestal, which would be used later on with a pulsed sample. This has not been done in the test beams but could certainly be done in the real experiment, for instance between two fills.

The asymmetry between odd and even chip channels is also a problem when working with test pulses. In the Beetle chips, the cross-talk is correlated to the parity (odd-even) of the channels. This is a known feature of the Beetles, which comes from the way they internally handle the signals. Beam data will of course exhibit the same effect, but the fixed pattern of the test pulses enhances this asymmetry. Indeed, the ϕ-sensor pattern is modulo 12 and the test pulse pattern is typically modulo 32. Hence the global pattern is modulo 96, which is an even number. If one wants to measure a sensor effect, one will look at channels with a modulo 12 pattern. Eventually, a sensor cross-talk estimation

will be computed only on odd or only on even Beetle channels, resulting in an enhanced sensitivity.

Thus, the cross-talk estimation through test pulses data, which is in principle easy and clean, is somehow biased. For these drawbacks we have done an attempt to estimate the cross-talk correction parameters from beam data which should be closer to the real experimental configuration.

3.3.5 Measuring the cross-talk with real data

Working with beam data, it is necessary to provide some selection mechanism that will, as much as possible, select one-strip clusters and ensure that the corresponding aggressed channels have not been hit in the same event. Therefore, we cut on the signal, on the aggressed channel amplitude and also on all their closest neighbor amplitudes. Another obvious requirement in the selection of pairs is that both channels share the same common mode conditions, e.g. they belong to the same analog link and the same sector of a sensor (the latter being ensured by the former).

(a) Closest neighbors in the chip channel layout.

(b) Closest neighbors in the ϕ sensor strip layout.
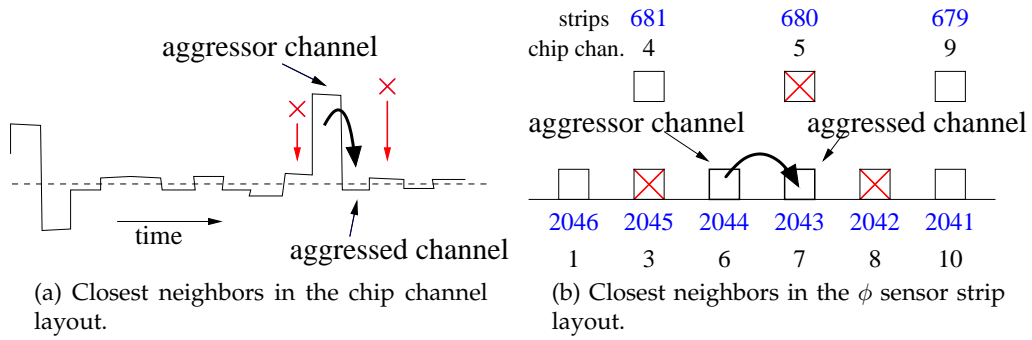
Figure 3.11: Closest neighbor definition. The channel marked with 'X' are the closest neighbors of the aggressor-aggressed channel pair.

The selection of aggressor-aggressed channel pairs requires two thresholds: a minimum signal amplitude which defines whether a channel contains a signal, and a maximum amplitude which is applied on the aggressed channel candidate and its closest neighbors (as defined in Fig. 3.11). Deciding which neighbors are considered (sensor or chip neighbors) depends on what is studied. For an r sensor, one uses on the chip channel neighbors scheme, whereas for a ϕ sensor both schemes are applied. Cutting around the aggressor channel minimizes the risk of charge sharing (see below), and the cut on the aggressed channel neighbors makes sure none of them have been hit by a particle in the same event, which would create an unwanted cross-talk contribution.

Charge sharing between strips occurs when a particle triggers more than one strip, i.e. some of the created electron-hole pairs are collected by two different strips. From the point of view of the cross-talk, this effect (when present) will always tend to increase the aggressed channel amplitude. Moreover, it represents a non-Gaussian contribution to the aggressed channel distribution[4], which we show in Fig. 3.12. In the following, charge sharing effects are neglected.

---

[4]The intrisic noise has a Gaussian distribution and the cross-talk essentially shifts this distribution to a non-zero mean but does not change much the shape.
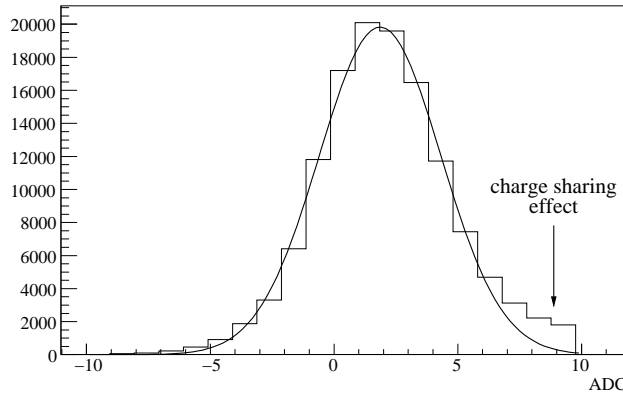
Figure 3.12: Aggressed channel distribution in beam data. The non Gaussian tail is due to charge sharing. The cut on the aggressed channel removes amplitude above 10 ADC.

3.4 Cross-talk studies with beam data

During the year 2006, two test-beams periods at CERN were dedicated to the VeLo. They are called Alignment Challenge and Detector Commissioning (ACDC) 2 and 3 (ACDC1 was a preliminary test). Applying the technique presented above, we quantify the amount of cross-talk that can be seen in the collected data, disentangling the different cross-talk sources by selecting some specific ϕ-sensor channel pairs, as described in Sec. 3.3.3.

3.4.1 ACDC2 beam data

The data acquisition chain of the August 2006 test beam used a full analog readout. Six sensors (3 r and 3 ϕ) were connected, but each had only one sector (512 strips) read out. The signals were lead through 60 m compensated cables to six different TELL1 boards. The data were then sent to a single computer performing the event reconstruction. Details on the test-beam setup can be found in Ref. [29].

Readout cross-talk

The observed readout cross-talk is roughly symmetric between forward (the aggressor channel is n, the agressed one is $n + 1$) and backward (the aggressor channel is n, the agressed one is $n - 1$) cross-talk (see Fig. 3.13). Only the first neighbors are significantly affected with a negative amplitude of $\sim 4\%$. We recall here that the readout cross-talk is a combined effect of the analog-line cross-talk and the Beetle chip cross-talk. The former can only create forward cross-talk, and the latter can create any kind of cross-talk. They add up together and are sampled on one point of the analog signal. Because the observed cross-talk depends on the sampling phase, the observations reported here only represent a particular example of the readout cross-talk in a full acquisition chain. The results (4% in the first neighbors) are nonetheless believed to be representative of what one can expect in the real data.
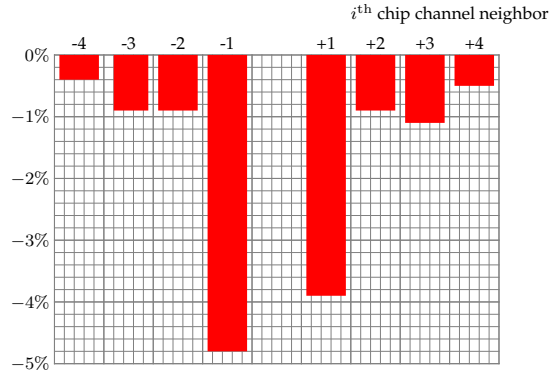
Figure 3.13: Forward and backward readout cross-talk observed in the ACDC2 test beam.

ϕ sensor cross-talk

To sort out the different ϕ sensor cross-talk sources, we analyzed all possible strip pairs that can be defined in Fig. 3.9, i.e. all the pairs which represent different topologies. The obtained values are shown in Fig. 3.14 for neighboring strips only. Pairs which are not shown have values compatible with zero unless they are first or second neighbors in the chip. Some points can be noticed:
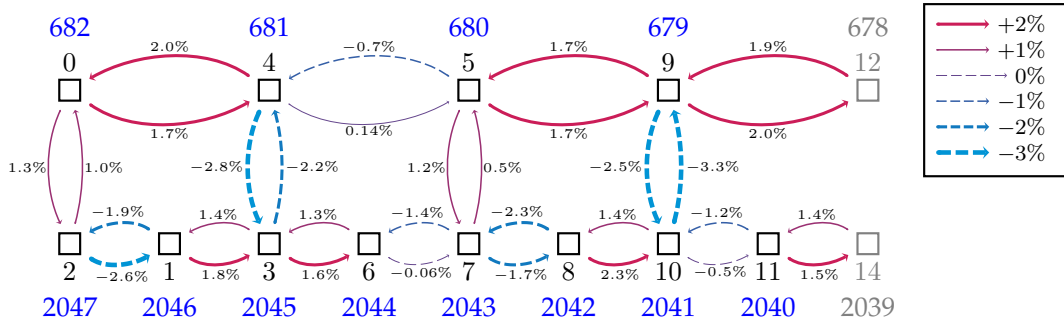
Figure 3.14: ϕ sensor first neighbor cross-talk observed in the ACDC2 test beam. The line width indicates the cross-talk amplitude. Dashed lines indicates negative values.

- Any negative cross-talk value (indicated with a dashed line) is associated to neighboring chip channel pairs, which are also sensitive to the readout cross-talk.

- All the non-neighboring chip channel pairs have a positive cross-talk of $\sim 2\%$. This is mainly the sensor contribution.

Second neighbor strips in a ϕ sensor (e.g. between chip channel 7 and 10) show that the 'long range' cross-talk is around -0.3%. These contributions can be neglected. Other 'exotic' cross-talk patterns have been tested. Pairs like (1,4) or (5,10) do not generate any cross-talk except when they happen to be first neighbors in the cable.

To summarize, a ϕ sensor and its analog-line show six main cross-talk categories corresponding to all the combinations depending on whether or not the two strips are first neighbors and whether or not the two associated chip channels are forward or backward

first neighbors. This simplified category scheme is possible because long range sensor and readout cross-talk are close to zero thus allowing the disentangling of the different cross-talk sources. Table 3.1 gives the values for these six categories for a typical ϕ sensor in the ACDC2 test beam. We notice that the difference between the two lines reported in this table is $\sim 2\%$, which is the estimated sensor contribution (top right number in the table). This indicates that the sensor effects add up with the readout effects in a rather linear way.

	forward	backward	non consecutive
neighbor	**−2.6%**	**−3.4%**	**2%**
non neighbor	**−3.9%**	**−4.8%**	**0.3%**

Table 3.1: The different ϕ sensor cross-talk categories and their values in one particular TELL1, computed as averages over all equivalent pair topologies.

3.4.2 ACDC3 beam data

During a second test beam in November 2006, six modules (12 sensors) were fully read out. The sensors were 300 μm thick, whereas the ACDC2 configuration had 200 μm thick sensors. The analog link cables were only 15 m long, and the frequency compensation was totally removed for such short cables, with the consequence that the corresponding bandwidth reduction was not corrected for. Hence we expected, and observed, that ACDC3 data have larger readout cross-talk. This shows that the analog frequency-compensation is working as expected when present and is necessary to compensate for the cable cross-talk. Despite this, the ACDC3 data present a good opportunity to study data with high cross-talk values.

Readout cross-talk

Figure 3.15 shows the cross-talk as a function of the signal channel offset, computed as averages over the analog link, including ϕ sensor neighboring strips. The readout contribution is most visible in the first neighbors (channel offsets $+1$ and -1). Their positive value is the sign of an under-compensated signal, which is expected as the analog corrections were removed. The difference between r and ϕ sensors is especially visible in the channel offset -2. It corresponds to chip channel pairs like $(2, 0)$ and $(7, 5)$ which — referring to Fig. 3.9— are an outer–to–inner strip cross-talk and are specific to ϕ sensors.

ϕ sensor cross-talk

Computing the different ϕ sensor cross-talk contributions for ACDC3 data is not as clean as for ACDC2 data. Indeed, the long range readout cross-talk being significant (see
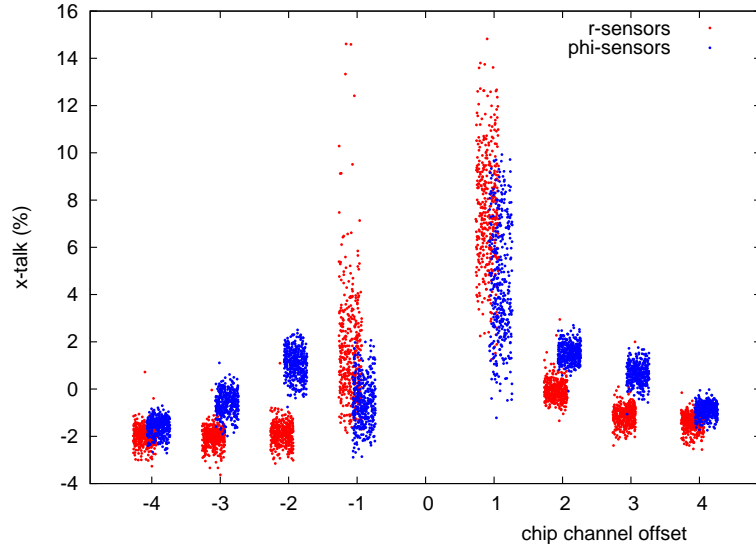
Figure 3.15: Readout cross-talk for ACDC3 test-beam data averaged on the contribution of each analog line. The red dots (slightly shifted to the left) represent cables associated with r sensors, whereas the blue dots (shifted to the right) represent cable associated with ϕ sensors.

Fig. 3.15), it spoils the measurement of the sensor cross-talk. We estimate the values presented below to have a $\sim 2\%$ uncertainty.

Taking one TELL1 as example, we compute a cross-talk between outer strips of $\sim 3.5\%$, a cross-talk between inner strips of $\sim 3\%$ and a cross-talk between an outer strip and inner routing line of $\sim 7\%$. For this particular category, all ϕ sensors show an amplitude between 4 and 8%. This appears to be the highest sensor cross-talk type and it is not present in ACDC2 data. It may be a side effect of the increased sensor thickness.

3.4.3 High-voltage dependence

This analysis has been done on a dedicated test-pulse dataset taken in the assembly laboratory with one module on which a variable high voltage was applied. Previous measurements had shown that the full depletion voltage for this sensor is 60 V. The variation of the different types of cross-talk is shown in Fig. 3.16. The readout cross-talk (computed on non-neighboring strips pairs) is almost insensitive to the high voltage. The cross-talk is reduced at the full depletion voltage, both for inner and outer strips. However, the inner-outer cross-talk, which depends on the double metal layer properties, increases at full depletion voltage.

It is rather intuitive to understand that the capacity between the strips increases at the depletion voltage, hence the cross-talk drops down. However, it is not understood why the cross-talk between the outer strips and the routing lines increases.

3.4.4 Expected cross-talk in the real experiment

To summarize, we have analyzed the cross-talk in two test-beam datasets. The readout contribution we expect in a full VeLo setup can be estimated from the ACDC2 dataset and is about -4% on both the forward and the backward chip channel neighbors, the contribution in the neighbors further away being negligible.
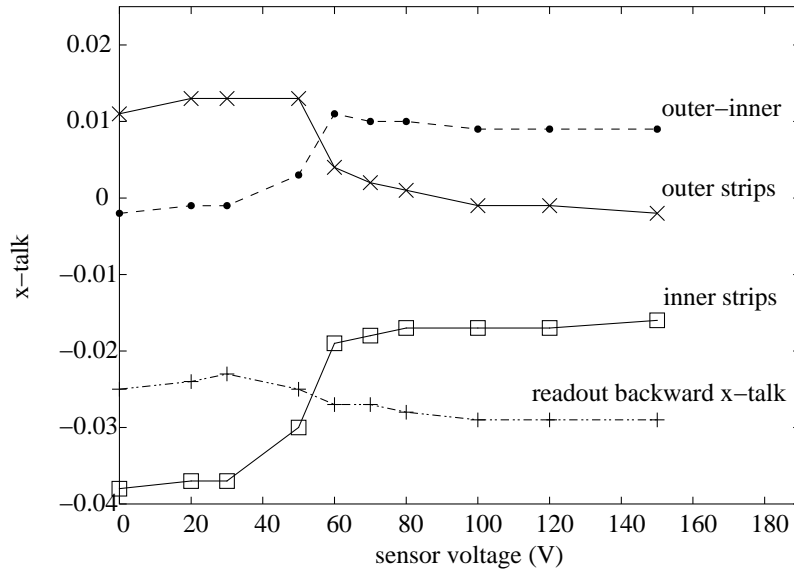
Figure 3.16: Cross-talk effects as a function of the sensor high-voltage. The cross-talk is measured on the first strip neighbors, between outer strips, between inner strips and between the outer strips and the overlying inner strip routing line. The first backward neighbor of the readout cross-talk is also shown.

The sensor contribution is estimated from the ACDC3 data because they include the increased sensor thickness (300 μm). The cross-talk contribution from the sensor is about 3% between the strips but a 7% contribution is coming from the capacity between the outer strip and the double metal layer on top of it.

3.5 The Finite Impulse Response algorithm

3.5.1 Algorithm description

A Finite Impulse Response (FIR) filter is quite often found in digital equipment [30, 31]. It is 'finite' because its response to an impulse ultimately settles to zero. This is in contrast to infinite impulse response filters which have internal feedback and may continue to respond indefinitely. FIR filters applied to the VeLo have already been foreseen (Refs. [32, 33]).

In the VeLo, the sensors, the Beetle and the analog cable have together a defined transfer function $H(z)$, where z is the frequency in the complex domain. This function is assumed to be time independent. Thus the response to an impulse $p(z)$ is

$$w(z) = H(z)p(z) \ . \tag{3.2}$$

A FIR filter is designed to apply a function $G(z)$ to the observed signal $w(z)$ such that $G(z)w(z) = p(z)$, which implies $G(z)H(z) = 1$. The functions can also be expressed in the time domain through a Fourier transform, and as the TELL1 boards work with a sampled time, the integral over the time becomes a sum over the channels. It means that the functions H and G are fully defined by a set of coefficients h_n and g_n, where the index n runs over the channels. Figure 3.17 gives a scheme of the different variables. The p_n are the amplitudes in the sensors, the w_n are the amplitudes as observed after the readout,

i.e. transformed by the function H, and thus the h_n are the cross-talk coefficients. The function G describes the correction (the g_n are the FIR coefficients) and the y_n are the amplitudes after the correction has been applied. Hence in the sampled time domain, Eq. 3.2 becomes:

$$w_n = \sum_{k=-\infty}^{\infty} h_k p_{k-n} \; ,$$

and similarly

$$y_n = \sum_{k=-\infty}^{\infty} g_k w_{k-n} \; .$$

Transforming the condition $G(z)H(z) = 1$ in the time domain is straightforward. The
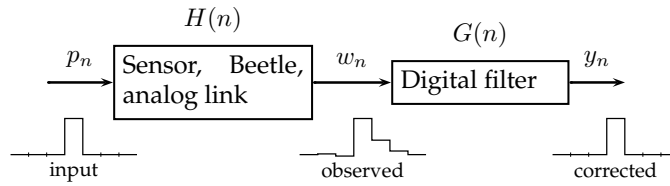
Figure 3.17: Transfer function of the acquisition chain.

Fourier transform of 1 is a Dirac function, or equivalently in a sampled time, the Kronecker delta function $\delta_n = 1$ if $n = 0$, 0 elsewhere. The convolution product is transformed in an infinite sum over the sampled time. Thus we write

$$\delta_n = \sum_{k=-\infty}^{\infty} g_k h_{k-n} \; . \tag{3.3}$$

To deal with finite series, we assume that the g_k will be non-zero only for a finite odd number N_G of terms around the bin n. Thus Eq. 3.3 becomes

$$\delta_n = \sum_{k=-\frac{N_G-1}{2}}^{\frac{N_G-1}{2}} g_k h_{k-n} \; , \tag{3.4}$$

which can be seen as an infinite series of equation, one for each value of n. These equations will have non-trivial solution only for values of n where the h_{k-n} are not zero for all terms in the sum. Assuming that the cross-talk (the effect of the function H) only extends N_H bins around the bin n (N_H is odd), the requirements on n for non trivial equations are:

$$-\frac{N_G-1}{2} - n \leq \frac{N_H-1}{2} \quad \text{first term in the sum is non-zero,}$$

$$\frac{N_G-1}{2} - n \geq -\frac{N_H-1}{2} \quad \text{last term in the sum is non-zero.}$$

This gives $N_G + N_H - 1$ equations, $-(N_G + N_H)/2 + 1 \leq n \leq (N_G + N_H)/2 - 1$, defining the relations between the g_n and h_n coefficients. These constraints can be re-written in one expression:

$$\sum_{k=-\frac{N_G-1}{2}}^{\frac{N_G-1}{2}} g_k \cdot h_{k-n} = \begin{cases} 0 & \text{if } -(N_H + N_G)/2 \leq n < 0 \\ 1 & \text{if } n = 0 \\ 0 & \text{if } 0 < n \leq (N_H + N_G)/2 \; . \end{cases} \tag{3.5}$$

This is a system of $N_G + N_H - 1$ equations which has N_G unknowns, the g_n. It can be written in a matrix form. We expand for the case $N_H = 3$ and $N_G = 3$. Defining

$$
H \equiv \begin{pmatrix} h_1 & 0 & 0 \\ h_0 & h_1 & 0 \\ h_{-1} & h_0 & h_1 \\ 0 & h_{-1} & h_0 \\ 0 & 0 & h_{-1} \end{pmatrix}, \qquad \Delta \equiv \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \qquad G \equiv \begin{pmatrix} g_{-1} \\ g_0 \\ g_1 \end{pmatrix},
$$

then

$$
\begin{aligned}
H \cdot G &= \Delta \\
H^T \cdot H \cdot G &= H^T \cdot \Delta \\
G &= (H^T \cdot H)^{-1} \cdot H^T \cdot \Delta
\end{aligned}
$$

gives the expression of the g_n as a function of the cross-talk coefficients. This method can trivially be extended to any N_G and N_H orders.

3.5.2   Algorithm implementation

We suggest here a possible implementation of a FIR. As seen before, a FIR filter is characterized by its orders N_H and N_G representing the number of channels which determine the impulse response (N_H) and the number of channels to correct for (N_G). In the VeLo setup, there is an additional aspect which we call the *granularity*. It specifies which system a given set of FIR coefficients is applied to. Possible systems are: the full VeLo, a single TELL1 board, or a single analog cable. Obviously, choosing a 'cable granularity' requires to compute more coefficient sets (one per analog link) but also allows a finer correction.

The currently implemented version relies on three coefficients per TELL1 acting on the first neighbors and the aggressor channel ($N_G = 3$)[5]. The version proposed in this document has a cable granularity. It assumes that the central coefficient is one (i.e. the aggressor channel amplitude is left untouched). This has shown to be a satisfying approximation and spares some computation and resources in the TELL1 boards. For study purposes, the implemented FIR software allows up to 8 coefficients around the central bin (i.e. a FIR of order $N_G = 9$, $N_H = 9$).

The actual TELL1 FIR implementation has to rely only on binary calculation. Thus some transformations of the expressions of Sec. 3.5.1 are required. Expanding the case of an order 5 FIR, one gets the corrected amplitude for channel n:

$$
\begin{aligned}
y_n &= g_{-2} \cdot w_{n-2} + g_{-1} \cdot w_{n-1} + g_0 \cdot w_n + g_{+1} \cdot w_{n+1} + g_{+2} \cdot w_{n+2} \\
&= \frac{1}{128} \left\{ (512 g_{-2}) \cdot \frac{w_{n-2}}{4} + (512 g_{-1}) \frac{w_{n-1}}{4} + (512 g_{+1}) \frac{w_{n+1}}{4} + (512 g_{+2}) \frac{w_{n+2}}{4} \right\} + w_n \\
&= \frac{1}{128} \left\{ K_{-2} \frac{w_{n-2}}{4} + K_{-1} \frac{w_{n-1}}{4} + K_{+1} \frac{w_{n+1}}{4} + K_{+2} \frac{w_{n+2}}{4} \right\} + w_n \,,
\end{aligned}
$$

where $g_0 = 1$ is assumed. The $K_i = 512 g_i$ are the values effectively stored in the TELL1 registers as signed 8-bit integers. As consequences, the smallest representable cross-talk amplitude is $1/512 \simeq 0.2\%$ and the maximum absolute value is $127/512 \simeq 25\%$. Dividing the measured amplitude w_k by 4 (i.e. dropping the two least significant bits) allows to work with smaller multiplication register and drop the amplitude fraction which is highly subject to noise fluctuations.

---

[5] At the time this document is written, the version presented here is being implemented.

3.5.3 Cross-talk correction with a FIR filter

The present section quantifies the FIR efficiency. The program `VeloXTalkComputer` measures the cross-talk while the program `VeloTELL1CableFIRFilter` [34] implements the algorithm described above. Measurements and corrections of the cross-talk are always applied on independent datasets.

Figure 3.18 shows for each analog link, the readout cross-talk before and after the FIR correction on ACDC2 beam data. All the corrected cross-talk values are below 1.5%.
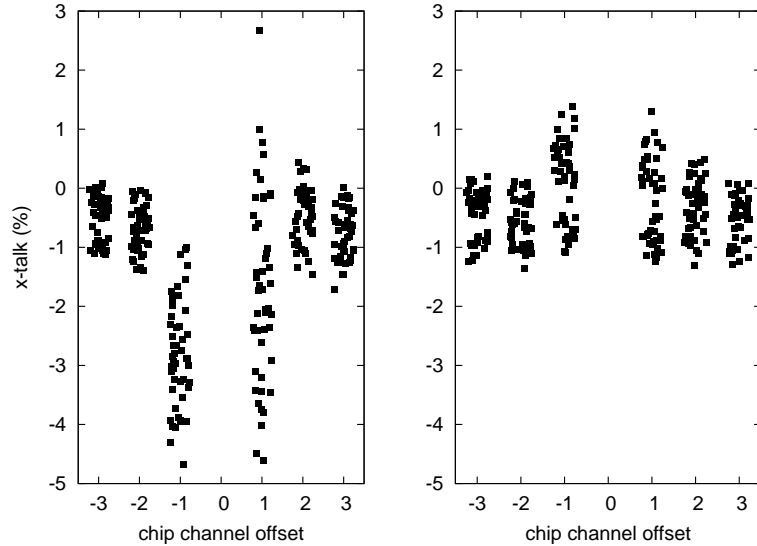
Figure 3.18: Readout cross-talk computed before (left) and after (right) the FIR correction, using ACDC2 data (1 r and 2ϕ sensors).

This can be interpreted as almost zero regarding the bit precision. However, if one tries to correct higher cross-talk amplitudes, like the one visible in ACDC3 data, we see (Fig. 3.19) that the correction is not perfect and remains correlated with the original cross-talk value.

Figure 3.20 shows the readout cross-talk in each analog link, after a correction based either on a TELL1 granularity (left plot) or on an analog cable granularity (right plot). As expected the per cable correction shows better results.

Yet, the important quantity for the LHCb physics is the cross-talk influencing the cluster reconstruction, i.e. at the strip and chip channel level. Figure 3.21 shows the chip channel first neighbor cross-talk distribution after the two types of granularity correction (per TELL1 and per cable). The main effect of the FIR correction is to center the cross-talk distribution, whatever the granularity or the sensor type. The distribution width, although slightly reduced by the per cable correction, is mainly left unchanged. This supports the idea that we reach here the limit where the sensor cross-talk dominates over the readout cross-talk. The width of the r sensor cross-talk distribution is more reduced than that of the other sensor type. This is expected, as the r sensor cross-talk at the strip level will globally behave like the readout cross-talk, hence it will be corrected by the FIR filter, down to a level of $\pm 2.6\%$ in the present case. We stress here that 3% of cross-talk, considering a MIP pulse of ~ 30 ADC counts, represents 1 ADC count. This is the smallest value measurable by the TELL1 and is about half of the typical noise level obtained during the test beams.

