

**Contribution to the alignment of the
LHCb tracking system and
measurement of the mean lifetimes of
 $B_{d(s)} \rightarrow D_{(s)}\pi$, $D_{(s)} \rightarrow K K \pi$**

Vincent Fave
of EPFL



Abstract

The LHCb experiment is one of the four large experiment of the Large Hadron Collider (LHC) situated at CERN, on the swiss-french border near Geneva. The LHCb detector is a single-arm spectrometer dedicated to the study of rare b -hadrons decays and to precision CP violation measurements. The LHCb experiment has so far collected 1fb^{-1} of data at a center of mass energy of 7 TeV.

This thesis addresses three topics related to the LHCb experiment. The first part concerns the alignment of the LHCb Inner Tracker with the first LHC data. Misalignments in the tracking system degrade the momentum measurement and flight distance determination of particles. Such quantities are vital for accurate lifetime and mass measurements. A standalone alignment of the Inner Tracker was performed using a method to stabilize the Inner Tracker alignment without the need of fixing elements. The Inner Tracker was aligned to a precision of $102 \pm 10 \mu\text{m}$, with a bias of $0 \pm 13 \mu\text{m}$.

In the second part, a first optimization of the selection cuts of the Same Side Kaon tagger has been performed using $D_s^+ \rightarrow \phi\pi^+$ 2010 data along with B_s^0 Monte Carlo data. The $D_s^+ \rightarrow \phi\pi^+$ channel was shown to be suitable for this purpose in the absence of a large B_s^0 calibration sample. Compared to the initial seection of kaon, tuned on Monte Carlo data, an improvement both in tagging power and mistag rate was observed.

In the third part, a method to fit the B_s^0 lifetime in $B_s^0 \rightarrow D_s^- \pi^+$ and B^0 lifetime in $B^0 \rightarrow D^- \pi^+$ was developed using a global decay time acceptance function and per-event decay time uncertainty estimates. The precision of the fits result is found to be dominated by systematic uncertainty due to the precision to which the acceptance function can be modeled. The average B_s^0 lifetime is measured to be

$$\tau_{B_s^0} = 1.515 \pm 0.015(\text{stat})_{-0.039}^{+0.041}(\text{sys}) \text{ ps}$$

This result has an accuracy comparable to the world average provided by the Particle Data Group [1] using $B_s^0 \rightarrow D_s X$ decays

$$\tau_{B_s^0} = 1.425 \pm 0.041 \text{ ps}$$

In the case of the B^0 , the average lifetime is measured to be

$$\tau_{B_d^0} = 1.513 \pm 0.020(\text{stat})_{-0.037}^{+0.033}(\text{sys}) \text{ ps}$$

It is expected that the precision of these measurements can be improved by using per-event acceptances in place of a global function.

Keywords: LHCb, Alignment, Flavour Tagging, B_s^0 meson lifetime

Abstract

L'expérience LHCb est l'une des trois grandes expériences du Grand Collisionneur de Hadrons (LHC), situé au CERN, à la frontière entre la Suisse et la France près de Genève. Le détecteur LHCb est un spectromètre à un seul bras dont le but est la mesure des désintégrations rares des hadrons b ainsi que la mesure précise de la violation CP. L'expérience LHCb a enregistré un total de 1fb^{-1} de données à une énergie au centre de masse de 7 TeV.

Cette Thèse aborde trois sujets liés à l'expérience LHCb. La première partie de ce travail porte sur l'alignement du trajectographe interne de l'expérience LHCb avec les premières données du LHC. Un désalignement du système des trajectographes dégrade la mesure de la quantité de mouvement et de la distance de vol. Ces grandeurs sont essentielles pour réaliser des études précises de masses ou de durée de vie. Un alignement du trajectographe interne à l'aide d'une méthode n'utilisant pas d'éléments fixes comme contraintes a été réalisé. Cet alignement atteint une résolution de $102 \pm 10 \mu\text{m}$, avec un biais de $0 \pm 13 \mu\text{m}$.

Dans la seconde partie, une première optimisation des coupures de sélection des kaons utilisés par le 'Same Side Kaon Tagging' a été réalisée à l'aide de données 2010 pour la désintégration $D_s^+ \rightarrow \phi\pi^+$ ainsi que de données Monte Carlo B_s^0 . Il est démontré que le canal de désintégration $D_s^+ \rightarrow \phi\pi^+$ permet d'obtenir une amélioration à la fois dans l'efficacité de l'étiquetage et dans la réduction du taux d'identifications erronées.

Dans la troisième partie, une méthode qui utilise une fonction d'acceptance globale et un traitement d'erreurs par événements pour mesurer la durée de vie des mésons B_s^0 et B^0 est développée. Il est montré que la résolution de ces mesures est dominée par l'erreur systématique causée par la limite de précision avec laquelle la fonction d'acceptance peut

être modélisée. Le résultat de la mesure de temps de vie moyen du méson B_s^0 est

$$\tau_{B_s^0} = 1.515 \pm 0.015(\text{stat})_{-0.039}^{+0.041}(\text{sys}) \text{ ps}$$

Ce résultat est d'une précision comparable à la moyenne mondiale mise à disposition par le Particle Data Group [1] pour $B_s^0 \rightarrow D_s X$

$$\tau_{B_s^0} = 1.425 \pm 0.041 \text{ ps}$$

Pour le méson B^0 , le temps de vie Le résultat de la mesure de temps de vie moyen est

$$\tau_{B_d^0} = 1.513 \pm 0.020(\text{stat})_{-0.037}^{+0.033}(\text{sys}) \text{ ps}$$

La précision de ces mesures peut être améliorée en utilisant des acceptances par événement à la place d'une fonction globale.

Mots-clé: LHCb, Alignement, Etiquetage de saveur, Temps de vie du meson B_s^0

Declaration

This dissertation is the result of my own work, except where explicit reference is made to the work of others, and has not been submitted for another qualification to this or any other university.

Vincent Fave

Acknowledgements

Cette thèse n'aurait pas pu être accomplie sans l'aide et le soutien des nombreuses personnes avec qui j'ai eu la chance de travailler durant ces quatre années. Je tiens à commencer par remercier tous les membres du LPHE, dont nos deux secrétaires Erika et Esther, pour la bonne ambiance de travail du 6^{ème}. Je remercie également tous les collègues de la Collaboration LHCb avec lesquels j'ai eu le plaisir de travailler.

I want to give my thanks and gratitude to my thesis director, Dr. Minh-Tâm Tran for his help and support. Also, I really appreciate the freedom I had during my thesis under his guidance. I thank Frédéric Blanc for his great availability and his contagious good mood. I also want to tell my great appreciation to Matthew Needham, with his ability to seemingly being able to answer almost any software question I could think of, and also for coaching me during the TED alignment.

I thank Louis Nicolas for helping me sort out the messy alignment software upgrades, devious recursive option files and for being a great teacher during my LHCb software beginnings. I also have fond memories of the discussions I had with Adlène Hicheur and his patience in teaching me the alignment ropes.

A big thank to my very good friend Johan Solo for the numerous lunches, coffees, C++ tricks, perl tricks, inside jokes and overall fun discussion we had during our thesis years.

A very special thanks to Paul Szczypka for all the help and support during my B_s^0 lifetime analysis, for the multiple corrections and advices and his support through late evenings and weekends at the end of this thesis.

I also want to thank my family for their support during my many years as a student.

Finally I thank my wife Celia with all my heart for her love, patience and support during these four thesis years.

Contents

1	Introduction	1
2	LHCb Experiment	5
2.1	The Large Hadron Collider	5
2.1.1	LHC Experiments and Accelerator Chain	6
2.2	The LHCb Experiment	8
2.2.1	The Subdetectors	9
2.3	Tracking System	10
2.3.1	Vertex Locator	11
2.3.2	Dipole Magnet	13
2.3.3	Tracker Turicensis	13
2.3.4	Inner Tracker	13
2.3.5	Outer Tracker	16
2.4	Particle Identification System	16
2.4.1	RICH detectors	17
2.4.2	Calorimeters	21
2.4.3	Muon System	24
2.4.4	Combined PID Likelihood	25
2.5	Trigger System	25

2.5.1	Level-0 Trigger	25
2.5.2	High Level Trigger	27
2.6	LHCb Software	29
3	Theoretical Introduction	31
3.1	Standard Model of Particles	31
3.1.1	Fundamental Interactions	31
3.1.2	Discrete Symmetries	32
3.2	Gauge Sector	33
3.2.1	Abelian example with $U(1)$	33
3.2.2	Non-Abelian Gauge Theory	36
3.2.3	Spontaneous Symmetry Breaking and Higgs Mechanism	39
3.2.4	Electroweak Theory	42
3.3	Leptons Sector	45
3.3.1	Fermion Masses	47
3.4	Quark Sector	48
3.4.1	CKM Matrix	50
3.5	CKM Matrix and CP violation	52
3.6	CP Violation in Mesons	54
3.6.1	CP Violation in Decay	58
3.6.2	CP Violation the Mixing	59
3.6.3	CP Violation by Interference	59
4	Tracking and Alignment at LHCb	61
4.1	Track Reconstruction	61
4.1.1	Track Types	61

4.1.2	Pattern Recognition	62
4.1.3	Track Fitting	64
4.2	Alignment Strategy	64
4.2.1	VELO Internal Alignment	65
4.2.2	IT and OT Internal and Relative Alignment	65
4.2.3	VELO to T Station Relative Alignment	66
4.2.4	TT Alignment from VELO to T Station Alignment and Long Tracks	66
4.2.5	RICH1 and RICH2, Calorimeters and Muon Stations Alignment	67
4.3	Alignment Theoretical Framework	67
4.3.1	Regularisation	69
4.3.2	Alignment Analysis Method	70
5	IT Alignment Simulation Studies	71
5.1	Beam-Gas Simulation Studies	71
5.2	Event and Track Selection	71
5.2.1	Weak Modes Study Methodology	72
5.2.2	Weak Modes Analysis Results	74
5.3	Alignment Stabilization using Eigenvalues	76
6	Alignment of the Inner Tracker with Data	81
6.1	Alignment using the June 2009 TED Data	81
6.1.1	TED Track Reconstruction	82
6.1.2	Event and Track Selection	82
6.1.3	Alignment Sequence	83
6.2	Alignment using the October 2009 TED Data	95
6.2.1	Event and Track Selection	95

6.2.2	Alignment Sequence	96
6.3	Conclusion on Inner Tracker Alignment with TED Data	103
7	Same Side Kaon Flavor Tagging Calibration with $D_s^+ \rightarrow \phi \pi^+$	105
7.1	Tagging Study Introduction	105
7.2	Selection of Same Side Tagging Kaons	106
7.3	Same Side Kaon Control Channels	109
7.4	$D_s^+ \rightarrow \phi \pi^+$ as Control Channel for the SSK Tagger	109
7.4.1	$D_s^+ \rightarrow \phi \pi^+$ Selection	110
7.4.2	Comparison of $D_s^+ \rightarrow \phi \pi^+$ vs $B_s^0 \rightarrow D_s^- \pi^+$ using MC10 Data	112
7.4.3	Comparison of $D_s^+ \rightarrow \phi \pi^+$ MC10 vs $D_s^+ \rightarrow \phi \pi^+$ Data	120
7.5	Optimization of the SSK Selection Cuts	122
7.6	Performance on $B_s^0 \rightarrow D_s^- \pi^+$ Data	125
7.7	Calibration of the Mistag Estimate	126
7.7.1	Using $D_s^+ \rightarrow \phi \pi^+$ with NNet for SSK Tagger Calibration	129
7.8	Conclusions	131
8	B_s^0 Lifetime Measurement from $B_s^0 \rightarrow D_s^- \pi^+$	135
8.1	$B_s^0 \rightarrow D_s^- \pi^+$ Event Reconstruction and Selection	136
8.1.1	Trigger Selection	136
8.1.2	Offline Pre-selection	137
8.1.3	$B_q \rightarrow D_q X$ Offline Selection	137
8.2	Studies on Monte Carlo Simulated Events	139
8.2.1	Fitter Description	139
8.2.2	Monte Carlo Event Samples	140
8.2.3	Mass Distributions	141

8.3	Decay Time Distributions and Acceptance	145
8.3.1	Decay Time Acceptance	146
8.3.2	Per-Event Decay Time Uncertainty	149
8.4	Heavy and Light States	152
8.5	Systematic Uncertainties and Fitter Bias	153
8.5.1	Systematic Uncertainties due to the Mass Models	153
8.5.2	Systematic Uncertainty from Acceptance Model	155
8.5.3	Systematic Uncertainty from $\Delta\Gamma_s/\Gamma_s$	155
8.5.4	Other Systematic Uncertainties	156
8.5.5	Summary of the Systematic Uncertainties	157
8.5.6	Fitter Bias	158
8.6	Results of the B_s^0 Lifetime Measurement in the $B_s^0 \rightarrow D_s^- \pi^+$ Channel . . .	158
9	B^0 Lifetime Measurement from $B^0 \rightarrow D^- \pi^+$	161
9.1	$B^0 \rightarrow D^- \pi^+$ Event Reconstruction and Selection	162
9.2	Studies on Monte Carlo Simulated Events	162
9.2.1	Monte Carlo Event Samples	162
9.2.2	Mass Distributions	163
9.3	Decay Time Distributions and Acceptance	166
9.3.1	Decay Time Acceptance	166
9.3.2	Per-Event Decay Time Uncertainty	167
9.4	Systematic Uncertainties and Fitter Bias	168
9.4.1	Systematic Uncertainties due to the Mass Models	168
9.4.2	Systematic Uncertainty from Acceptance Model	170
9.4.3	Other Systematic Uncertainties	170

9.4.4	Summary of the Systematic Uncertainties	171
9.5	Results of the B^0 Lifetime Measurement in the $B^0 \rightarrow D^- \pi^+$ Channel . . .	172
10	Conclusions	175
A	CV	177
	References	179

Chapter 1

Introduction

One of the most interesting feature of the weak interactions is that they allow the transition from one type of quark to an other, a feature called the non conservation of quark flavour. All of the flavour transitions are done via the exchange of charged W bosons, which restricts the transitions to quarks having a ± 1 electric charge difference. Flavour-changing through the neutral Z boson is not observed, which means that there is no neutral flavour-changing current. In the language of the Standard Model, this means that a u quark can be converted into a d, s or b quark. The same holds true when replacing the u quark by a c or t quark. One can translate this into simple the equations where the q' quarks are the weak interaction quark states and the q quarks are the flavour states.

$$\begin{aligned}u' &= V_{ud}d + V_{us}s + V_{ub}b \\c' &= V_{cd}d + V_{cs}s + V_{cb}b \\t' &= V_{td}d + V_{ts}s + V_{tb}b.\end{aligned}$$

The V_{ij} terms are a parametrization of the relative amplitudes for the transition from an i quark to a j quark. The 9 V_{ij} parameters form a 3×3 matrix, called the Cabibbo Kobayashi Maskawa (CKM) matrix. The V_{ij} matrix elements are free parameters of the standard model. The role of the CKM matrix is explicitly visible in the expression of

the charged weak current

$$\begin{aligned}
 J_\mu^{c.c.} &= \sum_{color} \begin{pmatrix} \bar{u} \\ \bar{c} \\ \bar{t} \end{pmatrix} \gamma_\mu \frac{(1 - \gamma_5)}{2} \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} \\
 &= \sum_{color} \begin{pmatrix} \bar{u} \\ \bar{c} \\ \bar{t} \end{pmatrix} \gamma_\mu \frac{(1 - \gamma_5)}{2} V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}.
 \end{aligned} \tag{1.1}$$

In Equation 1.1 the V_{CKM} matrix connects the weak interaction q' quarks to the flavour q quarks states. In the absence of neutral flavour-changing currents this matrix has to be unitary, introducing constraints on the V_{ij} parameters. In particular, sets of relations between the V_{ij} , $i \neq j$ parameters called the unitarity triangles. The LHCb experiment studies in particular one unitarity triangle,

$$V_{ub}V_{us}^* + V_{cb}V_{cs}^* + V_{tb}V_{ts}^* = 0. \tag{1.2}$$

In Equation 1.2 the sum of three complex numbers has to be zero. If one represent those numbers as lines in a plane, they have to close up to form a triangle, hence the unitarity triangle name. One can reparametrise this triangle such that one vertex of the triangle is at (0,0) and another one at (1,0). In that case the position of the last vertex contains the only degree of freedom. Measuring the position of this point is an important way of experimentally testing the consistency of the Standard Model.

The most interesting features about this triangle is that if it has a non-zero area (non-degenerate triangle) then the Standard Model has CP violation. Measuring the amount of CP violation and testing the unitarity of the CKM matrix is one important goal of the LHCb experiment. Most of these studies are done using B^0 and B_s^0 mesons decays, in particular, in this work I will present the average lifetime of the B_s^0 Heavy and Light states. This measurement is performed using the $B_s^0 \rightarrow D_s^- \pi^+$ decays, with $D_s^- \rightarrow K^- K^+ \pi^-$. An other part of this thesis related to b -physics is the flavour tagging presented in Chapter 7. The B^0 and B_s^0 have the particular feature of oscillating between their particle and anti-particle states, that is $B^0 \Leftrightarrow \bar{B}^0$ and $B_s^0 \Leftrightarrow \bar{B}_s^0$, respectively. Flavour tagging is a set of methods allowing to determine the state of the B^0 or B_s^0 at its production. These flavour tagging methods, also called taggers, are divided in two categories, the Same Side (SS) taggers and the Opposite Side (OS) taggers. The former

use particle produced along the B meson of interest to tag its initial state while the later make use of the non-signal B meson to tag the B signal state at production. In the SS case, the tagging particle is a pion for B^0 and a kaon for B_s^0 mesons, respectively. This work presents a novel calibration method for the Same Side kaon tagger, especially suitable for the low statistics available at the beginning of the LHCb experiment.