Figure 3.22, taken from Ref. [35], is based on ACDC3 data and shows the mean value
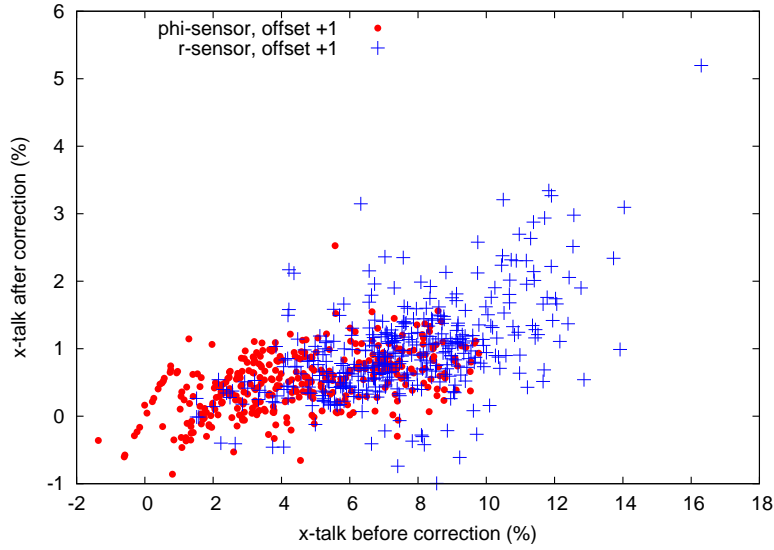
Figure 3.19: Corrected readout cross-talk versus initial readout cross-talk for the cable first neighbor with ACDC3 data. Each dot is an analog link. The blue crosses represent ϕ sensors, the red circles represent r sensors.
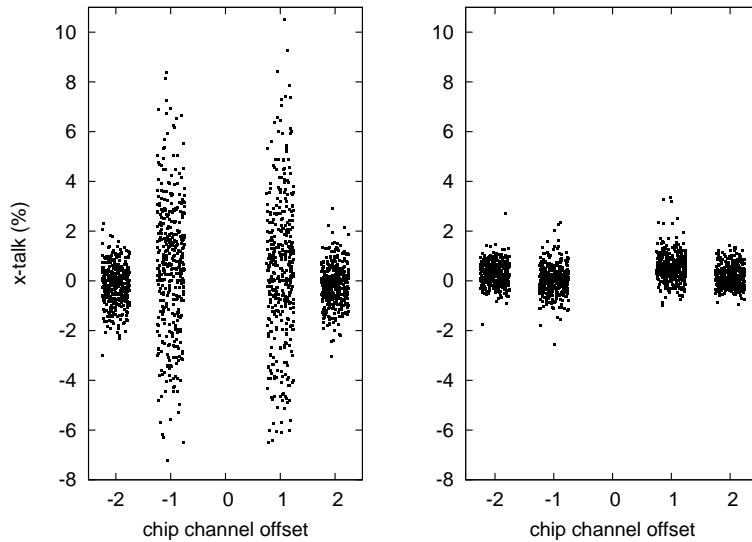
Figure 3.20: Remaining readout cross-talk after a FIR correction per TELL1 board (left) and a per analog cable (right).

(a) Per channel readout forward cross-talk in an r sensor.

(b) Per channel readout forward cross-talk in a ϕ sensor.
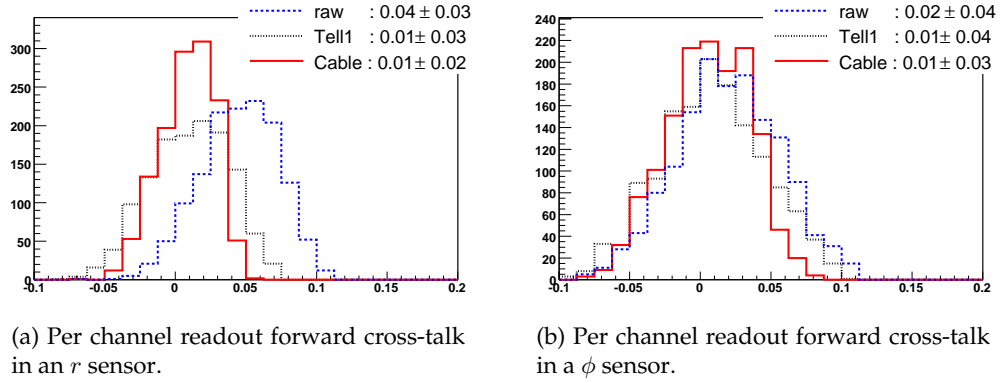
Figure 3.21: Distribution of the first neighbor readout cross-talk before any correction (dashed line), and after a correction per TELL1 board (dotted line) or per analog link (solid line). The mean values and the RMS of the distributions are also given.

of the track residuals as a function of the ϕ coordinate. Figure 3.22a was not corrected for any cross-talk, whereas Fig. 3.22b shows the same data with the cross-talk removed according to the above algorithms ($N_H = N_G = 5$, per-cable granularity). The gap visible in Fig. 3.22a is a consequence of the analog line cross-talk. Indeed, as mentioned in the r sensor description (see Sec. 2.2.1), the routing lines happen to sometimes reverse the readout order. For instance, the first sector of an r sensor is read out from strip 128 to 511 and then from 127 to 0. We sketch in Fig. 3.23 the readout scheme of an r sensor. Because the analog line cross-talk mainly creates forward cross-talk, the reconstructed clusters will be shifted in a direction related to the readout order. We see that, after having applied the FIR (Fig. 3.22b), the gap between the two readout directions is compatible with zero. Two things can be outlined: first, a cross-talk amplitude of $\sim 8\%$ in the readout, like in the ACDC3 data, produces a visible effect in the tracking. Second, the FIR filter is able to reduce the cross-talk influence at a level were this gap is not visible anymore.

The cross-talk correction in a ϕ sensor can be addressed in two ways. None of them has been implemented or tested because a cross-talk in the -3% to $+4\%$ range is considered as low enough (again, it represents less than one sigma of the test beam observed noise).

The first possible way is to correct for the cross-talk visible in the readout and, after the channel reordering, to use a second FIR, which will then see the strip layout. Yet this is hard to calibrate as it requires to systematically disentangle the different cross-talk sources. Moreover it possibly represents a non-affordable factor two in the TELL1 resources occupied by the FIR algorithm.

The other approach is based on the observation that any strip pairs which are first neighbors in the sensor, are fourth neighbor or less in the cable (see Fig. 3.9). Hence, if one defines (and manages to calibrate) a FIR filter of order $H_N = 9$ and $N_G = 9$, and if one changes the coefficient set for each single chip channel, one gets a FIR which corrects the sensor and the readout cross-talks in one go. This may look very attractive, but the issue here is that we break one of the FIR filter hypotheses, namely that the $H(z)$ and $G(z)$ functions should be invariant under time translation. Indeed the correction would not be the same for channels 4 and 10. But one can go from one channel to another with a time translation. As a conclusion, the FIR formalism does not apply in such a case, and

(a) Without correction for the cross-talk

(b) With correction for the cross-talk
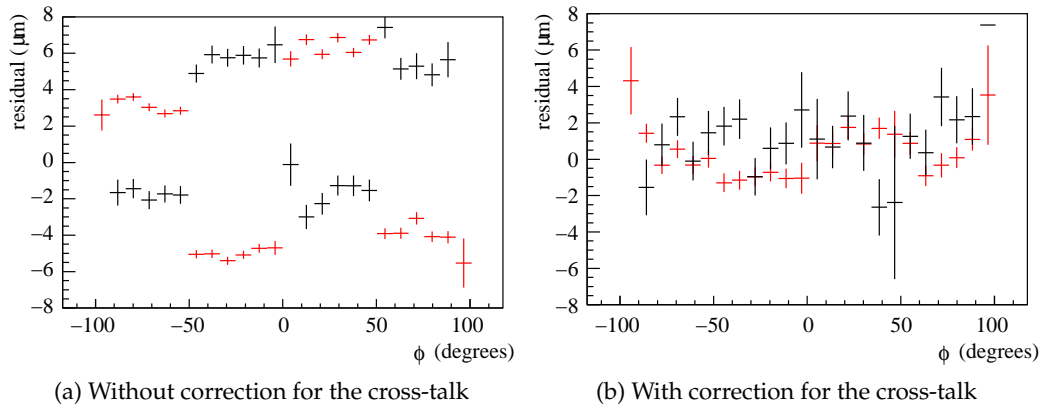
Figure 3.22: Mean value of the track residuals in a r sensor after having applied the alignment procedures [36, 37]. The red and black colors are correlated to Fig. 3.23. Figure taken from Ref. [35].
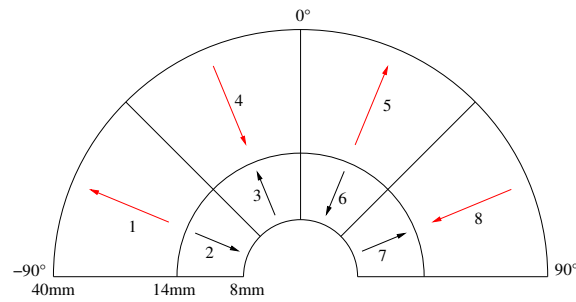
Figure 3.23: Sketch of the readout order of an r sensor. The numbers indicate the readout order, the arrows shows the readout the directions. The red and black colors are correlated to Fig. 3.22.

therefore the way to invert the cross-talk coefficients into FIR coefficients is unknown.

3.6 Conclusion

This analysis was an attempt for a systematic study of the cross-talk occurring along the VeLo acquisition chain. Different ways of computing the cross-talk have been considered, but the chosen method is the mean amplitude of the aggressed channel divided by the mean aggressor channel amplitude, an estimator which has shown to be robust. Looking at the test beam data, we conclude that a symmetric FIR filter (i.e. acting on both time direction) is required to cope with the non-zero backward time cross-talk. This model has been implemented in a software tool available to the collaboration.

Test-beam data cross-talk

The cross-talk in both ACDC test-beam datasets has been measured, the readout contribution is highly variable but is of the order of -5% for ACDC2 and $+8\%$ for ACDC3. This higher value for ACDC3 is due to the removal of the cable compensations. The sensor contributions are of the order of 2% and 3% for ACDC2 and ACDC3 respectively. ACDC3 ϕ sensors show an outer to inner cross-talk of the order of 5 to 7%; this phenomenon is not visible in ACDC2 data, but the sensor thicknesses in the two test beams are different.

FIR correction

The FIR filter is a digital correction applied on the data to correct for the cross-talk occurring in the readout. It is software based and can therefore be quickly adapted to correct for a noisy link or an increased radio frequency pickup. FIR filter corrections can be applied with a per TELL1 or with a per analog cable granularity, the latter showing slightly better results. A FIR of orders 3 is the minimum to correct for the main cross-talk sources. An order 5 FIR is safer and keeps the TELL1 resource needs to a reasonable level. After correction, the cross-talk distribution at the chip channel level decreases from $(4 \pm 3)\%$ to $(1 \pm 2)\%$ for r sensors. The ϕ sensors are less responsive to the correction, going from $(2 \pm 4)\%$ to $(1 \pm 3)\%$. Moreover, the cross-talk correction has a clear impact on the track residuals, especially on r sensors, where these residuals can be reduced to $2~\mu$m.

Chapter 4

The LHCb Monte Carlo

THE Monte Carlo technique [38] is widely used in particle physics and other fields necessitating simulations based on random numbers. It consists in modeling a phenomenon with a set of probability density functions (pdf) and generating random variables following these pdfs. At its very base, the Monte Carlo approach requires a random number generator able to produce a uniform distribution in a finite range, e.g. $[0, 1]$. The simplest generator is based on the action of the modulo operator, but, as the need of longer generator period and better randomness increased, more sophisticated algorithms appeared, among which the Mersenne Twister generator [39] used in Sec. 6 of the present document. Once a flat random distribution is available, one can generate any multi-dimensional distribution, e.g. with the accept-reject methods, the simplest, but also the less efficient. Thus, provided that the distributions are normalized, they are suitable to describe any physical process one would like to model. Note that random number generators always require one or more numbers as seed to start their sequence, and that once the seed is fixed, the sequence looks random but is fully predictable and repeatable.

This chapter begins with a description of the simulation framework and details the successive applications used in the simulation steps (Sec. 4.1). We then describe in Sec. 4.2 the algorithms used in the event reconstruction: the tracking, the vertexing, the triggers and the tagging algorithms. Finally we give in Sec. 4.3 a technical reference of the different datasets used in the present analysis.

4.1 Simulation framework

Most of the LHCb software is based on the `Gaudi` framework [40], which is implemented in `C++` (an object-oriented programming language). `Gaudi` monitors a linear succession of algorithms, which define the program flow, and also provide an interface to the so-called Transient Event Store (TES). The latter is a memory buffer, cleared on every event (hence the name), and allows the algorithms to share objects along the simulation or the reconstruction process.

We describe hereafter the successive steps involved in the LHCb event simulation and reconstruction process.

Event generation: The generation of the pp collision is handled by `Pythia` [41], which simulates the QCD (Quantum Chromo Dynamic) effects occurring in the collision process as well as the parton shower and hadronization steps. Its outcome is a set of particle four-vectors. The decay of these particles is handled by `EvtGen` [42].

The latter is configured with a set of allowed decays and their associated branching fractions, and generates the kinematics of the decay products. In addition, the recent generation of simulated data also benefits from `PHOTOS` [43], which is used by `EvtGen` and generates radiative photons. All the programs mentioned here are integrated and monitored in the LHCb framework through `Gauss` [44], a `Gaudi` based application.

Detector simulation: The particles are propagated through the magnetic field and the detector elements (e.g. the calorimeters). The interaction with the material (energy deposition and multiple scattering) is simulated by `Geant 4` [45]. This phase of the simulation is also handled by `Gauss`. The output is mainly a large ($\mathcal{O}(10^3)$) number of simulated particles, described with their hits and their energy deposition in the detector.

Digitization: Running the output of the simulation through `Boole` [46] allows to emulate the detector response. Hits are added to the sub-detectors to simulate spillover events (residual hits from a previous events, due to a capacitance which has been recently triggered). Then the energy depositions are converted into electronic responses. Finally, the L0 hardware-based trigger is simulated. After this, the simulated data are supposed to be identical to the real one, and the subsequent software makes no difference between the two cases (except of course for the Monte Carlo truth part).

Online processing: Real or simulated data are then processed by `Moore` [47], the LHCb HLT software[1]. It will be run in a dedicated online farm and, in case of a positive trigger decision, the event will be written on disk and made available for the offline processing.

Reconstruction: The first stage of the offline reconstruction is performed by the `Brunel` application [48]. It consists of track reconstruction and particle identification. The latter consists of building for each track a likelihood function based on information from the RICH detectors, the calorimeters and the muon chambers.

Physics analysis: At this stage, one can test the characteristics of a specific decay channel reconstruction. This is done through the `DaVinci` application [49], which allows the user to define a series of algorithms and cuts to be applied on every event. These are done mostly through toolkits which, being used by many users, ease the study of the reconstruction systematics.

Distributed analysis: The LHC expects 15 `Petabytes` of data per year. To handle this, the LCG (LHC Computing Grid) [50], or simply the grid, has been developed. It allows the data from the LHC experiments to be distributed around the globe, according to a hierarchical model with four levels. A primary backup will be recorded on tape at CERN, the "Tier-0" center of LCG. After initial processing, the data will be distributed to a series of Tier-1 centers, which are large computer farms. These resources will then be made available on demand for the Tier-2 and 3 centers, the former being computing centers consisting of one or several collaborations, the latter are local clusters, e.g. in a University Department. The job monitoring on the grid is performed by `DIRAC` (Distributed Infrastructure with Remote Agent Control) [51]. Eventually, on the very top of this software pyramid is `Ganga` [52], a

---

[1]At the time this document is written, the `Moore` application is going through a re-tuning process. This is why we dot not report its efficiency in the present analysis.

python program that provides functions such as splitting a job into multiple sub-jobs and their submission on various processing platforms, among which, a local computer and the grid.

4.2   Event reconstruction

We give in this section a short description of the main algorithms used for the event reconstruction.

4.2.1   Track reconstruction

The track reconstruction aims to combine a set of detector hits into a parametrized curve describing a particle trajectory. This is a challenging task due to the large track multiplicity $\mathcal{O}(10^2)$ of a typical event. The tracking is of major importance in LHCb and plays a central role in the association of particle identification objects (e.g. RICH rings and calorimeter clusters) with the measured particles.

The LHCb offline tracking is based on a fitting procedure, which is used throughout the pattern recognition. The latter aims at assigning the correct clusters to the tracks, while the former accurately determines the tracks parameters. A detailed description of the LHCb track finding algorithms can be found in Ref. [53] and the latest tracking performance is summarized in Ref. [54].

Track fitting

The track fitting is based on the Kalman formalism [55], a method mathematically equivalent to a χ^2 minimization. It allows the computation of the track parameters and their covariances, starting from a small number of clusters and progressively adding hits to it. It computes the track χ^2 in an iterative way, avoiding a global fit which would require large matrix inversions.

The very first step is to find a seed, that is, a collection of three aligned hits. The track is then extrapolated (generally in the forward direction, but the fit procedure can also be applied backward). It seeks hits lying in a cone, which aperture accounts for the possible multiple scattering; should a hit lie inside the cone, it is added to the track parametrization and the new χ^2 is computed. After all the possible hits have been found, the track is smoothed by removing outlier clusters and re-fitted.

Pattern recognition

The pattern recognition task is to find all the possible tracks while keeping a low ghost rate. The algorithms differ depending on the type of track they reconstruct. LHCb distinguishes five types of tracks (Fig. 4.1) depending on the detector region they go through:

VeLo tracks are reconstructed only in the VeLo and are crucial for the reconstruction of the primary vertices. They also represents the only information available on the backward part of the events.

T tracks are reconstructed in the T stations only and used in the RICH 2 reconstruction.

Upstream tracks are made of VeLo tracks that have been successfully extrapolated to the TT, but not further. As they do not go through the magnet, their momentum
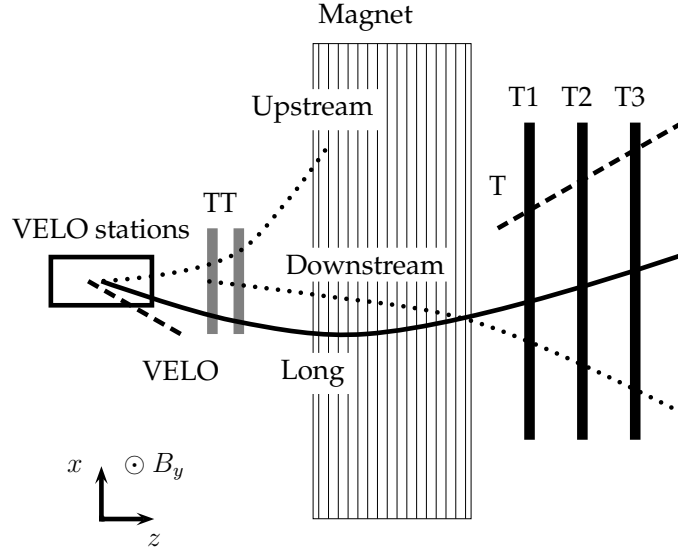
Figure 4.1: Sketch of the different track types in the bending plane xz.
Figure taken from Ref. [56].

is poorly measured. They are used mainly for the RICH 1 reconstruction and for
specific studies (like the one presented in Chapter 7).

Long tracks are reconstructed from an upstream track and a T track matched across the
   magnetic field. Hence, their momentum and other track parameters are accurately
   known. This makes them the most useful for physics analysis. The efficiency of the
   pattern recognition for long tracks is shown in Fig. 4.2.

Downstream tracks only traverse the TT and T stations and are not associated to a VeLo
   track. They allow the reconstruction of K_S^0 decays outside the VeLo sensitive region.

   We quote here the results from Ref. [54]. The tracking software gives for long tracks a
reconstruction efficiency of 91.4% with a ghost rate of 14.6%. For these tracks a momen-
tum resolution of 4.2 per mille is obtained after the Kalman filter based fit.
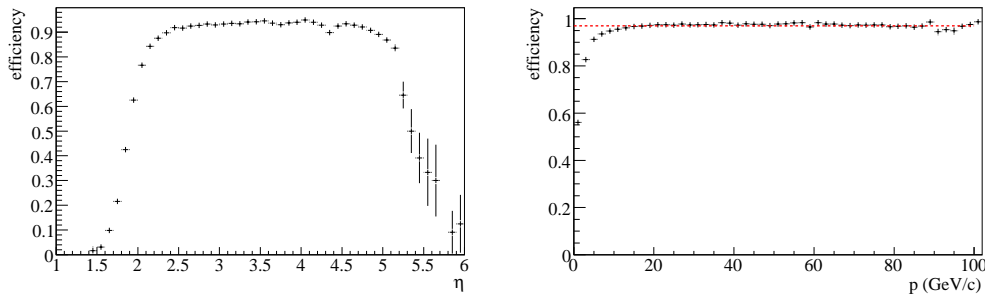
Figure 4.2: Track finding efficiency as a function of the pseudorapidity
η (left) and of the track momentum (right). Figures taken from Ref. [54].

4.2.2 Primary vertex fitting

The primary vertex fitting procedure differs depending on the context. Two reconstruction algorithms are used at the trigger level and are described in Ref. [57]. The third one is used for the offline reconstruction. It is based on two steps, the first one being the vertex seeding and the second one the vertex fitting. We summarize here the results from Ref. [58].

The seeding procedure is meant to provide initial positions for the vertices. These positions are assumed to lie on the z axis and are computed by analytical clusterization (rather than histogramming). All possible tracks are used for the reconstruction, namely VeLo, upstream and long tracks. The seeding starts with one cluster per track, computed as the point of closest approach on the z axis. Nearby clusters are then merged together. What remains at the end of this procedure is a small number of z clusters which will be used as seeds for the fitting.

The fit minimizes the weighted track impact parameters with respect to the vertex seed. Once a new position is found, the track with the largest Impact Parameter Significance (IPS) is discarded from the fit until all the tracks IPS are below 4.

The obtained resolutions for the primary vertices are $\sigma_z = 59$ μm, $\sigma_x = 8$ μm and $\sigma_y = 10$ μm. A systematic bias of about 10 μm is observed in the z direction. The averaged reconstruction efficiency is 97.5%. Typical track multiplicities for the primary vertices range from 40 to 120.

4.2.3 Triggers

At a luminosity of 2×10^{32} cm^{-2}s^{-1}, the rate of visible interactions, i.e. with at least two tracks that can be reconstructed by the detector, is expected to be slightly above 10 MHz. The rate at which one can write events on a physical medium, considering the available CPU resources, is around 2 kHz. This means only one event in five thousands can be kept on average. The aims of the LHCb trigger system is to apply this reduction factor in the most efficient way for the studied physics. In particular, the trigger will exploit the fact that b-flavoured hadrons are heavy and long lived.

The trigger system is two fold. The first level, Level-0 (L0), is implemented in hardware. Its main goal is to select high transverse energy particles using partial detector information. The output rate of the L0 trigger is 1 MHz. The selected events are then send to a dedicated PC farm of about 1800 CPU, which performs the event building and runs the High Level Trigger (HLT). The HLT exploits the L0 information and refines the selection, using for example the large impact parameters which are a characteristics of the B-decay products. We briefly describe here the two trigger levels. The interested reader can refer to [57, 59, 60].

The Level-0 trigger

The L0 trigger has three functions, namely the selection of high transverse energy (E_T) particles, neutral or charged, the rejection of complex or busy events, which are difficult and time consuming in terms of reconstruction, and the rejection of beam-halo events. Different sub-detectors are involved; each one processes its information and sends its decision to the Level-0 Decision Unit (L0DU) which gives a final L0 decision. The sub-detector parts are:

- The pile-up system, which is a component of the VeLo, made of two silicon disks positioned upstream with respect to the spectrometer ($z < 0$). It detects multiple interaction events with an efficiency of about 60% and a purity of 95%.

- The Level-0 calorimeter trigger, which outputs to the L0DU the highest-E_T hadron, electron, photon and π^0 candidates, the total E_T and the event multiplicity.

- The Level-0 muon trigger, which sends to the L0DU the highest-p_T candidates.

Figure 4.3 shows the L0 efficiencies on the different channels. We anticipate the results of Chapter 5, and give the L0 efficiency on $B_s \rightarrow D_s^- \pi^+$ events which are selected by the $B_s \rightarrow D_s h$ selection: $(44.2 \pm 0.3)\%$.
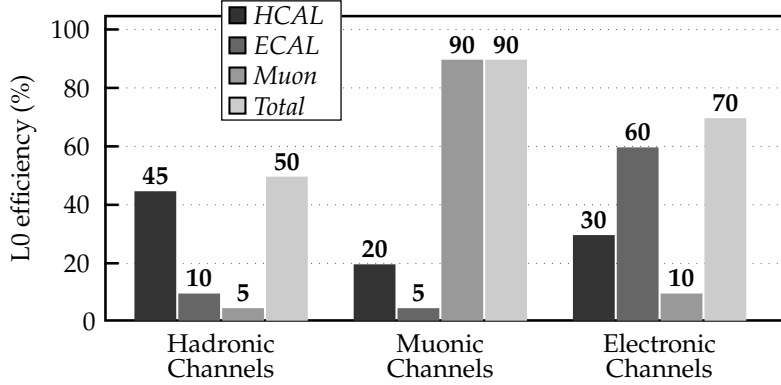
Figure 4.3: Typical Level-0 efficiencies for hadronic, muonic and electronic channels. Figure taken from Ref. [59].

The High Level Trigger

The HLT is based on the concept of independent 'trigger alleys', a set of dedicated streams that exploit and refine the Level-0 triggering information. The current implementation of the HLT has four alleys: muon, muon+hadron, hadron, and ECAL. After these, the sequence includes inclusive and exclusive selections. A schematic of the trigger decision flow is shown in Fig. 4.4. Only the inclusive and the exclusive selections perform a full reconstruction (tracking and vertexing) of the events. The inclusive selections typically reconstruct particles like D_s or J/ψ mesons, whereas exclusive selections reconstruct full decay chains. The $B_s \rightarrow D_s h$ selection presented in this document will become one of the HLT exclusive selections.

4.2.4   Flavour tagging

Flavour tagging aims to determine the flavour of a reconstructed B meson at its production time, namely, whether it contained a b quark or a $\bar{\text{b}}$ quark. This technique is required for most of the CP asymmetry measurements and flavour oscillation studies. The outcome of the tagging algorithms is simply one of the b, $\bar{\text{b}}$ or unknown states. Relevant variables to characterize its performance over a sample are the tagging efficiency ε_{tag}, that is, the probability that a B is tagged, the wrong tag fraction ω_{tag} and the tagging power ε_{tot}. They are defined from N_{R}, N_{W} and N_{U}, the number of correctly tagged, incorrectly tagged and untagged B mesons tags respectively:

$$\varepsilon_{\text{tag}} = \frac{N_{\text{R}} + N_{\text{W}}}{N_{\text{R}} + N_{\text{W}} + N_{\text{U}}} \ , \qquad \omega_{\text{tag}} = \frac{N_{\text{W}}}{N_{\text{R}} + N_{\text{W}}} \ , \qquad \varepsilon_{\text{tot}} = \varepsilon_{\text{tag}}(1 - 2\omega_{\text{tag}})^2 \ . \quad (4.1)$$
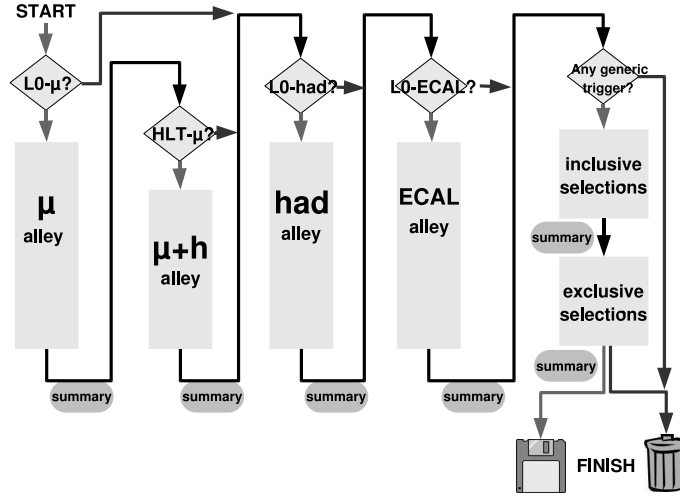
Figure 4.4: The High Level Trigger flow-chart. Figure taken from Ref. [59].
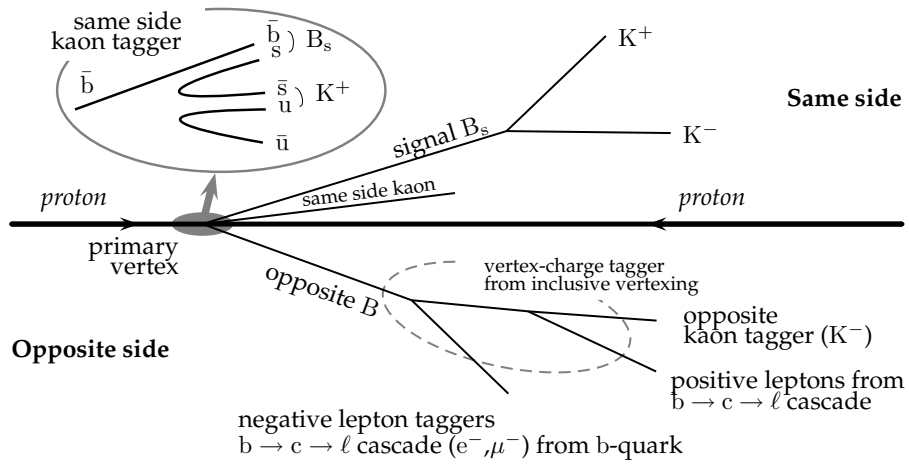
Figure 4.5: Illustration of the different sources of information available to tag the initial flavour of a signal B candidate, here $B_s \to K^- K^+$. Figure taken from Ref. [61].

Figure 4.5 sketches the different tagging techniques used in LHCb. We briefly give here a description of the different techniques. A detailed description of the tagging algorithms is given in Refs. [61, 62].

Single particle taggers of the opposite-side type use the charge of a lepton from the semileptonic b decay or of a kaon from the b $\rightarrow$ c $\rightarrow$ s decay chain to tag the flavour of the B meson.

Vertex charge taggers are based on an inclusive reconstruction of the opposite b-hadron secondary vertex.

Kaon same-side tagger: when a B_s mesons is produced, an additional $\bar{s}$ quark is available from the fragmentation and will form a charged kaon in about one half of the cases. The charge of this kaon can be used to infer the B_s flavour.

Pion same-side tagger: the same method as for the same-side kaon tag can be applied to the B_d, though the large pion combinatorics spoils the performance.

The tagging algorithms, when applied, test all the different taggers and, with neural nets, assign each of them a probability of being correct. The final outcome is a decision computed as a combined probability of all the taggers of being correct. We anticipate here the results of Chapter 5 and show in Table 4.1 the outcome of the neural-net decision algorithm on the selected $B_s \rightarrow D_s^- \pi^+$ events.

$\mathbf{B_s \rightarrow D_s^- \pi^+}$		selected	selected and triggered
ε_{tag}	(%)	65.97 ± 0.16	69.85 ± 0.23
ω_{tag}	(%)	36.64 ± 0.21	34.97 ± 0.29
ε_{eff}	(%)	4.73 ± 0.10	6.31 ± 0.17

Table 4.1: Tagging efficiencies, wrong tag fractions and total efficiencies for $B_s \rightarrow D_s^- \pi^+$ Monte Carlo events filtered with the $B_s \rightarrow D_s h$ selection, before and after the L0 trigger. Statistical errors are also shown.

4.3   Data sample and Monte Carlo technicalities

At the time this document is written, the detector performances are well established with the DC 2004 (DC04) data. During the year 2006, the collaboration started the DC 2006 (DC06) data generation. The latter, among other things, was based on a new magnetic-field map that more closely reflects measurements made in the real magnet; tilted detector elements[2]; and a lot of dead material like cables, cooling elements and frames were added to the description of the geometry.

The direct consequence of this more realistic description was an overall decrease in the reconstruction performance. On the other hand, the collaboration is constantly working on improving the reconstruction algorithms, which are still evolving rapidly.

We illustrate this phenomenon with the tagging performances quoted in Table 4.1. For the $B_s \rightarrow D_s h$ selection, the total efficiency is $\sim 3\%$ lower than the one it used to be

---

[2]It happens that the beam is not parallel to the floor and the heavy detector elements (e.g. the calorimeters) must be vertical, thus "tilted" by 3.2 mrad with respect to the z axis.

in the DC04 generation. This simply comes from the fact that the tagging algorithms did not yet accommodate to the very last version of the reconstruction (`Brunel` v31).

Table 4.2 shows the different DC06 Monte Carlo samples used in the analyzes of this document. All the datasets were analyzed with `DaVinci` v19r9, the `FlavorTagging` package v7r3p1 and some private algorithms.

Stat.	Decay	Event Type (see Ref. [63])	`Gauss`	`Boole`	`Brunel`	`DaVinci` stripping	`Brunel` reproc.
Samples used in the $B_s \to D_s h$ analysis (Chapter 5)							
1.8 M	$B_s \to D_s^- \pi^+$	13'264'001	v25r7	v12r10	v30r14	v19r7	v31r12
41 k	$\Lambda_b \to D_s^- p$	15'264'001	v25r7	v12r10	v30r14	–	–
34 k	$\Lambda_b \to \Lambda_c^- \pi^+$	15'164'001	v25r7	v12r10	v30r14	–	–
22 M	inclusive $b\bar{b}$	10'000'000	v25r7	v12r10	v30r14	v19r7	v31r17
1.7 M	min. bias events passing the L0 trig.	30'000'000	v25r7	v12r4	v31r10	–	–
Samples used in the analysis of the secondary interactions (Chapter 7)							
12.5 M	$pp \Rightarrow D_s X$	23'263'002	v25r7	v12r10	v30r14	–	–
700 k	min. bias events	30'000'000	v25r10	v12r10	v30r17	–	–

Table 4.2: List of the data samples used in this document. The amount of statistics and the different program versions used to generate the data are given. A *minimum-bias* sample contains at generator level all kinds of events that can occur in a pp-collision in natural proportions. An inclusive $b\bar{b}$ sample is made of events containing a $b\bar{b}$ quark pair.

Chapter 5

The $B_s \rightarrow D_s^- \pi^+$ offline selection
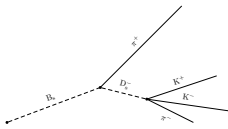
The decay channels $B_s \rightarrow D_s^- \pi^+$ will be used to extract the physics parameters Δm_s, $\Delta \Gamma_s$ and the wrong tag dilution. Simulation studies based on Monte Carlo samples show that a total of 155 k $B_s \rightarrow D_s^- \pi^+$ events are expected to be triggered (L0 only), reconstructed and selected in 2 fb^{-1} of data.

BEFORE any measurement can be done on the data, one needs to reconstruct and select the appropriate decays. We present in this chapter the selection of the $B_s \rightarrow D_s^- \pi^+$ decays[1]. The L0 trigger, the reconstruction of the tracks, the vertexing of the PV and the tagging algorithms are those described in Chapter 4. The $B_s \rightarrow D_s^- \pi^+$ Cabibbo-favored decay has a large branching fraction $(3.3 \pm 0.5) \times 10^{-3}$ and can therefore be reconstructed cleanly. It is thus expected that most of the LHCb measurements of (or limits on) B_s branching fractions will be normalized to the $B_s \rightarrow D_s^- \pi^+$ branching fraction. The $B_s \rightarrow D_s^- \pi^+$ decay is flavour specific (i.e. the B_s flavour at decay is known from the charge of its decay products), thus this high-statistics channel is well suited for Δm_s measurements. It can also be used to extract information on other B_s–$\overline{B}_s$ mixing parameters such as $\Delta \Gamma_s$ and $|q/p|$ (see Chapter 1.2.1). Moreover, it can be used as a control channel in time-dependent analysis of B_s decays to non flavour-specific modes, to measure the experimental dilutions due to imperfect flavour tagging and proper-time resolution.

The selection of the $B_s \rightarrow D_s^- \pi^+$ decay channel presented here is based on the LHCb public note [64], where the $B_s \rightarrow D_s^- \pi^+$ decay channel was studied together with the $B_s \rightarrow D_s^\mp K^\pm$ decay channel under the name of the $B_s \rightarrow D_s h$ selection. Yet the latter is based on the DC04 generation of data. We therefore adapted the $B_s \rightarrow D_s h$ selection to filter the $B_s \rightarrow D_s^- \pi^+$ decays and assess the performance with the DC06 data.

In this document we consider only the $K^- K^+ \pi^-$ final states of the D_s^- decay, including both non-resonant and resonant contributions ($\phi \pi^-$ and $K^{*0} K^-$) in an inclusive way. The mass and helicity angle cuts specific to the resonant intermediate states of the D_s decay are not used in the selection described here.

Figure 5.1 shows the topology of a $B_s \rightarrow D_s^- \pi^+$ decay. As the B_s has a lifetime of ~ 1.5 ps and an average momentum of ~ 135 GeV/c, it typically flies 1.1 cm before decay. This produces a displaced secondary vertex. The D_s meson has a lifetime of ~ 0.5 ps and therefore produces a displaced tertiary vertex.

---

[1]Unless specified otherwise, charge conjugate modes are implied throughout this document.
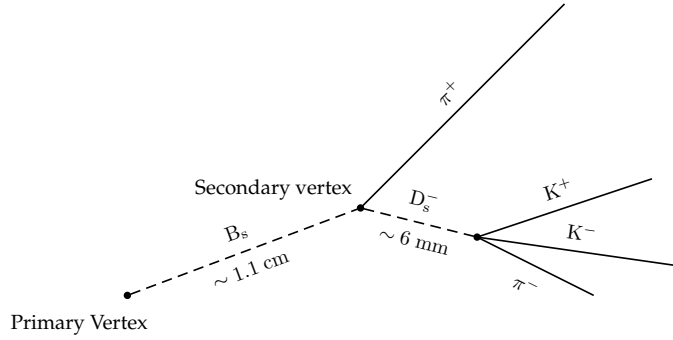
Figure 5.1: Topology of a $B_s \rightarrow D_s^- \pi^+$ decay.

The remainder of this chapter is organized as follows: Section 5.1 describes the variables used for the $B_s \rightarrow D_s^- \pi^+$ selection. Section 5.2 gives the yields and the background levels, based on explicit assumptions for the branching fractions. Finally, Section 5.3 outlines some properties of the selections like the mass resolution or the time distribution.

5.1 $B_s \rightarrow D_s^- \pi^+$ selection

The $B_s \rightarrow D_s^- \pi^+$ selection mainly relies on topology cuts and on PID requirements. Table 5.1 summarizes the values of the cuts and Figs. 5.3, 5.4, 5.3, 5.5 and 5.6 show for each variable, the distribution of the signal (black histograms) and of the $b\bar{b}$ background (dashed red histograms), both selected with looser cuts. All the histograms are arbitrarily normalized.

For comparison purposes, the only difference with the selection presented in Ref. [64] is the enlarged the mass window to account for the slightly worse mass resolution observed. And the PID cuts on all the pions and kaons to reject mis identification coming from protons and muons. The former greatly reduce the $\Lambda_b \rightarrow D_s^- p$ and $\Lambda_b \rightarrow \Lambda_c^+ \pi^-$ background.

K and π selection

To ensure a good momentum resolution, every reconstructed long track[2] momentum (p) has to be larger than 2 GeV/c (see Fig. 5.2a). To avoid selecting prompt track, an IPS bigger than three with respect to each primary vertex is also required (see Fig. 5.2b) (At the luminosity of 2×10^{-32} cm^{-2}s^{-1} about 35% of the events have two or more primary vertexes).

The other cuts are related to the particle identification (PID). They are based on a quantity called Delta Log Likelihood (DLL). For a given reconstructed particle p, $\Delta\mathcal{L}_{p_1 p_2}$ quantifies the hypothesis that this particle is of type p_1 rather than p_2. By convention, such cuts are applied only if the track is assumed to be of type p_1, e.g. $\Delta\mathcal{L}_{pK}$ will not be applied to test a kaon, but $\Delta\mathcal{L}_{Kp}$ will be used to test whether the kaon is kept.

The track will be used as a kaon if it satisfies $\Delta\mathcal{L}_{K\pi} > 2$ (see Fig. 5.3a), and it will be used as a pion if $\Delta\mathcal{L}_{\pi K} > -10$ (see Fig. 5.3b). Hence, a track can be tagged simultaneously as both π and K. Additional PID cuts are applied to ensure that the particle is neither a

---

[2]Long tracks are reconstructed tracks that leave at least 3 r and 3 ϕ associated clusters in the VeLo and at least 1 x and 1 stereo hit in each the tracking stations (T1-T3).

muon nor a proton (see Table 5.1) their distributions are shown in the four bottom plots of Fig. 5.3.

The bachelor is any pion whose transverse momentum [3] (p_T) is larger than 600 MeV/c (see Fig. 5.4).

(a) Momentum   (b) Impact parameter significance

Figure 5.2: Distributions of the different variables for all the particles from the decay (pions and kaons). The red dashed histograms are background events, the black, continuous histograms are signal events. The cut values are shown.

D_s reconstruction

To reconstruct $D_s^- \to K^+K^-\pi^-$ candidates, every combination of $K^+K^-\pi^-$ tracks is tested. The transverse momentum of each D_s daughter (p_T) is required to be larger than 300 MeV/c (see Fig. 5.5a). Only D_s candidates with a vertex fit χ^2 lower than 16 are kept (see Fig. 5.5b), they must also have a transverse momentum larger than 2 GeV/c (see Fig. 5.5c) and should have an invariant mass within 21 MeV/c^2 of the D_s mass (see Fig. 5.5d). Finally, to ensure a well displaced secondary vertex, a four sigma significance is required on the D_s impact parameter (see Fig. 5.5e) with regard to each primary vertex.

B_s reconstruction

A B_s is reconstructed out of a D_s and a bachelor track. The following conditions must be fulfilled by each candidate: the value of the reconstructed invariant mass should be within a ±70 MeV/c^2 window around the B_s mass (see Fig. 5.6a). The χ^2 of the B_s decay vertex must be smaller than 9 (see Fig. 5.6b). To ensure that the reconstructed B_s does come from a primary vertex, the smallest impact parameter significance with respect to each primary vertex has to be smaller than 4 (see Fig. 5.6c). The flight distance significance from this primary vertex must be larger than 12 (see Fig. 5.6d) and the cosine of the B_s pointing angle has to be larger than 0.99997 (see Fig. 5.6e). Pointing angle here means the angle between the particle's reconstructed momentum and its flight direction, defined by the B_s decay vertex and its associated primary vertex. Moreover, it is required that the reconstructed vertex of the B_s is more upstream than the D_s vertex, i.e. $z_{D_s} > z_{B_s}$ (see Fig. 5.6f).

---

[3]p_T is computed with respect to the z axis ($p_T \equiv \sqrt{p_x^2 + p_y^2}$) and not to the B-meson direction. This approximation is possible because the B meson tends to fly along the z axis.
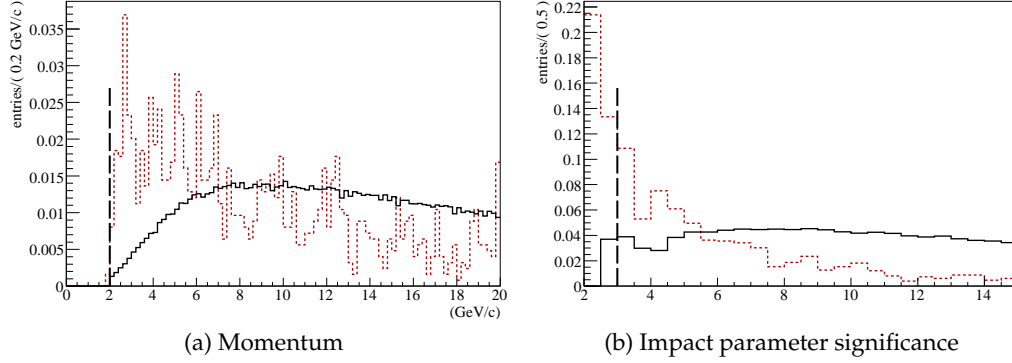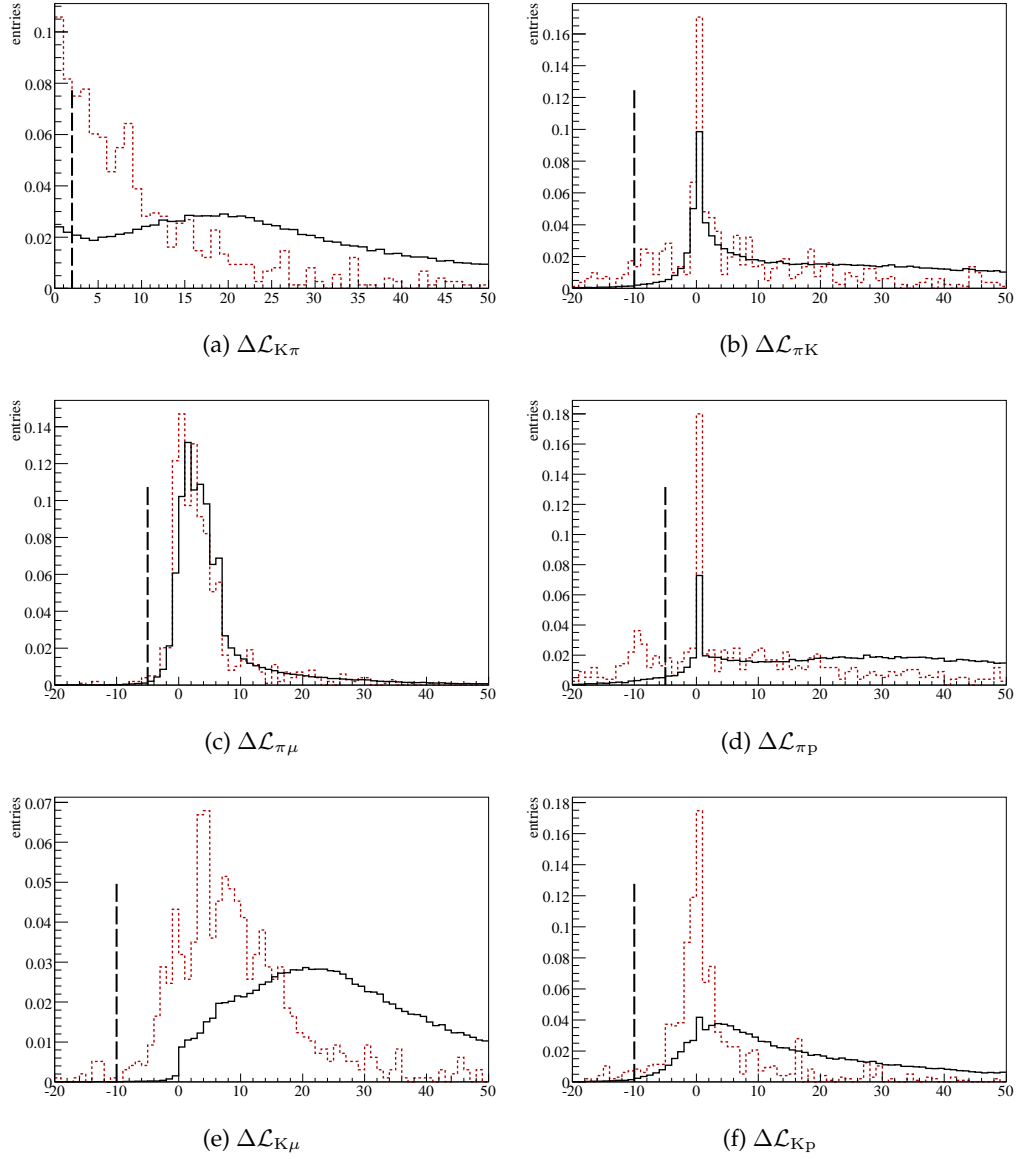
Figure 5.3: Distributions of the different variables for all the particles from the decay (pions and kaons). The red dashed histograms are background events, the black, continuous histograms are signal events. The cut values are shown.
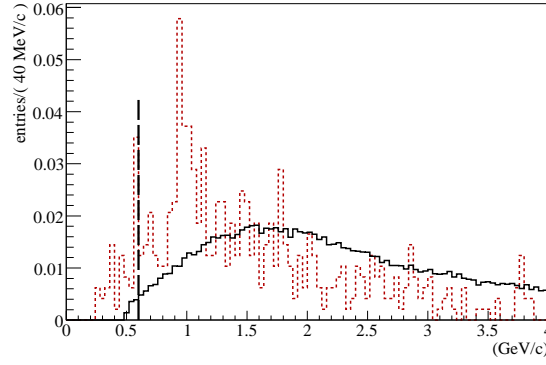
Figure 5.4: Transverse momentum of the bachelor. The red dashed histogram is background events, the black, continuous histogram is signal events. The cut value is shown.

Cuts on all the kaons and pions		
p	>	$2 \text{ GeV}/c$
IPS	>	3
$\Delta\mathcal{L}_{\pi K}$	>	-10
$\Delta\mathcal{L}_{\pi p}$	>	-5
$\Delta\mathcal{L}_{\pi\mu}$	>	-5
$\Delta\mathcal{L}_{K\pi}$	>	2
$\Delta\mathcal{L}_{Kp}$	>	-10
$\Delta\mathcal{L}_{K\mu}$	>	-10
Cuts on the bachelor		
$p_{T,h}$	>	$600 \text{ MeV}/c$
Cuts on the D_s and its daughters		
p_T for the D_s daughters	>	300
p_{T,D_s}	>	$2 \text{ GeV}/c$
IPS_{D_s}	>	4
D_s vertex χ^2	<	16
ΔM_{D_s}	$\pm$	$21 \text{ MeV}/c^2$
Cuts on the B_s		
IPS_{B_s}	<	4
B_s vertex χ^2	<	9
ΔM_{B_s}	$\pm$	$70 \text{ MeV}/c^2$
FS_{B_s-PV}	>	12
$z_{D_s} - z_{B_s}$	>	0 mm
$\cos\theta_p$	>	0.99997

Table 5.1: Cut values for the $B_s \rightarrow D_s^- \pi^+$ selection.

(a) Transverse momentum of the D_s daughters

(b) χ^2 of the D_s vertex

(c) Transverse momentum of the D_s

(d) D_s mass

(e) Impact parameter significance of the D_s
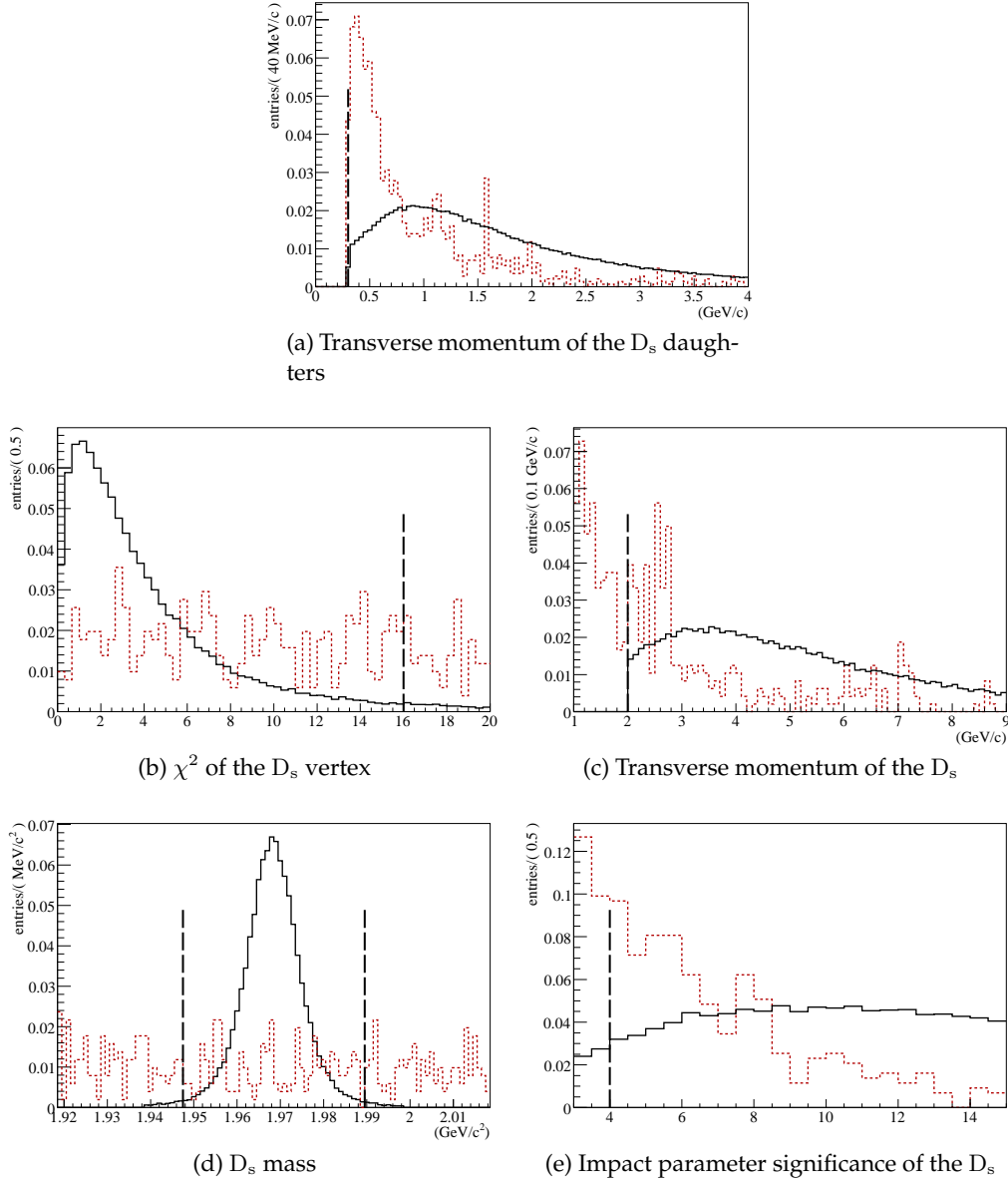
Figure 5.5: Distributions of the variables related to the D_s and its decay products. The red dashed histograms are background events, the black, continuous histograms are signal events. The cut values are shown.

(a) B_s mass

(b) Vertex χ^2

(c) Impact parameter significance

(d) Flight significance

(e) Cosine of the angle between the B_s flight direction and line joining the B_s and the D_s vertices

(f) Signed distance in z between the D_s vertex and the B_s vertex ($z_{D_s} - z_{B_s}$)

Figure 5.6: Distributions of the B_s related variables. The red dashed histograms are background events, the black, continuous histograms are signal events. The cut values are shown.

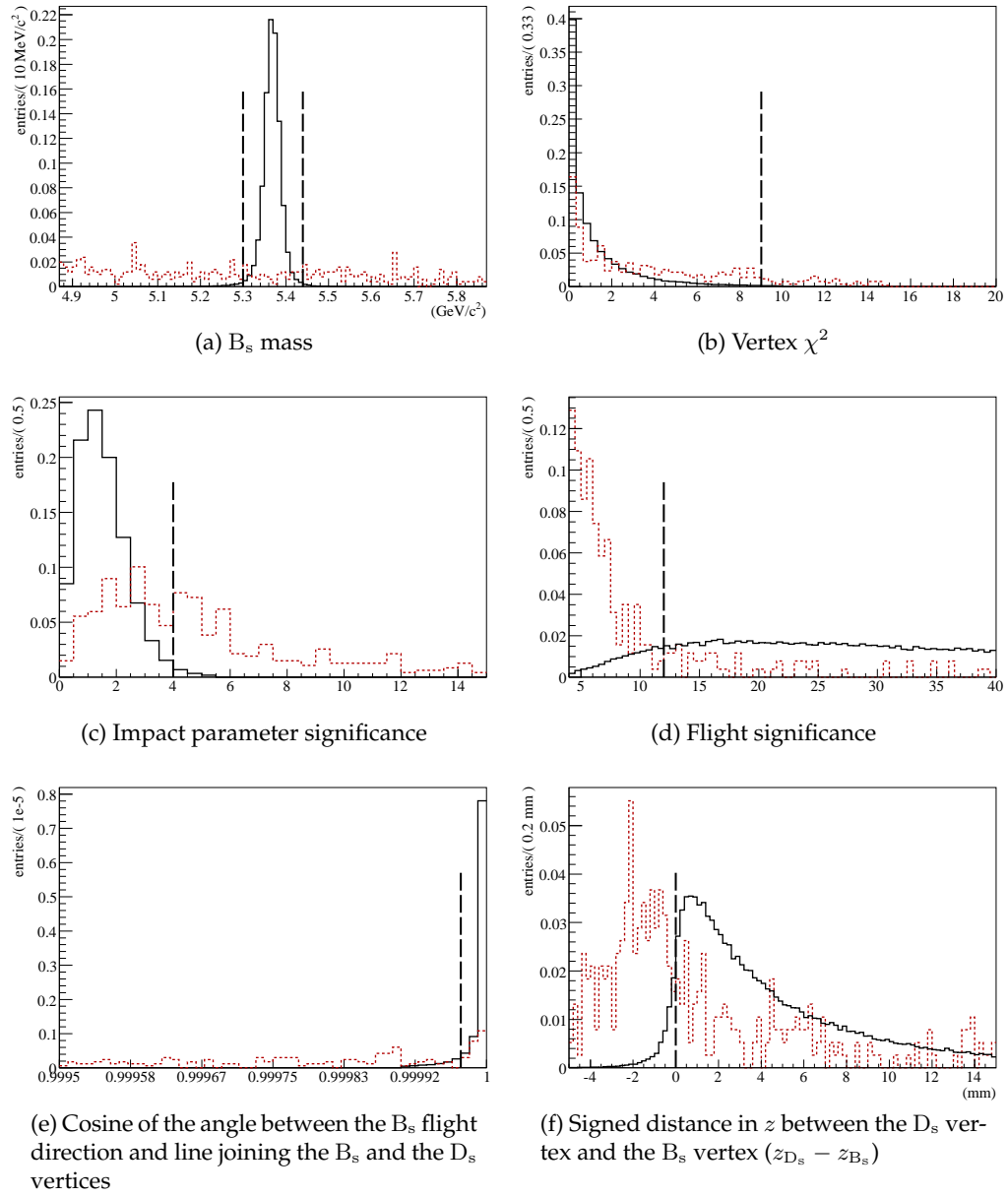### 5.1.1   Signal efficiencies and event yields

We summarize in Table 5.2 the selection efficiency obtained on a statistics of 1.8 M signal events. A breakdown of the efficiency, based on DC04 data, can be found in [64].

	efficiency (%)
offline selection	0.927 ± 0.004
L0 trigger for selected events	44.2 ± 0.3
offline + L0-trigger selection	0.409 ± 0.002

Table 5.2: Efficiency of selection and L0 trigger on $B_s \rightarrow D_s^- \pi^+$ signal events. The quoted errors are statistical.

5.2   Signal yield and background level

Since the yield estimates rely on the b-hadron production cross section and on branching fractions, we give below an explanation of the values used for this calculation.