Chapter 4, 5 and 6 of this thesis present my work on the LHCb Inner Tracker alignment. The Inner Tracker (IT) is a particular part of the LHCb tracking system placed around the beam-pipe, after the LHCb dipole magnet. The IT covers only 2% of the tracking surface but contains as much as 20% of the tracks. It is a silicon micro-strip detector with an intrinsic hit resolution of $57\mu m$, and has a total of 12 tracking planes. The alignment of the LHCb tracking detectors is of utmost importance for any physics measurement as it underpins the precision on the particle momentum, and therefore of their mass and lifetime. This thesis concentrates on the alinement of the Inner Tracker using track from beam-dump events to provide a first alignment with real tracks. This is a very important step, as it allowed to run on first collision data with a reasonably well aligned detector as well as providing a solid basis for further improvements in alignment. The method presented in this thesis has the particularity of not fixing any tracking plane degree of freedom within the IT to provide constraints, relying instead on the removal of problematic eigenvalues and corresponding eigenvectors from the alignment solution.

Chapter 2

LHCb Experiment

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a 27 km proton-proton collider. It is situated in the 100 m underground tunnel previously used by the LEP collider at CERN near Geneva in Switzerland. The European Organization for Nuclear Research (CERN) is an international organization whose purpose is to operate the world's largest particle physics laboratory. The nominal energy of the proton-proton collisions is of $\sqrt{s} = 14$ TeV in the centre of mass energy but is now running at $\sqrt{s} = 7$ TeV. The prime goal of the LHC and its experiments is to test the Standard Model of particle physics and search for new physics.

Even with the rather small curvature of the 27 km accelerator, the very high momentum of the accelerated protons require high magnetic fields to bend their trajectories. The required 8.34 T magnetic field is generated by Nb-Ti superconducting magnets cooled to the temperature of 1.9 K.

The two proton beams circulate in opposite directions in separate beampipes. The two beams only cross at the four interaction points where the beampipes merge, allowing collisions to happen. The beams are structured in a way that allows a maximum of $N_{bx} = 2808$ proton bunches in the 3564 available bunch slots. The slots left empty correspond to the SPS and LHC kicker rise time and also the LHC dump kicker rise time. Each of these LHC bunches contains approximately 10^{11} protons. The bunch crossing frequency is 40 MHz and this frequency sets the requirement for all the detector front-end electronics processing speed. At the LHCb experiment, the average bunch-

bunch collision rate is close to 30 MHz.

The number of inelastic pp collisions at LHCb depends on three parameters : the total pp cross-section, the number of protons in the LHC rings, the crossing frequency, the beam profiles and overlap. This number can be computed over a period of time by the following relation :

$$N_{pp} = \sigma_{pp}^{\text{inel}} \int_{t_0}^t \mathcal{L} dt, \quad (2.1)$$

where $\mathcal{L}$ is the instantaneous luminosity and the $\sigma_{pp}^{\text{inel}}$ being the inelastic pp cross-section at the given collision energy. The instantaneous luminosity characterizes the collision rate. It depends on the transverse size of the beams at the collision point, the number of protons per bunch and crossing frequency. The LHC design luminosity is $10^{34} \text{ cm}^{-2}\text{s}^{-1}$. The LHCb experiment uses a lower luminosity to limit the number of collisions per crossing as the LHC design luminosity corresponds to an average of 20 pp collisions. This would lead to track multiplicities too high to be dealt with, as well as radiation damage to the LHCb detectors, in particular the vertex detector which is much closer to the beam than any other detector at the LHC. For these reasons, the nominal LHCb luminosity is limited. The actual running luminosity is close to the nominal luminosity with an average close to $2 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$. This luminosity corresponds to an average of 2.5 collisions per crossing, which is much higher than the nominal value of 0.5.

2.1.1 LHC Experiments and Accelerator Chain

The LHC makes use of a chain of smaller accelerators as shown in Figure 2.1. The total chain contains five accelerators, the LINAC2, the Proton Synchrotron Booster, the Proton Synchrotron, the Super Proton Synchrotron and the LHC itself. The chain is composed of older accelerators that were upgraded to match the LHC requirements.

The LHC has a total of four collision points, hosting a total of six experiments. The six experiments are :

1. **The TOTEM experiment** [2] is designed to measure the total pp cross section and luminosity simultaneously. This experiment covers the forward region in pseudo-rapidity $3.1 > |\eta| > 6.5$. The detector is composed of two parts placed several hundreds of meters either side of the CMS interaction point. A third part of the detector is integrated inside the CMS detector system.

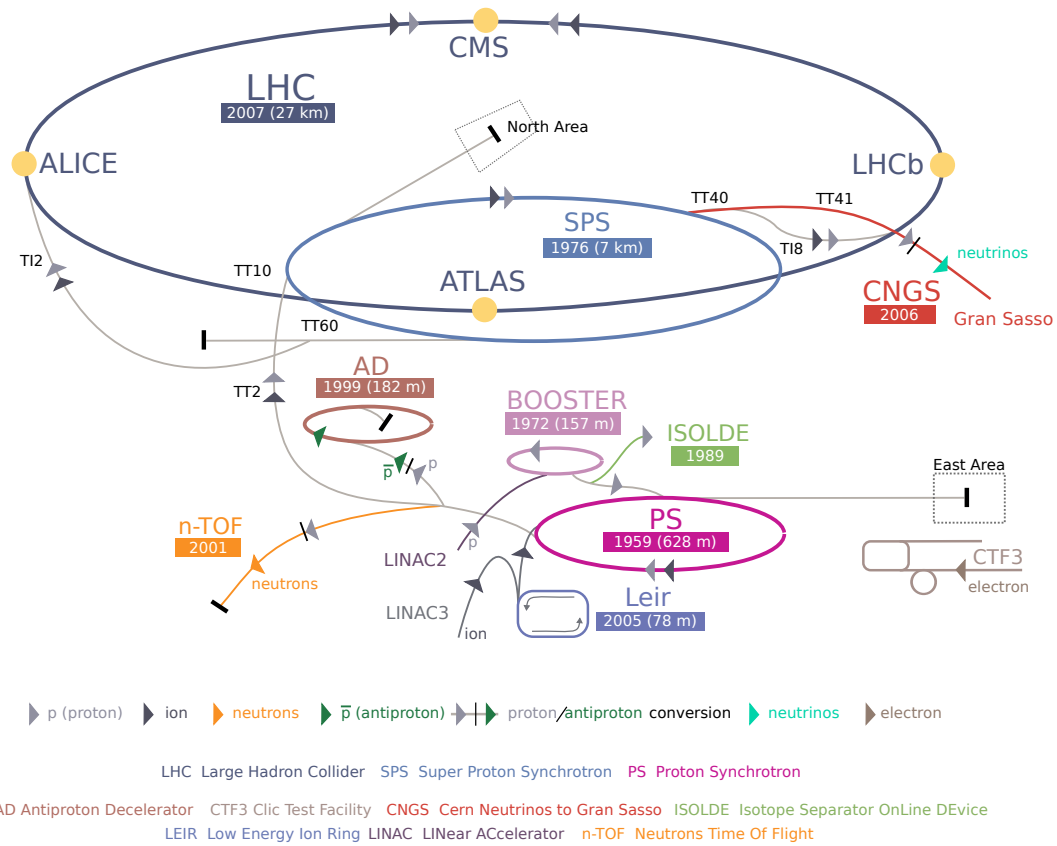


Figure 2.1: LHC Experiments and Accelerator Chain.

2. **The LHCf experiment** [3] is a forward experiment similar to TOTEM, but its goals are to confirm hadronic models at very high energy.
3. **ALICE** [4] studies quark-gluon plasma and strong interactions produced by very high energy $Pb - Pb$ ion collisions. This experiment is designed to deal with the extreme track multiplicities produced in heavy ion collisions.
4. **LHCb** [5] is dedicated to the study of b quark physics, in particular CP -violation processes and rare decays. This detector is a single-arm forward spectrometer. It will be detailed in Section 2.2.
5. **ATLAS** [6] and **CMS** [7] are the two LHC general-purpose experiments. They share the same goals of searching for the Higgs boson and physics beyond the standard model. While these two experiments are designed for pp collisions, they are also able to operate during ion collision runs.

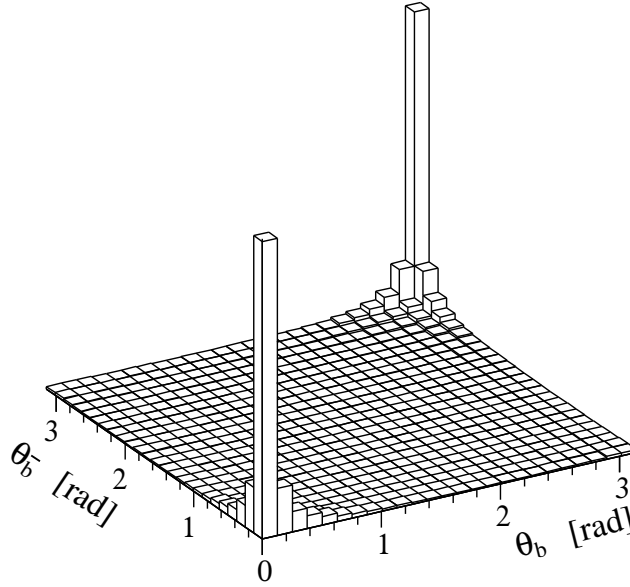


Figure 2.2: Polar angle of b - and $\bar{b}$ - hadrons obtained with the **Pythia**[9] event generator.

2.2 The LHCb Experiment

The Large Hadron Collider beauty (LHCb) experiment is dedicated to the precise study of rare decays and CP -violating decays [8], most of which is done with B mesons. This thesis will be focused on the study of the hadronic decay $B_s^0 \rightarrow D_s^- \pi^+$ and $B^0 \rightarrow D^- \pi^+$. In particular, a measurement of the B_s^0 and B^0 mesons lifetimes are done in these two hadronic decays in Chapters 8 and 9.

Most of the B hadrons are produced at small angles with respect to the beam axis with both b quarks being close to each other as shown in Figure 2.2. The LHCb design makes use of this feature by using a single-arm forward spectrometer with an opening angle of 10 mrad to 300 mrad in the horizontal plane and to 250 mrad in the vertical plane. Expressed in terms of pseudo-rapidity, this corresponds to an acceptance of $1.9 < \eta < 4.9$.

The different subdetectors of LHCb are described in the following pages. The detailed description of the LHCb subdetectors before their construction may be found in the Technical Design Reports (TDR). As these documents were written before the construction of the subdetectors, some of the information may be obsolete. An updated document has since been written [10]. The following sections are structured as follows :

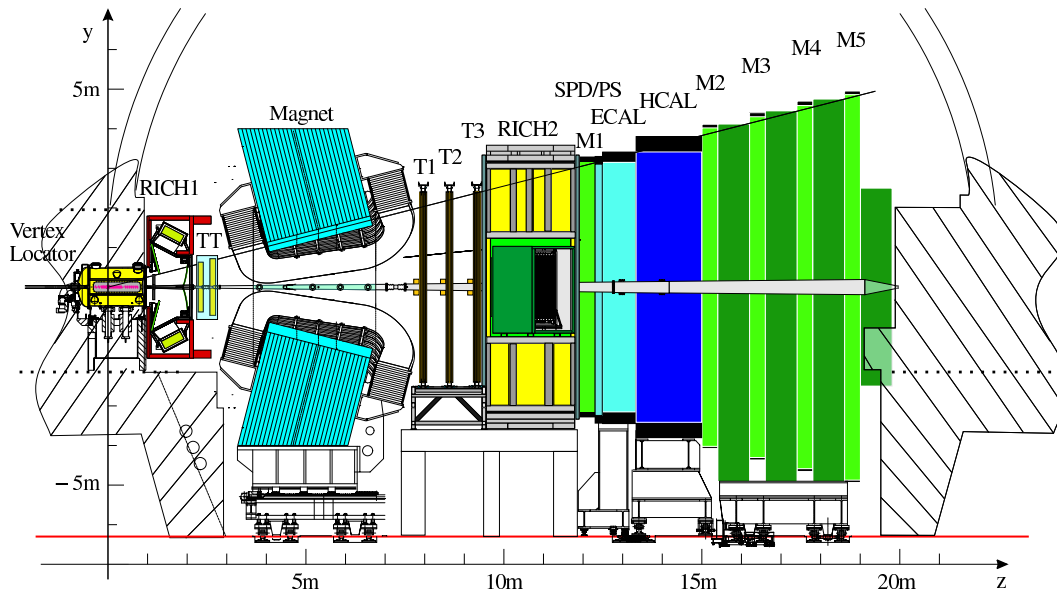


Figure 2.3: Lateral view of the LHCb single-arm spectrometer with its coordinate system.

first the subdetectors are briefly presented in Subsection 2.2.1, then more detailed descriptions are given for the different parts of the tracking system in Section 2.3 as these detectors are the most important for this thesis work, in particular for the alignment Chapters 4, 5 and 6. Particle identification system are described in Section 2.4.

2.2.1 The Subdetectors

The LHCb detector can be divided in five different subsystems with specific roles as visible in Figure 2.3 :

1. **The Vertex Locator (VELO)** is a silicon-strip vertex detector used for the reconstruction of vertices and track seeding. It also contains a pile-up veto counter used to reject pile-up events. All the Velo detector parts are used in the trigger.
2. **RICH Counters (RICH)** : are two RICH detectors, one aerogel/gas RICH counter placed before the LHCb magnet and one gaseous RICH counter placed after. They are used for particle identification, in particular for the separation of pions, kaons and protons.
3. **Tracking System and Magnet** are the two components of the LHCb spec-

trometer. This system provides trajectory information and momenta for charged particles.

4. **Electromagnetic and Hadronic Calorimeters (ECAL and HCAL)** : these detectors are used for the separation of electrons, photons, π^0 and hadrons as well as the measurement of the energy of these particles. The two calorimeters are also used in the trigger.
5. **The Muon Detector** is made of a succession of iron layers and tracking chambers. It is used in the trigger and offline reconstruction to identify muons, and is also used in the opposite side muon b-flavour tagging.

2.3 Tracking System

The LHCb tracking system reconstructs the trajectories of charged particles as they travel through the detector. The electric charge to momentum ratio q/p is retrieved from the trajectories' curvature induced by the LHCb dipole magnet. The tracking system is made of four tracking subdetectors, two before the LHCb dipole magnet and two after. The two subdetectors in the region before the magnet are the VELO and the TT (Tracker Turicensis). Both are silicon micro-strip detectors. The two subdetectors in the region after the LHCb dipole magnet are the IT (Inner Tracker) and OT (Outer Tracker). As their names indicate, the first one is positioned close to the beampipe where the track occupancy is high, while the second one covers the lower occupancy region further away from the beampipe. The Inner Tracker uses silicon strip technology, suitable for the high occupancy environment, while the Outer Tracker is a straw-tube gas detector. These two subdetectors are made of three tracking stations, each containing four detection layers, providing a total of 12 detection layers in the after-magnet region. As the momentum resolution at LHCb is dominated by multiple scattering for particles below 80 GeV¹, special attention has been dedicated to minimizing the tracking system material budget.

¹This thesis uses the unit convention where $\hbar = c = 1$

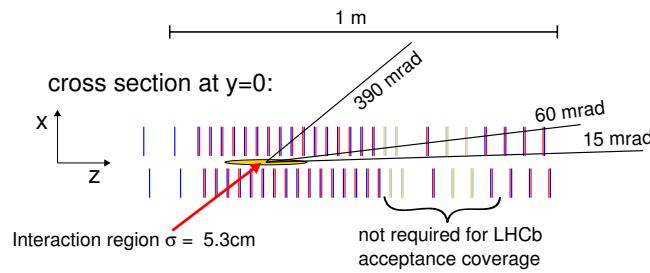


Figure 2.4: Layout of the 21 silicon sensors inside the VELO.

2.3.1 Vertex Locator

The roles of the VELO is to provide precise measurement of the charged tracks close to the interaction region and to find the primary vertices. These charged tracks are used to find and reconstruct primary and secondary vertices. The VELO covers both the forward and backward regions, and can therefore use backward track information for the primary vertices and track reconstruction. The forward region consist in the area covered by the main body of the LHCb detector, while the backward region is only covered by a part of the VELO. The forward tracks reconstructed by the VELO are used as seeds for the pattern recognition algorithms in track reconstruction. The VELO subdetector is used also with other subdetectors for triggering on possible signal events.

The VELO sensitive area is made of silicon-strip sensors placed very close to the beam axis and perpendicular to it. The whole detector is contained in a vacuum vessel which is separated from the LHC primary vacuum by a thin aluminium foil. The vacuum vessel uses a Roman pot design which allows the two halves of VELO sensors to move away from the beam during the LHC injection phase. This is done to avoid possible radiation damage to the sensors by the LHC beam. Once the beam is stable, the sensors are moved into their closed position for data taking.

The separation from the primary LHC vacuum is ensured by thin aluminium RF-foils that act both for primary to secondary vacuum separation and as radio-frequency protection for the VELO electronics and wake-field guide for LHC beams. The VELO sensor layout and vacuum vessel are shown in Figures 2.4 and 2.5, respectively.

The VELO has 21 silicon sensor planes spread along the beamline. The silicon sensors are $300\,\mu\text{m}$ thick half disks and provide a measurement of either the radial (r) or the angular (ϕ) coordinate. The r -sensors have 4×512 semi-circular strips centered around the beam axis to measure the r coordinate of the track to the centre, while the ϕ -sensors

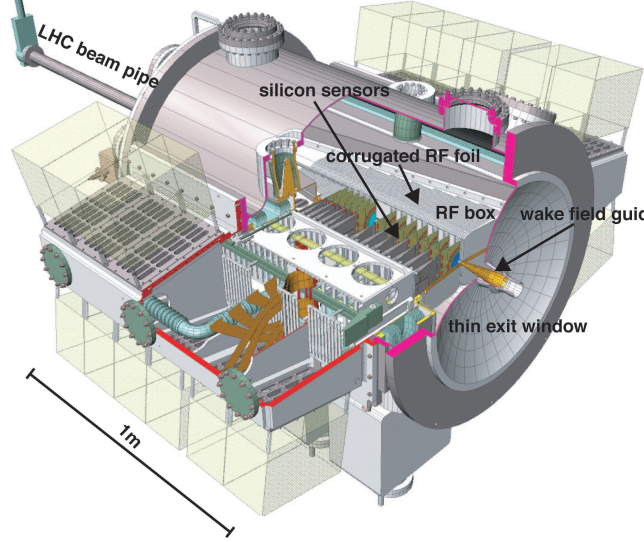


Figure 2.5: Layout of the VELO vacuum vessel with the LHC beam pipe and the VELO sensors.

have 683 (inner region) + 1365 (outer region) radial strips to measure the azimuthal coordinate ϕ .

The r and ϕ sensors are paired together using a 2 mm thick carbon fiber support structure to form a *module*. Each pair of modules separated by 1.5 cm in z form a *station*. The gap in z between the two halves of a station allows an overlap region between the sensors, a very valuable feature for VELO alignment. Each of the VELO stations can be seen as a 2D plane able to provide the r and ϕ coordinates for a track hit. This 2D point measurement is then combined with the z -axis information from the plane position to form 3D points. In addition to its 21 stations, the VELO has four r -sensors placed in the backward region. These form the pile-up VETO stations and are used by the first-level hardware trigger.

With its geometry and sensor parameters, the VELO achieves a hit resolution between 5 and 25 μm depending on the track angle and the strip pitch, which varies between 40 μm and 100 μm from the inner part of the sensors outwards. The primary vertex offline reconstruction resolution is of 40 μm in x and y .

2.3.2 Dipole Magnet

The LHCb dipole magnet is made of two trapezoidal coils, which have a 45° bending on the two transverse sides. The two coils are contained inside an iron yoke and made of aluminium wires. The role of the LHCb dipole magnet is to bend the trajectories of charged particles, permitting the measurement of the q/p ratio with the help of the tracking system. It is placed close to the interaction point but after the VELO, first ring imaging Cherenkov (RICH) subdetectors and Tracker Turicensis. The magnet has a maximal field strength of 1.1 T and an integrated field strength of 4 Tm. To correct for possible left-right asymmetries, the magnet polarity can and is reversed on regular basis, providing data samples of roughly the same size for physics analysis.

2.3.3 Tracker Turicensis

The TT is a silicon strip detector and covers an area of approximately $1.3 \times 1.5 \text{ m}^2$ with four tracking layers, leading to a total active region of 8.4 m^2 . It is built with two stations, TTa and TTb, each containing one x layer and one stereo layer. These two stations are separated by a distance of 27.0 cm along the z axis. The x layers have vertical strips (X layer) while the stereo layers are rotated around the z axis with an angle of -5° (U layer) for TTa and $+5^\circ$ (V Layer) for TTb. The layers are composed of ladders, containing 7 sensors above and below the beampipe and 14 sensors for the two sides of the beampipe. There are two 7-sensor ladders for the TTa and TTb stations layers, while there are seven 14-sensor ladders for TTa and eight 14-sensor for TTb. The detailed sensor geometrical layout of the TT stations are visible in Figure 2.6. In total, the TT subdetector has 896 strip sensors. These sensors have a pitch of $183 \mu\text{m}$ and 512 readout strips. Their size is of 9.44 cm in length, 9.64 cm in width for a thickness of $500 \mu\text{m}$.

The spatial resolution of this subdetector is of $50 \mu\text{m}$. For p_T measurements with VELO and TT only, the resolution on p_T for high momentum tracks is of order 10 – 40 %.

2.3.4 Inner Tracker

The IT is a subdetector covering the high occupancy region around the beampipe in the after-magnet region. It uses silicon strip technology while the OT uses straw tubes. The region covered by the IT is limited to 2 % of the total IT+OT area, but detects 20 % of

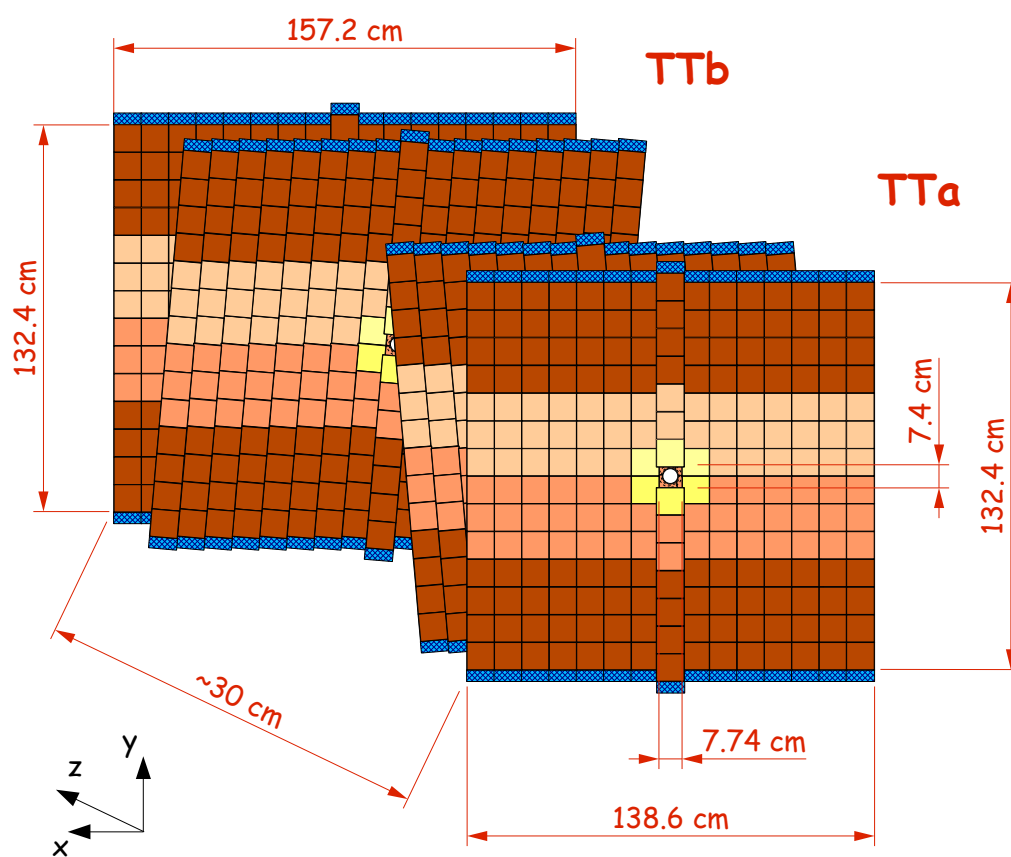


Figure 2.6: Layout of the four TT sensor layers.

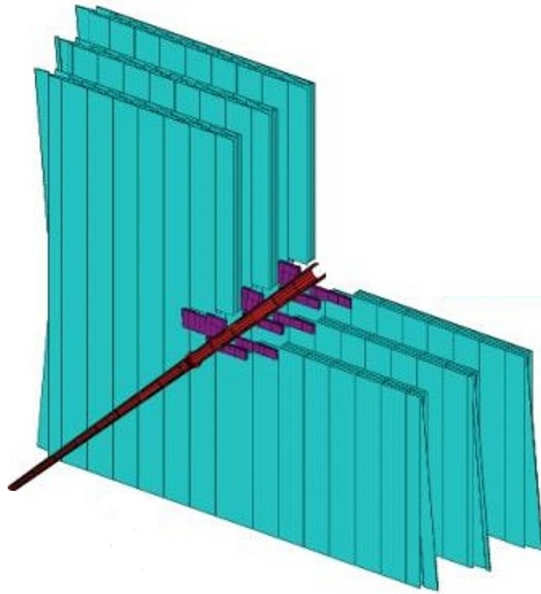


Figure 2.7: Layout of the IT boxes (in purple) and the OT (in blue) in the three tracking stations.

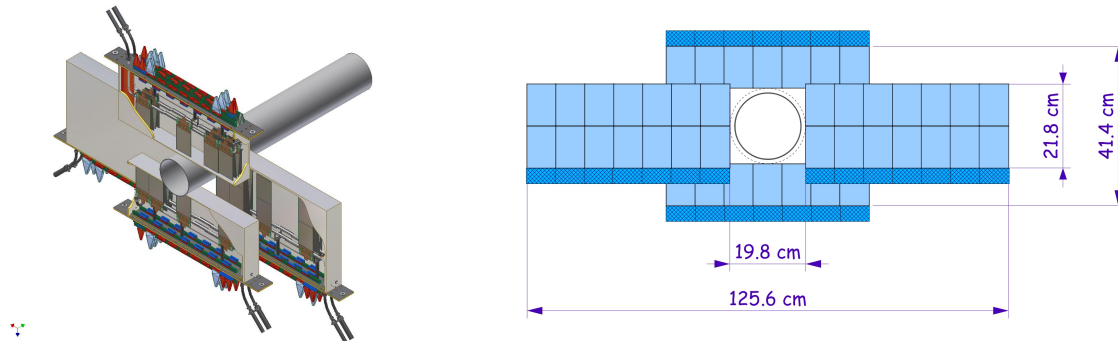


Figure 2.8: Inside IT boxes layout (left), layout of the IT boxes around the beam-pipe (right).

the tracks due to the high track occupancy around the beam-pipe.

The IT is part of the three T stations (tracking stations) and is composed of four boxes placed around the beampipe with a cross shape. Each of these boxes contain four detection layers, the two external ones being vertical layers while the two central ones are stereo layers. The boxes each have overlaps with two other boxes, allowing for the relative alignment of the boxes within a tracking station, as shown in Chapter 4. The detector layout is shown in Figures 2.7 and 2.8.

Each of the active layers in the IT boxes is made of seven sensor ladders. These ladders contain one silicon sensor for the top and bottom boxes while the ladders on each side of the beampipe contain two silicon sensors bonded together. The silicon sensors have a length of 11.0 cm, a width of 7.6 cm and a thickness of $410\ \mu\text{m}$ (A- and C-side) or $320\ \mu\text{m}$ (Top and Bottom). The length is limited by the available size of monocrystal silicon, while the thickness is chosen as a compromise between minimizing the material budget while keeping a signal-to-noise ratio of at least 12.

The IT sensors have 384 strips with a $200\ \mu\text{m}$ pitch. The detector has a single hit resolution of $57\ \mu\text{m}$. The spatial resolution is not the same for vertical and stereo sensors. The latter have a larger effective pitch due to the $\pm 5^\circ$ angle rotation, lowering their resolution compared to the vertical ones.

2.3.5 Outer Tracker

The OT is the second component of the T stations, as shown in Figure 2.7. It complements the IT for the tracking in the region situated after the LHCb magnet. The OT is similar in layout to the IT, with three stations, each of which containing four layers. As in the IT, the two external layers are vertical x layers while the two central ones are stereo layers tilted with an angle of $\pm 5^\circ$. It covers an area of $29\ \text{m}^2$ per station and has a spatial resolution of $\sim 200\ \mu\text{m}$.

The OT is exposed to much lower particle flux compared to the IT. For this reason, the design choices are driven by a need of keeping a low cost, to cover a large area while retaining a good spatial resolution. The technology chosen for the OT is gas filled straw tubes. The gas mix is 30 % CO_2 and 70 % Ar. It gives a fast drift time smaller than 50 ns, keeping it below two LHC bunch crossings ($2 \times 25\ \text{ns}$).

2.4 Particle Identification System

The physics goals of LHCb with b hadrons require excellent particle identification capabilities, in particular $K - \pi$ separation and muon identification. The muon information is crucial both for triggering and for event reconstruction. The particle identification quality is also a deciding factor for flavour tagging which allows the study of neutral b -meson oscillations. The requirements of particle identification are met by three different

subdetectors : The RICH counters, the calorimeters and the muon system.

2.4.1 RICH detectors

The LHCb experiment has two Ring Imaging Cherenkov (RICH) subdetectors, one placed between the VELO and the TT while the second is found after the T stations. The RICH detectors use the rings [11] of Cherenkov light produced when a charged particle passes through the radiator material with a velocity in the medium greater than the speed of light. Each event appears as a collection of rings in the RICH detectors. These rings are reconstructed by a pattern recognition algorithm and are then used to extract the Cherenkov emission angle θ_c . The velocity v of the particles can then be recovered using the relation

$$\cos \theta_c = \frac{c}{nv}, \quad (2.2)$$

where, n stands for the refractive index of the radiator medium. Combining the velocity information with the momentum information from the tracking system, the mass of the radiating particles can be found. Particles can then be identified by their mass.

Both RICH detectors operate following the same basis. Charged particles pass through the RICH radiator material and emit Cherenkov light rings. The light rings are then reflected and focused by spherical mirrors to photons detectors. With the help of these spherical mirrors, the photon detectors are placed outside the LHCb acceptance, lowering the material budget. Sets of secondary plane mirrors are used to make the detectors more compact.

The polar angle of the tracks are strongly correlated with their momentum, as seen in Figure 2.9. For this reason, RICH1 is used to identify low momentum tracks while RICH2 is used to identify high momentum tracks.

RICH1

The RICH1 detector is situated just after the VELO. It uses two consecutive radiator medium, the first one being silicon aerogel tiles while the second one is C_4F_{10} gas. The tiles are 5 cm thick and have a refractive index of $n = 1.03$. They are placed around 1 m from the interaction point. The gas ($n = 1.0014$) radiator is contained in a gas-tight vessel that also contains the mirror system. The RICH1 subdetector uses hybrid

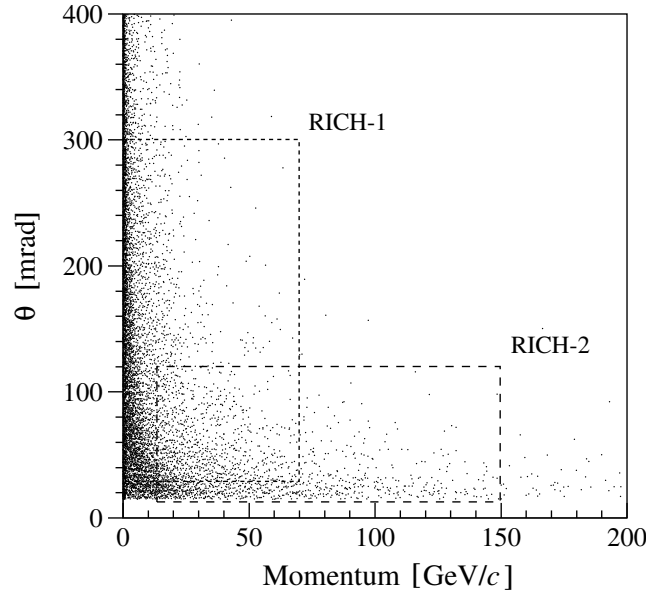


Figure 2.9: Correlation between polar angle of tracks and their momentum.

photodetectors HPDs that are sensitive to magnetic fields, and is therefore shielded from the LHCb magnet.

The Cherenkov light rings are focused to the HPDs by a set of four 2.7 m spherical mirrors. These mirrors are within the LHCb acceptance and are therefore coated carbon fiber mirrors to maintain a low material budget. A set of plane mirrors then reflect the light rings to the HPDs. The HPDs are separated from the C_4F_{10} gas vessel by quartz windows. A diagram of the RICH 1 subdetector is shown in Figure 2.10.

RICH2

This second Cherenkov subdetector is situated after the T stations and has much larger dimensions than the RICH1. It uses only one radiator medium, CF_4 gas. As in the RICH1, the gas is contained in a gas-tight vessel. It has a refractive index of $n = 1.0005$. The mirror system is nearly identical to the one used for RICH1. There is a first set of spherical mirrors to focus the light rings and a second set of plane mirrors to send the light rings to the HPDs. The spherical mirrors are composed of a total of 56 hexagonal mirrors forming two spherical arrays. The plane mirrors are composed of a total of forty square mirror tiles of side 437 mm. A diagram of the RICH2 subdetector layout is presented in Figure 2.11.

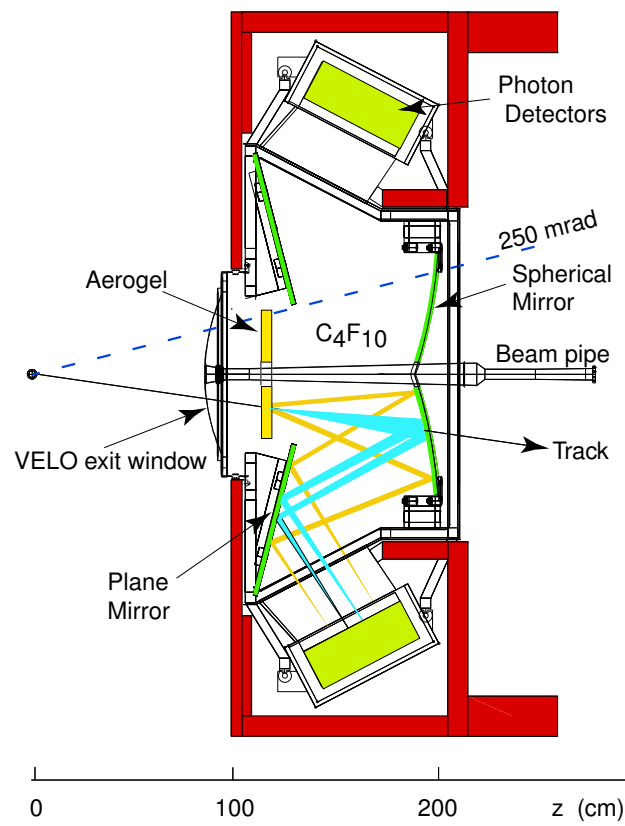


Figure 2.10: Layout of the RICH1 detector.