Table 5.3 shows the different branching fractions used for the present selection. Most values come from the PDG [65]. However, the branching fractions for $B_s \rightarrow D_s^- \pi^+$ does not come from the PDG and is explained below:

Two measurements of $\mathcal{B}(B_s \rightarrow D_s^- \pi^+)$ are now available. The first one comes from the CDF collaboration, which has measured the following ratio [66]:

$$\frac{\mathcal{B}(B_s \rightarrow D_s^- \pi^+)}{\mathcal{B}(B_d \rightarrow D^- \pi^+)} = \frac{N(B_s)}{N(B_d)} \frac{\varepsilon(B_d)}{\varepsilon(B_s)} \frac{f_d}{f_s} \times \frac{\mathcal{B}(D^- \rightarrow K^+ \pi^- \pi^-)}{\mathcal{B}(D_s^- \rightarrow \phi \pi^-) \mathcal{B}(\phi \rightarrow K^- K^+)} \quad (5.1)$$

$$= 1.13 \pm 0.08(\text{stat.}) \pm 0.05(\text{syst.}) \pm 0.15(\text{bf}) \pm 0.17(\text{pf}) ,$$

where the two last errors come from the D, D_s and ϕ branching fractions and the hadronization fractions respectively. f_s stands for the B_s production fraction, $\varepsilon(B_s)$ for the CDF's detection efficiency of the $B_s \rightarrow D_s^- \pi^+$ decay and N is the number of selected events. Multiplying this value by $\mathcal{B}(D_s^- \rightarrow \phi \pi^-) = (4.4 \pm 0.6)\%$ [66] and $\mathcal{B}(B_d \rightarrow D^- \pi^+) = (2.83 \pm 0.17) \times 10^{-3}$ [66] and combining all the errors quadratically, we get:

$$\mathcal{B}(B_s \rightarrow D_s^- \pi^+)_{\text{CDF}} \times \mathcal{B}(D_s^- \rightarrow \phi \pi^-) = (1.40 \pm 0.27) \times 10^{-4} . \quad (5.2)$$

The second measurement comes from a recent Belle study [67, 68] and is a direct estimate of the branching fraction:

$$\mathcal{B}(B_s \rightarrow D_s^- \pi^+) = \{3.40^{+0.33}_{-0.31}(\text{stat.}) \pm 0.40(\text{syst.}) \quad (5.3)$$
$$\times \, ^{+0.46}_{-0.45}(f_s) \, ^{+0.33}_{-0.28}(\mathcal{B}(D_s \rightarrow \phi \pi))\} \times 10^{-3} .$$

As for the CDF value, we compute:

$$\mathcal{B}(B_s \rightarrow D_s^- \pi^+)_{\text{Belle}} \times \mathcal{B}(D_s^- \rightarrow \phi \pi^-) = (1.53 \pm 0.27) \times 10^{-4} . \quad (5.4)$$

The two values given in (5.2) and (5.4) are independent because the correlation between the original measurements (5.1) and (5.3) has been removed. They are compatible within 1 sigma. Their weighted average divided by $\mathcal{B}(D_s^- \rightarrow \phi \pi^-) = (4.5 \pm 0.4)\%$ [65], leads to:

$$\mathcal{B}(B_s \rightarrow D_s^- \pi^+) = (3.3 \pm 0.5) \times 10^{-3} .$$

Which is the value used in this document.

Decay	value	estimated as	Ref.
$B_s \to D_s^- \pi^+$	$(3.3 \pm 0.5) \times 10^{-3}$	see text	[66, 68, 65]
$B_d \to D^- \pi^+$	$(2.68 \pm 0.13) \times 10^{-3}$		[65]
$\Lambda_b \to D_s^- p$	$(3.3 \pm 0.5) \times 10^{-3}$	$= \mathcal{B}(B_s \to D_s^- \pi^+)$	
$\Lambda_b \to \Lambda_c^+ \pi^-$	$(8.8 \pm 3.2) \times 10^{-3}$		[65]
$\Lambda_c^+ \to pK^- \pi^+$	$(5.0 \pm 1.3) \times 10^{-2}$		[65]
$D_s^+ \to K^+ K^- \pi^+$	$(5.3 \pm 0.8) \times 10^{-2}$		[65]
$D^+ \to K^- \pi^+ \pi^+$	$(9.51 \pm 0.34) \times 10^{-2}$		[65]
f_s	$(10.8 \pm 1.2) \times 10^{-2}$		[65]
f_d	$(40.0 \pm 1.0) \times 10^{-2}$		[65]
f_{Λ_b}	8%		[69]

Table 5.3: Branching fraction values assumed in the calculation of the yields and background levels. f_s, f_d and f_{Λ_b} refer to the B_s, B_d and Λ_b production fractions respectively.

Yield computation

The yield for a decay channel i is computed as

$$Y_i = L_{\text{int}} \times \sigma_{b\bar{b}} \times 2 \times f_i \times \text{BR}_i^{\text{vis}} \times \varepsilon_i \ ,$$

where $L_{\text{int}} = 2$ fb^{-1} is the integrated luminosity after 10^7 seconds of data taking at an average instantaneous luminosity of 2×10^{32} cm^{-2}s^{-1}, and $\sigma_{b\bar{b}} = 500$ μb is the expected $b\bar{b}$ production cross-section. The factor 2 stands for the two b quarks and f_i is the probability for a b-quark to hadronize in the involved B meson. BR_i^{vis} is the total visible branching fraction –defined as the product of all the different branching fractions involved in the decay chain– of the decay channel i. ε_i is the selection efficiency.

Thus, assuming 2 fb^{-1} per year, we compute a yield after the L0 trigger of:

$$\text{Yield}(B_s \to D_s^- \pi^+) = 155 \text{ k} \pm 1 \text{ k (stat.)} \pm 37 \text{ k (syst.)} \ , \tag{5.5}$$

where the systematical error is computed based on the errors shown in Table 5.3 (note that it does not include an error on $\sigma_{b\bar{b}}$).

5.2.1 Background from minimum-bias events

The $B_s \to D_s^- \pi^+$ offline selection has been tested on 1.74 M events passing the L0 trigger. One background event, reconstructed with tracks coming from b-hadron decays, happens to be selected in the wide mass window, which represents a raw selection rate of 0.63 Hz. We show in Fig. 5.7 the topology of this candidate and the particles used to reconstruct the $B_s \to D_s^- \pi^+$ decay. Two errors are done in the reconstruction: the bachelor track is mis-identified and D_s is formed as a random track combination. We checked, for all the variables of the $B_s \to D_s^- \pi^+$ selection, that it is not possible to reject this very event without loosing a substantial part of the signal. Should this event not be an unlucky statistical fluctuation, the selection will have to be tightened. We have considered that more statistical significance is required to make a decision, moreover, the event contains true b hadrons and should therefore be included in the estimation of the $b\bar{b}$ background (see below).
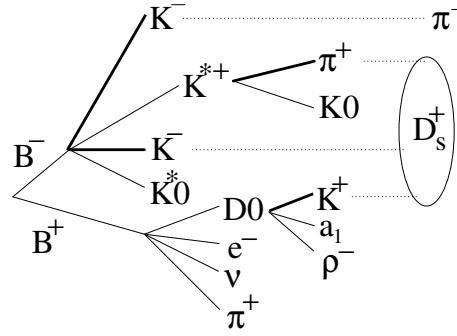
Figure 5.7: Sketch of the particles used to reconstruct the selected minimum bias event.

5.2.2   Background from $b\bar{b}$ events

Events containing a $b\bar{b}$ quark pair are expected to be the main source of background since they have a displaced vertex and are thus easily accepted by the triggers. A total of 22 M of $b\bar{b}$ events were analyzed, corresponding to 96 seconds at the nominal luminosity of 2×10^{32} cm^{-2}s^{-1}. To cope with the fact that very few $b\bar{b}$ events pass the selection, we artificially increase the statistics by enlarging the B_s mass window to ± 500 MeV$/c^2$ (whereas the nominal cut is ± 70 MeV$/c^2$). Moreover, no trigger decision is required during the selection. We assume that this is conservative for the computation of the B/S ratio.

In this sample, 13 $B_s \rightarrow D_s^- \pi^+$ decays were correctly identified from the $b\bar{b}$ sample. Given the branching fractions used for the generation and the total signal efficiency reported in Eq. 5.2, 10 ± 3 events were expected.

In the inclusive $b\bar{b}$ sample 43 non-signal events are reconstructed as $B_s \rightarrow D_s^- \pi^+$ and pass the loose mass window. These candidates are described below using the information from the Monte Carlo truth:

38 of them are reconstructed from a peaking background and do actually not pass the tight mass window cut (± 70 MeV$/c^2$), hence they are not counted in the B/S estimation. However, as some of the mass distribution of the involved decay channels have tails which may spread in the signal window, we separately estimated their B/S. Most of these studies have been reported in Ref. [64] and we will assume that the values are still valid. 3 events are specific decays which happens to fall in the signal mass window and we include these in the specific background computation, we mention each of them in the list below. The two last events are due to a wrong reconstruction.

13 are reconstructed from a $B_s \rightarrow D_s^{(*)-} \rho^+$ decay. One of them falls in the signal window and is counted in the B/S value. The $B_s \rightarrow D_s^- \rho^+$ channel has been studied separately, its B/S value is lower than 2×10^{-3} at 90% CL [64].

11 events come from a $B_s \rightarrow D_s^{*-} \pi^+$ decay. This channel has been studied on a dedicated sample. Its B/S contribution is lower than 4×10^{-3} at 90% CL [64].

4 events, 2 containing a $B_d \rightarrow D^- \pi^+$ decay and 2 others with a $B_d \rightarrow D^{*-} \pi^+$ decay were reconstructed. The mis-identification of one of the pion D daughters as a kaon creates a fake D_s. One of the $B_d \rightarrow D^- \pi^+$ decay passes the tight mass window cut and is included in the B/S estimation.

4 events come from a mis-reconstructed $\Lambda_b \to (\Lambda_c^+ \to pK^-\pi^+)\pi^-$ decay. Though its mass ($5624 \text{ MeV}/c^2$) is much larger than the $5419.6 \text{ MeV}/c^2$ upper value of the B_s mass window, the mis-identification of a true proton as a kaon brings the reconstructed B_s mass in the tight mass requirement. One out of these four fall into the tight mass window. It has been included in the estimation of the $b\bar{b}$ background.

4 events come from $B_s \to D_s^{(*)-}\ell^+\nu$ modes. Again, a dedicated study shows a B/S upper limit of 2×10^{-2} at 90% CL [64].

2 events are $B_d \to D_s^{(*)-}\pi^+$ decays. Though it is not possible to ignore this channel *a priori*, we assume it has the same reconstruction efficiency as $B_s \to D_s^-\pi^+$ events except for the B_s mass window cut. Hence, a Gaussian mass distribution with the same sigma ($17 \text{ MeV}/c^2$) as the signal mass spectrum, will have 16% of its surface in the tight B_s mass window. The $B_d \to D_s^-\pi^+$ and $B_d \to D_s^{*-}\pi^+$ branching fractions, according to Ref. [65] are respectively 200 and 100 times lower than the $B_s \to D_s^-\pi^+$ branching fraction. This allows to compute a limit to the B/S contribution of 1.6×10^{-3}.

1 event comes from a $B_s \to D_s^{*-}K^{*+}$ decay. The two excited states, as well as the bachelor mis-identification add up to lower the reconstructed B_s mass. Hence we assume this event can be discarded from the B/S computation.

1 event comes from a $B_s \to D_s^- a_1^+$ decay, where the bachelor has been reconstructed out of the $a_1^+ \to \rho^+\pi^0$ charged pion daughter. Hence the reconstructed B_s mass misses at least twice the π^0 mass.

1 event comes from a $B_d \to D^{*-}\rho^+$ decay. The missed π^0 and the B_d mass together lower the reconstructed B_s mass.

2 events are not associated to any physical backgrounds. One is a ghost, i.e. a random association of hits which forms a track; one is a so-called combinatorial event, i.e. a random association of tracks which happens to be reconstructed as a $B_s \to D_s^-\pi^+$ decay.

Thus 2 combinatorial (downscaled by the mass ratio) and $1(\Lambda_b \to \Lambda_c^+\pi^-) + 1(B_d \to D^-\pi^+) + 1(B_s \to D_s^{*-}\rho^+) = 3$ specific events (which are not downscaled by the mass ratio) remain for the B/S computation. Separate limits are given. For the combinatorial background contribution, the central value is 0.016 ± 0.004 and the 90% CL interval according to [70] is:

$$\frac{B^{\textbf{combinatorial}}}{S}_{(\text{B}_s \to \text{D}_s^-\pi^+)} \in [\textbf{0.004}, \textbf{0.05}] \textbf{ at 90\% CL,}$$

whereas for the specific contribution, the central value is 0.17 ± 0.04 and the confidence interval is:

$$\frac{B^{\textbf{specific}}}{S}_{(\text{B}_s \to \text{D}_s^-\pi^+)} \in [\textbf{0.06}, \textbf{0.4}] \textbf{ at 90\% CL.}$$

5.2.3 Specific B background contributions

Table 5.4 shows the contributions of some specific backgrounds studied with dedicated Monte Carlo samples. These contributions are computed for a tight B_s mass window ($\pm 70 \text{ MeV}/c^2$) and prior to any trigger.

Decay channel		B/S
$B_d \to D^- \pi^+$	$=$	0.060 ± 0.015
$\Lambda_b \to D_s^- p$	$=$	0.018 ± 0.003
$\Lambda_b \to \Lambda_c^+ \pi^-$	$\in$	$[0.01, 0.05]$ at 90% CL

Table 5.4: Specific background contributions for $B_s \to D_s^- \pi^+$.

5.3   Properties of selected signal events

5.3.1 B_s mass resolution

Figure 5.8 shows the reconstructed B_s mass distributions for the signal and the main background sources. The histograms are normalized to 2 fb^{-1} of data. Performing a double Gaussian fit on the signal mass spectrum gives a mass resolution of

$$\langle \sigma \rangle_{B_s} = \left(0.77 \times 13.5^2 + 0.33 \times 24^2\right)^{\frac{1}{2}} = 17 \text{ MeV}/c^2 .$$
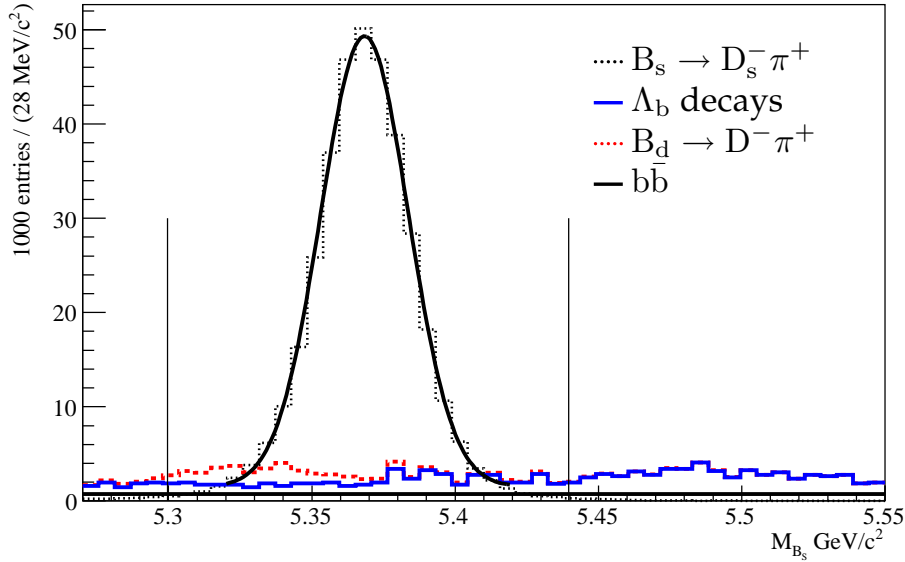
Figure 5.8: Reconstructed B_s with the $B_s \to D_s^- \pi^+$ selection. No trigger is applied. The upper limit of the estimated combinatorial $b\bar{b}$ background is represented as a flat distribution (continuous black line). The blue histogram is the reconstructed $\Lambda_b \to D_s^- p$ and $\Lambda_b \to \Lambda_c^+ \pi^-$ decays. The dashed red histogram represents the $B_d \to D^- \pi^+$ events. All the backgrounds are stacked together. The histograms are normalized to 2 fb^{-1} of data.

5.3.2   Proper-time reconstruction

The proper-time τ of a B_s meson can be estimated from its reconstructed mass m and momentum $\vec{p}$ through the relation

$$\tau = m \frac{\vec{p} \cdot \vec{FD}}{|\vec{p}|^2} ,$$

where $\vec{\mathrm{FD}}$ is the B_s meson flight distance vector (i.e. the vector between the production vertex and its decay vertex). The proper-time value and its error estimate are extracted by a dedicated algorithm[4].

Proper-time error and pull

Figure 5.9 shows the proper-time error estimate and the proper-time pull $\frac{\tau^{\mathrm{rec}} - \tau^{\mathrm{true}}}{\sigma_\tau}$ for $B_s \to D_s^- \pi^+$ events.

The estimate of the error on the reconstructed B_s lifetime has a mean value of 36 fs and a most probable value of 33 fs. The pull distribution is fitted with a double Gaussian with a common mean. The average sigma is $\langle \sigma \rangle^2 = f_1 \sigma_1^2 + (1 - f_1) \sigma_2^2 = (1.11 \pm 0.03)^2$. This shows that the errors are underestimated. The mean is $\mu = 0.036 \pm 0.004$ showing a bias in the proper-time reconstruction.
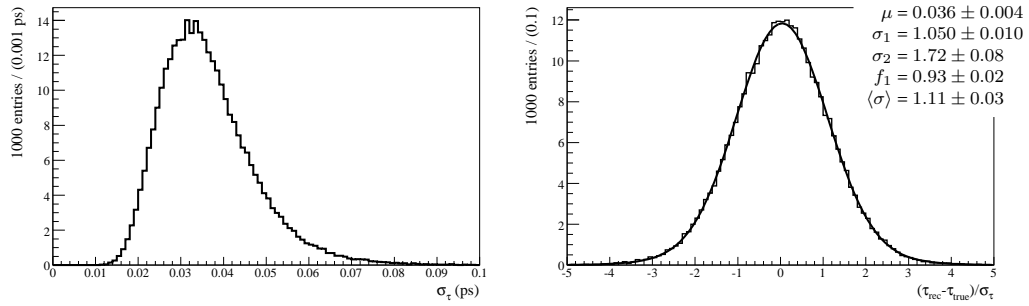

Figure 5.9: Error estimate (left) and pull (right) on the reconstructed proper time for signal events passing the $B_s \to D_s^- \pi^+$ selection (no trigger).

Reconstruction of the B_s–$\overline{B}_s$ oscillation

Figure 5.10 shows the time-dependent B_s decay rates at different stages of the reconstruction. The statistics correspond to 0.5 fb^{-1} of data. The Δm_s Monte Carlo input value is 20 ps^{-1}, while the $\Delta\Gamma_s$ value is 0.068 ps^{-1}. The wrong tag fraction is 36.6%.

The plots in Fig. 5.10 are obtained from reconstructed $B_s \to D_s^- \pi^+$ signal events without background. The B_s flavour comes either from Monte Carlo information (Fig. 5.10a and 5.10b) or from the tagging algorithm (Fig. 5.10c and 5.10d). The flavour at decay is deduced from the reconstructed bachelor charge. The shaded histograms represent the events which have mixed (initial B_s decaying to an $D_s^+ \pi^-$ state or initial $\overline{B}_s$ decaying to $D_s^- \pi^+$). The other histogram represents all events.

Figure 5.10a shows the distribution of the true proper time for the Monte Carlo events. Figure 5.10b is the reconstructed proper time with the tagging information taken from the Monte Carlo. Figure 5.10c shows the true proper-time for tagged events. It also shows the distribution for all the events (tagged and untagged). Thus some wrong tag is present, reducing the oscillation amplitude. Figure 5.10d shows the reconstructed propertime with the tagging algorithm applied. Figure 5.10e shows the reconstructed propertime with the tagging algorithm and the L0-trigger applied.

---

[4]`LifeTimeFitter` introduced in Ref. [71]
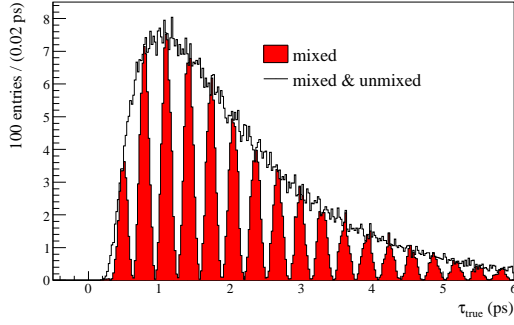
(a) True proper-time distribution, tagging based on the Monte Carlo.
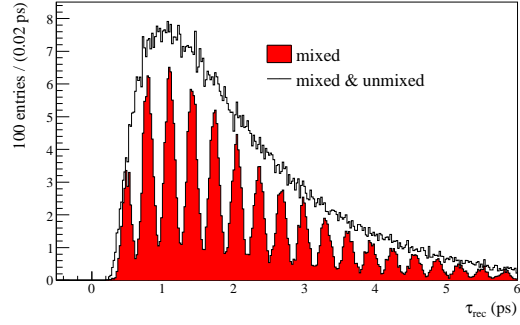
(b) Reconstructed proper-time distribution, tagging based on the Monte Carlo.
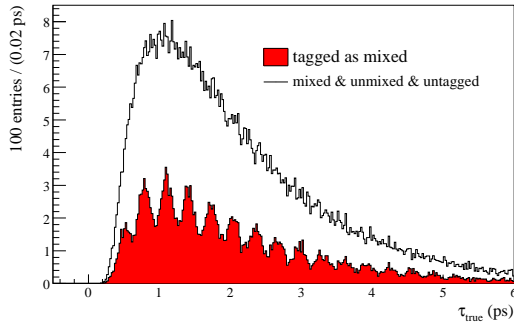
(c) True proper-time distribution for tagged events.
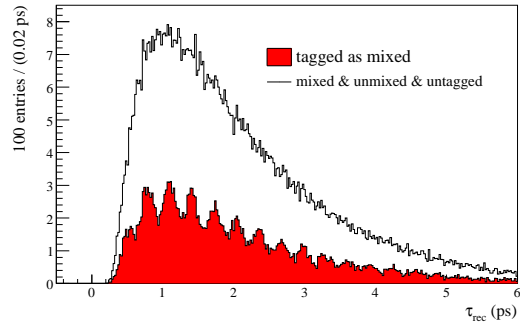
(d) Reconstructed proper-time distribution for tagged events.
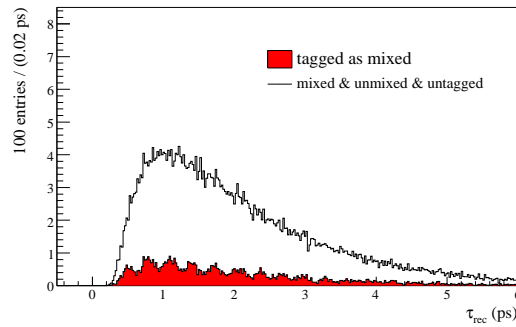
(e) Reconstructed proper-time distribution for tagged and L0-triggered events.

Figure 5.10: Proper-time distributions of selected $B_s \to D_s^- \pi^+$ signal events corresponding to 0.5 fb^{-1} of data at different stages of the full simulation and reconstruction.

Chapter 6

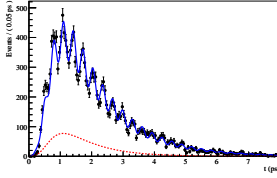# Statistical sensitivity to the $\mathrm{B_s}$–$\overline{\mathrm{B}}_\mathrm{s}$ Mixing Parameters

We use the results from the previous chapter as inputs to a fast Monte Carlo study of the $\mathrm{B_s}$–$\overline{\mathrm{B}}_\mathrm{s}$ oscillation frequency and other parameters related to the $\mathrm{B_s}$-meson system. We discuss the model used to simulate the B-meson decay rates and the detector effects, and assess the statistical sensitivity of the extracted parameters.

$\mathcal{W}$E extract the LHCb statistical sensitivities to the $\mathrm{B_s}$–$\overline{\mathrm{B}}_\mathrm{s}$ mixing parameters by fitting fast Monte Carlo samples of events parametrized using the full simulation as inputs. Each sample is generated according to the model and then analyzed to extract back the parameters of interest. The generation is performed with the RooFit toolkit [72] and fitting with the MINUIT minimization package [73].

Each toy experiment generates, based on a unique random seed, an amount of data fluctuating according to a Poisson distribution of mean 150 k, which is the expected statistics after the L0 trigger, assuming 2 fb^{-1} of acquired data. Yet, as the model is meant to be compatible with the data, the outcome of such a study only tests whether the model provides an unbiased estimator of the extracted parameters, and the sensitivity one can expect from such a parametrization.

The value of each parameter is estimated according to the maximum likelihood, which we briefly describe: Given a set of parameters $\boldsymbol{\alpha}$, a set of observables $\boldsymbol{x}$ and a probability density function (pdf) $f(\boldsymbol{x}; \boldsymbol{\alpha})$, the probability for one event i to be in an interval $\mathrm{d}\boldsymbol{x_i}$ is

$$\mathrm{P}(\boldsymbol{x_i} \in [\boldsymbol{x_i}, \boldsymbol{x_i} + \mathrm{d}\boldsymbol{x_i}]) = f(\boldsymbol{x_i}; \boldsymbol{\alpha})\mathrm{d}\boldsymbol{x_i} \,,$$

The simultaneous probability for all events is the product over all i of the above expression. Moreover, looking at f as a function of the $\boldsymbol{\alpha}$ with the $\boldsymbol{x_i}$ as parameters, one expects the probability to be the highest for the nominal value of $\boldsymbol{\alpha}$. Therefore, one defines the likelihood function

$$\mathcal{L}(\boldsymbol{\alpha}) = \prod_i f(\boldsymbol{x_i}; \boldsymbol{\alpha}) \tag{6.1}$$

and the value of $\boldsymbol{\alpha}$ which maximizes $\mathcal{L}$ is an estimator of $\boldsymbol{\alpha}$ and is called the maximum likelihood (ML) estimator. From the statistical point of view, maximum likelihood estimators are *asymptotically* unbiased, normally distributed and have minimum

variance [74]. For finite samples, the maximum likelihood is proved to have these optimal properties only when the parent distribution f is of the exponential form (i.e. $f(\boldsymbol{x}; \boldsymbol{\alpha}) = \exp(a(\boldsymbol{\alpha})b(\boldsymbol{x}) + c(\boldsymbol{\alpha}) + d(\boldsymbol{x}))$.

6.1   Resolution effects from first principles

The probability for a B meson to decay into a final state f is given by the rates of Eqs. 1.23 and 1.24 in Sec. 1.2.3:

$$P(\text{B} \to f) \propto e^{-\Gamma_s t} \left\{ \cosh \frac{\Delta \Gamma_s}{2} t + D \cos(\Delta m_s t) \right\} , \tag{6.2}$$

where D is the experimental dilution due to imperfect tagging. To account for the resolution, we use a Gaussian model, which is convoluted with the decay rate, and for the simplicity of the notation, we drop the subscript 's' of Δm_s, Γ_s and $\Delta \Gamma_s$. Thus, taking into account a Gaussian resolution of width σ the above equation becomes:

$$P(\text{B} \to f) \propto \int_0^\infty e^{-\Gamma t'} \left\{ \cosh \frac{\Delta \Gamma}{2} t' + D \cos(\Delta m t') \right\} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{t-t'}{\sigma}\right)^2} dt' , \tag{6.3}$$

This expression can be integrated analytically, and to simplify the final expression, we will perform the convolution product over $]-\infty, \infty[$, whereas it should in principle be performed over the $[0, \infty[$ range. As a consequence, the result we obtain is valid only for positive and 'large' proper time, i.e. $t \gg \sigma$. Note that, experimentally, this condition can be considered to be verified, because the selection efficiency is almost zero at small proper times[1]. Considering the following equality:

$$\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^\infty e^{-\Gamma t'} e^{\alpha t'} e^{-\frac{1}{2}\left(\frac{t-t'}{\sigma}\right)^2} dt' = e^{\frac{1}{2}(\alpha-\Gamma)^2 \sigma^2} e^{(\alpha-\Gamma)t} , \tag{6.4}$$

where α is any real number, we write:

$$\int_{-\infty}^\infty e^{-\Gamma t} \cosh(\frac{\Delta \Gamma}{2} t) e^{-\frac{1}{2}\left(\frac{t-t'}{\sigma}\right)^2} dt' = \frac{1}{2} \int_{-\infty}^\infty e^{-\Gamma t'} \left(e^{\frac{\Delta \Gamma}{2} t'} + e^{-\frac{\Delta \Gamma}{2} t'} \right) e^{-\frac{1}{2}\left(\frac{t-t'}{\sigma}\right)^2} dt'$$
$$= \frac{1}{2} \left[e^{\frac{1}{2}(\frac{\Delta \Gamma}{2}-\Gamma)^2 \sigma^2} e^{(\frac{\Delta \Gamma}{2}-\Gamma)t} + e^{\frac{1}{2}(-\frac{\Delta \Gamma}{2}-\Gamma)^2 \sigma^2} e^{(-\frac{\Delta \Gamma}{2}-\Gamma)t} \right]$$
$$= e^{\frac{1}{2}(\Gamma^2 + \frac{\Delta \Gamma}{2}^2)\sigma^2} e^{-\Gamma t} \cosh(\frac{\Delta \Gamma}{2}(t - \Gamma \sigma^2)) \tag{6.5}$$

and

$$\int_{-\infty}^\infty e^{-\Gamma t} D \cos(\Delta m \, t) e^{-\frac{1}{2}\left(\frac{t-t'}{\sigma}\right)^2} dt' = \frac{1}{2} D \int_{-\infty}^\infty e^{-\Gamma t'} \left[e^{i\Delta m t'} + e^{-i\Delta m t'} \right] e^{-\frac{1}{2}\left(\frac{t-t'}{\sigma}\right)^2} dt'$$
$$= e^{\frac{1}{2}(\Gamma^2 - \Delta m^2)\sigma^2} e^{-\Gamma t} D \cos(\Delta m(t - \Gamma \sigma^2)) . \tag{6.6}$$

---

[1]Unless one defines a lifetime unbiased selection, which is not the case of the $B_s \to D_s^- \pi^+$ selection.