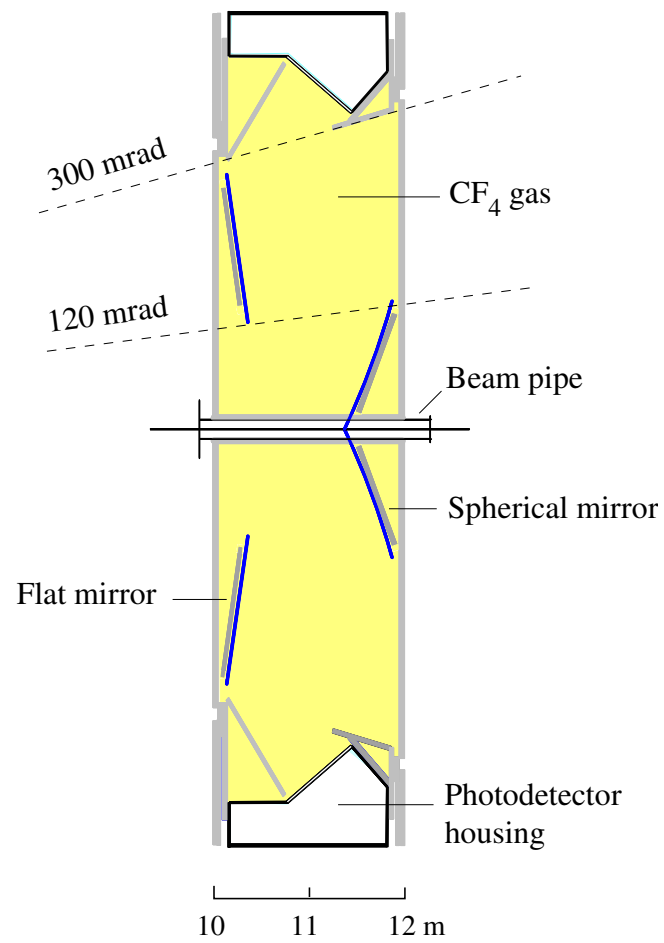


Figure 2.11: Layout of the RICH2 detector.

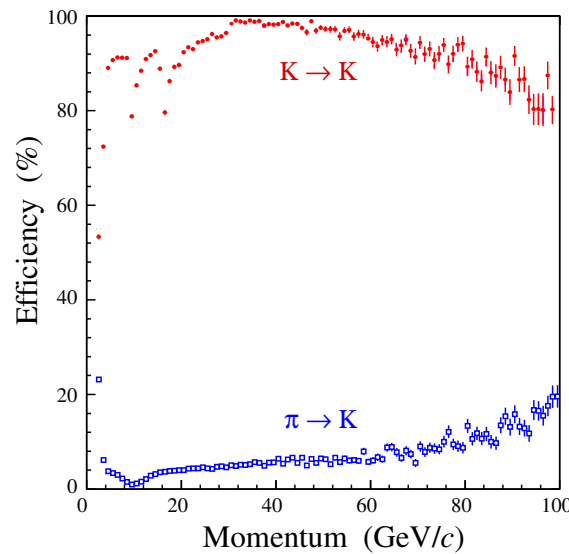


Figure 2.12: Kaon identification efficiency and pion misidentification rate versus the momentum of the particles.

Particle Identification with RICHes

The particle identification with the two RICHes uses the following method. For each charged reconstructed track, a mass hypothesis is made. Using the hypothesis, the probability distribution for photon detection in the RICH detectors pixels is computed. This distribution is then compared with the observed photon pattern detected and a likelihood is determined. The mass hypothesis for the track is then varied to maximize the likelihood value. Using this method, the kaon identification efficiency reaches $\sim 88\%$ for an average pion misidentification rate of $\sim 3\%$, as shown in Figure 2.12[11].

2.4.2 Calorimeters

The calorimeters are used to measure the energy of photons, π^0 , electrons and hadrons. These subdetectors also give position information on these particles. Finally, one of the important roles of the calorimeters is to differentiate electrons and photons from hadrons. The calorimetry system is able to deliver fast raw information, a feature used to trigger on possible signal events.

The calorimetry system is composed of three parts, the pad/preshower subdetector, the electromagnetic calorimeter and the hadronic calorimeter. The best criterion to



Figure 2.13: SPD scintillating pad (left), ECAL modules in their three different granularities configurations (right).

separate electromagnetic and hadronic showers is to look at the longitudinal profile of these showers. In terms of design, this is accomplished by having the Scintillating Pad Detector (SPD) located in front of the whole system. It is followed by a first layer of lead absorber and the Preshower (PS) component. The Electromagnetic Calorimeter (ECAL) then measures the energy of photons and electrons, followed by the Hadronic Calorimeter (HCAL) which measures the energy of hadrons.

All LHCb calorimeters use wavelength shifting optical fibers coiled in scintillating blocks to conduct collected light from the scintillators to the photomultipliers. The photomultiplier gains are tuned with the distance to the beampipe such that the E_T resolution stays constant over all of the LHCb acceptance. The segmentation of the calorimeters is optimized to average the occupancy. This leads to three regions with different granularities for SPD/PS and ECAL, and two for the HCAL.

The Pad/Preshower Calorimeter (SPD/PS)

These two components consist of a lead layer sandwiched between the cell layers of the SPD and the PS. The lead layer has a thickness of 15 mm which corresponds to $2.5 X_0$, where X_0 is the radiation length. The SPD and PS separate photons from electrons as the latter lose energy in a shower when crossing the lead layer. The SPD provides particle identification to the L0 trigger. In particular, photons with an energy in the 20 – 50 GeV range have a misidentification probability of around 1.0 %. The PS subdetector provides electron to pion discrimination to the L0 trigger. With a pion rejection rate of 92 % the electron detection efficiency is of 99.6 %

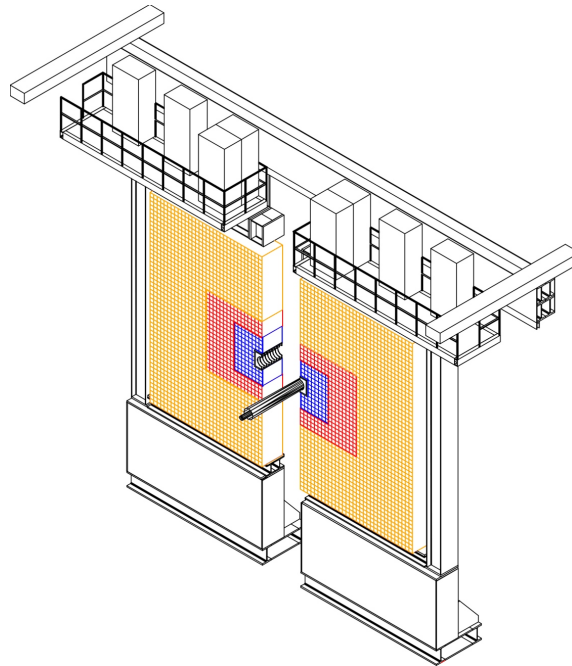


Figure 2.14: ECAL granularity layout, with 9 scintillating cells in blue, 4 scintillating cells in red, 1 scintillating cells in orange.

Electromagnetic Calorimeter (ECAL)

The electromagnetic calorimeter is designed to absorb the whole electromagnetic shower from high energy electrons. The showers are initiated by the lead layer sandwiched by the SPD and the PS. The ECAL has a total of $25 X_0$. The system is composed of modules made of lead layers of 2 mm thick altered with 4 mm thick scintillating plates. The modules come in three different kinds that are used to provide high granularity close to the beampipe and a low one in the outer regions by containing respectively 9 – 4 – 1 scintillating cells. The inner and middle section modules are read by 144 fibers while the outer modules are read by 64 fibers. The layout of the ECAL is shown in Figure 2.14. The energy resolution of the ECAL is

$$\sigma_E/E = 10\%/\sqrt{E} \oplus 1\%, \quad \text{with } E \text{ in GeV.} \quad (2.3)$$

In Equation 2.3, the first term reflects the statistical fluctuations from the shower while the second term is the systematic uncertainty from calibration.

Hadronic Calorimeter (HCAL)

The HCAL aim is mainly to be used in the trigger and to provide particle identification information. Its total thickness is 1.5 m, corresponding to $5.6 X_0$, which is not enough to fully contain the hadronic showers. For this reason, it can only provide an estimate of the hadron energy with a resolution limited to,

$$\sigma_E/E = 80\%/\sqrt{E} \oplus 10\%, \quad \text{with } E \text{ in GeV.} \quad (2.4)$$

As for the ECAL, the granularity is higher in the center than in the outer region. This is done by having squared cell size in two formats, 131.3 mm and 262.6 mm, respectively.

2.4.3 Muon System

The LHCb muon system is made of five muon stations. Four stations (M2-M5) are placed after the calorimeters, while the M1 station is placed before the SPD subdetector. The calorimeters are used as filter for the muons, as they stop every other type of charged particle. The M1 muon station uses a combination of a triple-GEM gas detector for the central region and multi-wire proportional chambers for the outer region, whilst M2-M5 use multi-wire proportional chambers. The muon stations are tiled with four different detector pad sizes. This is to ensure a relative p_T resolution constant over the full detector surface. The detector pads are interleaved with 80 cm thick iron absorbers. Combined with the calorimeter thickness, this only allow muons with a momentum greater than 6 GeV to fly through all five muon stations.

Muons can be used to tag the initial B -meson flavour in semileptonic decays and are present in many CP-sensitive decays like $B^0 \rightarrow J/\psi K_S^0$ and $B_s^0 \rightarrow J/\psi \phi$. The muon detectors are used in the L0 to trigger on high p_T muons. This high p_T trigger enriches the fraction of heavy flavours in the triggered events, use of a standalone muon reconstruction and p_T measurement. This standalone muon tracking requires hits in the five muon stations and is implemented in the L0 hardware-based trigger. The muon track reconstruction efficiency is 95 % and the p_T measurement resolution is 20 %.

Muons from semileptonic b -decays are used to tag neutral B mesons with their charge. This method is called opposite side muon tagging. In terms of the general LHCb tracking, the importance of muons requires the offline tagging to be highly efficient. The muon information is added using track segments from a T station that are propagated to the

muon subdetector. If hits are found where expected in the muons stations, the muon hypothesis for the track is confirmed.

2.4.4 Combined PID Likelihood

The information from all three particle identification are combined into a single log likelihood difference. The difference is calculated between a PID (Particle ID) hypothesis and a pion hypothesis for each charged track. It may be expressed as follows :

$$\Delta \ln \mathcal{L}_{x\pi} = \ln \mathcal{L}_x - \mathcal{L}_\pi = \ln \left[\frac{\mathcal{L}_x}{\mathcal{L}_\pi} \right]. \quad (2.5)$$

The $\mathcal{L}_x$ contains the combination of the RICH calorimeter and muon station information before PID hypothesis . The delta log likelihood (DDL) between two particle hypothesis x and y is simply obtained by taking the difference between the two DLL,

$$\Delta \ln \mathcal{L}_{xy} = \Delta \ln \mathcal{L}_{x\pi} - \Delta \ln \mathcal{L}_{y\pi}. \quad (2.6)$$

2.5 Trigger System

The purpose of the trigger system is to deal with the ~ 10 MHz visible interaction rate at LHCb. Such an interaction rate produces a raw data flow too large to be stored for offline analysis. Furthermore, the expected $b\bar{b}$ production rate at LHCb is much lower than the total rate, at 100 kHz. This $b\bar{b}$ rate is reduced when only considering events with at least one b -hadron having its decay products within the LHCb acceptance. The goal of the trigger system is to reduce the 10 MHz rate to a manageable level for offline storage while retaining as much as possible interesting heavy flavour decays. This task is performed using two levels of triggers [12] : the Level-0 Trigger (L0) and the High-Level Trigger (HLT), which are both presented in the following two subsections. A general layout of the trigger system is shown in Figure 2.15

2.5.1 Level-0 Trigger

The L0 trigger reduces the event rate from about 10 MHz to 1.0 MHz, which is the designed rate for the whole detector readout. In order to cope with the 40 MHz bunch

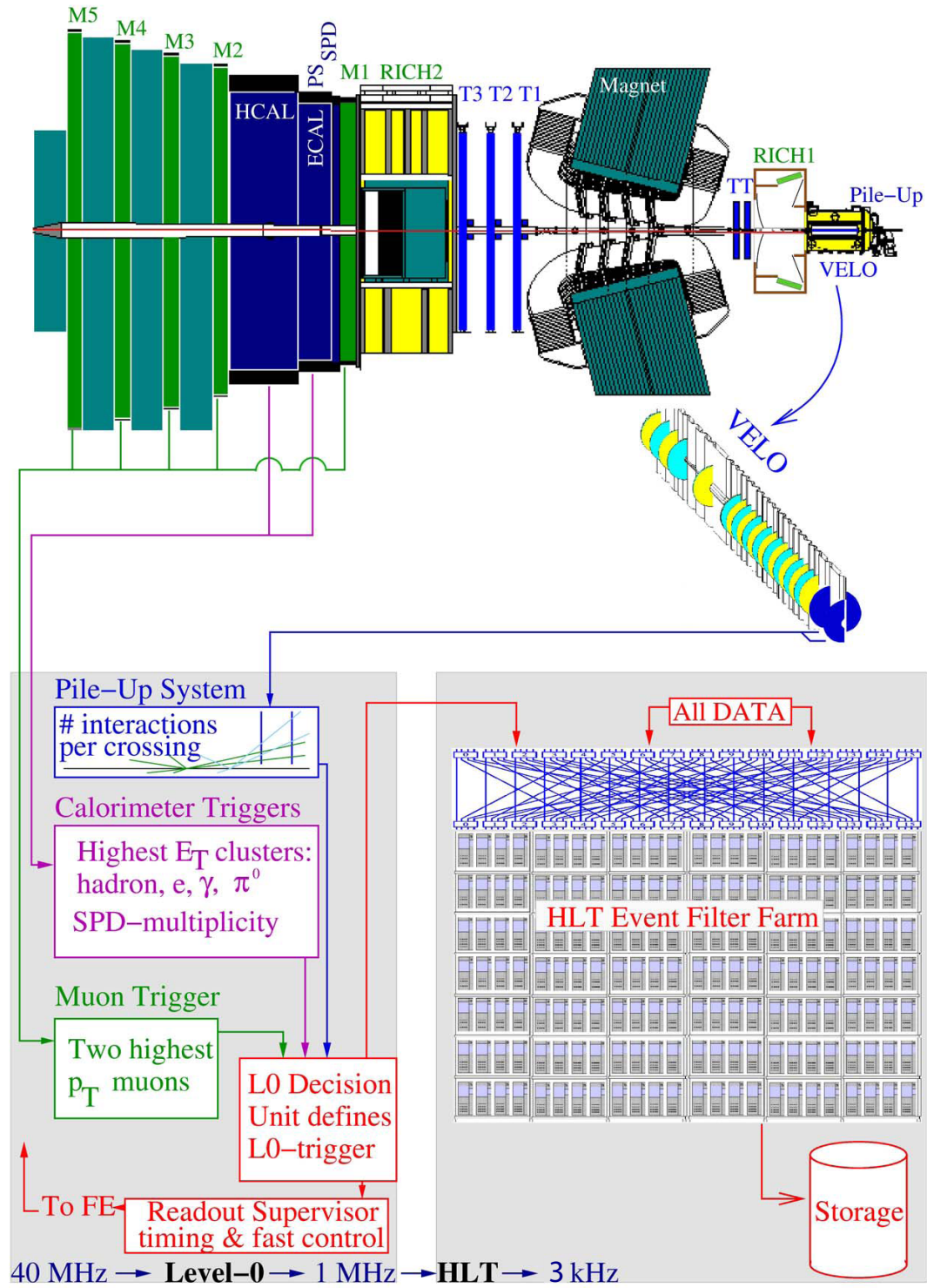


Figure 2.15: LHCb experiment trigger schematics.

crossing rate, the L0 trigger operates with partial information from the LHCb detector. This trigger level leverages the characteristics of b -hadron decays to retain them. These are consequences of the b -hadron large masses, such as decay products with large transverse momentum p_T and energy E_T . A potential problem is that multiple pp interaction events can lead to higher probabilities of trigger decisions based more on combinatorics than real b hadron decays. These types of events have the other disadvantage of using a large amount of bandwidth and computing resources during the reconstruction. To limit this, the number of pp interactions during one event is also used as an input information for the trigger. The L0 trigger makes its decision using the following :

1. **The L0 Calorimeter Trigger** searches for the highest E_T electrons, photons, neutral pions and hadrons in the event. A candidate is sent to the decision unit (L0DU) if an electron, photon ou neutral pion is reconstructed with $E_T > 2.5$ GeV or if a hadron is reconstructed with $E_T > 3.5$ GeV. The decision taken by the L0DU is vetoed if there are more than 280 clusters in the SPD or less than 5.0 GeV of total energy in the calorimeters.
2. **The L0 Muon Trigger** selects the two highest p_T muons in the event. A candidate is send to the L0DU only if one of the muon has a $p_T > 1.2$ GeV or if the sum of their p_T is greater than 1.5 GeV.
3. **The Pile-Up System** uses a system of four VELO r -sensors which are a part of the VELO subdetector. It provides track multiplicities and interaction number for the considered event. Its role in the L0 is to veto events with more than 112 particles.