Hence Eq. 6.3 becomes:

$$
\begin{aligned}
P(\mathrm{B} \to f) \quad \propto \quad & \int_{-\infty}^{\infty} \mathrm{e}^{-\Gamma t'} \left\{ \cosh \frac{\Delta\Gamma}{2} t' + D \cos(\Delta m t') \right\} \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{1}{2}\left(\frac{t-t'}{\sigma}\right)^2} \mathrm{d}t' \\
= \quad & \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\Gamma\widetilde{t} - \Gamma^2\sigma^2} \left\{ \mathrm{e}^{\frac{1}{2}(\Gamma^2 + \frac{\Delta\Gamma^2}{2})\sigma^2} \cosh(\frac{\Delta\Gamma}{2}\widetilde{t}) + \right. \\
& \left. D \mathrm{e}^{-\frac{1}{2}\Delta m^2\sigma^2} \mathrm{e}^{\frac{1}{2}\Gamma^2\sigma^2} \cos(\Delta m \, \widetilde{t}) \right\} \\
= \quad & \frac{\mathrm{e}^{\frac{1}{2}(\frac{\Delta\Gamma^2}{2} - \Gamma^2)\sigma^2}}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\Gamma\widetilde{t}} \left\{ \cosh(\frac{\Delta\Gamma}{2}\widetilde{t}) + \right. \\
& \left. \mathrm{e}^{-\frac{1}{2}(\frac{\Delta\Gamma^2}{2} + \Delta m^2)\sigma^2} D \cos(\Delta m \, \widetilde{t}) \right\} \,, \quad (6.7)
\end{aligned}
$$

where we have defined $\widetilde{t} \equiv t - \Gamma\sigma^2$. Keeping only the time dependence, we write:

$$
\widetilde{P} \quad \propto \quad \mathrm{e}^{-\Gamma\widetilde{t}} \left\{ \cosh(\frac{\Delta\Gamma}{2}\widetilde{t}) + \widetilde{D} \cos(\Delta m \, \widetilde{t}) \right\} \,, \quad (6.8)
$$

where

$$
\widetilde{D} \equiv \mathrm{e}^{-\frac{1}{2}(\frac{\Delta\Gamma^2}{2} + \Delta m^2)\sigma^2} D \,. \quad (6.9)
$$

We notice that, apart from a time shift $\Gamma\sigma^2$ and an extra dilution factor of $\mathrm{e}^{-\frac{1}{2}(\frac{\Delta\Gamma^2}{2} + \Delta m^2)\sigma^2}$, Eq. 6.8 is exactly the same as Eq. 6.3, the decay rate with no resolution model.

The value for σ, quoted in Chapter 5 is 0.036 ps, it represents a time shift of $\Gamma\sigma^2 \simeq 1$ fs. Therefore we do not expect any sensitivity to this shift, even with the most optimistic scenario for the amount of collected data. However, the additional dilution factor due to the resolution is equal to $\mathrm{e}^{-\frac{1}{2}(\frac{\Delta\Gamma^2}{2} + \Delta m^2)\sigma^2} \simeq 82\%$, assuming $\Delta m = 17.8$ ps^{-1}, $\sigma = 0.036$ ps and $\Delta\Gamma \simeq 0$. Hence the overall dilution is expected to be significantly affected by the proper-time resolution. Note however that, in the $\mathrm{B_s} \to \mathrm{D_s^-}\pi^+$ selection, the numerical value of D is ~ 0.26, and therefore this is still the dominant factor which reduces the oscillation amplitude.

We conclude from this discussion that a modeling similar to the one of Eq. 6.3 is neither sensitive to the width of the resolution σ, nor to the actual value of the dilution D. But one is sensitive to $\widetilde{D}$ or, if the scale factor is provided to D.

6.2 Model description

We describe the likelihood function used to generate and fit the $\mathrm{B_s}$–$\overline{\mathrm{B}}_\mathrm{s}$ oscillations. Its full expression is the product over all the events:

$$
\begin{aligned}
\mathcal{L} \quad = \quad & \frac{(N^{\mathrm{sig}} + N^{\mathrm{bkg}})^{N^{\mathrm{obs}}}}{N^{\mathrm{obs}}!} \times \\
& \prod_{i}^{N^{\mathrm{obs}}} \left[r^{\mathrm{sig}} f_m^{\mathrm{sig}}(m_i) f_\tau^{\mathrm{sig}}(t_i|\sigma_i, q_i) + (1 - r^{\mathrm{sig}}) f_m^{\mathrm{bkg}}(m_i) f_\tau^{\mathrm{bkg}}(t_i|\sigma_i, q_i) \right], \quad (6.10)
\end{aligned}
$$

where N^{obs} is the actual number of observed event in the sample (Poisson distributed) such that the number of expected events by the model is $N^{\mathrm{exp}} = N^{\mathrm{sig}} + N^{\mathrm{bkg}}$, the signal probability is $r^{\mathrm{sig}} = \frac{N^{\mathrm{sig}}}{N^{\mathrm{sig}} + N^{\mathrm{bkg}}}$, the superscripts sig and bkg account for the signal and the background description respectively, the subscript m and τ stand for the mass and the proper-time description. m_i is the measured value of the mass, t_i and σ_i are the measured proper time and its error estimate. q_i is related to the tagging information (see Sec. 6.2.2).

6.2.1   Mass models

The signal mass is described by a simple Gaussian with width equal to the reconstructed B_s mass resolution (17 MeV/c^2). The background is modeled with a decreasing exponential function. Explicitly, for an event i:

$$f_m^{\text{sig}}(m_i) = \frac{1}{\sqrt{2\pi}\sigma_{B_s}} e^{-\frac{1}{2}\left(\frac{m_i - M_{B_s}}{\sigma_{B_s}}\right)^2}, \qquad (6.11)$$

$$f_m^{\text{bkg}}(m_i) = \frac{1}{\kappa^{\text{bkg}}} e^{-\kappa^{\text{bkg}} m_i}. \qquad (6.12)$$

The background mass description is rather simple. The real data will of course require a more detailed description, which will include the peaking backgrounds like $B_d \rightarrow D^- \pi^+$. However, in the present model, the only quantity that matters is the amount of background in the signal window, parametrized by the background-over-signal (B/S) ratio. We tried the same fit with a more realistic mass model which included, within their selected proportions the $B_s \rightarrow D_s^- \rho^+$, $B_s \rightarrow D_s^{*-} \pi^+$ and $B_d \rightarrow D^- \pi^+$ low mass modes. Apart from a slight increase in the processing time, no difference could be observed. The influence of these modes will come from the description of their proper-time distribution, yet this is not included in the present study. Hence, for the sake of simplicity, we use an exponential distributed mass background with a parameter $\kappa^{\text{bkg}} = -1$. Moreover, we arbitrarily chose the wide mass window to be ± 0.3 GeV/c^2, such that the available background statistics in the sidebands is roughly equal to the signal statistics.

We show in Fig. 6.1 one example of generated mass distribution. The signal contribution is plotted with a red dashed line.
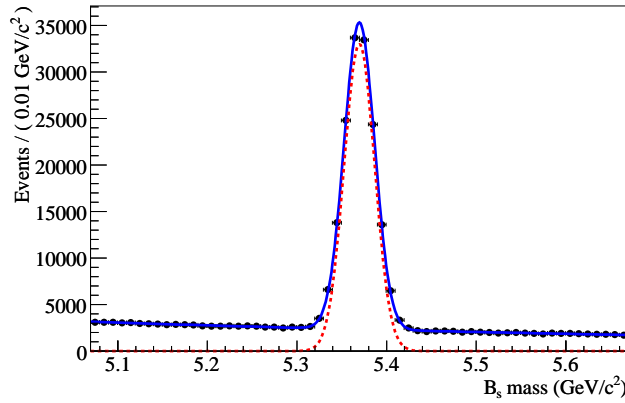
Figure 6.1: Simulated B_s mass distribution.

6.2.2   Signal proper-time model

We express the $B_s \rightarrow D_s^- \pi^+$ signal pdf for an event i as:

$$f_\tau^{\text{sig}}(t_i | \sigma_i, q_i) = \frac{1}{\mathcal{N}_S} A(t_i) \left\{ P(t_{\text{true}}, q_i) \otimes R(t_{\text{true}} - t_i, \sigma_i) \right\} E(\sigma_i) \qquad (6.13)$$

where t_i and σ_i are the reconstructed proper time and its error estimate respectively, and q_i is defined below.

The function P is the B_s time-dependent decay rate, R is the resolution function, and A describes the proper-time dependent acceptance (i.e. the time dependence of the selection efficiency). The function E is an empirical description of the distribution of the error-estimate. Finally, $\mathcal{N}_S$ is a normalization factor.

True proper-time model

The modeling of the true decay rate uses the expression:

$$P(t_{\text{true}}, q_i; \Gamma, \Delta\Gamma, \Delta m, \omega) = \mathrm{e}^{-\Gamma t_{\text{true}}} \left\{ \cosh \frac{\Delta\Gamma}{2} t_{\text{true}} + q_i(1 - 2\omega) \cos \Delta m t_{\text{true}} \right\}, \quad (6.14)$$

where t_{true} is the true proper time of the B meson. It is of course not accessible and will be integrated away in the convolution product with the resolution function (see below). The parameter q_i includes both the flavour tagging information and the decay product charges. It is set to $+1$, -1 or 0 depending on whether the event has mixed, or not or whether it is untagged. ω is the wrong tag fraction, leading to the dilution $D = 1 - 2\omega$. The parameters Γ, $\Delta\Gamma$ and Δm follow their usual definition, namely they represent the B_s-meson average decay width, the decay-width difference, and the oscillation frequency respectively.

Resolution model

The resolution function is a single Gaussian distribution,

$$R(t_{\text{true}} - t_i, \sigma_i; s) = \frac{1}{\sqrt{2\pi} s \sigma_i} \mathrm{e}^{-\frac{1}{2}\left(\frac{t_{\text{true}} - t_i}{s\sigma_i}\right)^2}, \quad (6.15)$$

to be convoluted with the decay rate parametrization (Eq. 6.14). The width of the Gaussian is set to σ_i multiplied by the factor s to account for a global scaling of the proper-time error estimates. We refer to Fig. 5.9, which shows the pull of selected $B_s \to D_s^- \pi^+$ events, i.e. the full simulation result. It is fitted with a double Gaussian, however, the contribution of the second Gaussian is only 7%. We use therefore a single Gaussian model, but with the real data, the resolution model will be more complex [75].

Acceptance function

We call 'acceptance' the following function:

$$A(t_i; \alpha) = \frac{\alpha t_i^3}{1 + \alpha t_i^3}, \text{ if } t_i > 0, \ 0 \text{ elsewhere.} \quad (6.16)$$

The acceptance function represents the proper-time dependence of the selection efficiency. Indeed, most of the selections studied by the LHCb collaboration (and $B_s \to D_s^- \pi^+$ is no exception here), cut on variables which bias the original proper-time distribution (e.g. cutting on the impact parameter significance rejects small proper-time events). Strictly speaking, the acceptance should be part of the resolution model as it accounts for 'the detector effect on the proper-time reconstruction'. Thus one would write an expression like $P \otimes R$, where the resolution R would have a complicated proper-time dependence. As this is too complicated, one splits R in a Gaussian resolution and an acceptance function. Yet, it is not fully clear whether the acceptance should be described as a function of the true or the reconstructed proper time (i.e. whether it should be part or not of

the convolution product). Internal discussions within LHCb have suggested that the a parametrization as a function of the reconstructed proper time is preferable, which is the way we choose.

The actual acceptance description we have chosen is rather simple. Looking at Ref. [64], we see that the proper time dependence of the acceptance can have a more complex shape. We did a pragmatic choice here: other studies [56] have shown, and we have tested, that the simultaneous extraction of Γ and an acceptance function which has a dependence in the high proper-time is very hard. The simple solution is of course to fix the acceptance, or to use a model which does not involve too strong a correlation with Γ. We choose the latter solution for it is not clear whether the HLT still creates a high proper-time dependence in the selection efficiency. Yet we do not expect the acceptance parametrization to have an impact on the sensitivity to Γ, provided that it is not extracted simultaneously. Finally, note that a method has been studied [76] which allows the extraction of the acceptance function without the help of any Monte Carlo inputs. We did not use it here, mainly because is it heavy in terms CPU resources.

Distribution of the error estimates

To ensure that biases like the one reported in [77] will not be a problem, we multiply the description of the decay rate and of the background proper-time parametrization by their parametrization of the error-estimate distribution.

$$
E = \begin{cases}
\frac{(\sigma-\sigma_0)^n}{2\times 10^{-8}\alpha_\sigma + (\sigma-\sigma_0)^n} e^{-\kappa_\sigma(\sigma-\sigma_0)} & \text{if } \sigma > \sigma_0 \\[2mm]
0 & \text{if } \sigma \le \sigma_0 \,.
\end{cases}
\tag{6.17}
$$

We stress here that this distribution has no physical content, we simply chose a description which can match the outcome of the full simulation results. The numerical factor 2×10^{-8} scales α_σ such that the nominal value is around 1. We show in Fig. 6.2 the full simulation distribution of the error estimates. It is fitted with the function E. As the parameter σ_0 depends on the range of the actual observed data, we cannot fit it with the ML technique[2] and we will keep it fixed during the fits.
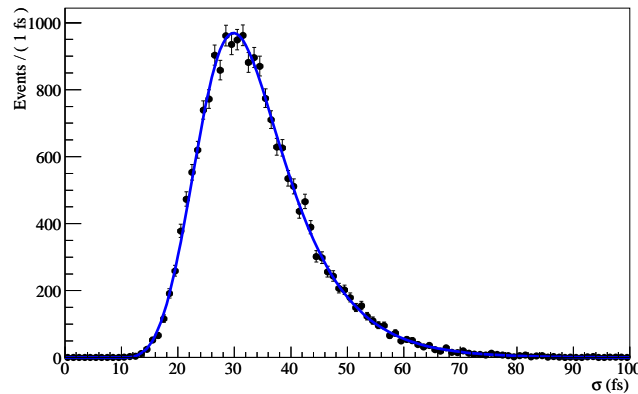

Figure 6.2: Distribution of the error estimates of selected $B_s \rightarrow D_s^- \pi^+$ events. We show the outcome of the fit performed with the function E.

---

[2]It is a known fact that the ML technique is a bad estimator when the range of observed data actually depends on one of the estimated parameter [74].

Normalization

The factor $\mathcal{N}_S$ in Eq. 6.13, accounts for the normalization over all the variables. We give it here for completeness. Note that we do not integrate P over σ as it is a conditional variable of this pdf.

$$\mathcal{N}_S = \int_{t_{\min}}^{t_{\max}} A(t) \sum_{q=-1}^{+1} P(t|\sigma;q)\mathrm{d}t \int_{\sigma_{\min}}^{\sigma_{\max}} E(\sigma)\mathrm{d}\sigma \; . \tag{6.18}$$

Due to the presence of the acceptance term, we were not able to compute an analytical integration of the above expression. Because of this, the integration is handled numerically, and therefore, between finite limits.

6.2.3 Background proper-time model

The proper-time distribution of the background is modeled with a simple exponential. The other terms are analogous to the signal description. Explicitly, we use:

$$\mathcal{L}_\tau^{\mathrm{bkg}}(t_i|\sigma_i, q_i) = \frac{1}{\mathcal{N}_B} A(t_i) \left\{ P(t_{\mathrm{true}}; a) \otimes R(t_{\mathrm{true}} - t_i, \sigma_i) \right\} E(\sigma_i) \tag{6.19}$$

where

$$P(t_{\mathrm{true}}; a) = \frac{1}{a} \mathrm{e}^{-at_{\mathrm{true}}} \; , \tag{6.20}$$

and R, A and E have the same parametrization as for the signal (Eqs. 6.15, 6.16 and 6.17 respectively). Note though, that the actual parameter values used to generate the data, differ for each term between the signal and the background.

The normalization of the background is also performed numerically, it is computed (by analogy with the signal) as:

$$\mathcal{N}_B = 3 \int_{t_{\min}}^{t_{\max}} A(t) P(t|\sigma;q)\mathrm{d}t \int_{\sigma_{\min}}^{\sigma_{\max}} E(\sigma)\mathrm{d}\sigma \; , \tag{6.21}$$

where the factor 3 stands for the normalization over q.

To check that an exponential-time description is satisfactory, we computed the Fast Fourier Transform (FFT) of reconstructed and tagged $B_s \rightarrow D_s^- \pi^+$ and $B_d \rightarrow D^- \pi^+$ events, the latter being selected with a looser selection to increase the statistics. We also tried to compute the FFT of $B_s \rightarrow D_s^- \rho^+$ events, however, these were generated and reconstructed with the DC04 data and must therefore be compared with caution to the DC06 samples. We show in Fig. 6.3 the magnitude of the FFT transforms. Here 20 ps^{-1} is the value used for Δm during the data generation process. A peak at this value is clearly visible for the $B_s \rightarrow D_s^- \pi^+$ events. It is also present in the $B_s \rightarrow D_s^- \rho^+$ data, although with much less significance. The $B_d \rightarrow D^- \pi^+$ data, on the other hand, show no clear peak, nor at 0.5 ps^{-1} (the value of Δm_{d}), neither at 20 ps^{-1}. We conclude that an exponential distribution is certainly satisfactory to describe the combinatorial background, but the specific sources will have to be explicitly introduced in further analysis, probably some of these background will share the same value for Δm and thus contribute to the latter sensitivity.

6.3 Toy modeling

Most of the inputs come, or are inspired from the full Monte Carlo simulation described in Chapter 5. The initial value of each parameter is given in Table 6.1. We explain some
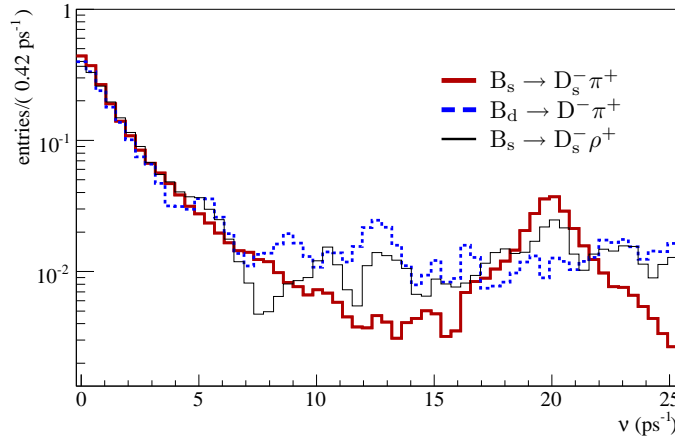
Figure 6.3: Fast Fourier Transform applied on tagged $B_s \to D_s^- \pi^+$ (red), $B_d \to D^- \pi^+$ (dashed blue) and $B_s \to D_s^- \rho^+$ (black) events. The selection criterion have been slightly relaxed to increase the statistics. The $B_s \to D_s^- \rho^+$ events come from the DC04 data, but no real change is expected with the DC06 data. See the text for details.

of them hereafter. Each toy experiment is based on a number of signal events fluctuating according to a Poisson distribution with a mean value equal to the expected annual yield of 150 k. The B/S value taken as input is 0.2, it corresponds to a naive addition of the expectation for the $B_s \to D_s^- \pi^+$ specific and combinatorial backgrounds. The value used for the B_s mass and its average lifetime are $M_{B_s} = 5.3696\ \mathrm{GeV}/c^2$, and $1/\Gamma = 1.46$ ps.

The parameter values for the signal distribution of the error-estimate are taken from a fit on the full MC events. The background error-estimate distribution is meant to be close but different than the signal, such that one cannot factorize out of the likelihood the effect of the error estimates.

The parametrization of the signal acceptance comes from a fit using the DC04 data and the background acceptance, again, meant to be close but different from the signal.

6.4   Extracted parameters

6.4.1   Fit procedure

Due to the large correlation between certain parameters, it is not possible to fit them all simultaneously. Therefore, we perform the following successive fits:

Mass-distribution fit: fixing all the time related dependencies, we start fitting the mass distributions in the total mass window, leaving the following parameters free:

$$N^{\mathrm{sig}},\ N^{\mathrm{bkg}},\ M_{B_s},\ \sigma_{B_s},\ \kappa^{\mathrm{bkg}}\ .$$

The above parameters are then fixed. Note that $N^{\mathrm{sig}}/(N^{\mathrm{sig}} + N^{\mathrm{bkg}})$ fixes the B/S value.

Error-distribution fits: as the signal and background error-estimate distributions are very similar, it is hard to extract them simultaneously, thus, we perform separate fits for

	Name	Value	Origin
Global parameters			
Yield	N^{sig}	150 k	Full MC
Background over signal ratio	B/S	0.2	arbitrary value
Mass distribution			
B_s mass	M_{B_s}	5.3696 GeV/c^2	arbitrary value
B_s mass resolution	σ_{B_s}	17 MeV/c^2	Full MC
Slope	κ^{bkg}	$-1(\mathrm{GeV}/c^2)^{-1}$	arbitrary value
Signal time-distribution			
Scale factor	s^{sig}	1.1	Full MC
Acceptance	α^{sig}	1.4	based on Ref. [64]
Decay width difference	$\Delta\Gamma$	0.07 ps^{-1}	arbitrary value
B_s oscillations frequency	Δm	17.77 ps^{-1}	Ref. [78]
Mean B_s lifetime	$1/\Gamma$	1.46 ps	arbitrary value
Wrong-tag fraction	ω	0.37	Full MC
Background time-distribution			
Scale factor	s^{bkg}	1.5	arbitrary value
Acceptance	α^{bkg}	1.1	arbitrary value
Slope	a	-2 ps^{-1}	arbitrary value
Signal error-estimate distribution			
	$\kappa_{\sigma}^{\mathrm{sig}}$	120 ps^{-1}	Full MC
	$\alpha_{\sigma}^{\mathrm{sig}}$	1.175 ps^{-1}	Full MC
	n^{sig}	4.6	Full MC
	$\sigma_{0}^{\mathrm{sig}}$	8.8×10^{-3} ps	Full MC
Background error-estimate distribution			
	$\kappa_{\sigma}^{\mathrm{bkg}}$	140 ps^{-1}	arbitrary value
	$\alpha_{\sigma}^{\mathrm{bkg}}$	1.4 ps^{-1}	arbitrary value
	n^{bkg}	4	arbitrary value
	$\sigma_{0}^{\mathrm{bkg}}$	1×10^{-2} ps	arbitrary value

Table 6.1: Input values of the fast Monte Carlo experiments.

the signal and the background. We fit the error-estimate distribution of the background in the sidebands[3] with the following free parameters:

$$\alpha_\sigma^{\mathrm{bkg}}, \; \kappa_\sigma^{\mathrm{bkg}}, \; n^{\mathrm{bkg}} \; .$$

Then, keeping the above parameters fixed, we extract the signal error distribution in the tight mass window, the proportion of signal and background distribution being fixed by the B/S value. The free parameters in this third fit are:

$$\alpha_\sigma^{\mathrm{sig}}, \; \kappa_\sigma^{\mathrm{sig}}, \; n^{\mathrm{sig}} \; ,$$

which are then fixed in subsequents fits.

Sideband fit: we fit the time distribution of the background on the sidebands, the free parameters are:

$$a, \; \alpha^{\mathrm{bkg}} \; ,$$

which are then fixed in subsequents fits.

Signal-window fit: the extraction of the physics parameters is done together with the signal acceptance in the signal mass window. The free parameters are

$$\Delta m, \; 1/\Gamma, \; \Delta\Gamma, \; \omega, \; \alpha^{\mathrm{sig}} \; .$$

6.4.2   Fit results

We present the results based on a initial set of 190 toy experiments. 166 converged and correctly computed the second derivatives matrix (i.e. the errors according to HESSE), out of which 121 managed to follow the likelihood profile to compute most of the MINOS errors. Still in some cases, MINOS fails to compute one side of the error of a parameter. In such cases, we excluded the faulty parameter from the statistics, but kept the other parameter results. This explains why all the plots and results below are not based exactly on the same statistics.

We summarize in Table 6.2 all the fit outcome, and we show the distributions of the residuals and the pulls for Δm (Fig. 6.4), $1/\Gamma$ (Fig. 6.5), $\Delta\Gamma$ (Fig. 6.6) and ω (Fig. 6.7).

The only parameters which are biased to more than one sigma are the one related to the error-estimate distributions. Note that, due to the size of the relative error of the fitted parameters, the distributions still look like they correctly describe the data. Furthermore, as the distributions do not contain any physics, we consider that these biases can be ignored.

The other problem comes from the background time-distribution description. Indeed, Table 6.2 shows that the pulls of a and α^{bkg} are too large, but looking at their pull distribution (Figs. 6.8 and 6.9) shows that the latter have been wrongly estimated because the distributions are not Gaussian. We did not further investigate this problem, again the relative errors are of the order of 0.02 and thus, the data are still fitted in a way which is satisfactory.

---

[3]The signal mass window is defined as $|m - M_{B_s}| < 70 \; \mathrm{MeV}/c^2$, whereas the sidebands are defined as $80 \; \mathrm{MeV}/c^2 < |m - M_{B_s}| < 300 \; \mathrm{MeV}/c^2$.