2.5.2 High Level Trigger

The LHCb detector is read out at 1 MHz in the LHT. This rate is still too high for offline storage. A second level of trigger is therefore needed to reduce the rate to 2 approximately kHz. This trigger, unlike the L0, can use the whole detector information on the event, allowing for more developed selection criteria. The HLT is implemented in software and is run on a computer farm. Because it is software-based, this trigger level trigger is much more flexible than the L0 and can be modified to follow the physics priorities of the experiment.

The HLT is composed of two sublevels, the HLT1 and the HLT2. The first one

is designed to confirm the L0 decision and to refine it, making use of partial event reconstruction. The second uses fully reconstructed events that are then handed to the inclusive or exclusive selections.

The HLT1 Trigger

The HLT1 trigger is a group of parallel sequences of algorithms corresponding to each of the L0 trigger type. These algorithm sequences are called trigger 'alleys'. Some events are selected by several L0 triggers and as a consequence go through several HLT1 alleys. The algorithm sequences seek to confirm the L0 decision at each step by using incremental information from the detector. If the candidate provided by the L0 is not confirmed at one step of the sequence, it is discarded. On the contrary, when a candidate successfully goes through the complete HLT1 sequence, it is retained as a candidate for the HLT2 trigger. At this stage of the trigger, the event rate is of order 10 kHz.

The HLT2 Trigger

The HLT2 takes as input the candidates selected by the HLT1 trigger. As for HLT1, it is software-based but more specialized. The HLT2 trigger is driven by the needs of specific physics analyses. It is composed of two groups of selections, the inclusive group and the exclusive group. If an event is selected by any member of these two groups, a final positive trigger decision is reached and the event is saved on tape for offline analysis. Because the HLT1 trigger has reduced considerably the event rate, the HLT2 can perform a full event reconstruction. The HLT2 uses looser cuts than the offline selections, as it does not have access to full detector calibration and alignment information, and also to allow offline sensitivity studies on these cuts. The first step in the HLT2 trigger is the selection of tracks to form composite particles such as J/ψ , D_s , D^0 , ϕ and K^* . These intermediate particles are then used by the exclusive and inclusive triggers to reach their decisions.

1. **The inclusive trigger** selects events containing particles likely to be produced in b -hadron decays. These particles include resonances such as J/ψ , ϕ or D^* or muon particles.
2. **The exclusive trigger** is used to select event with specific decays. This trigger follows the needs of LHCb physics analyses, by having dedicated selections for

requested signal and control channels.

2.6 LHCb Software

The LHCb software is developed within the **Gaudi** framework. This framework is a well-structured C++ object-oriented architecture providing the general needs of the LHCb software components. The LHCb software uses specific components for different tasks. The most important components for Monte Carlo generation and physics analysis are presented here :

1. **Gauss** [13] is the Monte Carlo generation application. It performs the event generation in two steps. The first step is the pp collision generation using the **Pyhtia** program [9]. This generation step produces particles with their corresponding four-vectors. These particles are then decayed, using an adapted version for LHCb of the **EvtGen** [14] program in the case of b -hadrons. The second step is the simulation of the interaction between the detector and the generated particles. This step includes the interaction of the particles with the detector material, the description of the magnetic field and the loss of energy by radiation emissions. The program used to perform this second step is **Geant4** [15]. The detector description contains the geometry of the detector elements, their chemical composition and a map of the field produced by the LHCb magnet. An accurate description of the detector is a requirement for a correct simulation of particle energy loss by radiation and interaction with the material and for multiple scattering. The accuracy of the detector model is crucial for the sensitive regions, as the digitalization steps depends on it.
2. **Boole** is the digitalization application. It takes as input the Monte Carlo particle hits in the sensitive areas of the detector and simulates the response of the different subdetectors to these hits as well as the response of the readout chain electronics. Possible alteration of the signal by both the preceeding and following bunch crossings, called spillover, is introduced. Finally, the L0 trigger is simulated. The output of the **Boole** application is Monte Carlo data having the same format as the raw data from real pp collisions.
3. **Moore** [16] executes the high level trigger code. **Moore** either run in the online trigger farm processing online data from the DAQ system, or offline starting from

real data or from the output of the detector digitization application **Boole**.

4. **Brunel** [17] is the event reconstruction application. It uses hits provided by the **Boole** application or raw data to reconstruct tracks. Using pattern recognition, it then computes the momenta of these tracks and extracts the PIDs using information provided by the particle identification system presented in Section 2.4. The **Brunel** output is then stored in a ROOT [18] format used for physics analysis.
5. **DaVinci** [19] is the physics analysis application. It is used to select the reconstructed events of interest. This is done by building selections using particle information such as their PID, momentum, track parameters, and so on. Vertices are fitted with tracks of several particles, leading to additional available information to further refine selections. Vertex-related information includes the flight distance of particles, the quality of the vertex fit and so forth.

Chapter 3

Theoretical Introduction

3.1 Standard Model of Particles

The Standard Model (SM) of particles describes the elementary particles and their interactions. Elementary particles are divided in two sets, fermions and bosons. In the SM, fermions are matter particles and have spin $\frac{1}{2}$. The elementary fermions are the quarks and leptons. Interactions between them are mediated by spin 1 gauge bosons. The gauge bosons are the photon, the $W^\pm$, the Z^0 and the 8 gluons. The SM predicts the existence of an additional boson called the 'Higgs' boson, which has not yet been observed.

The fermions come in three generations of increasing mass, with one pair of quarks and one pair of leptons in each of them. The number of generations and the masses of fermions are not predicted by the SM. Experimental evidence and theoretical arguments only support the existence of three generations so far.

3.1.1 Fundamental Interactions

The Standard Model describes three of the four fundamental interactions. Each of these interactions are carried by boson particles, the photon for the electromagnetic interaction, the $W^\pm$ and Z^0 for the weak interaction and the gluons for the strong interaction. In this theoretical chapter a quick introduction to gauge theories will be made, followed by an application to the theory of electroweak interactions, first the

	Gen I	Gen II	Gen III	Gauge bosons
Quarks	$2.4 \pm 0.9 \text{ MeV}$	$1.27^{+0.07}_{-0.11} \text{ GeV}$	$171.2 \pm 2.3 \text{ GeV}$	
	$q = \frac{2}{3}$ $s = \frac{1}{2}$ u	$q = \frac{2}{3}$ $s = \frac{1}{2}$ c	$q = \frac{2}{3}$ $s = \frac{1}{2}$ t	$q = 0$ $s = 1$ γ
	up	charm	top	photon
	$4.8 \pm 1.2 \text{ MeV}$	$105.0^{+25}_{-35} \text{ MeV}$	$4.2^{+0.17}_{-0.07} \text{ GeV}$	
	$q = -\frac{1}{3}$ $s = \frac{1}{2}$ d	$q = -\frac{1}{3}$ $s = \frac{1}{2}$ s	$q = -\frac{1}{3}$ $s = \frac{1}{2}$ b	$q = 0$ $s = 1$ g
	down	strange	bottom	gluon
Leptons	0.511 MeV	105.7 MeV	$1776.82 \pm 0.16 \text{ MeV}$	$91.1876 \pm 0.0021 \text{ GeV}$
	$q = -1$ $s = \frac{1}{2}$ e	$q = -1$ $s = \frac{1}{2}$ μ	$q = -1$ $s = \frac{1}{2}$ τ	$q = 0$ $s = 1$ Z^0
	electron	muon	tau	weak boson
	$< 2.5 \text{ eV}$	$< 170 \text{ KeV}$	$< 18 \text{ MeV}$	$80.398 \pm 0.025 \text{ GeV}$
	$q = 0$ $s = \frac{1}{2}$ ν_e	$q = 0$ $s = \frac{1}{2}$ ν_μ	$q = 0$ $s = \frac{1}{2}$ ν_τ	$q = \pm 1$ $s = 1$ $W^\pm$
	electron neutrino	muon neutrino	tau neutrino	weak boson

Table 3.1: SM table of elementary particles.

gauge part, then the coupling to leptons and quarks. In this latter part, the origin of CP violation in the SM will be shown.

3.1.2 Discrete Symmetries

Discrete symmetries are introduced as operators transforming states in a way that leaves all physical observables invariant. In particular, there are three important discrete symmetries for quantum field theory, **C**, the charge conjugation operator, **P**, the parity operator, and **T** the time reversal operator.

3.2 Gauge Sector

The Standard Model is a gauge theory, that is its Lagrangian is invariant under a set of local gauge transformations. It allows to describe interactions through the exchange of massless bosons, as in the case of the electromagnetic interaction with the photon and the strong interaction with the gluons. The gauge invariance of the Lagrangian is always linked to a symmetry group, for example $U(1)$ for the electromagnetism, as presented in Section 3.2.1. In the case of larger symmetry groups than $U(1)$, like $SU(2)$ or $SU(3)$, it will be shown in Section 3.2.2 that the gauge bosons have self-interaction terms in the form of 3 and 4 boson vertices. A problem will arise with the fact that mass terms for the gauge bosons are not gauge invariant, even though massive gauge bosons are observed in the case of the weak interactions. This will be addressed by the Higgs mechanism in Section 3.2.3, where it will be shown that spontaneously breaking the gauge symmetry group allows for massive gauge bosons. Finally, the theory of electroweak interactions will be introduced in Section 3.2.4.

3.2.1 Abelian example with $U(1)$

Let us start with the Lagrangian for a free Dirac field,

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi. \quad (3.1)$$

This Lagrangian is invariant under the global transformation

$$\psi \rightarrow e^{iq\alpha}\psi, \quad \bar{\psi} \rightarrow e^{-iq\alpha}\bar{\psi}, \quad (3.2)$$

using Noether theorem, this implies the existence of a conserved current

$$j^\mu = q\bar{\psi}\gamma^\mu\psi. \quad (3.3)$$

The quantity q may be identified as the electric charge, j^μ is therefore the electromagnetic current for a fermion. This shows the relation between the global invariance under $U(1)$ transformations and the conservation of the electromagnetic current j^μ . If one promotes the global $U(1)$ gauge invariance to a local $U(1)$ gauge invariance,

$$\psi \rightarrow e^{iq\alpha(x)}\psi, \quad \bar{\psi} \rightarrow e^{-iq\alpha(x)}\bar{\psi}. \quad (3.4)$$

Writing invariant terms under the local gauge transformation is simple for terms containing no derivatives, such as the $\bar{\psi}m\psi$ part of the Lagrangian Equation 3.1. Problems arise when trying to build invariant terms containing derivatives. The partial derivative along a normalized directional vector n^μ is,

$$n^\mu \partial_\mu \psi = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} [\psi(x + \epsilon n) - \psi(x)] \quad (3.5)$$

It can be seen that since $\psi(x + \epsilon n)$ and $\psi(x)$ transform differently, this definition of the derivative is of little use given that the difference between the two terms will not be gauge invariant. To be able to take the derivative in a meaningful way, that is taking the difference of $\psi(x)$ at neighbouring points, an additional factor must be added to the derivative 3.5 to compensate the difference in phase transformation from one point to another. It may be done by introducing a scalar quantity $U(x_2, x_1)$ depending on the two points considered. It is defined as to follow the transformation

$$U(x_2, x_1) \rightarrow e^{i\alpha(x_2)} U(x_2, x_1) e^{-i\alpha(x_1)} \quad (3.6)$$

and also the one defined in Equation 3.4. When $x_2 = x_1$ the transformation is set as $U(x_2, x_1) = 1$. Moreover $U(x_2, x_1)$ can be written as a phase function $U(x_2, x_1) = \exp(i\phi(x_2, x_1))$ continuous with the positions x_2 and x_1 . For this reason $\psi(x_2)$ and $U(x_2, x_1)\psi(x_1)$ transform the same way, allowing to build a meaningful derivative

$$n^\mu D_\mu \psi = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} [\psi(x + \epsilon n) - U(x + \epsilon n, x)\psi(x)] \quad (3.7)$$

which is called the covariant derivative. To get an explicit definition of D_μ , an expression for $U(x + \epsilon n, x)$ is needed. This can be done by using the continuity of $U(x + \epsilon n, x)$ with the positions $x + \epsilon n$ and x , as $U(x + \epsilon n, x)$ may be expanded,

$$U(x + \epsilon n, x) = 1 - iq\epsilon n^\mu A_\mu(x) + \mathcal{O}(\epsilon^2). \quad (3.8)$$

Here a constant q is extracted and may be identified with the constant q introduced in Equation 3.3 when considering $U(1)$; and a vector field A_μ is introduced as the coefficient of the infinitesimal displacement ϵn^μ . The covariant derivative is explicitly written as,

$$D_\mu \psi(x) = \partial_\mu \psi(x) + iqA_\mu(x)\psi(x). \quad (3.9)$$

Inserting Equation 3.8 into 3.6 gives the way A_μ transforms under the local gauge transformation

$$A_\mu(x) \rightarrow A_\mu(x) - \frac{1}{q} \partial_\mu \alpha(x). \quad (3.10)$$

The covariant derivative may be used to rewrite an invariant Lagrangian under the local gauge transformation, leading to,

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi, \quad (3.11)$$

which may be expanded using $D_\mu = \partial_\mu + iqA_\mu$,

$$\bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi - qA_\mu \bar{\psi}\gamma^\mu \psi = \mathcal{L}_{\text{free}} - j^\mu A_\mu, \quad (3.12)$$

where the new Lagrangian is explicitly the sum of the free Lagrangian and a new interaction term between the Dirac field and the vector Gauge field A_μ . The interaction term also contains j^μ , the electromagnetic current of Equation 3.3.

To complete this invariant Lagrangian under local $U(1)$ transformations one has to add the kinematical term for the field A_μ . The kinematical term only depends on A_μ and its derivatives, in this case $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. This tensor is called the field-strength and can always be derived directly from the commutator of the covariant derivatives,

$$[D_\mu, D_\nu]\psi = iq(\partial_\mu A_\nu - \partial_\nu A_\mu)\psi = iqF_{\mu\nu}\psi. \quad (3.13)$$

The final Lagrangian for the interacting fermion ψ with a vector field A_μ may then be written as

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \quad (3.14)$$

The gauge invariance does not allow to give a mass term to the field A_μ , as it would have the form

$$\mathcal{L} = \frac{1}{2}m^2 A_\mu A^\mu, \quad (3.15)$$

which is not gauge invariant. This means that without an additional mechanism, gauge theories cannot have massive gauge bosons.

In this example it has been shown that by asking the invariance under local gauge

transformations an interaction term has to be introduced through the covariant derivative. It also fixed the form of the kinematical term for the vector field mediating the interaction.

3.2.2 Non-Abelian Gauge Theory

Promoting a global gauge invariance to a local gauge invariance can be done for larger groups than $U(1)$, with more complicated transformations than the phase rotation. The main difference with the Abelian case is the presence of several gauge bosons and the presence of interaction terms between them, in the form of 3- and 4-bosons vertices. The Standard Model symmetry group is $U(1) \times SU(2) \times SU(3)$, where the local gauge invariance under the $U(1) \times SU(2)$ part leads to the electroweak interactions and the local gauge invariance under $SU(3)$ leads to the strong interaction. These two parts commute and may be treated separately. The gauge invariant Lagrangian under $SU(2)$, called the Yang-Mills Lagrangian will be derived.