Parameter	Average fitted value	Relative error	Bias relative to the error	Pull
N^{bkg}	(178.5 ± 0.4) k	2.4×10^{-3}	0.17	
N^{sig}	(150.0 ± 0.4) k	2.5×10^{-3}	6×10^{-3}	
κ^{bkg}	(-1.000 ± 0.012) ps^{-1}	0.012	0.027	
M_{B_s}	(5.36966 ± 0.00005) GeV/c^2	10^{-5}	0.023	
$\sigma_{\text{B}_\text{s}}$	(17.00 ± 0.05) MeV/c^2	2.8×10^{-3}	-0.05	
$\kappa_\sigma^{\text{bkg}}$	(141.0 ± 0.7) ps^{-1}	0.005	1.4	
$\alpha_\sigma^{\text{bkg}}$	(1.12 ± 0.16) ps^{-1}	0.14	0.7	
n^{bkg}	3.987 ± 0.027	0.007	-0.5	
$\kappa_\sigma^{\text{sig}}$	(120.3 ± 0.8) ps^{-1}	0.006	0.4	
$\alpha_\sigma^{\text{sig}}$	1.62 ± 0.22	0.14	2.0	
n^{sig}	4.529 ± 0.027	0.006	-2.6	
α^{bkg}	1.097 ± 0.021	0.019	-0.16	
a	(-2.00 ± 0.03) ps^{-1}	0.016	-0.12	
α^{sig}	1.395 ± 0.010	0.007	-0.5	
$\Delta\Gamma$	(0.08 ± 0.03) ps^{-1}	0.4	0.5	
Δm_s	(17.770 ± 0.008) ps	4×10^{-4}	-0.04	
$1/\Gamma$	(1.452 ± 0.013) ps	0.009	-0.6	
ω	0.370 ± 0.003	0.009	-0.16	$-2\sigma \quad -\sigma \quad 0 \quad +\sigma \quad +2\sigma$

Table 6.2: Fit result for all the extracted parameters. Parameters fitted simultaneously are grouped together. For each parameter p, one fits the distribution of the estimate returned by the fit p_{fitted} with a Gaussian of mean μ_p and sigma σ_p. The 'fitted value' is $\mu_p \pm \sigma_p$, the relative error is $|\sigma_p/\mu_p|$, the relative bias is $(\mu_p - p_{\text{true}})/\sigma_p$. The last column sketches the pull width and its central value extracted with a Gaussian fit.
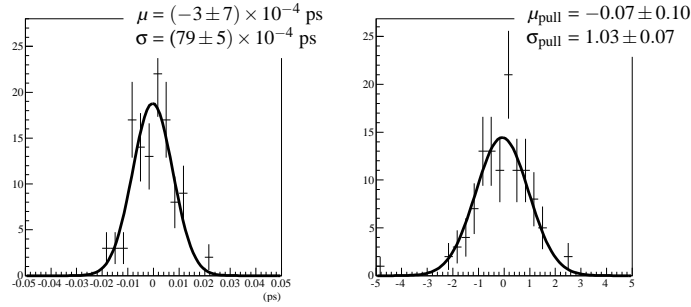
Figure 6.4: Residual (left) and pull (right) of the B_s–$\overline{B}_s$ oscillation frequency, Δm.
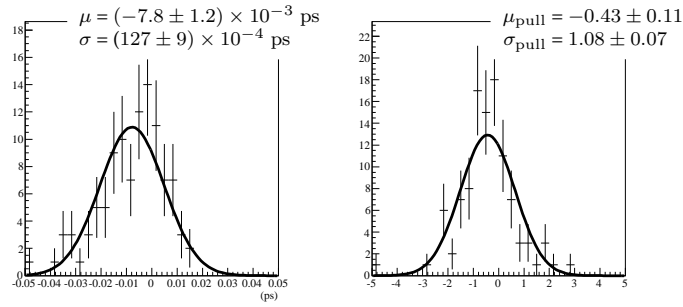

Figure 6.5: Residual (left) and pull (right) of the B_s average lifetime, $1/\Gamma$.
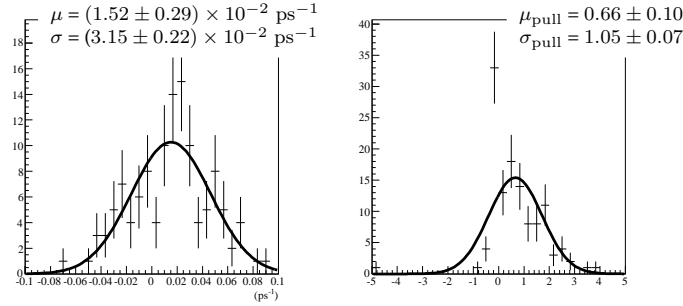

Figure 6.6: Residual (left) and pull (right) of the decay width difference, $\Delta\Gamma$.
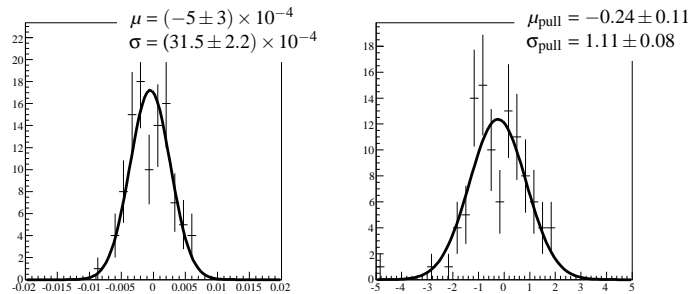

Figure 6.7: Residual (left) and pull (right) for the wrong-tag fraction, ω.
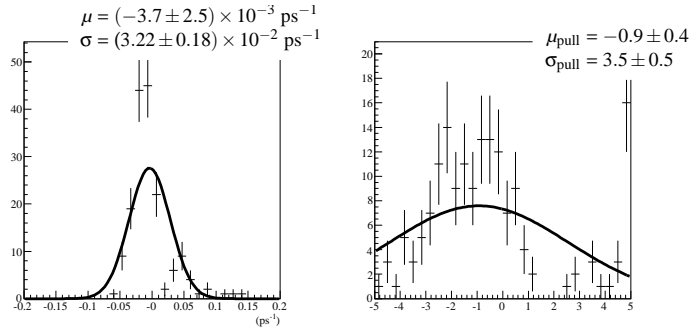
Figure 6.8: Residual (left) and pull (right) of the background time-distribution slope, a.
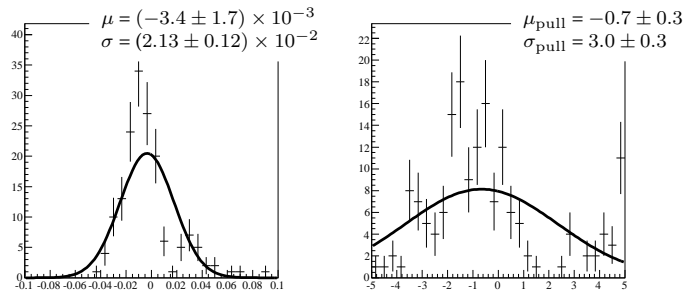

Figure 6.9: Residual (left) and pull (right) of the background acceptance function, α^{bkg}.

6.4.3 Additional tests

We performed the same analysis with a B/S ratio of 0.4, i.e. increased by a factor 2. No sizable effect could be seen.

As an exercise, we generated some background and used the results of the full Monte Carlo simulation for the signal. We then applied the fit procedure. We show the mass distribution in Fig. 6.10, and the proper-time distributions for mixed, unmixed and all events are shown in Figs. 6.11a, 6.11b and 6.11c respectively. All the extracted parameters are compatible within their errors to the fit inputs.

6.5 Outlook

Using the $B_s \rightarrow D_s^- \pi^+$ decay channel, we quantified the statistical sensitivity to some parameters of the B_s–$\overline{B}_s$ system with an unbinned fit of the B_s-mesons tagged proper-time distributions. The model accounts for the mass distributions, a simple background modeling and the detector resolution. We show in Table 6.3, the statistical sensitivities to the main extracted-parameters. It is based on a yield of 150 k triggered and selected events and a B/S ratio of 0.2.
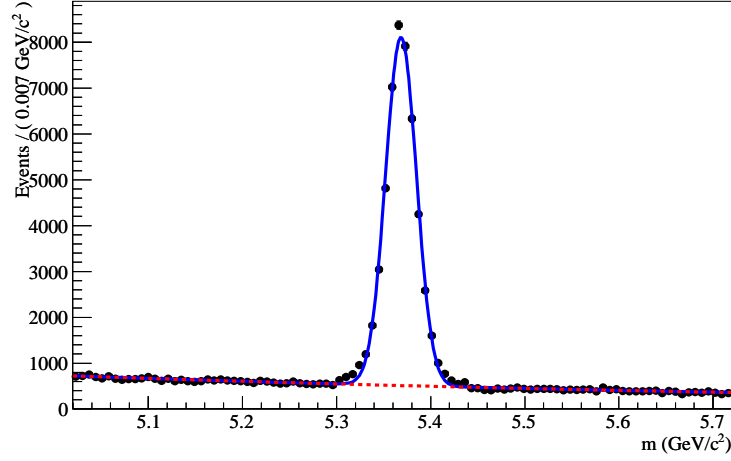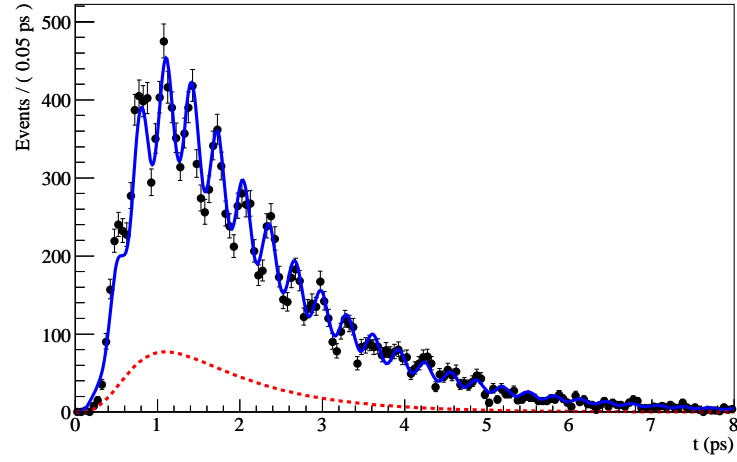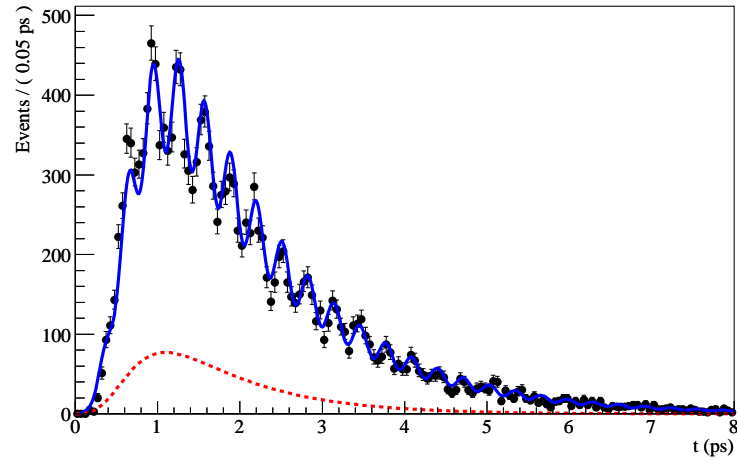
Figure 6.10: Fitted mass distribution. The signal (in blue) comes from the fully simulated Monte Carlo, the background (dashed red) is generated according to the model. The statistics correspond to 0.5 fb^{-1} of data.

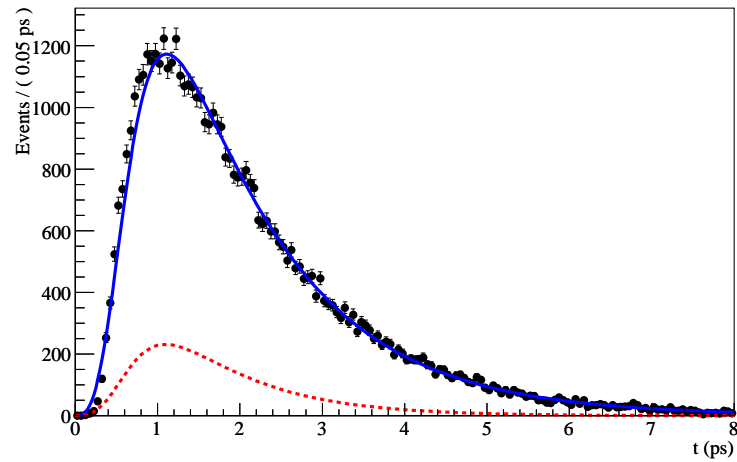| Parameter	Statistical sensitivity	Relative sensitivity
$\Delta\Gamma_s$	0.03 ps^{-1}	40%
Δm_s	0.008 ps^{-1}	4×10^{-4}
$1/\Gamma_s$	0.013 ps	9×10^{-3}
ω	0.3%	9×10^{-3}

Table 6.3: Statistical sensitivities to some of the parameters of the B_s-meson system extracted with $B_s \rightarrow D_s^- \pi^+$ decays.

(a) Fitted proper-time distribution for mixed events.

(b) Fitted proper-time distribution for unmixed events.

(c) Fitted proper-time distribution for mixed, unmixed and untagged events.

Figure 6.11: Fitted proper-time distribution. The signal (in blue) comes from the fully simulated Monte Carlo, the background (dashed red) is generated according to the model. The statistics correspond to 0.5 fb^{-1} of data.

Chapter 7

Length scale calibration
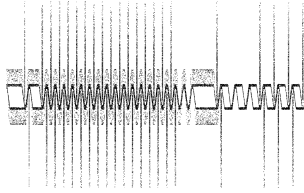
We introduce a method to calibrate the length scale based on the reconstruction of secondary interactions occurring in the VeLo material. Using the survey measurements as reference, we show that we quickly assess the length scale at the 6×10^{-5} level.

$\mathcal{A}$ simple way to calibrate the proper time scale of a detector is to compare a measured lifetime with its known value. For instance LHCb could use the B_d lifetime as a reference: the mean B_d proper time can be measured with a high statistical precision in channels like $B_d \rightarrow (J/\psi \rightarrow \mu^- \mu^+)(K^{*0} \rightarrow K^+ \pi^-)$. Yet such a method is ultimately limited by the knowledge of the B_d lifetime itself. Its current value is 1.530 ± 0.009 ps [65], and therefore it cannot be used to determine the proper-time scale with an accuracy better than 0.6%. The proper time is computed as $\frac{mL}{p}$, where m is the particle mass, L its flight distance measured with a vertex detector and p its momentum. The calibration of the momentum (and hence indirectly of the masses) will be done through the observation of narrow resonances like J/ψ or Υ [1]. The masses of these resonances are known to relative precisions of 4×10^{-6} and 3×10^{-5} respectively [65]. It is therefore expected that the momentum calibration will quickly reach comparable accuracies. Hence, the systematic error on the length, i.e. the length scale, will be the dominant factor in the evaluation of the proper-time systematics, and we address in this chapter a way to assess the length scale.

The method presented here, used by BaBar already [79], involves the reconstruction of secondary interactions occurring in the VeLo material. By comparing the observed spatial distribution with the known detector geometry, one can measure some of the detector elements and hence assess a length scale and its precision. The procedure needs of course to take into account the uncertainty due to the imperfect knowledge of the detector element sizes and positions.

Clearly, this method is entirely based on the reliability of the survey information; if one scales all survey measurements by an arbitrary factor this method will not see any disagreement. To resolve such ambiguities, one has to use the mass measurement of sharp resonances such as J/ψ or Υ [2] or rely on the measurement of a known lifetime.

---

[1]The momentum calibration first depends on the magnetic field calibration. Its full calibration will be a difficult task, yet a simple global factor accounting for a first order correction is relatively straightforward.

[2]In a two-body decay, the invariant mass is reconstructed as $M_{12}^2 = (E_1 + E_2)^2 - (\vec{p_1} + \vec{p_2})^2$ and therefore

However this method is sensitive to any scale factor picked up somewhere in the reconstruction process. Although is is actually hard to imagine how such an effect could occur, one still needs to assess a precision level. Applying the secondary interaction method allows to constrain all the detector reconstruction process (alignment, tracking, vertexing) to the survey scales in a way which is independent of the momentum calibration.

The secondary interaction method holds in particular for the distance between two VeLo sensors. If one is able to reconstruct the secondary interactions occurring in these sensors, one can measure the distance between them. It is important to see here that, whatever the software alignment procedure does to the sensor positions, we compare the reconstructed distance (which is sensitive to the effects of the alignment) to what is known from the survey measurements (which is not affected by the alignment). Consequently a modification of the distance scale introduced by the reconstruction or alignment software can be detected.

As the tracking and hence the vertexing explicitly uses the sensor positions to build tracks and vertices, one can reasonably question whether the above method is independent of the metrology. The survey information are only used to initialize the alignment procedure. Once the latter has converged, the sensor positions used for the reconstruction become independent of the survey measured values, and secondary interactions may be used to cross-check the module positions.

The rest of this chapter contains a description of secondary interactions studied with the Monte Carlo truth. It then explains in Sec. 7.2 the cuts applied to reconstruct these events and gives the expected selection yield and efficiency. Section 7.3 describes the model used to fit the sensor positions and also tries to extract the RF-foil shape (Sec. 7.4). We then estimate the precision of the survey measurements (Sec. 7.5) and end this chapter by giving the LHCb z scale sensitivity.

7.1   Secondary interaction characteristics

A secondary interaction is defined as the collision between a particle and the detector material. Such event can be produced by any secondary particle coming from a pp interaction, a scattered proton emerging from the beam, or a beam Bremsstrahlung photon.

From a geometrical point of view, the VeLo sensors are well suited for the study of secondary interactions. The silicon has a non-negligible cross-section and the sensors are perpendicular to the beam axis, allowing the reconstruction of many well defined z positions. Moreover, they are spread over a long distance with respect to the involved scale (the flight distance of a B meson) and hence the relative error is small.

Assuming a perfect reconstruction, we describe the characteristics of the secondary interactions extracted from a 100 k minimum bias events, i.e. events simulated as close as possible to what the LHCb detector will effectively see before any trigger or selection is applied.

It appeared, at the beginning of this study, that the secondary interactions, simulated in the LHCb framework do not include a simple tag to recognize them as such. Thus we had to design a 'selection', which we describe below, to filter true Monte Carlo secondary interactions. We will assume in the rest of this document that this selection has a 100% efficiency.

The selection requires true vertices located inside the VeLo geometrical volume but outside the beam region (8 mm $< \rho <$ 42 mm). As most of these vertices are decays in

---

contains a $\cos\theta_{12}$ dependence, which, as an angle, is sensitive to the ratio of the x or y scale over the z scale. However this method has the drawback to correlate the momentum calibration with the length scales.

flight, a very effective criterion is to require a proton in the decay products. One is then left with secondary interactions and some $\Lambda \to p\pi^-$ decays, which we reject by looking at the PID of the particles. After applying these cuts on the Monte Carlo truth, we plot a top view of the VeLo with the vertex positions (Fig. 7.1). The sensors and the RF foil are clearly visible. The small insert at the top of the figure shows that each module consists of an r and a ϕ sensor, and that the RF has left and right parts.
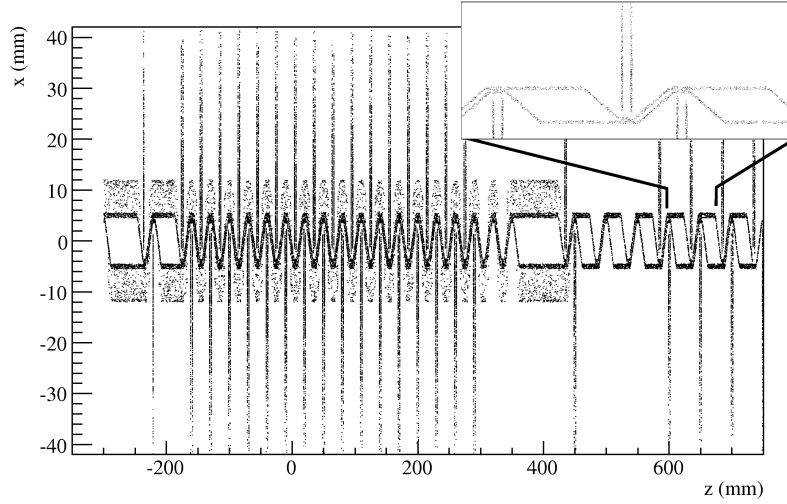

Figure 7.1: Top view of the VeLo region, showing the true positions of all the secondary interaction vertices with $8 \text{ mm} < \rho < 42 \text{ mm}$.

The selection finds at least one secondary interaction in $\sim 40\%$ of the minimum bias events. The momentum distribution of the mother particle is plotted in Fig. 7.2a; in 80% of the cases, the momentum is smaller than $10 \text{ GeV}/c$. Figure 7.2b shows the multiplicity of all outgoing particles as well as that of the charged tracks in the LHCb geometrical acceptance[3], with a mean values of 9 and 1.2 respectively. This big difference shows that the particles produced in a secondary interaction are only weakly boosted along the beam axis. This is even more obvious in Fig. 7.2c, which shows the distribution of the mean and maximum angle computed between each pair of outgoing tracks. The peak visible at 2.1 rad ($120°$) is due to secondary interactions producing three tracks with a total momentum of zero; the drop after this peak is present because only 2 tracks interaction can have a bigger mean angle. Fig. 7.2d shows the proportions of particles emerging from a secondary interaction. This is biased towards the protons as one of them is required by the selection.
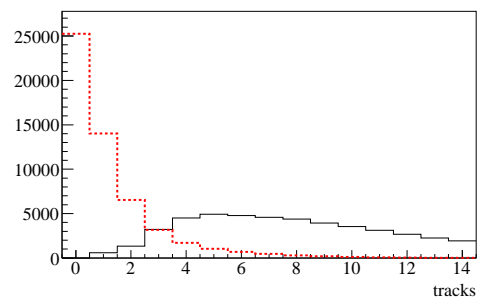
7.2 Secondary interaction reconstruction

We describe in this section the selection of secondary interactions in the VeLo material (without using the MC truth). As not enough minimum bias data were available to perform a fit of the sensor positions, we analyzed a sample of 12.5 M events containing each a D_s decay. The D_s vertex represents here an enhanced source of background in case of the mis-identification of one of the daughters as a proton, but this will not bias the sensor positions and we expect a similar or slightly better performance on a minimum bias

---

[3]We consider here an acceptance of 310 mrad in the xz plane and 255 mrad in the yz plane.
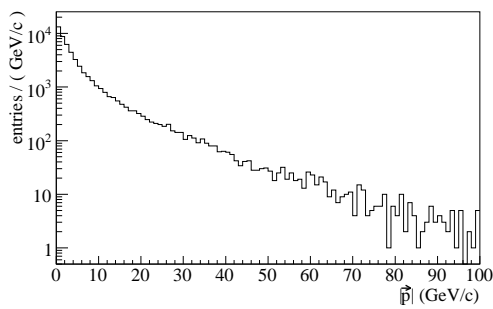
(a) Momentum of the particle initiating a secondary interaction.

(b) Multiplicity of all particles (black) and of charged particles in the LHCb acceptance (dashed red).
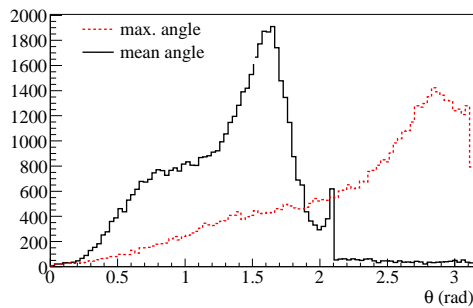
(c) Mean (black) and maximum (dashed red) angle between all pairs of tracks in a secondary interaction.

(d) Proportions of particle emerging from a secondary interaction producing at least one proton.

Figure 7.2: Properties of true secondary interactions.

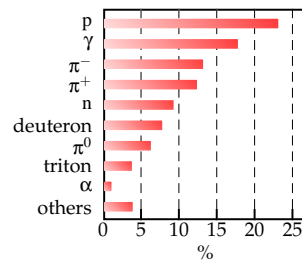sample. In fact no background, neither from combinatorics nor from particle decays will bias the sensor positions. Indeed, none of these fake secondary interaction vertices will occur in a particular z region. Thus the selection aims at filtering good quality vertices while keeping a high efficiency rather than a high purity. With real data, statistics will not be a problem, so one may decide to tighten the cuts to increase the selection purity.

7.2.1 Description of the selection

Figure 7.3 shows the relevant distributions used in the selection, both for true secondary interactions and for background candidates obtained with a loose selection. The algorithm designed to reconstruct and select the secondary interactions uses only long and upstream tracks which have a χ^2/dof fit smaller than 3. It then successively takes as seed each of the tracks identified as protons ($\Delta\mathcal{L}_{\text{p}\pi} > 1$), and forms a vertex with as many tracks as possible, still the χ^2/dof of the vertex has to be smaller than 3 (see Fig. 7.3a). In Fig. 7.3b, we plot the proton–pion DLL. Figure 7.3c shows the distribution of the smallest IPS with respect to all the PV. A cut on this variable is actually very powerful to enhance the significance ($S/\sqrt{B+S}$) of the selection. However, it has the drawback to reject a large fraction of the interactions occurring in the 150–250 mm z range as their outgoing tracks tend to point back to the primary interaction region ($z \simeq 0$). Cutting on values as high as 30 (which is the outcome of a significance optimization), leaves almost no sensitivity to the sensors in this region. We have chosen a value of 4, at the cost of a larger background, but an enhanced z-scale sensitivity. Figure 7.3d shows the distribution of the biggest angle between each pair of outgoing tracks. Cutting on this variable has two purposes: ensuring a better vertex resolution, and removing candidates built out of clones. This occurs when the same track is used for the reconstruction of two particles. We plot on Fig. 7.3e the distribution error estimate of the z vertex position, which is also a very powerful discriminant between true secondary interactions and random combinations. The cut value is fixed at 4. In addition, for two-tracks vertices we use the invariant mass to reject the $\Lambda \to \text{p}\pi^-$ and $\text{K}^0_\text{S} \to \pi^+\pi^-$ decays. A cut on the track momentum (> 2 GeV/c) is also applied. Although efficient to guarantee the track quality, it has no real discriminant power between signal and background. The selection cuts are summarized in Table 7.1.

7.2.2 Secondary interaction selection efficiency

To estimate the selection efficiency, we use a minimum bias sample of 700 k events. Applying the cuts from Table 7.1 we select 570 candidate vertices, of which $(24 \pm 2)\%$ are true secondary interaction occurring in the VeLo material and $(7\pm1)\%$ are true secondary interaction actually coming from a sensor.

The sensor position extraction described in the next sections is based on 25 k reconstructed secondary interactions. This represents, with the selection efficiency quoted above, an initial statistics of 32 M events. Assuming that the triggers do not modify the selection efficiency (which is a rough approximation), and assuming an event acquisition rate of 2 kHz, 32 M events represent 4.5 hours of data taking. This shows that nor the statistics, neither the time to collect them will be an issue for the study of the secondary interactions.

7.2.3 Properties of reconstructed secondary interactions

Most of the reconstructed vertices (97%) have only two tracks, this because the mean number of tracks in the acceptance is 1.2 (see Fig. 7.3d). We show in Fig. 7.4 the invariant

(a) Vertex χ^2/dof

(b) $\Delta\mathcal{L}_{\mathrm{p}-\pi}$ distribution

(c) Impact parameter significance

(d) Maximum opening angle between the tracks
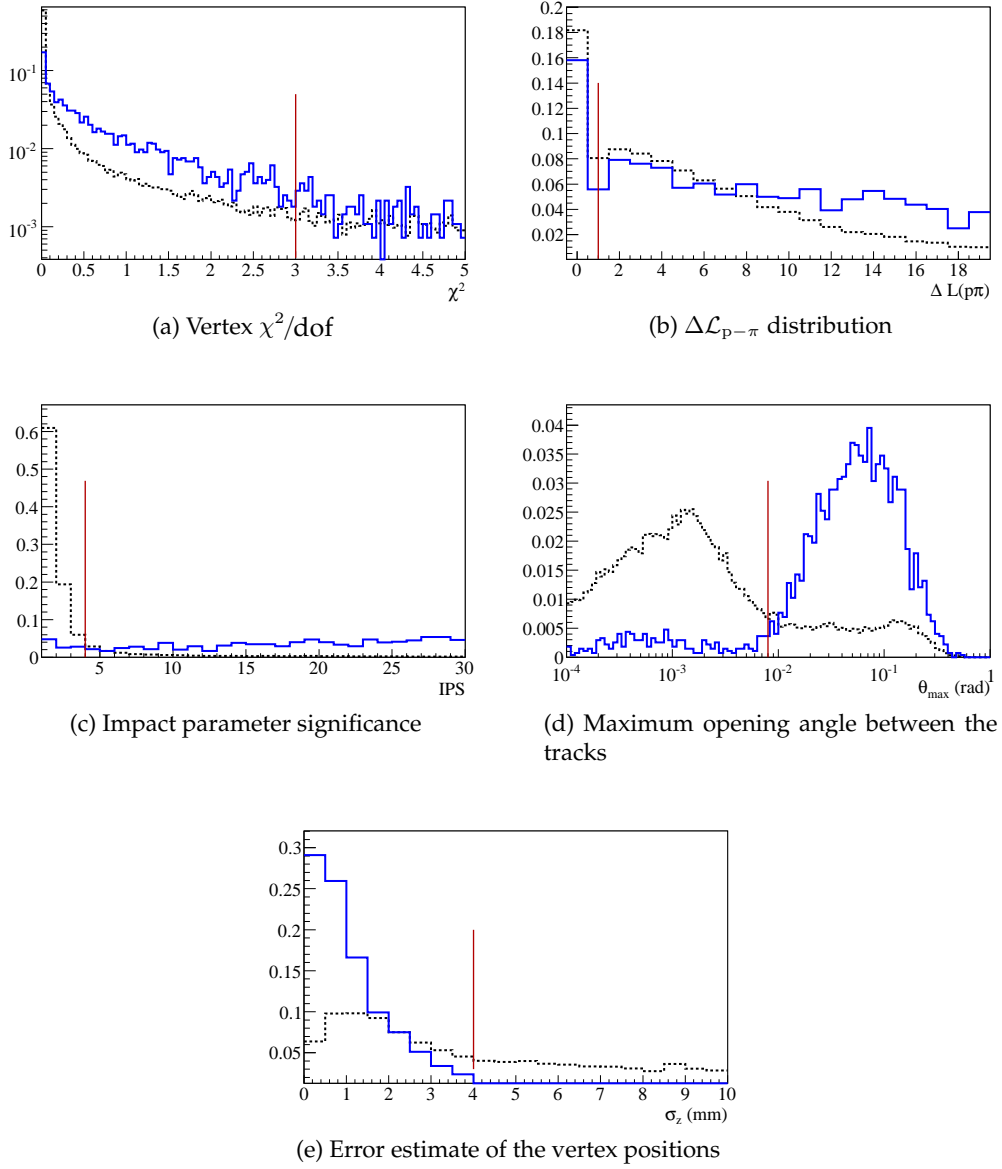
(e) Error estimate of the vertex positions

Figure 7.3: Distributions of the variables relevant for the selection of secondary interactions. The blue histograms are for true secondary interactions while the dashed black histograms are for minimum bias background events selected with a loose vertex reconstruction. All the distributions have been normalized.

Track selection		
χ^2/dof	$<$	3
p	$>$	$2\ \mathrm{GeV}/c$
IPS	$>$	4

Proton identification		
$\Delta\mathcal{L}_{\mathrm{p}\pi}$	$>$	1

Vertex selection		
χ^2/dof	$<$	3
$\max(\theta_{ij})$	$>$	8 mrad
σ_z	$<$	4 mm
z	$\in$	$[0, 700]$ mm
ρ	$\in$	$[6, 42]$ mm

Invariant mass requirements (2 prong vertices only)		
$\lvert M_{\mathrm{p}\pi} - M_\Lambda \rvert$ and $\lvert M_{\pi\mathrm{p}} - M_\Lambda \rvert$	$>$	$15\ \mathrm{MeV}/c^2$
$\lvert M_{\pi\pi} - M_{\mathrm{K}^0_\mathrm{S}} \rvert$	$>$	$10\ \mathrm{MeV}/c^2$

Table 7.1: Secondary interaction selection cuts.

mass distribution of all the selected vertices. The peak at $1.88\ \mathrm{GeV}/c^2$ is made of reconstructed proton pairs. Figure 7.5 shows the distribution of the vertices in the xz plane. The region $z > 300$ mm is quite clean and the RF-foil structure is well visible. We show on Fig. 7.6 the projection on the z axis. Its global shape represents a balance between two main effects: on one hand, the average sensor surface crossed by the collision cone increases with z, but on the other hand, forward sensors have not enough stations after them to reconstruct good quality tracks and vertices. Moreover, as mentioned earlier, the cut on the IPS tends to reduce the statistics in the 150–300 mm region. In the negative z region, the selection is overwhelmed by combinatorics and no module shape is visible, therefore we cut this region out where, according to the Monte Carlo truth, no interaction should be seen as they are boosted in the wrong direction.

We show in Fig. 7.7 the projection of the reconstructed vertices on the x axis. The two very narrow peaks are due to the RF-foil structure visible in the high-z region of Fig. 7.5. The two smaller innermost peaks are the projection of the RF-foil shape in its narrow regions, whereas the two broader external peaks are the projection of the RF-foil shape in the z ranges between the modules. The left-right asymmetry can be explained by the magnetic field polarity and the fact that one proton is required in the selection. Thus the average electric charge of the vertices is $+1$ (most of the events are $\mathrm{p}\pi^+$ or $\mathrm{p}\pi^-$). Positive tracks reconstructed on the left VeLo side (positive x) tend to fly on the left side of the detector and are kicked on the right side by the magnetic field. Whereas positive tracks reconstructed on the right side travel in the right side and are kicked out of the detector acceptance by the magnet. The projection of Fig. 7.7 could be used, among other things, to resolve a resolution model along the x axis and, should that be of any use, to check the distance between the two foils.
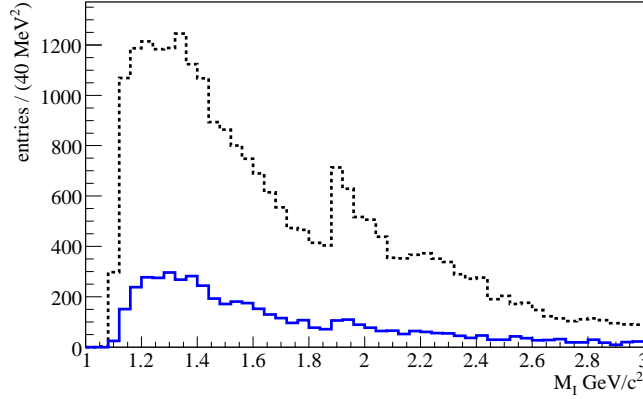
Figure 7.4: Invariant mass distribution for selected secondary interactions. The blue continuous line represents the true secondary interactions. The dashed black line is the full selected statistics.

7.2.4   Selection biases

Figure 7.8a shows, for true secondary interactions, the z position bias computed as $z_{\text{rec}} - z_{\text{true}}$. Figure 7.8b shows the corresponding pull $\left(\frac{z_{\text{rec}} - z_{\text{true}}}{\sigma_z} \right)$. We see that the reconstruction has a systematic shift of 46 μm, i.e. the reconstructed vertices tend to be displaced in the forward direction. To investigate this effect, we plotted the bias as a function of the true vertex z position (Fig. 7.9). We conclude that no obvious pattern is visible. We then plotted the bias as a function of the extrapolation distance, defined as the distance in z between the true position of the secondary interaction and that of the closest module in the forward direction. In other words, this is the distance to the first module that may have been hit by the daughters. A sketch of two extrapolation distances is given in Fig. 7.10. By bins of 1 mm in the extrapolation distance, we perform a Gaussian fit of the pull distribution. We show their central values and associated errors in Fig. 7.11.

The plot deserves some explanations. First of all, the point visible at an extrapolation of 0 is the result of secondary interactions that occurred in a sensor, we see that these are almost not biased (1.6σ effect). Then no point is visible until -6 mm because that is the closest distance (in z) that the RF foil gets to the modules. With such small extrapolation, the distance between the two tracks hitting the sensor is ~ 300 μm (assuming an opening angle of 60 mrad), hence the two hits are separated by about 3 strips. The VeLo clusters being made out of 4 strips at most, it is likely that the two hits get sometimes merged in a single cluster and hence that the tracks are slightly pulled together. This in turn, attracts the reconstructed vertex toward the sensor, hence a positive bias for small extrapolation distances, which we see in the bins at -6 mm and -7 mm. A similar explanation holds for the bias at -15 mm. Indeed, a secondary interaction occurring in a left-side module can encounter, 15 mm further, the right-side module as its first measurement point, leading to the same phenomenon of vertex attraction if the tracks are not separated well enough.

However, due to the limited available statistics, the effects described above have no more than 2σ significance. Hence from the above explanation, we retain the following: secondary interactions when occurring in a sensor are not biased but, it is likely that a particle interacting just in front of a sensor will have its reconstructed vertex shifted towards the sensor. A way to remove this bias would be to reconstruct the secondary interactions and, after the module located just after the interaction has been identified, to
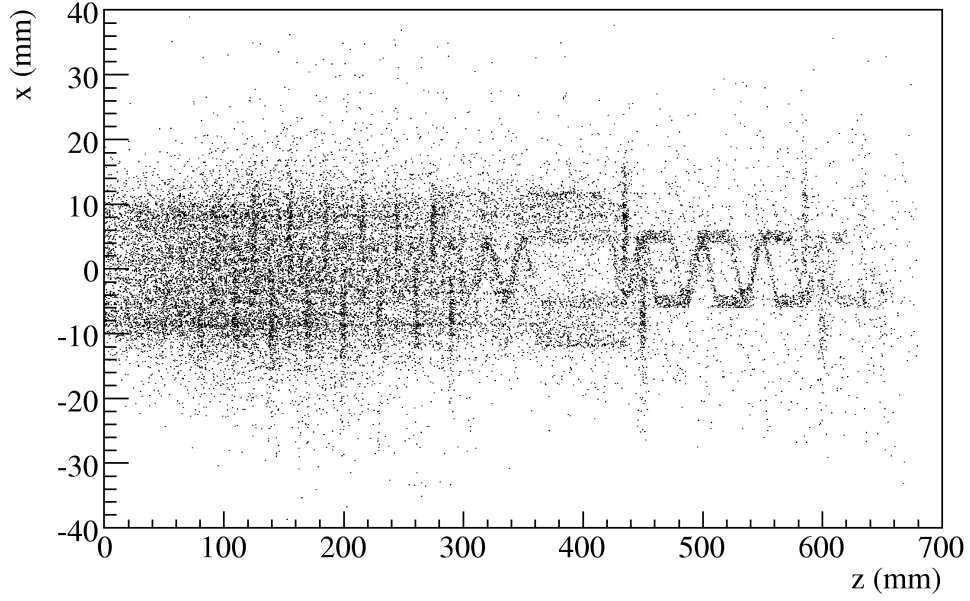
Figure 7.5: Reconstructed secondary interactions projected on the xz plane. The bin width in each direction is 300 μm (equal to the VeLo sensor thickness).
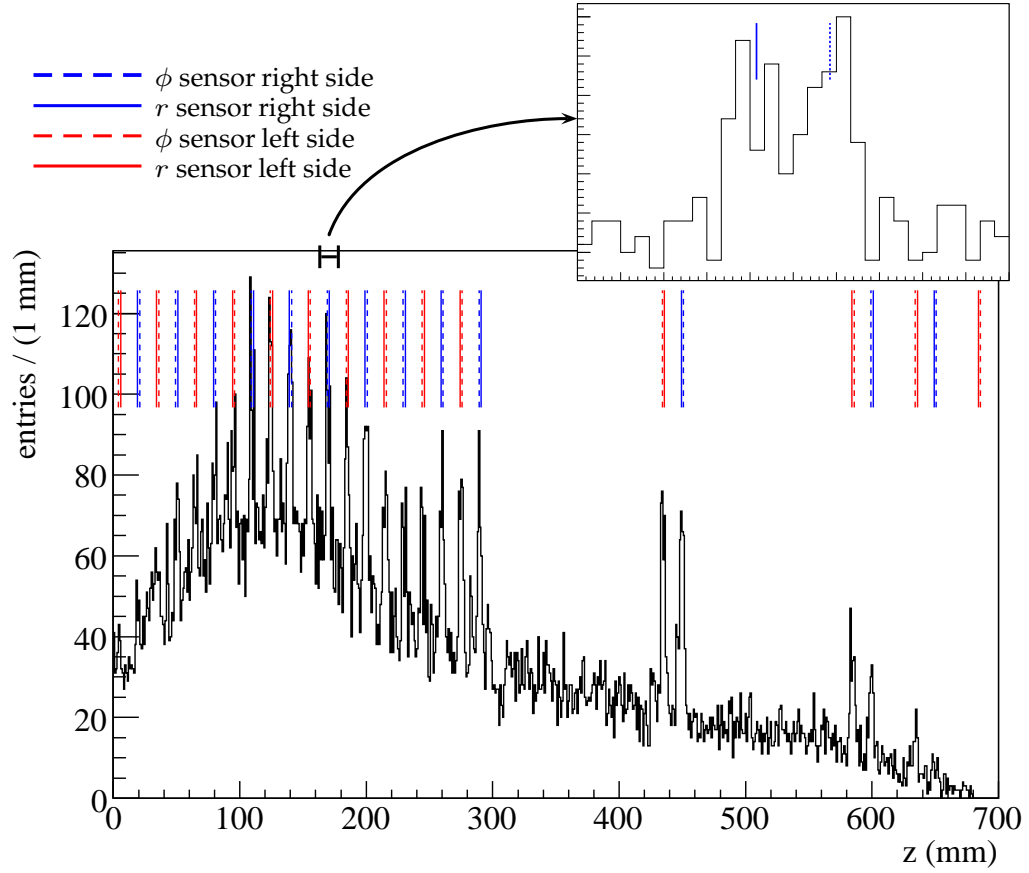

Figure 7.6: Reconstructed secondary interactions projected on the z axis. The bin width is 1 mm. The insert on the top right shows one specific module with a bin width of 300 μm.
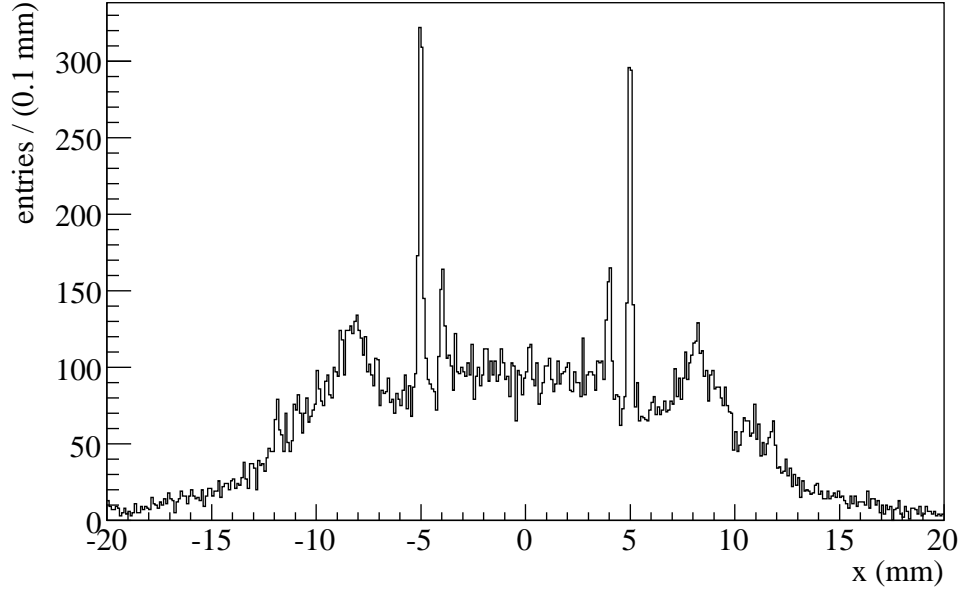
Figure 7.7: Projection or the reconstructed vertices on the x-axis. The bin width is 100 μm.

(a) Vertex z-position resolution
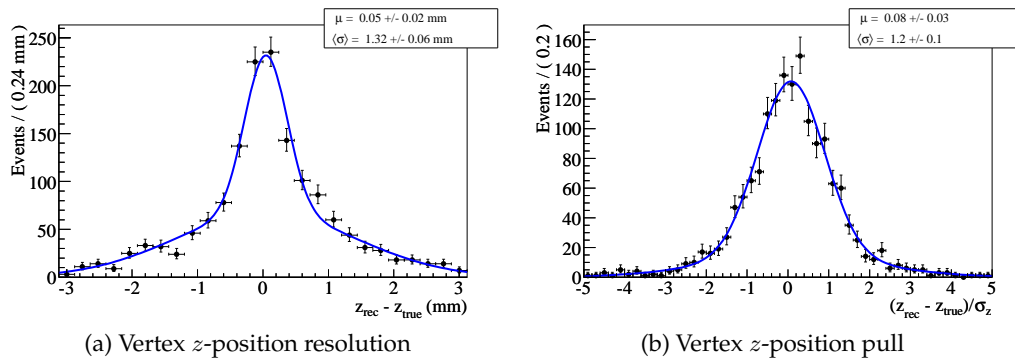
(b) Vertex z-position pull

Figure 7.8: Resolution and pull of selected true secondary interactions. The two fits are done with a double Gaussian with a common mean. The average sigma value is computed as $\langle\sigma\rangle = \left(f_1\sigma_1^2 + (1-f_1)\sigma_2^2\right)^{\frac{1}{2}}$.
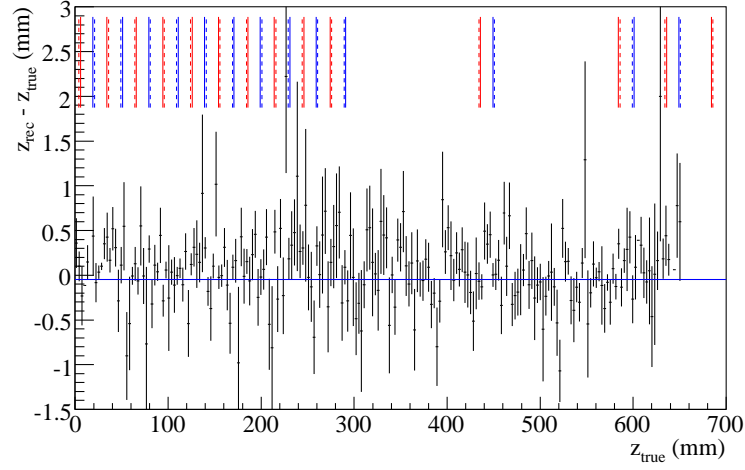
Figure 7.9: Bias in the z direction along the z axis. The error bars are defined as RMS/$\sqrt{N}$. The bin width is 3 mm, about one module width. A line as the bias mean value (46 μm) is drawn.
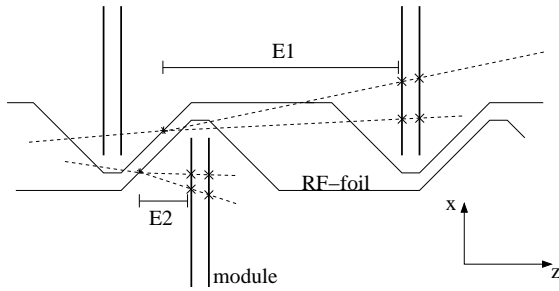

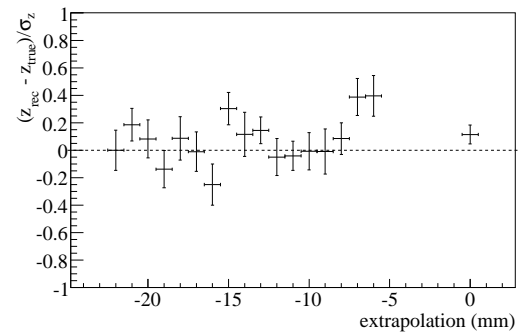Figure 7.10: Sketch of two extrapolation distances, $E1$ and $E2$.

Figure 7.11: Central value of the vertex-position pull versus the extrapolation distance. A line at 0 is drawn.

re-run the tracking and the vertexing, ignoring all the hits in this module. This has not been done in the present analysis mainly for time reasons.

7.3   Extraction of the sensor positions

A fit of the secondary interaction vertex distribution along the z axis is performed to determine the sensor positions. In the present Monte Carlo study, we compare the measured positions to the true sensor positions. However, in the real data analysis, one will use the survey measurements as reference.

The implemented fit uses an event-per-event resolution. We used, for the signal, a double Gaussian because Fig. 7.8 suggests that the pull has non Gaussian tails. Because the error estimate is correlated with the z vertex position (which is expected as forward vertices have less stations to measure their tracks), we included a numerical approximation of the local error.

7.3.1   Model description

A sensor shape is described along the z direction by a step function θ defined as

$$\theta_{z^0,T}(z) \;\; = \;\; \left\{ \begin{array}{ll} 1/T & \text{if } z \in [z^0, z^0 + T] \\ 0 & \text{elsewhere,} \end{array} \right. \tag{7.1}$$

convoluted with the detector resolution R modeled as the sum of two Gaussians:

$$R_{s_1,s_2,F}(z|\delta z) \;\; = \;\; \frac{1}{\sqrt{2\pi}\delta z}\left[\frac{F}{s_1}\mathrm{e}^{-\frac{1}{2}\frac{z^2}{(s_1\delta z)^2}} + \frac{1-F}{s_2}\mathrm{e}^{-\frac{1}{2}\frac{z^2}{(s_2\delta z)^2}} \right] \,, \tag{7.2}$$

where δz is the estimated error on z. The two parameters s_1 and s_2, called scale factors, account for a possible global mis-estimation of the errors. Thus a sensor probability density function (pdf) is

$$\begin{aligned} S_{z^0,T,s_1,s_2,F}(z|\delta z) \;\; = \;\; & \theta(z) \otimes R(z) = \int_{-\infty}^{\infty} \theta(z')R(z-z')\mathrm{d}z' \\ = \;\; & F\left\{ \Gamma\left[\frac{z-z^0}{\sqrt{2}s_1\delta z}\right] + \Gamma\left[\frac{-z+z^0+T}{\sqrt{2}s_1\delta z}\right] \right\} \\ & +(1-F)\left\{ \Gamma\left[\frac{z-z^0}{\sqrt{2}s_2\delta z}\right] + \Gamma\left[\frac{-z+z^0+T}{\sqrt{2}s_2\delta z}\right] \right\} \,, \end{aligned} \tag{7.3}$$

where $\Gamma(x)$ is the error function defined as $\Gamma(x) = \frac{2}{\sqrt{\pi}}\int_0^x \mathrm{e}^{\left(-\frac{1}{2}t^2\right)}\mathrm{d}t$. A module is then described as the sum of two sensors separated by a distance Δ

$$M_{z^0,T,\Delta,r,s_1,s_2,F}(z|\delta z) \;\; = \;\; rS_{z^0,T,s_1,s_2,F}(z|\delta z) + (1-r)S_{z^0+\Delta,T,s_1,s_2,F}(z|\delta z) \,, \tag{7.4}$$

were r is the relative fraction of the two sensors. We sketch in Fig. 7.12 the different parameters assuming a perfect resolution.

The background is modeled by a linear function

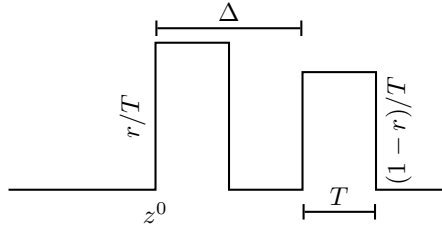$$B(z|\delta z) \;\; = \;\; \frac{1}{N_b}(bz+1) \,, \tag{7.5}$$

Figure 7.12: Illustration of the different parameters describing a module. A perfect resolution is assumed.

where N_b is a normalization term[4]. The total join pdf of the background and a module is multiplied by the pdf of the error estimate $E(\delta z)$, which we model by a numerical approximation[5]. The likelihood function becomes:

$$\mathcal{L}_k(z_k^0, T, \Delta, r_k, b_k, f_k, s_1, s_2, F) = \prod_{z_i \in \mathbb{R}_k} \{f_k M_k(z_i|\delta z_i)+$$
$$(1 - f_k)B_k(z_i|\delta z_i)\} E_k(\delta z_i), \qquad (7.6)$$

where f_k is the relative fraction of between the pdf of the module M_k and the pdf of the background B_k; and $\mathbb{R}_k$ is defined as a region of ± 1.5 cm around the module nominal central position ($z^0 + 0.5(\Delta + T)$); k is the module index, and $\{z_i, \delta z_i\}$ are the measured vertex z positions and error estimates. Eventually, the full likelihood is defined as

$$\mathcal{L} = \prod_k \mathcal{L}_k(z_k^0, T, \Delta, r_k, b_k, f_k, s_1, s_2, F). \qquad (7.7)$$

To minimize the likelihood, we use the RooFit toolkit, based on the ROOT framework [80, 72] which serves as interface to the MINUIT minimization package [73]. The latter implements the MIGRAD algorithm, which we use for the minimization and the HESSE algorithm, which computes numerically the full second derivative matrix of the function and inverts it. It therefore estimates the parameter errors as parabolic errors.

7.3.2 Fit procedure

The fitting is done in two phases. First, one fixes the resolution parameters s_1, s_2, F to the values describing the pull from Fig. 7.8b and perform a separate maximization of $\mathcal{L}_k$ (Eq. 7.6) for each module M_k. This first step mainly allows to begin the full fit with meaningful values and errors. The second phase is a simultaneous unbinned likelihood fit of $\mathcal{L}$ (Eq. 7.7), with free scale factors (s_1, s_2, F). During all the fits, we keep fixed at their nominal values the sensor thicknesses T and the distance Δ between the two sensors in each module. In this second phase, the number of free parameters is $4N + 3$, where $N = 20$ is the number of fitted modules[6] and 3 stands for s_1, s_2 and F. 83 is a large number of free parameters, yet, by construction, they are only correlated through the scale factors, hence the fit manages to converge. Table 7.2 summarizes the fit parameters, and their state during the fit.

---

[4]It happens that any scale factor put in a Gaussian resolution model vanishes in the normalization in case of an order one polynomial. Thus, there is not point in adding a resolution model to the background. This is not true anymore for a second order polynomial.

[5]We use here the so-called `RooKeysPdf` to approximate the data.

[6]Out of a total of 42 modules in the VeLo.

Parameter description		state	initial value
Module position	z_k^0	free	true position
Space between the module sensors	Δ	fixed	2 mm
silicon thickness	T	fixed	300 μm
Module sensor amplitude fraction	r_k	free	0.5
Background slope	b_k	free	0
Background fraction	f_k	free	0.2
First scale factor	s_1	free in the 2$^{\text{nd}}$ fit	0.8
Second scale factor	s_2	free in the 2$^{\text{nd}}$ fit	2
Resolutions fraction	F	free in the 2$^{\text{nd}}$ fit	0.7

Table 7.2: Fit parameters. The subscript k indicates the module number.

7.3.3   Fit results

The fit successfully estimates the position z_k^0 of 20 modules along the z axis, situated between $z = 50$ mm and $z = 600$ mm. The largest difference $z_{k\,\text{fitted}}^o - z_{k\,\text{true}}^0$ is 170 μm, which is less than the sensor thickness. We show in Fig. 7.13 the pull distribution for all the modules $(z^0_{\text{fitted}} - z^0_{\text{true}})/\sigma_z$, where σ_z is the estimated error on $z_{k\,\text{fitted}}^0$, and in Fig. 7.14 two examples of fitted modules.
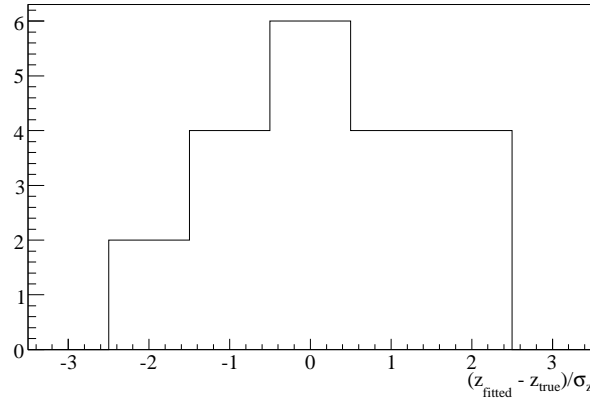

Figure 7.13: Pull of the module positions. Fitting the distribution with a Gaussian gives a mean value of 0.4 ± 0.7 and a sigma of 1.8 ± 0.9.

The pull as well as the biases show that the fit is working well, still, two things indicate that the model used is not fully optimal. First, the Δ_k –the distances between the r and ϕ sensor in each module– were held fixed in the above fit. Let as free parameters, they converge to a value which is on average 400 μm too big, but as the two sensors in a module are shifted in opposite directions (i.e. the z_k^0 are biased as well), this has no effect on the z-scale estimated value. Furthermore, we did the z-scale calibration with the Δ_k as free parameters. It has also no influence on the error of the estimated z scale because the loss in precision (the sensor positions are more spread) gets compensated by the doubled number of measured positions; indeed, if the Δ_k are free, it means one actually fits independently every sensor position instead of every module (made of an r
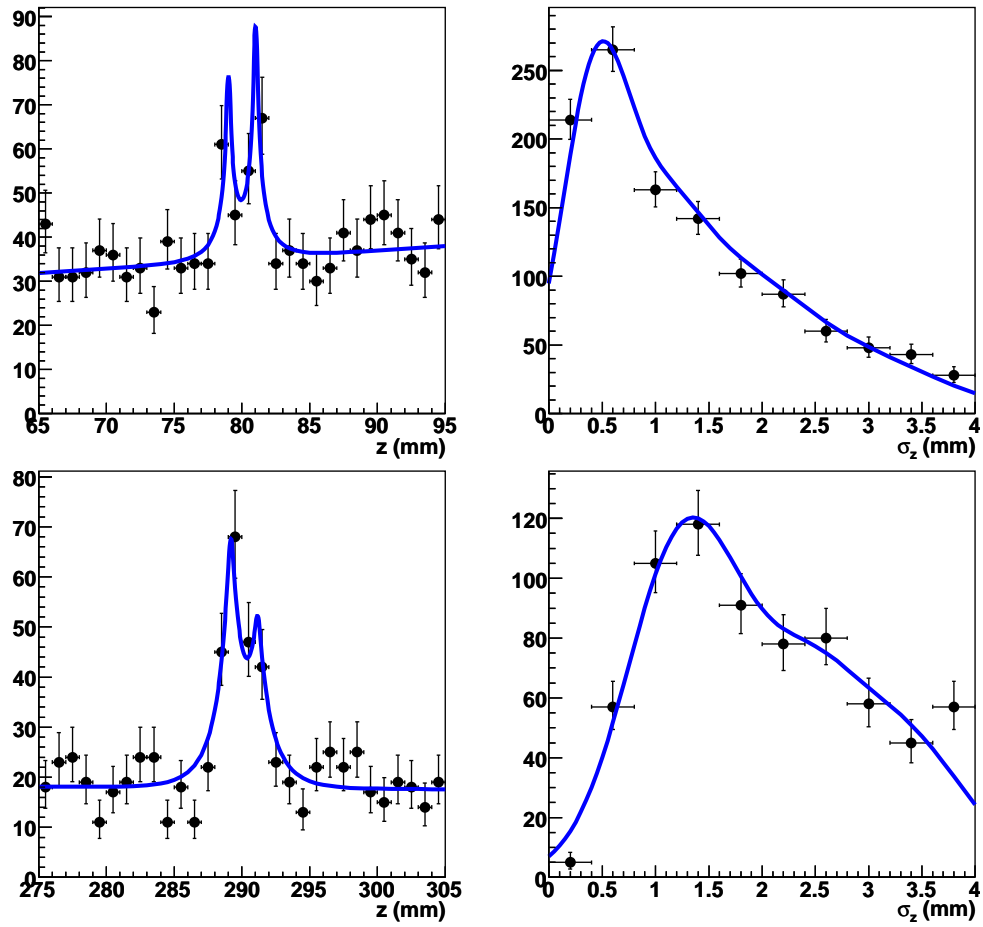
Figure 7.14: Two examples of fitted module positions and their corresponding error distribution. The curves on top of the error estimates are numerical approximations.

and a ϕ sensor). The second point, indirectly related to the module widths, is the fact that the scale factors converge to $\langle \sigma \rangle = \left[0.6 \times 0.34^2 + (1 - 0.6) \times 0.9^2 \right]^{\frac{1}{2}} = 0.62 \pm 0.06$. This is a ten sigma discrepancy with the values extracted from the pull distribution shown in Fig. 7.8b. Again, this effect is symmetric around the module central position and does therefore not influence the z-scale estimate. The two effects mentioned above could be a consequence of the sensor 'vertex attraction' described in Sec. 7.2.4, but this remains unverified at the moment.