Let us consider a doublet of Dirac fields,

$$\psi = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \end{pmatrix}. \quad (3.16)$$

The free Lagrangian for this doublet of Dirac field may be written as

$$\mathcal{L}_{\text{free}} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi. \quad (3.17)$$

With the global gauge transformation

$$\psi(x) \rightarrow \exp\left(\frac{i}{2}\vec{\sigma} \cdot \vec{\alpha}\right)\psi(x). \quad (3.18)$$

In Equation 3.18 the $\vec{\sigma}$ are the Pauli matrices,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (3.19)$$

As in the $U(1)$ case, the global gauge symmetry is promoted to a local symmetry by asking the Lagrangian in 3.17 to be invariant under the gauge transformation 3.18 but with $\vec{\alpha}$ as an arbitrary function $\vec{\alpha}(x)$ of x . The procedure to construct a Lagrangian

invariant under this local symmetry is

$$\psi(x) \rightarrow V(x)\psi(x), \quad \text{where } V(x) = \exp\left(\frac{i}{2}\vec{\sigma} \cdot \vec{\alpha}\right). \quad (3.20)$$

This is similar to what has been done for the $U(1)$ case, but complications arise from the non-Abelian character of the case at hand, as there are now three orthogonal non-commuting local symmetries. The first step is to define the covariant derivative. To keep the calculations as generic as possible, everything will be done for the case of an arbitrary transformation $V(x)$, which can be replaced by the $SU(2)$ case as shown in Equation 3.20. The procedure is identical to the one used in the Abelian case, starting with the introduction of a scalar transformation

$$U(x_2, x_1) \rightarrow V(x_2)U(x_2, x_1)V^\dagger(x_1). \quad (3.21)$$

As before $U(x_2, x_2) = 1$ is set. For neighbouring points x_2 and x_1 , that is $U \simeq 1$, the expression for $U(x_2, x_1)$ can be written as an expansion in terms of the generators of $SU(2)$.

$$U(x + \epsilon n, x) = 1 + ig\epsilon n^\mu b_\mu^i \frac{\sigma^i}{2} + \mathcal{O}(\epsilon^2). \quad (3.22)$$

Where g is a constant, as q was before, and will later be related to the coupling constant of the interaction. Putting this expansion into Equation 3.7 leads to the covariant derivative expression for the $SU(2)$ symmetry case.

$$D_\mu = \partial_\mu - igb_\mu^i \frac{\sigma^i}{2} = \partial_\mu I + igB_\mu, \quad \text{where } B_\mu = b_\mu^i \frac{\sigma^i}{2} \text{ and } I \text{ is the identity.} \quad (3.23)$$

By definition, the covariant derivative transforms as

$$D_\mu \psi \rightarrow D'_\mu \psi' = V(D_\mu \psi). \quad (3.24)$$

It is possible to find the way the vector field B_μ transforms using 3.24,

$$\begin{aligned} D'_\mu \psi' &= (\partial_\mu + igB'_\mu)\psi' \\ &= V(\partial_\mu \psi) + (\partial_\mu V)\psi + igB'_\mu(V\psi) \\ &= V(\partial_\mu \psi + igB_\mu \psi). \end{aligned} \quad (3.25)$$

From Equation 3.25 it is possible to write

$$\begin{aligned} B'_\mu &= \frac{i}{g} \left[(\partial_\mu V) V^{-1} - ig V B_\mu V^{-1} \right] \\ &= V \left[B_\mu + \frac{i}{g} V^{-1} \partial_\mu V \right] V^{-1}. \end{aligned} \quad (3.26)$$

If the scalar transformation V is now replaced explicitly by its expression for the $SU(2)$ case, the explicit expression for B'_μ can be written

$$B'_\mu = b_\mu^i \frac{\sigma^i}{2} \rightarrow b_\mu^i \frac{\sigma^i}{2} + \frac{1}{g} (\partial_\mu \alpha^i) \frac{\sigma^i}{2} + i \left[\alpha^i \frac{\sigma^i}{2}, b_\mu^j \frac{\sigma^j}{2} \right]. \quad (3.27)$$

It may be observed that while the two first terms are very similar to the ones in Expression 3.10, the third one is new and comes directly from the non-commutativity of the local transformations. As in the $U(1)$ case, by promoting the global gauge invariance to a local one, we were forced to introduce a covariant derivative. When rewriting the free Dirac Lagrangian with these covariant derivatives,

$$\mathcal{L} = \bar{\psi} (i\gamma^\mu D_\mu - m) \psi = \mathcal{L}_{free} - g \bar{\psi} \gamma^\mu B_\mu \psi. \quad (3.28)$$

Finally, as for $U(1)$, a kinematical term is needed for the vector field B_μ . This term is derived from the commutator of the covariant derivative

$$[D_\mu, D_\nu] = -ig F_{\mu\nu}^i \frac{\sigma^i}{2}. \quad (3.29)$$

The expression of $F_{\mu\nu}^i$ can be expanded as

$$\begin{aligned} F_{\mu\nu}^i \frac{\sigma^i}{2} &= \partial_\mu b_\nu^i \frac{\sigma^i}{2} - \partial_\nu b_\mu^i \frac{\sigma^i}{2} - ig \left[b_\mu^i \frac{\sigma^i}{2}, b_\nu^j \frac{\sigma^j}{2} \right] \\ &= \partial_\mu b_\nu^i \frac{\sigma^i}{2} - \partial_\nu b_\mu^i \frac{\sigma^i}{2} + g \epsilon^{ijk} b_\mu^i b_\nu^j \frac{\sigma^k}{2}. \end{aligned} \quad (3.30)$$

The infinitesimal form of the transformation of the field strength tensor under $SU(2)$ is

$$F_{\mu\nu}^i \frac{\sigma^i}{2} \rightarrow F_{\mu\nu}^i \frac{\sigma^i}{2} + \left[i\alpha^i \frac{\sigma^i}{2}, F_{\mu\nu}^j \frac{\sigma^j}{2} \right], \quad (3.31)$$

which is not gauge invariant. However, it is possible to form an invariant combination of the three field strengths,

$$\mathcal{L} = -\frac{1}{2}\text{Tr}\left[\left(F_{\mu\nu}^i \frac{\sigma^i}{2}\right)^2\right] = -\frac{1}{4}\left(F_{\mu\nu}^i\right)^2. \quad (3.32)$$

In Equation 3.32, $\text{Tr}(\dots)$ stands for the trace. The term expressed in 3.32 is a kinetic term for the vector field B_μ and is properly gauge invariant. This term presents differences with the one obtained in the $U(1)$ case as, when expanded,

$$\begin{aligned} \text{Tr}(F_{\mu\nu}F^{\mu\nu}) &= \text{Tr}\left[(\partial_\mu B_\nu - \partial_\nu B_\mu + ig[B_\mu, B_\nu])(\partial^\mu B^\nu - \partial^\nu B^\mu + ig[B^\mu, B^\nu])\right] \\ &= \text{Tr}\left[(\partial_\mu B_\nu - \partial_\nu B_\mu)(\partial^\mu B^\nu - \partial^\nu B^\mu)\right] \\ &\quad + ig \text{Tr}\left[[B_\mu, B_\nu](\partial^\mu B^\nu - \partial^\nu B^\mu) + (\partial_\mu B_\nu - \partial_\nu B_\mu)[B^\mu, B^\nu]\right] \\ &\quad - g^2 \text{Tr}\left[[B_\mu, B_\nu][B^\mu, B^\nu]\right]. \end{aligned} \quad (3.33)$$

The two last lines of Equation 3.33 involve terms containing three instances of the vector field B_μ . In practical terms, this implies the presence of three- and four-boson vertices using the same coupling constant involved in fermion-bosons interactions. These self-interacting terms of the gauge bosons fields are a consequence of the non-Abelian nature of the $SU(2)$ local gauge symmetry we required.

In the Standard model, the $SU(2)$ symmetry leads to the weak interaction gauge bosons $W^\pm$ and Z^0 . However, there is still a problem as the weak interaction gauge bosons are massive, and mass terms for gauge bosons are not gauge invariant as seen in Equation 3.15. For these gauge bosons to become massive an other mechanism, called spontaneous symmetry breaking, is needed.

3.2.3 Spontaneous Symmetry Breaking and Higgs Mechanism

Higgs Mechanism : $U(1)$ example

In order to have massive gauge bosons as observed in weak interactions, it is necessary to introduce spontaneous symmetry breaking. We will start by giving a simple Abelian example with the $U(1)$ symmetry. Let us consider a complex scalar field ϕ coupled to

itself and to the electromagnetic field,

$$\mathcal{L} = |D_\mu \phi|^2 - \frac{1}{4}(F_{\mu\nu})^2 - V(\phi). \quad (3.34)$$

Here the covariant derivative D_μ is the one derived in 3.9, that is $D_\mu = \partial_\mu + iqA_\mu$. The Lagrangian 3.34 is invariant under local $U(1)$ gauge transformation,

$$\begin{aligned} \phi(x) &= e^{iq\alpha(x)}\phi(x), \\ A_\mu(x) &= A_\mu(x) - \partial_\mu\alpha(x). \end{aligned} \quad (3.35)$$

The potential term $V(\phi)$ is chosen as

$$V(\phi) = -\mu^2\phi^*\phi + \frac{\lambda}{2}(\phi^*\phi)^2, \quad \text{where } \lambda > 0. \quad (3.36)$$

When $\mu^2 > 0$, the complex scalar field ϕ gets a vacuum expectation value and the $U(1)$ symmetry will be spontaneously broken. The minimum of $V(\phi)$ is at

$$\langle\phi\rangle = \phi_0 = \sqrt{\frac{\mu^2}{\lambda}}. \quad (3.37)$$

The Lagrangian 3.34 is then expanded around the vacuum state 3.37. We start by decomposing the complex scalar field ϕ

$$\phi(x) = \phi_0 + \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2). \quad (3.38)$$

The potential term $V(\phi)$ introduced in 3.36 may then be rewritten as

$$V(\phi) = -\frac{1}{2\lambda}\mu^4 + \mu^2\phi_1^2 + \mathcal{O}(\phi_i^3). \quad (3.39)$$

Here the ϕ_1 boson has a mass $m_1 = \sqrt{2}\mu$ and the ϕ_2 is massless and is called a Goldstone boson. We now use this result to rewrite the kinematical term $|D_\mu\phi|^2$ using the expansion of the complex scalar field ϕ ,

$$\begin{aligned} |D_\mu\phi|^2 &= \frac{1}{2}(\partial_\mu\phi_1)^2 + \frac{1}{2}(\partial_\mu\phi_2)^2 \\ &+ q^2\phi_0^2 A_\mu A^\mu + \sqrt{2}q\phi_0 A_\mu \partial^\mu\phi_2 + \dots \end{aligned} \quad (3.40)$$

The term that interests us in Equation 3.40 is the $q^2\phi_0^2 A_\mu A^\mu$ term. This term is a mass term for the gauge boson, and is a consequence of the spontaneous symmetry breaking.

We might write it as a mass term for the photon,

$$\mathcal{L}_{\gamma Mass} = \frac{1}{2}m_\gamma^2 A_\mu A^\mu, \quad \text{where } m_\gamma = \sqrt{2}q\phi_0. \quad (3.41)$$

It is to be observed that the number of degrees of freedom before the symmetry breaking and after is the same. Before it we had a complex scalar field and a massless vector field for a total of 4 degrees of freedom. After the symmetry breaking, we have one real scalar field and one massive vector field for a sum of 4 degrees of freedom. The way the symmetry is broken depends on the representation used for the scalar field. In the example of the breaking of $U(1)$, the ϕ_2 component of the complex scalar field does not appear explicitly in the Lagrangian. This massless scalar boson is called a Goldstone boson and is 'absorbed' by the vector field when it acquires its mass as a longitudinal polarisation mode. A detailed explanation for this mechanism may be found in [20].

Higgs Mechanism : $SU(2)$ example

In the standard model, the weak interactions are described by $SU(2)$, with its tree generators leading to three massless bosons as seen in 3.30. It is possible to break the $SU(2)$ symmetry to have massive weak interaction bosons. The procedure is similar to the $U(1)$ example. Detailed calculations may be found in [20] and [21]. Breaking $SU(2)$ using a complex scalar doublet,

$$\phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}, \quad \text{where } \phi_1 = \frac{\phi_{1,1} + i\phi_{1,2}}{\sqrt{2}}, \quad \phi_2 = \frac{\phi_{2,1} + i\phi_{2,2}}{\sqrt{2}}. \quad (3.42)$$

This complex scalar doublet may be rewritten as

$$\phi = \frac{1}{\sqrt{2}} \exp\left(\frac{i\sigma^i \xi_i}{v}\right) \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \quad (3.43)$$

where the three ξ_i scalar fields will be the Goldstone bosons and h the Higgs boson. The Lagrangian is

$$\mathcal{L} = |D_\mu \phi|^2 - V(\phi) - \frac{1}{4}F_{\mu\nu}^i F^{\mu\nu i}, \quad (3.44)$$

where the expanded terms of the Lagrangian are

$$\begin{aligned} F^{\mu\nu i} &= \partial^\mu A^{\nu i} - \partial^\nu A^{\mu i} - g\epsilon^{ijk} A^{\mu j} A^{\nu k} \\ D_\mu \phi &= \left(\partial_\mu - ig \frac{\sigma^i}{2} A_\mu^i \right) \phi \\ V(\phi) &= -\mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \end{aligned} \quad (3.45)$$

After the symmetry breaking, the kinematical term for ϕ in the Lagrangian will become

$$|D_\mu \phi|^2 = \frac{1}{2} (\partial_\mu h) (\partial^\mu h) + \frac{g^2 v^2}{8} A_\mu^i A^{\mu i} + \frac{g^2}{8} h h A_\mu^i A^{\mu i} + \frac{g^2 v}{4} h A_\mu^i A^{\mu i}. \quad (3.46)$$

In this kinematic term 3.46 the $\frac{g^2 v^2}{8} A_\mu^i A^{\mu i}$ is the mass term for the gauge bosons and is the same for the three of them. The h field is the Higgs field and is coupled to the gauge bosons through 3 and 4 boson vertices. The mass term for the Higgs boson comes from the potential term in 3.45,

$$-V(\phi) = \frac{\mu^2}{2} (v + h)^2 + \frac{1}{4} (v + h)^4 = \mu^2 h^2 + \mathcal{O}(h^3). \quad (3.47)$$

This gives us the mass of the Higgs boson, $m_h = \sqrt{2}\mu$. As in the $U(1)$ example, the number of degrees of freedom is conserved. The three scalars ξ_i became massless and disappear from the Lagrangian, while the three massless gauge bosons became massive. In addition the kinematics for the field ϕ contains coupling terms between the Higgs and the gauge fields.

3.2.4 Electroweak Theory

We will now build the theory of electroweak interactions using what has been presented in 3.2.2 and 3.2.3 to derive it. First, we saw in 3.2.3 that breaking $SU(2)$ will provide three massive gauge bosons, but no massless one. As we are interested in electroweak interactions and as this includes the electromagnetism, there will be a massless gauge boson in the photon. For this reason the correct symmetry group for the electroweak interactions is $SU(2) \times U(1)_Y$. A scalar field ϕ is introduced to break the symmetry, and is set to receive a charge $Y = \frac{1}{2}$ called the hypercharge, under the $U(1)_Y$ symmetry. This gives the scalar field ϕ the following gauge transformation for the $SU(2) \times U(1)_Y$

symmetry,

$$\phi \rightarrow \exp\left(\frac{i}{2}\alpha^a\sigma^a + i\beta Y\right) \quad \text{where } Y = \frac{1}{2}. \quad (3.48)$$

If the vacuum expectation value of the field ϕ is chosen as

$$\langle\phi\rangle = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}, \quad (3.49)$$

then each generator of $SU(2) \times U(1)_Y$ which does not leave the vacuum state $\langle\phi\rangle$ invariant, a Goldstone boson will be present. These massless bosons will become the longitudinal modes of the massive gauge bosons through the Higgs mechanism. The vacuum state $\langle\phi\rangle$ will be invariant if

$$\exp(i\alpha G)\langle\phi\rangle = \langle\phi\rangle. \quad (3.50)$$

That is, for infinitesimal transformations,

$$(1 + i\alpha G)\langle\phi\rangle = \langle\phi\rangle \implies G\langle\phi\rangle = 0. \quad (3.51)$$

The generators σ^a and Y of $SU(2) \times U(1)$ acting on the vacuum state will not let it invariant,

$$\begin{aligned} \sigma^1\langle\phi\rangle &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{v}{\sqrt{2}} \\ 0 \end{pmatrix} \\ \sigma^2\langle\phi\rangle &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} -i\frac{v}{\sqrt{2}} \\ 0 \end{pmatrix} \\ \sigma^3\langle\phi\rangle &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} -\frac{v}{\sqrt{2}} \\ 0 \end{pmatrix} \\ Y\langle\phi\rangle &= \begin{pmatrix} Y\frac{v}{\sqrt{2}} \\ 0 \end{pmatrix}. \end{aligned} \quad (3.52)$$

However, using the Gell-Mann-Nishijima relation $Q = T_3 + Y$, where $T_3 = \frac{1}{2}\sigma^3$ and $Y = \frac{1}{2}$, we get a combination of generators corresponding to the electric charge which leaves the vacuum state $\langle\phi\rangle$ invariant. The theory will therefore contain a massless

vector boson and three massive vector bosons. The covariant derivative is

$$D_\mu \phi = \left(\partial_\mu - ig A_\mu^a \frac{\sigma^a}{2} - ig' Y B_\mu \right). \quad (3.53)$$

In Equation 3.53 A_μ^a and B_μ are $SU(2)$ and $U(1)_Y$ gauge bosons. Since $SU(2)$ and $U(1)_Y$ commutes, there are two coupling constant, g and g' . The gauge bosons mass terms are obtained by taking the relevant terms of $|D_\mu \langle \phi \rangle|^2$,

$$\begin{aligned} \mathcal{L}_{\text{gauge masses}} &= \frac{1}{2} \begin{pmatrix} 0 & v \end{pmatrix} \left(g A_\mu^a \frac{\sigma^a}{2} + ig' Y B_\mu \right) \left(g A^{a\mu} \frac{\sigma^a}{2} + ig' Y B^\mu \right) \begin{pmatrix} 0 \\ v \end{pmatrix} \\ &= \frac{v^2}{8} \left[g^2 (A_\mu^1 + i A_\mu^2)(A_\mu^1 - i A_\mu^2) + (-g A_\mu^3 + g' B_\mu)^2 \right]. \end{aligned} \quad (3.54)$$

The expression of Equation 3.54 corresponds to three massive vector bosons and a mass-less one, defined as linear combinations of the A_μ^a and B_μ fields,

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}} (A_\mu^1 \mp i A_\mu^2) \quad \text{with mass} \quad m_w = \frac{v}{2} g \\ Z_\mu^0 &= \frac{1}{\sqrt{g^2 + g'^2}} (g A_\mu^3 - g' B_\mu) \quad \text{with mass} \quad m_z = \frac{v}{2} \sqrt{g^2 + g'^2} \\ A_\mu &= \frac{1}{\sqrt{g^2 + g'^2}} (g' A_\mu^3 + g B_\mu) \quad \text{with mass} \quad m_A = 0. \end{aligned} \quad (3.55)$$

The Covariant derivative 3.53 may be rewritten in terms of the mass eigenstate fields, using the following definitions : $T^\pm = \frac{1}{2}(\sigma^1 \pm i\sigma^2)$ and $T^3 = \frac{1}{2}\sigma^3$

$$\begin{aligned} D_\mu &= \partial_\mu - i \frac{g}{\sqrt{2}} (W_\mu^+ T^+ + W_\mu^- T^-) \\ &\quad - i \frac{1}{\sqrt{g^2 + g'^2}} Z_\mu^0 (g^2 T^3 - g'^2 Y) \\ &\quad - i \frac{gg'}{\sqrt{g^2 + g'^2}} A_\mu (T^3 + Y). \end{aligned} \quad (3.56)$$

The electromagnetic coupling constant q may be identified as

$$q = \frac{gg'}{\sqrt{g^2 + g'^2}}. \quad (3.57)$$

A weak mixing angle θ_w can be defined as the rotation angle needed to go from the basis (A^3, B) to (Z^0, A) . This will allow us to simplify the expression of the covariant

derivative 3.56 in terms of the mass eigenstate fields,

$$\begin{pmatrix} Z^0 \\ A \end{pmatrix} = \begin{pmatrix} \cos \theta_w & -\sin \theta_w \\ \sin \theta_w & \cos \theta_w \end{pmatrix} \begin{pmatrix} A^3 \\ B \end{pmatrix}. \quad (3.58)$$

Using this weak mixing angle and the relations 3.59 linked to it,

$$\cos \theta_w = \frac{g}{\sqrt{g'^2 + g^2}} \quad \sin \theta_w = \frac{g'}{\sqrt{g'^2 + g^2}} \quad g' = g \tan \theta_w. \quad (3.59)$$

The covariant derivative 3.56 takes the form

$$\begin{aligned} D_\mu = & \partial_\mu - i \frac{g}{\sqrt{2}} (W_\mu^+ T^+ + W_\mu^- T^-) \\ & - i \frac{g}{\cos \theta_w} Z_\mu^0 (T^3 - \sin^2 \theta_w Q). \\ & - i q A_\mu (Q) \end{aligned} \quad (3.60)$$

In the end, the coupling of the electroweak gauge bosons is determined by only two parameters, q and θ_w , the electron charge and the weak mixing angle respectively. The masses of the weak bosons are not independent as $m_w = m_z \cos \theta_w$, which leaves only three free independent parameters for the electroweak gauge sector.