We now aim to extract the matching between the fitted module positions and their true values. We have a set of fitted module positions $z^0_{k\,\mathrm{fit}}$ with their estimated errors σ_k and their corresponding true values $z^0_{k\,\mathrm{true}}$. We perform a linear fit of the expression $\xi z + B$ to the data points $\left\{ z^0_{k\,\mathrm{true}},\ z^0_{k\,\mathrm{fit}} \pm \sigma_{zk} \right\}$. In this expression, ξ is the scale factor between the detector geometry taken from the Monte Carlo truth and the reconstructed positions, whereas B represents an offset between these two quantities. If the reconstructed z scale perfectly matches the true scale, we expect $\xi - 1 = 0$ and $B = 0$. We solve the above system with the least squares method [74] and assume the σ_z are independent (hence the standard variances on ξ and B can be computed). Figure 7.15 shows the fit result, we get

$$\xi - 1 = (1 \pm 7) \times 10^{-5}, \tag{7.8}$$
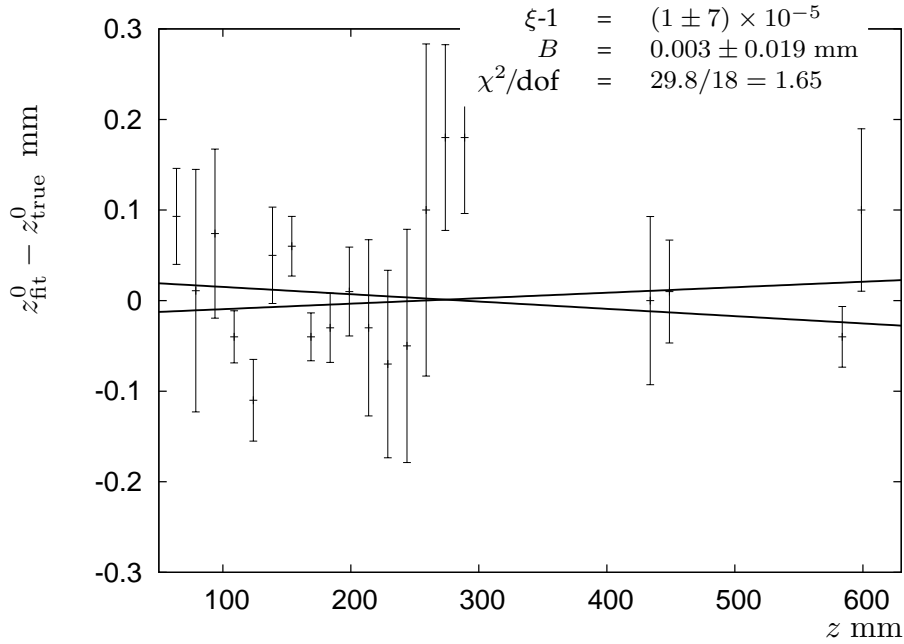
$$B = (0.003 \pm 0.019)\ \mathrm{mm}. \tag{7.9}$$

Figure 7.15: Module position residuals along the z axis. A linear fit is performed through the points, we plot the one sigma standard-deviations $((\xi \pm \delta\xi) - 1)z + B \mp \delta B$.

The statistical uncertainties $\delta\xi$ and δB are the sensitivities to the two kind of z scale mis-match. These sensitivities are directly correlated to the amount of analyzed statistics, which, in the above case, corresponds to 25 k reconstructed secondary vertices.

7.4 Extraction of the RF-foil positions

Figure 7.5 shows that the forward z region is very clean, mainly because almost all the combinatorics, created by decays in flight and primary vertex tracks, occur in the low z region. We therefore tried to fit the positions of the RF-foil corrugations, visible in the xz plane. To do so, we keep only reconstructed interactions in the central band ($|x| < 3$ mm) and the forward region ($450 < z < 600$ mm). We perform a $-\pi/4$ rotation of the xz plane, such that the RF-foil corrugations become perpendicular to the new basis. We have:

$$z' = z\cos(-\pi/4) - x\sin(-\pi/4) \; ,$$
$$x' = z\sin(-\pi/4) + x\cos(-\pi/4) \; ,$$

Figure 7.16 shows the projection of the reconstructed vertices on the z' axis. The RF-
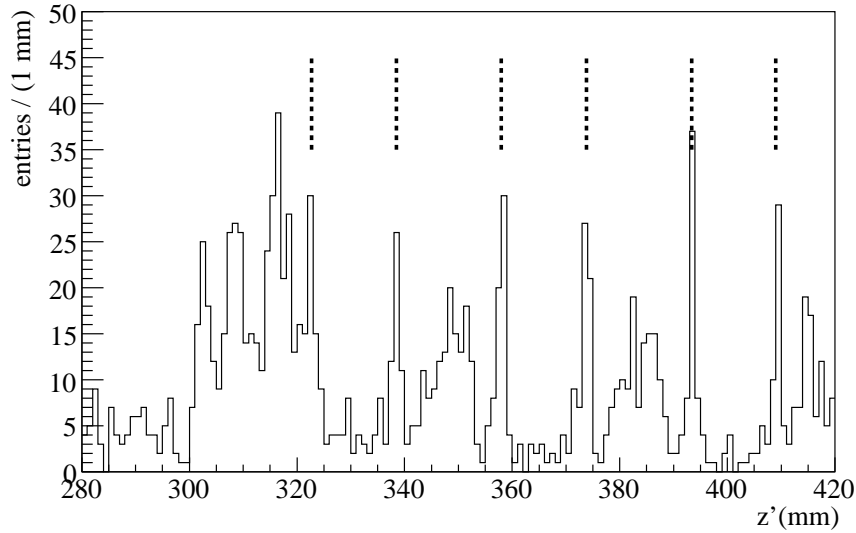

Figure 7.16: Reconstructed secondary interactions projected on the z' axis, which is perpendicular to the RF-foil corrugations. The positions of these corrugations are indicated with the dashed lines.

foil peaks are visible but mixed with the rotated sensor projections. We then fit each of these, in a narrow window, with a simple Gaussian and a linear local background (no errors are taken into account). We give two examples of fitted positions in Fig. 7.17. Eventually, we repeat the fit procedure of Sec. 7.3.2 to extract the matching between the reconstructed positions with respect to the nominal positions. The fit results are shown in Fig. 7.18. We obtain a scale factor of $\xi^{\mathrm{RF}} = 0.9975 \pm 0.0012$, which deviates from 1 at the two-sigma level. The method sensitivity is a factor 10 smaller than the fit of the sensor positions. On top of that, the positions and shape of the RF-foil corrugations (which is most probably not exactly flat) is much less accurately known than that of the sensors positions.

7.5 Survey measurements of the VeLo modules

The VeLo metrology is a non trivial task performed by the survey team. It consists of measuring 19 points per module face, among which ten are on the silicon itself. The two
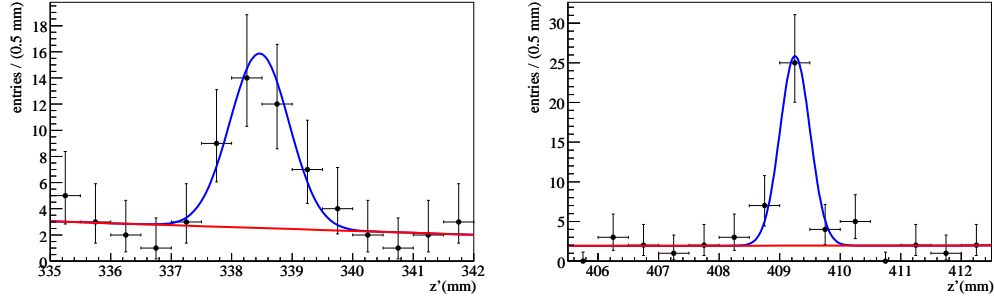
Figure 7.17: Two examples of fitted RF-foil positions.
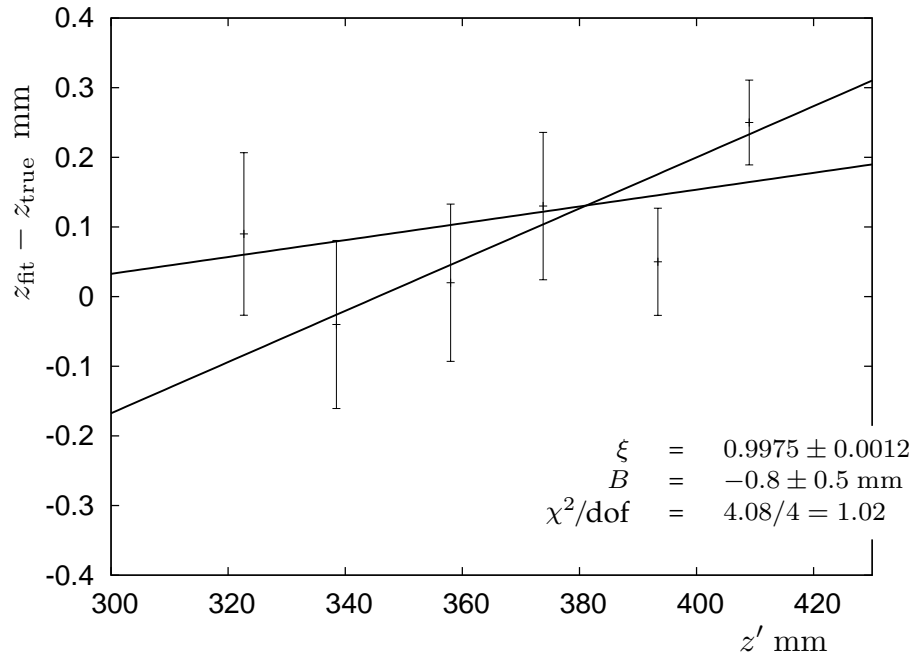

Figure 7.18: Deviation of the fitted RF-foil corrugations with respect to
their nominal positions.

VeLo halves were surveyed independently, in a clean room, with an optical measuring machine [81] at room temperature. We discuss hereafter the statistical precision of these measurements and then comment about the temperature effects.

7.5.1 Survey statistical error

In the case of the VeLo, the survey precision is not limited by the measurement device itself, which can easily reach accuracies of the order of 1 μm, but rather by the assessment of what is measured. At these levels, one can reasonably start to ask whether, for example, the target surface is well defined, whether its thickness is known, or what was the apparatus calibration process. The difficulty is how to merge in a single z position a set of ten points describing a non-flat surface[7]. We decided to consider a $\sigma_z^S = 100$ μm error on the module positions. This value comes from Ref. [81] and accounts for the possible tilt of the modules, due for example to the repeated attachment and removal of the cables.

To estimate the uncertainty on the z scale measured by the surveyors, we perform a trivial linear fit of all the module measured positions versus themselves, assuming an independent uncertainty of 100 μm for each measurement[8]. By construction the slope ξ^S is equal to 1 and holds no information, but its standard-deviation $\delta\xi^S$ can be computed and represents the 'statistical' error on the z scale. We get, considering each VeLo half separately before combining their results in quadrature[9] and with 21 modules in each half:

$$\delta\xi^S \;=\; 6 \times 10^{-5}\,. \tag{7.10}$$

This value represents an absolute lower limit on the z scale sensitivity, when calibrated on the VeLo geometry.

7.5.2 Systematic effects

As stressed already, if the survey measurements are wrong by some factor, the secondary interaction method described above will not be sensitive to it. One could think that temperature effects will scale the VeLo lengths as the module bases are in aluminum. But, as described in Sec. 2.2.1, the VeLo modules are held, relative to each other by a carbon fiber (CF) structure which constrains the global z scale. The actual value of the linear expansion coefficient for CF varies as a function of the exact composition and fabrication process of the CF, but typical values range from -2×10^{-6} K^{-1} to 0 and thus, a temperature offset of $30°$ with respect to the room temperature is required to produce a dilatation of the VeLo which would have the same amplitude as the survey measurements statistical fluctuations (Eq. 7.10). In the VeLo setup, the module bases are actively maintained at room temperature, and this implies also a monitoring of the different VeLo temperatures. Therefore a variation of more than a few degree is rather unlikely. We conclude that no effect from the temperature should be expected.

7.6 The LHCb length scale sensitivity

As mentioned in the introduction of this chapter, the propertime systematics has a contribution from the length scale, which will, in particular, affect the measurement of Δm_s.

---

[7]The VeLo sensors have shown deviations of 14 μm from a perfect plane.

[8]Although a bit strange, this approach is the exact analogy of what has been done in Sec. 7.3.3.

[9]To accommodate the fact that the two VeLo halves were surveyed separately.

We showed in this chapter that a collected statistics of ~ 30 M minimum bias events is sufficient to constrain the z scale perceived by the detector to an accuracy equal to the survey measurement precision. Ten times this statistics would tighten the constraint to a negligible fraction of the survey measurement accuracy.

Hence, provided a statistics of about 300 M minimum bias events, the LHCb length scale sensitivity can be assessed at the 6×10^{-5} level, moreover no effect from the temperature is expected.

We presented in this chapter a method to constrain and cross-check the software reconstruction quality to the survey measurements with a precision that quickly hits the survey accuracy. This method uses secondary interactions occurring in the sensor material to perform a radiography of the VeLo, allowing the comparison of the known detector lengths to the reconstructed picture. As side effect, it provides insight to the resolution model in the x and z directions.

Conclusion

We presented in this document various studies that can be divided in three main parts.

The signals routed from the LHCb vertex locator go through different analog electronic boards, devices and transmission lines. Among these, the sensors themselves, the Beetle chips and the copper cables. All of these elements are known to create cross-talk in ways which can be space or time related. The cross-talk coming from the analog line itself has been studied separately in the laboratory and an analog board (the Driver card) has been designed to cancel the cable cross-talk. The analysis of the system performance shows that the cable cross-talk after correction is below 5%. The readout cross-talk observed after a full acquisition chain is a superposition of all the various couplings occurring in the readout. Still, a careful selection of the analyzed channels allows to disentangle the different cross-talk sources. The sensors have two contributions: one is the cross-talk between sensor strips and is estimated to be $(3 \pm 1)\%$, while the other comes from the inner-to-outer cross-talk and has a mean amplitude of $(7 \pm 2)\%$. The analog lines, when compensated, have an absolute contribution of $(4 \pm 1)\%$ acting in the first chip channel neighbors. The design of a FIR filter, implemented at the software level in the TELL1 boards, allows to refine the readout cross-talk correction, individually on each analog line. The net result is a final readout cross-talk level reduced to $(1 \pm 2)\%$.

The study of $B_s \rightarrow D_s^- \pi^+$ decays allows the measurement of the B_s oscillations frequency as well as other parameters of interest, like the wrong-tag fraction. These decays will thus represent one of the important control channels for the LHCb analyses. We presented the selection of the $B_s \rightarrow D_s^- \pi^+$ events, using fully simulated Monte Carlo events. We showed that the expected yield after the L0-trigger is 155 k, assuming 2 fb^{-1} of collected data. The background level in this selection has been studied with generic background samples, and the potentially dangerous sources have been analyzed separately with dedicated samples. The overall combinatorial background level for this selection is expected to correspond to a background-over-signal ratio in the range $[0.004, 0.05]$ at 90% CL. The contribution from specific modes like $\Lambda_b \rightarrow D_s^- p$ or $B_d \rightarrow D^- \pi^+$ is expected to be in the range $[0.06, 0.4]$ at 90% CL. To assess the statistical sensitivity to the B_s oscillations, we used fast parametrized Monte Carlo samples. We introduced in the model a parametrization of the signal and of the background, including a detector resolution and a parametrization of the proper-time error estimates. We show that the statistical sensitivity to the B_s oscillations is $\sigma(\Delta m_s)/\Delta m_s = 4 \times 10^{-4}$ and that the wrong-tag fraction can be extracted with an absolute precision of 0.3%, provided that the resolution scale factor is known.

We then addressed one of the systematics involved in the measurements of any proper-time related quantity. We showed that the reconstruction of secondary interactions occurring in the VeLo material can be used to extract the sensor positions and assess the length scale by comparing these positions to the survey measurements. This method has a statistical error which quickly becomes negligible with respect to the error of the metrology measurements. We estimated the latter to correspond to a relative error on the length

scale of 6×10^{-5}. This is an absolute limit to the accuracy of this method and represents a negligible contribution with respect to the statistical error expected on Δm_s with 2 fb^{-1} of data.

To conclude, the trials with the test beams, in which the author was involved, have shown that LHCb is able to reconstruct the particles passing trough the VeLo, and that the acquisition chain is working with good performances. The results from the other sub-detectors are similar. Thus one can be confident that LHCb will be ready for the first data taking, which will start by the end of the Summer 2008. At this moment, one of the very first goals of the collaboration will be the measurement of Δm_s as this will represent a major proof that the LHCb detector is functional.

Bibliography

[1] G. BRANCO *et al.*, "CP Violation". Oxford University Press, New York, 1999.

[2] S. WEINBERG, "A Model of Leptons". Phys. Rev. Lett., volume 19, 1264–1266, 1967.

[3] A. SALAM, "Weak and Electromagnetic Interactions". Elementary Particle Theory, 1968.

[4] L. WOLFENSTEIN, "Parametrization of the Kobayashi–Maskawa Matrix". Phys. Rev. Letter, volume 51, 1945–1947, 1983.

[5] R. ALEKSAN, "Determining the CP violating phase γ". Z. Phys. C54, p. 653-659, 1992.

[6] S. COHEN *et al.*, "$\gamma + \phi_s$ sensitivity studies from combined $B_s \rightarrow D_s^- \pi^+$ and $B_s \rightarrow D_s^\mp K^\pm$ samples at LHCb". LHCb 2007-041, 2007.

[7] CERN public web pages, `http://www.cern.ch`.

[8] LHC Collaboration, P. LEFÈVRE *et al.* , "The Large Hadron Collider: conceptual design". CERN-AC-95-05-LHC, 1995.

[9] LHCb Collaboration, S. AMATO *et al.* , "LHCb Technical Proposal", CERN/LHCC 98-4, 1998.

[10] P. NASON *et al.*, "Bottom Production". hep-ph/0003142v2, 2001.

[11] LHCb Collaboration, ANTUNES-NOBREGA *et al.* , "LHCb reoptimized detector design and performance, Technical Design Report". CERN/LHCC 2003-030, 2003.

[12] LHCb Collaboration, A. A. ALVES *et al.* , "The LHCb Detector at LHC". Submitted to The Journal of Instrumentation (JINST), 2008.

[13] LHCb Collaboration, P. BARBOSA-MARINHO *et al.* , "Vertex Locator Technical Design Report". CERN/LHCC 2001-0011, 2001.

[14] P. TURNER *et al.*, "R-measuring sensor". EDMS Nr. 401568, 2004.

[15] P. TURNER *et al.*, "Phi-measuring sensor". EDMS Nr. 401568, 2005.

[16] Mechanical Design of the Vertex Locator, `http://www.nikhef.nl/pub/departments/mt/projects/lhcb-vertex/`.

[17] M. J. KRAAN *et al.*, "Structural Analysis Of The Vacuum Vessel For The LHCb Vertex Locator (VELO)". MT-VELO 04-1, EDMS 432626, 2004.

[18] LHCb Collaboration, S. AMATO *et al.* , "LHCb magnet: Technical Design Report". CERN/LHCC 2000-007, 1999.

[19] LHCb Collaboration, A. F. BARBOSA *et al.* , "Inner Tracker Technical Design Report". CERN/LHCC 2002-29, 2002.

[20] LHCb Collaboration, P. BARBOSA-MARINHO *et al.* , "LHCb Outer Tracker". CERN/LHCC 2001-24, 2001.

[21] LHCb Collaboration, S. AMATO *et al.* , "LHCb RICH". CERN/LHCC 2000-0037, 2000.

[22] LHCb Collaboration, S. AMATO *et al.* , "LHCb Calorimeters". CERN/LHCC 2000-0036, 2000.

[23] S. LOCHNER *et al.*, "The Beetle Reference Manual". LHCb-2005-105, 2006.

[24] G. HAEFELI *et al.*, "The LHCb DAQ interface board TELL1". Nucl. Instrum. Methods Phys. Res., A, volume 560, 494–502, 2006.

[25] J. BOREL *et al.*, "Measurements on the VeLo analog link system". LPHE 2005-013, LHCb 2005-030, 2005.

[26] F. LEGGER *et al.*, "TELL1: a common readout board for LHCb". LPHE 2004-13, LHCb 2004-100, 2004.

[27] D. JONES, "Common Mode Correction Algorithms for the LHCb Velo". LHCb-2007-108, 2007.

[28] A. VLADIMIRESCU, "The Spice Book". ISBN 0-471-60926-9, 1994.

[29] K. AKIBA *et al.*, "ACDC VELO phase-2 Testbeam Setup". LHCb 2008-007, to be published, 2008.

[30] J. O. SMITH, "Introduction to Digital Filters", `http://ccrma.stanford.edu/~jos/filters/`.

[31] M. KUNT, "Traitement numérique des signaux", volume XX of *Traité d'électricité*. Presse Polytechnique, 1999.

[32] L. LOCATELLI, "Direct search for Higgs boson in LHCb and contribution to the development of the Vertex Detector". Ph.D. thesis, EPFL, CERN-THESIS-2007-073, 2007.

[33] L. EKLUND, "LHCb Vertex Locator, overview and recent progess". Nucl. Inst. Meth. A, volume 569, 2006.

[34] VETRA – VeLo Tracking software, `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/vetra/`.

[35] S. VIRET *et al.*, "Performance of the LHCb Vertex Detector Alignment Algorithm determined with Beam Test Data". Submitted to NIM A.

[36] S. VIRET *et al.*, "LHCb VELO software alignment (Part I)". LHCb 2005-101, 2005.

[37] S. VIRET *et al.*, "LHCb VELO software alignment (Part II)". LHCb 2007-067, 2007.

[38] T. SJÖSTRAND, "Monte Carlo Generators". Technical Report, hep-ph/0611247, 2006.

[39] M. MASUMOTO *et al.*, "A 623-dimensionally equidistributed uniform pseudorandom number generator". ACM Transactions on Modeling and Computer Simulation, 1998.

[40] LHCb collaboration, "GAUDI, LHCb Data Processing Applications Framework". Technical Report, User Guide, 2001, `http://proj-gaudi.web.cern.ch/proj-gaudi/`.

[41] T. SJÖSTRAND, "High-energy-physics event generation with PYTHIA 6.1". Technical Report, hep-ph/0010017, 2001, `http://home.thep.lu.se/~torbjorn/`.

[42] D. LANGE, "The EvtGen particle decay simulation package". Nucl. Inst. Meth. A 462:152–155, 2001.

[43] P. GOLONKA et al., "The tauola-photos-F environment for the TAUOLA and PHOTOS packages". CERN-TH-2003-287, hep-ph/0312240v1, 2003.

[44] I. BELYAEV et al., "Gauss, LHCb Simulation Program". Technical Report, 2005, `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/gauss/`.

[45] S. AGOSTINELLI et al., "GEANT4 – a simulation toolkit". Technical Report, Nucl. Inst. Meth. A 506:250–303, 2003, `http://geant4.web.cern.ch/geant4/`.

[46] BOOLE – The LHCb digitization program, `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole/`.

[47] MOORE – LHCb High Level Trigger application, `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/moore/`.

[48] BRUNEL – LHCb Event Reconstruction application, `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/`.

[49] DAVINCI – Analysis software for the LHCb experiment, `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/`.

[50] J. KNOBLOCH et al., "Technical Design Report". Technical Report, lCG-TDR-001, CERN-LHCC-2005-024, 2005, `http://lcg.web.cern.ch/LCG/`.

[51] DIRAC – Distributed Infrastructure with Remote Agent Control, `https://lhcb-comp.web.cern.ch/lhcb-comp/DIRAC/`.

[52] GANGA, `http://ganga.web.cern.ch/ganga/`.

[53] J. VAN TILBURG, "Track simulation and reconstruction in LHCb". Ph.D. thesis, NIKHEF, CERN-THESIS-2005-040, 2005.

[54] M. NEEDHAM, "Performance of the LHCb Track Reconstruction Software". LPHE 2008-001, LHCb 2007-144, 2007.

[55] R. E. KALMAN, "A New Approach to Linear Filtering and Prediction Problems". Trans. ASME, volume D82, 1960.

[56] L. FERNÁNDEZ, "Exclusive trigger selections and sensitivity to the B_s–$\overline{B}_s$ mixing phase at LHCb". Ph.D. thesis, EPFL thesis 3613, CERN thesis 2006-042, 2006.

[57] LHCb Collaboration, R. A. NOBREGA et al. , "Trigger Technical Design Report". CERN/LHCC 2003-31, 2003.

[58] M. KRASOWSKI et al., "Primary vertex reconstruction". LHCb 2007-011, 2007.

[59] E. RODRIGUES, "The LHCb Trigger system". LHCb 2006-065, 2007.

[60] P. KOPPENBURG *et al.*, "HLT Exclusive Selections Design and Implementation". LPHE 2005-011, LHCb 2005-015, 2006.

[61] M. CALVI *et al.*, "Flavour Tagging Algorithms and Performances in LHCb". LHCb 2007-058, 2007.

[62] M. CALVI *et al.*, "LHCb Flavour Tagging Performance". LHCb-2003-115, 2003.

[63] G. CORTI *et al.*, "Monte Carlo Event Type Definition Rules". LHCb 2005-034, 2007.

[64] J. BOREL *et al.*, "The $B_s \to D_s^- \pi^+$ and $B_s \to D_s^\mp K^\pm$ selections". LHCb 2007-017, 2007.

[65] W.-M. YAO *et al.*, "Particle Data Group", 2006, `http://pdg.lbl.gov`.

[66] CDF Collaboration, A. ABULENCIA *et al.* , "Measurement of the Ratios of Branching Fractions $B(B_s^0 \to D_s^- \pi^+ \pi^+ \pi^-)/B(B_d^0 \to D^- \pi^+ \pi^+ \pi^-)$ and $B(B_s^0 \to D_s^- \pi^+)/B(B_d^0 \to D^- \pi^+)$". PRL 98, 061802, 2007.

[67] K. KINOSHITA, "Results from $\Upsilon(5S)$ at Belle: strange beauty and other beasts". Lake Louise Winter Institute 2008, `http://thwaite.phys.ualberta.ca/ indico-0.94/conferenceOtherViews.py?view=standard&confId=0`.

[68] R. LOUVOT *et al.*, "Study of the decay $B_s \to D_s \pi$ and search for $B_s \to D_s K$ at the $\Upsilon(5S)$". Belle Note #1036, unpublished.

[69] F. LEGGER, "Reconstruction of the decays $\Lambda_b \to \Lambda(1115)\gamma$ and $\Lambda_b \to \Lambda(1670)\gamma$ at LHCb". LPHE 2006-002, LHCb 2006-012, 2006.

[70] FELDMAN *et al.*, "Unified approach to the classical statistical analysis of small signals". Phys. Rev. D, 57 3873, 1998.

[71] G. RAVEN, "Selection of $B_s \to J/\psi\phi$ and $B^+ \to J/\psi K^+$". CERN-LHCb-2003-118, 2003.

[72] W. VERKERKE *et al.*, "RooFit User Manual v2.07", 2006, `http://www.root.cern. ch`.

[73] F. JAMES, "MINUIT Function Minimization and Error Analysis". CERN, v94.1, 1994.

[74] F. JAMES, "Statistical Methods in Experimental Physics". World Scientific, 2006.

[75] P. VANKOV *et al.*, "Proper-Time Resolution Modelling". LHCb 2007-055, 2007.

[76] V. GLIGOROV *et al.*, "Monte Carlo Independent Lifetime Fitting at LHCb in Lifetime Biased Channels". LHCb 2007-053, 2007.

[77] G. PUNZI, "Comments on Likelihood fits with variable resolution". PHYSSTAT2003, arXiv:physics/0401045v1, 2004.

[78] CDF Collaboration, A. ABULENCIA *et al.* , "Observation of B_s^0–$\overline{B}_s^0$ Oscillations". PRL 97, 242003, 2006.

[79] P. ROBBE, "Etude des désintégrations doublement charmées des mésons B avec l'expérience BaBar à SLAC". LAPP-T-2002-01, 2002.

[80] R. BRUN *et al.*, "ROOT, an Object-Oriented Data Analysis Framework". CERN, v5.16, 2006, `http://www.root.cern.ch`.

[81] P. SUTCLIFFE, "VELO Module Production - Final Module Metrology". LHCb 2007-087, 2006.

Curriculum Vitæ

Jérémie Borel

Laboratory for High-Energy Physics (LPHE)
Swiss Federal Institute of Technology (EPFL)
CH-1015 Lausanne-Dorigny

Born on August 11$^{\text{th}}$, 1976
Countries of citizenship: Switzerland

EDUCATION AND DEGREES

June 2008	Ph. D. in High-Energy Physics on *The analog readout of the LHCb vertex detector and study of the measurement of the* B_s *oscillation frequency*, Swiss Federal Institute of Technology (EPFL), Switzerland.
2006	CERN-JINR *2006 European School of High-Energy Physics*, Aronsborg, Sweden.
2004 – 2008	Ph. D. studies in High-Energy Physics at the LPHE (EPFL), Switzerland.
February 2004	Physics diploma thesis on *Contribution to the study of the analog link for the LHCb solid state detector (VeLo)*, Laboratory for High-Energy Physics (LPHE), EPFL.
1995 – 1997	Maturité Fédérale Scientifique, Lausanne.
1993 – 1994	Exchange student, Berlin.