3.3 Leptons Sector

As the covariant derivative for $SU(2) \times U(1)_Y$ has already been defined in 3.60, the coupling between the four gauge bosons, $W^\pm$, Z^0 and γ , and the leptons will only depend on the representation used for these. The fact that the $W^\pm$ bosons only couple to left-handed leptons has to be accounted for because it will require the use of a different representation for left-handed fermions and right-handed ones. Hopefully, these two components of the fermionic field are decoupled,

$$\bar{\psi} i \gamma^\mu \partial_\mu \psi = \bar{\psi}_L i \gamma^\mu \partial_\mu \psi_L + \bar{\psi}_R i \gamma^\mu \partial_\mu \psi_R. \quad (3.61)$$

For this reason the covariant derivative for the two component of the field can be different. The Lagrangian for the weak interaction of leptons is

$$\mathcal{L}_{\text{leptons}} = \bar{L}_L i \gamma_\mu D^\mu L_L + \bar{l}_R i \gamma_\mu D^\mu l_R. \quad (3.62)$$

In the Lagrangian 3.62, L_L^i stands for the doublet of left handed leptons and l_R^i for the singlet right handed lepton as there is no right-handed neutrinos observed coupling through weak interactions. The doublet is

$$L_L = \begin{pmatrix} l^i \\ \nu_l^i \end{pmatrix}_L = \left\{ \begin{pmatrix} e \\ \nu_e \end{pmatrix}_L, \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}_L, \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}_L \right\}, \quad (3.63)$$

while the right-handed singlet is

$$l_R = \{e_R, \mu_R, \tau_R\}. \quad (3.64)$$

Leptons of these singlets and doublets have the following electric charge Q , hypercharge Y and value for T_3 , Once the covariant derivatives in the Lagrangian 3.62 are explicitly

	Q	T_3	Y
l_L	-1	$-\frac{1}{2}$	$-\frac{1}{2}$
$\nu_{l,L}$	0	$+\frac{1}{2}$	$-\frac{1}{2}$
l_R	-1	0	-1

written using the expression of 3.60, it becomes,

$$\mathcal{L}_{\text{leptons}} = \bar{L}_L i \gamma_\mu \partial^\mu L_L + \bar{l}_R i \gamma_\mu \partial^\mu l_R + g(W_\mu^+ J_l^{\mu,+} + W_\mu^- J_l^{\mu,-} + Z_\mu^0 J_l^{\mu,0}) + e A_\mu J_{l,\text{em}}^\mu \quad (3.65)$$

where $J_l^{\mu,\pm}$, $J_l^{\mu,0}$ and $J_{l,\text{em}}^\mu$ correspond respectively to the weak charged current, the weak neutral current and the electromagnetic current for the leptons. The charged weak currents have the expressions

$$J_l^{\mu,+} = \frac{1}{\sqrt{2}} (\bar{l}_L \gamma^\mu \nu_{l,L}) \quad (3.66)$$

$$J_l^{\mu,-} = \frac{1}{\sqrt{2}} (\bar{\nu}_{l,L} \gamma^\mu l_L) \quad (3.67)$$

while the weak neutral current is of the form

$$J_l^{\mu,0} = \frac{1}{\cos \theta_w} \left[\frac{1}{2} \bar{\nu}_{l,L} \gamma^\mu \nu_{l,L} + (\sin^2 \theta_w - \frac{1}{2}) \bar{l}_L \gamma^\mu l_L + (\sin^2 \theta_w) \bar{l}_R \gamma^\mu l_R \right]. \quad (3.68)$$

3.3.1 Fermion Masses

One problem at this point is that it is not possible to form gauge invariant mass terms for the leptons, and in fact, any fermionic mass term will lead to issues regarding gauge invariance. Such a mass term would have the form

$$\bar{\psi}m\psi \rightarrow m(\bar{L}_L + \bar{l}_R)(L_L + l_R) = m(\bar{L}_L l_R)(\bar{l}_R L_L). \quad (3.69)$$

It can be shown that the term $\bar{L}_L l_R$ from equation 3.69 is not invariant under $SU(2)$,

$$\begin{aligned} \bar{L}_L &\rightarrow \exp\left(-\frac{i\alpha^i \sigma^i}{2}\right) \bar{L}_L, & \text{and } l_R &\rightarrow l_R. \\ \text{Therefore } \bar{L}_L l_R &\rightarrow \exp\left(-\frac{i\alpha^i \sigma^i}{2}\right) \bar{L}_L l_R. \end{aligned} \quad (3.70)$$

As it has been shown in 3.70 mass terms are not invariant under $SU(2)$, but it also the case for $U(1)_Y$. This result is also true for quarks. There is however an other way to introduce mass terms for the fermions that is gauge invariant. This is done by the coupling between the scalar Higgs field and the fermions. This coupling term is a Yukawa type of coupling and has the form

$$\mathcal{L}_{\text{Yuk}} = \lambda_l(\bar{l}_R \phi^\dagger L_L + \bar{L}_L \phi l_R). \quad (3.71)$$

This interaction term is invariant under $SU(2)$ since both ϕ and L_L are $SU(2)$ doublets.

If the Higgs field is inserted as $\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h \end{pmatrix}$ in equation 3.71 it becomes

$$\mathcal{L}_{\text{Yuk}} = -\lambda_l \frac{h}{\sqrt{2}}(\bar{l}_R l_L + l_R \bar{l}_L) - \lambda_l \frac{v}{\sqrt{2}}(\bar{l}_R l_L + l_R \bar{l}_L). \quad (3.72)$$

Making the identification $\lambda_l \frac{v}{\sqrt{2}} = m_l$, Equation 3.72 leads to

$$\mathcal{L}_{\text{Yuk}} = m_l \frac{h}{v}(\bar{l}_R l_L + l_R \bar{l}_L) - m_l(\bar{l}_R l_L + l_R \bar{l}_L). \quad (3.73)$$

Equation 3.73 contains two parts, one Higgs-fermion interaction term and a fermionic mass term. This procedure allows to produce masses both for the leptons (exception made of the neutrinos) and the quarks.

3.4 Quark Sector

As for the lepton sector, the interaction between the weak gauge bosons and the quarks is entirely determined by the covariant derivative 3.60. There is, however, a very distinctive feature present in the quarks sector which doesn't exist in the lepton sector : weak interactions mix the different generations of quarks. Also, while the weak interactions carried by the $W^\pm$ bosons are proportional to the coupling constant g for the leptons, it is smaller in the case of quarks. For those, the following has been observed,

$$\begin{aligned} e &\rightarrow \nu_e \sim g \\ d &\rightarrow u \sim g \cos \theta_c \\ s &\rightarrow u \sim g \sin \theta_c, \end{aligned} \quad (3.74)$$

where $\sin \theta_c$ has been measured to be 0.22. This mixing angle depends on the quarks considered, and it will be shown that there are several of these mixing angles between the different generations. As the charged weak current only couples to left-handed fermions, the representations for the quarks will be similar to the ones used for the leptons. The representations are $SU(2)$ doublets for the left-handed quarks and $SU(2)$ singlets for right-handed quarks. The doublets are

$$Q_L \equiv Q_L^i = \begin{pmatrix} u^i \\ d^i \end{pmatrix}_L = \left\{ \begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L \right\}, \quad (3.75)$$

while for the right-handed singlets we have

$$u_R \equiv u_R^i = \{u_R, c_R, t_R\} \quad d_R \equiv d_R^i = \{d_R, s_R, b_R\}. \quad (3.76)$$

Quarks of these singlets and doublets have the following electric charge Q , hypercharge Y and value for T_3 . The Lagrangian invariant under $SU(2) \times U(1)$ for the quarks is similar to the one for the leptons in 3.62,

$$\mathcal{L}_{\text{quarks}} = \bar{Q}_L i \gamma^\mu D_\mu Q_L + \bar{u}_R i \gamma^\mu D_\mu u_R + \bar{d}_R i \gamma^\mu D_\mu d_R. \quad (3.77)$$

	Q	T_3	Y
u_L	$+\frac{2}{3}$	$+\frac{1}{2}$	$+\frac{1}{6}$
d_L	$-\frac{1}{3}$	$-\frac{1}{2}$	$+\frac{1}{6}$
u_R	$+\frac{2}{3}$	0	$+\frac{2}{3}$
d_R	$-\frac{1}{3}$	0	$-\frac{1}{3}$

The explicit form of the covariant derivatives in the Lagrangian 3.77 is the one of 3.60, leading to the expanded expression

$$\begin{aligned}\mathcal{L}_{\text{quarks}} = & \bar{Q}_L i \gamma^\mu \partial_\mu Q_L + \bar{u}_R i \gamma^\mu \partial_\mu u_R + \bar{d}_R i \gamma^\mu \partial_\mu d_R \\ & + g(W_\mu^+ J_l^{\mu,+} + W_\mu^- J_l^{\mu,-} + Z_\mu^0 J^{\mu,0}) + e A_\mu J_{l,\text{em}}^\mu.\end{aligned}\quad (3.78)$$

The procedure is identical to the case of leptons, leading to the expressions for the weak charged currents

$$J_q^{\mu,+} = \frac{1}{\sqrt{2}}(\bar{u}_L \gamma^\mu d_L) = \frac{1}{\sqrt{2}}(\bar{u} \gamma^\mu (1 - \gamma_5) d), \quad (3.79)$$

$$J_q^{\mu,-} = \frac{1}{\sqrt{2}}(\bar{d}_L \gamma^\mu u_L) = \frac{1}{\sqrt{2}}(\bar{d} \gamma^\mu (1 - \gamma_5) u), \quad (3.80)$$

and also the expression for the weak neutral current

$$\begin{aligned}J_q^{\mu,0} = & \frac{1}{\cos \theta_w} \left[\bar{u}_L \gamma^\mu \left(\frac{1}{2} - \frac{2}{3} \sin^2 \theta_w \right) u_L + \bar{u}_R \gamma^\mu \left(-\frac{2}{3} \sin^2 \theta_w \right) u_R \right. \\ & \left. + \bar{d}_L \gamma^\mu \left(-\frac{1}{2} + \frac{1}{3} \sin^2 \theta_w \right) d_L + \bar{d}_R \gamma^\mu \left(\frac{1}{3} \sin^2 \theta_w \right) d_R \right].\end{aligned}\quad (3.81)$$

The expression of the weak currents is similar to the leptonic case, except that the mixing angle experimentally observed is missing from their expressions. For the leptons, this was not a problem as the three generations as independent terms in the Lagrangian were corresponding to the experimental reality. For the quarks, the situation in 3.79, 3.80 and 3.81 where the three generations are independent does not correspond to the observations. The way the mixing is introduced is through the possible difference between quark mass eigenstates and quark weak eigenstates. If these are different, mixing between the different generations is possible by the weak interactions. The quark masses are introduced in the same way it has been done for the leptons in 3.73 with a Yukawa

interaction with the Higgs boson,

$$\mathcal{L}_{\text{Yuk}} = -\lambda_u(\epsilon^{ab}\bar{Q}_{L,a}\phi_b^\dagger u_R) - \lambda_d(\bar{Q}_L\phi d_R) + h.c., \quad (3.82)$$

where the ϕ field is replaced by its value after spontaneous symmetry breaking $\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h \end{pmatrix}$, leading to mass terms for the quarks,

$$m_u = \frac{1}{\sqrt{2}}\lambda_u v \quad \text{and} \quad m_d = \frac{1}{\sqrt{2}}\lambda_d v. \quad (3.83)$$

3.4.1 CKM Matrix

In order to obtain the quark mixing terms in the weak currents, the weak eigenstate must be different from the mass eigenstate. To come back to example 3.74, this means having

$$\begin{aligned} d_w &= d \cos \theta_c + s \sin \theta_c \\ s_w &= -d \sin \theta_c + s \cos \theta_c. \end{aligned} \quad (3.84)$$

This will lead to the desired form for the charged weak current with, for example,

$$J_u^{\mu,-} = \frac{1}{\sqrt{2}}(\bar{u}_L \gamma^\mu d_L \cos \theta_c + \bar{u}_L \gamma^\mu s_L \sin \theta_c). \quad (3.85)$$

One has to remember that the weak coupling of the three generations of quarks derives directly from the substitution of the derivative by the covariant derivative as done in 3.77. This gives automatically the same weak coupling for all generation of quarks. Therefore, mixing cannot come from gauge invariance principle. However, the Higgs coupling is not derived from gauge invariance principle, and is not subject to the same limitations. In particular, the most general gauge-invariant term for the quark-Higgs coupling is, using notation from 3.75 and 3.76,

$$\mathcal{L}_{\text{Yuk}} = -\Lambda_u^{ij}(\epsilon^{ab}\bar{Q}_{L,a}^i\phi_b^\dagger u_R^j) - \Lambda_d^{ij}(\bar{Q}_L^i\phi d_R^j) + h.c.. \quad (3.86)$$

Redoing the calculation of the quark masses using 3.86 lead to the masses

$$M_u = \frac{1}{\sqrt{2}}\Lambda_u v \quad \text{and} \quad M_d = \frac{1}{\sqrt{2}}\Lambda_d v. \quad (3.87)$$

In the mass terms of 3.87, M_u , M_d , Λ_u and Λ_d are 3×3 matrices. Even though M_u and M_d are not Hermitian, they can be diagonalized using bi-unitary transformation of the form $U^\dagger \Lambda W = \Lambda_{\text{diag.}} \equiv D$. U and W are the unitary matrices which diagonalize $\Lambda \Lambda^\dagger$ and $\Lambda^\dagger \Lambda$ respectively. In summary we have

$$\begin{aligned} \Lambda_u \Lambda_u^\dagger &= U_u D_u^2 U_u^\dagger, & \Lambda_u^\dagger \Lambda_u &= W_u D_u^2 W_u^\dagger & \Rightarrow & \Lambda_u = U_u D_u W_u^\dagger, \\ \Lambda_d \Lambda_d^\dagger &= U_d D_d^2 U_d^\dagger, & \Lambda_d^\dagger \Lambda_d &= W_d D_d^2 W_d^\dagger & \Rightarrow & \Lambda_d = U_d D_d W_d^\dagger. \end{aligned} \quad (3.88)$$

In order to remove the U and W matrices from the Higgs coupling term, the following change of variables is used,

$$u_R^i \rightarrow W_u^{ij} u_R^j, \quad d_R^i \rightarrow W_d^{ij} d_R^j, \quad (3.89)$$

$$u_L^i \rightarrow U_u^{ij} u_L^j, \quad d_L^i \rightarrow U_d^{ij} d_L^j. \quad (3.90)$$

Upon applying the change of variable 3.89, the W matrix is eliminated from the Higgs coupling 3.86. Similarly, applying the change of variable 3.90 the U matrix is removed from 3.86. This means that the usual mass terms from the Higgs-quarks Yukawa coupling are recovered,

$$m_u^i = \frac{1}{\sqrt{2}} D_u^{ii} v, \quad m_d^i = \frac{1}{\sqrt{2}} D_d^{ii} v. \quad (3.91)$$

The changes of variables 3.89 and 3.90 are therefore a simple base change, allowing to return the quark fields to the mass eigenstates. The effect of the change of variables 3.89 and 3.90 on the weak interaction Lagrangian for quarks 3.77 has to be investigated. For the right-handed quarks singlets the Lagrangian term are

$$\mathcal{L}_{\text{quarks},R} = \bar{u}_R^i i \gamma^\mu D_\mu u_R^i + \bar{d}_R^i i \gamma^\mu D_\mu d_R^i. \quad (3.92)$$

The W_u and W_d matrices are eliminated from the term 3.4.1 since they commute with the covariant derivative, as each of the u_R^i and each of the d_R^i have the same coupling to the gauge fields. This means that applying the change of variable 3.89 to will leave it unchanged. Here is an explicit example with u_R^i quarks,

$$\bar{u}_R^i i \gamma^\mu D_\mu u_R^i \rightarrow \bar{u}_R^i i W_u^{\dagger ij} \gamma^\mu W_u^{jk} D_\mu u_R^k = \bar{u}_R^i i \gamma^\mu D_\mu u_R^i. \quad (3.93)$$

There will therefore be no mixing of right-handed quarks by the neutral weak current and the electromagnetic current.

In the case of the change of variable 3.90 for left-handed quarks, the situation is more complicated. The U_u and U_d matrices will be eliminated in terms of the form $\bar{u}_L^i i\gamma^\mu D_\mu u_L^i$ and $\bar{d}_L^i i\gamma^\mu D_\mu d_L^i$. The reason is the same as the one exposed in 3.93. This means that since the neutral weak current and the electromagnetic currents for left-handed only involve terms of this form, there will be no quark mixing by this two currents. The case of the charged weak current is different, as there are two different types of quarks involved, u_L^i and d_L^i . Explicitly, applying the change of variable 3.90 on the charged weak currents gives

$$J_q^{\mu,+} = \frac{1}{\sqrt{2}}(\bar{u}_L^i \gamma^\mu d_L^i) \rightarrow \frac{1}{\sqrt{2}}(\bar{u}_L^i \gamma^\mu U_u^{\dagger ij} U_d^{jk} d_L^k), \quad (3.94)$$

$$J_q^{\mu,-} = \frac{1}{\sqrt{2}}(\bar{d}_L^i \gamma^\mu u_L^i) \rightarrow \frac{1}{\sqrt{2}}(\bar{d}_L^i \gamma^\mu U_d^{\dagger ij} U_u^{jk} u_L^k). \quad (3.95)$$

The matrixed products $U_u^\dagger U_d$ and $U_d^\dagger U_u$ are not canceling in Equations 3.94 and 3.95. This implies quark mixing by the charged weak current, which is what we were looking for. The individual matrices U_u and U_d are not observable experimentally, only their product is. For this reason, we will rename the product as a unitary matrix W ,

$$V = U_u^\dagger U_d \quad V^\dagger = (U_u^\dagger U_d)^\dagger = U_d^\dagger U_u \quad (3.96)$$

This W matrix is known as the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix.

3.5 CKM Matrix and CP violation

The CKM matrix derived in Equation 3.96 is a 3-by-3 unitary matrix. It therefore has 9 independent parameters, 3 rotation angles and 6 phases. All phases but one can be removed by doing quark phase rotations, leaving only one overall phase. The CKM matrix has then 4 independent parameters, one complex phase and 3 angles. Using the

Particle Data Group parametrization [1],

$$\begin{aligned}
 V &= \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{3.97}
 \end{aligned}$$

where the c_{ij} and s_{ij} correspond to $\cos \theta_{ij}$ and $\sin \theta_{ij}$ respectively. In the case of s_{12} its value is close to $\sin \theta_c \simeq 0.23$. The measured magnitudes for the CKM matrix are,

$$W \simeq \begin{pmatrix} 0.9742 & 0.226 & 0.0036 \\ 0.0226 & 0.973 & 0.042 \\ 0.0087 & 0.041 & 0.9991 \end{pmatrix}. \tag{3.98}$$

It is usual to write an approximation of the CKM matrix using the Wolfenstein [22] parametrization,

$$W = \begin{pmatrix} 1 - \frac{\lambda^2}{2} - \frac{\lambda^4}{8} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda + (1 - 2(\rho + i\eta))A^2\frac{\lambda^5}{2} & 1 - \frac{\lambda^2}{2} - (1 + 4A^2)\frac{\lambda^4}{8} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 + (1 - 2(\rho + i\eta))A^2\frac{\lambda^4}{2} & 1 - A^2\frac{\lambda^4}{2} \end{pmatrix} + \mathcal{O}(\lambda^6). \tag{3.99}$$

The parameters in 3.99 are $\lambda \simeq \sin \theta_c \simeq 0.23$, $A \simeq 0.81$ and ρ, η are of $\mathcal{O}(1)$. Terms of $\mathcal{O}(\lambda^4)$ and $\mathcal{O}(\lambda^5)$ are important for the LHCb experiment as it has the sensitivity required to measure them.

The CP violation is contained within the η parameter as $\tan \delta = \eta/\rho$. As there are several possible conventions for the CP phase δ in the CKM matrix, the following convention-independent expression is very useful [20],

$$\bar{\rho} + i\bar{\eta} = -\frac{V_{ub}^*V_{ud}}{V_{cb}^*V_{cd}} \quad \bar{\rho} = \rho(1 - \frac{\lambda^2}{2}) \quad \bar{\eta} = \eta(1 - \frac{\lambda^2}{2}). \tag{3.100}$$

The unitarity of the CKM matrix is an important test of the Standard Model, as it has

been shown in 3.4.1 that this property is a consequence of the SM. Testing the unitarity of V is done in two ways. The first is called the weak universality, and derives from the property that rows and columns in a unitary matrix form orthonormal basis,

$$\sum_i |V_{ij}|^2 = \sum_j |V_{ij}|^2 = 1. \quad (3.101)$$

The second type of test is done on the unitary triangles, that is

$$\sum_i V_{ij}^* V_{ik} = 0 \quad \text{and} \quad \sum_k V_{ik} V_{jk}^* = 0. \quad (3.102)$$

Experimental results deviating from these relations would imply the existence of physics beyond the standard model.

3.6 CP Violation in Mesons

In this section, the different types of possible CP violation will be presented. The case will be restricted to neutral mesons M^0 and their CP conjugated states $\bar{M}^0$. It is assumed that the two states M^0 and $\bar{M}^0$ can only be distinguished by an internal quantum number Q . In addition, it is assumed that both the strong and electromagnetic interactions have $\Delta Q = 0$ and weak interaction has $\Delta Q \neq 0$. The Hamiltonian will be composed of three different parts,

$$\mathcal{H} = \mathcal{H}_{\Delta Q=0} + \mathcal{H}_{\Delta Q=1} + \mathcal{H}_{\Delta Q=2}. \quad (3.103)$$

In equation 3.103 the term $\mathcal{H}_{\Delta Q=0}$ stands for the strong and electromagnetic interactions, $\mathcal{H}_{\Delta Q=1}$ for the weak interaction and $\mathcal{H}_{\Delta Q=2}$ for the double weak interactions. Weak interaction will allow transitions from M^0 to $\bar{M}^0$ if there is an intermediate state I coupled to both M^0 and $\bar{M}^0$ by the weak interaction. The transition is then possible through

$$M^0 \rightarrow I \rightarrow \bar{M}^0 \quad \bar{M}^0 \rightarrow I \rightarrow M^0. \quad (3.104)$$

The super-weak interaction allows for direct transitions from $M^0 \rightarrow \bar{M}^0$ without going through an intermediate state I . Four different decay amplitudes may be defined for a

meson M^0 into a final state f ,

$$\begin{aligned} A_f &= \langle f | \mathcal{H} | M^0 \rangle & \bar{A}_f &= \langle f | \mathcal{H} | \bar{M}^0 \rangle \\ A_{\bar{f}} &= \langle \bar{f} | \mathcal{H} | M^0 \rangle & \bar{A}_{\bar{f}} &= \langle \bar{f} | \mathcal{H} | \bar{M}^0 \rangle. \end{aligned} \quad (3.105)$$

In these amplitudes $\mathcal{H}$ is the Hamiltonian for the weak interactions. The action of the CP Conjugation on the M and f states is to introduce phases

$$\begin{aligned} \text{CP} |M\rangle &= e^{+i\phi_M} & \text{CP} |f\rangle &= e^{+i\phi_f} \\ \text{CP} |\bar{M}\rangle &= e^{-i\phi_M} & \text{CP} |\bar{f}\rangle &= e^{-i\phi_f}. \end{aligned} \quad (3.106)$$

This implies $(CP)^2 = 1$. Also, ϕ_M and ϕ_f are unphysical, because the strong interaction has flavour symmetry. If the Hamiltonian $\mathcal{H}$ commutes with CP, then A_f and $\bar{A}_{\bar{f}}$ will only differ by a relative unphysical phase

$$\bar{A}_{\bar{f}} = e^{i(\phi_f - \phi_M)} A_f. \quad (3.107)$$

If the M^0 decays to the final state $f = \sum_i f_i$ and $\bar{M}^0$ decays to the final state $\bar{f} = \sum_i \bar{f}_i$, the time evolution of a $M^0 - \bar{M}^0$ mixed state $\Psi(t)$ is

$$|\Psi(t)\rangle = a(t)|M^0\rangle + b(t)|\bar{M}^0\rangle + c_i(t)|f_i\rangle + d_i(t)|\bar{f}_i\rangle. \quad (3.108)$$

Solving the Schrödinger equation,

$$i\hbar \frac{\partial}{\partial t} \Psi(t) = \mathcal{H} \Psi(t), \quad (3.109)$$

is not possible unless some additional assumptions are made :

- The initial state is mixing of only M^0 and $\bar{M}^0$.
- Only $a(t)$ and $b(t)$ are studied, not the $c_i(t)$.
- Considered t is much larger than typical strong interaction time scale.

In that case, the equation 3.110 may be rewritten as

$$i\hbar \frac{\partial}{\partial t} \psi(t) = \mathcal{H} \psi(t) \quad \text{with } \psi(t) = \begin{pmatrix} a(t) \\ b(t) \end{pmatrix}. \quad (3.110)$$

In such case, the matrix $\mathcal{H}$ takes the simple form

$$\mathcal{H} = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma}. \quad (3.111)$$

The constraints from **CPT**, **CP** and **T** are then added, using the following definitions for the **C**, **P** and **T** operators,

$$\mathbf{C}|M^0\rangle = -|\bar{M}^0\rangle \quad \mathbf{P}|M^0\rangle = -|M^0\rangle \quad \mathbf{T}|M^0\rangle = |M^0\rangle. \quad (3.112)$$

CPT invariance is assumed to hold, meaning that $M_{11} = M_{22}$ and $\Gamma_{11} = \Gamma_{22}$. Solving the Schrodinger equation assuming **CPT** invariance leads us to the mass eigenstates,

$$|M_1\rangle = p|M^0\rangle + q|\bar{M}^0\rangle \quad |M_2\rangle = p|M^0\rangle - q|\bar{M}^0\rangle \quad (3.113)$$

The full derivation of this result may be found in [23]. The eigenvalues corresponding to these two eigenstates are

$$\begin{aligned} M_1 - \frac{i}{2}\Gamma_1 &= M_{11} - \frac{i}{2}\Gamma_{11} + \frac{q}{p} \left(M_{12} - \frac{i}{2}\Gamma_{12} \right) \\ M_2 - \frac{i}{2}\Gamma_2 &= M_{11} - \frac{i}{2}\Gamma_{11} - \frac{q}{p} \left(M_{12} - \frac{i}{2}\Gamma_{12} \right). \end{aligned} \quad (3.114)$$

where the ratio q/p is

$$\frac{q}{p} = \pm \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}}. \quad (3.115)$$

The sign choice in 3.115 is equivalent to a label swap in 3.114, and so is just a convention matter. The positive signed solution in 3.115 is assumed from now. The difference in mass and lifetime between the two states are

$$\begin{aligned} \Delta M &= -\text{Re} \left(\frac{p}{q} (M_{12} - \frac{i}{2}\Gamma_{12}) \right) \\ \Delta \Gamma &= 2\text{Im} \left(\frac{p}{q} (M_{12} - \frac{i}{2}\Gamma_{12}) \right). \end{aligned} \quad (3.116)$$

If **CP** or **T** holds, then $M_{12} = M_{21}$ and $\Gamma_{12} = \Gamma_{21}$. This is equivalent to say that the ratio Γ_{12}/M_{12} is real, and 3.115 becomes $q/p = 1$. Also, in such non **CP** violation case,

the two mass eigenstates are orthogonal,

$$\langle M_1 | M_2 \rangle = |p|^2 - |q|^2 = 0. \quad (3.117)$$

In the more general case where **CP** violation is present, the time evolution of a neutral meson M_p starting as an initially pure mixture of M^0 and $\bar{M}^0$ is

$$\begin{aligned} |M_p^0(t)\rangle &= g_+(t)|M^0\rangle - \frac{q}{p}g_-(t)|\bar{M}^0\rangle \\ |\bar{M}_p^0(t)\rangle &= g_+(t)|\bar{M}^0\rangle - \frac{q}{p}g_-(t)|M^0\rangle, \end{aligned} \quad (3.118)$$

where the $g_{\pm}$ functions are,

$$g_{\pm}(t) = \frac{1}{2} \left(e^{-im_1 t - \frac{1}{2}\Gamma_1 t} \pm e^{-im_2 t - \frac{1}{2}\Gamma_2 t} \right). \quad (3.119)$$

To see where the **CP** violation occurs, it is most useful to look at the time dependent decay rate. To make the expression more compact, we define $x = \Delta M/\Gamma$ and $y = \Delta\Gamma/2\Gamma$, with $\Gamma = (\Gamma_1 + \Gamma_2)/2$. Using these definitions, the time-dependent decay rate for $M^0 \rightarrow f$ is

$$\begin{aligned} \frac{\frac{d}{dt}\Gamma(M_p^0(t) \rightarrow f)}{e^{-\Gamma t}N_f} &= \left(|A_f|^2 + \left|\frac{q}{p}\bar{A}_f\right|^2 \right) \cosh(y\Gamma t) \\ &+ \left(|A_f|^2 - \left|\frac{q}{p}\bar{A}_f\right|^2 \right) \cos(x\Gamma t) \\ &+ 2\text{Re} \left(\frac{q}{p}A_f^* \bar{A}_f \right) \sinh(y\Gamma t) \\ &- 2\text{Im} \left(\frac{q}{p}A_f \bar{A}_f^* \right) \sin(x\Gamma t), \end{aligned} \quad (3.120)$$

while the decay rate for $\bar{M}^0 \rightarrow f$ is

$$\begin{aligned} \frac{\frac{d}{dt}\Gamma(\bar{M}_p^0(t) \rightarrow f)}{e^{-\Gamma t}N_f} &= \left(\left|\frac{q}{p}A_f\right|^2 + |\bar{A}_f|^2 \right) \cosh(y\Gamma t) \\ &+ \left(\left|\frac{q}{p}A_f\right|^2 - |\bar{A}_f|^2 \right) \cos(x\Gamma t) \\ &+ 2\text{Re} \left(\frac{q}{p}A_f \bar{A}_f^* \right) \sinh(y\Gamma t) \\ &- 2\text{Im} \left(\frac{q}{p}A_f^* \bar{A}_f \right) \sin(x\Gamma t). \end{aligned} \quad (3.121)$$

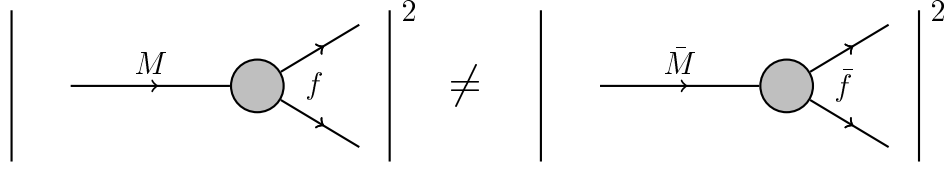


Figure 3.1: Illustration of direct **CP** violation.

Decay rates to the **CP** conjugate final state $\bar{f}$ are built the same way. In Equations 3.120 and 3.121, the N_f term is a time-independent normalization factor. Several types of terms can be identified and classified in terms of their physical contributions to the decay rates,

- Terms proportional to $|A_f|^2$ and $|\bar{A}_f|^2$ correspond to decays occurring without net oscillations. For charged mesons, they will be the only possible source of **CP** violation.
- Terms proportional to $|\frac{q}{p}A_f|^2$ and $|\frac{q}{p}\bar{A}_f|^2$ associate with decays after net oscillation happening. These terms can produce **CP** violation in the mixing.
- The $\sinh(y\Gamma t)$ and $\sin(x\Gamma t)$ terms are related to interferences between the two first cases. They will be a source of **CP** violation through interference between oscillation and decay.

3.6.1 CP Violation in Decay

CP violation in decay is also called direct **CP** violation. There must be little or no oscillations between the M^0 and $\bar{M}^0$ states occurring, that is $\Delta M = \Delta\Gamma = 0$. In such case, the time evolution is purely exponential and **CP** violation in decay will occur if,

$$|A_f| \neq |\bar{A}_{\bar{f}}| \quad (3.122)$$

This type of **CP** violation is the only one available for charged mesons. An illustration of this type of **CP** is shown in Figure 3.1 The corresponding **CP** asymmetry term for the charged meson example can be written,

$$\mathcal{A}_{f^\pm} = \frac{|\Gamma(M^- \rightarrow \bar{f}^-) - \Gamma(M^+ \rightarrow f^+)|}{|\Gamma(M^- \rightarrow f^-) + \Gamma(M^+ \rightarrow f^+)|} = \frac{|\bar{A}_{f^-}/A_{f^+}|^2 - 1}{|\bar{A}_{f^-}/A_{f^+}|^2 + 1} \quad (3.123)$$

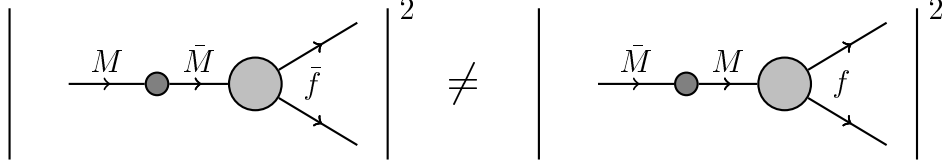


Figure 3.2: Illustration of **CP** violation in the mixing.

3.6.2 CP Violation the Mixing

Also called indirect **CP** violation, it is defined by

$$\left| \frac{q}{p} \right| \neq 1. \quad (3.124)$$

It occurs in flavour specific decays, where the final states f and $\bar{f}$ can be produced by either M^0 or $\bar{M}^0$ but not from both, as shown in Figure 3.2

$$M^0 \rightarrow f \text{ and } \bar{M}^0 \rightarrow \bar{f} \quad \text{or} \quad \bar{M}^0 \rightarrow f \text{ and } M^0 \rightarrow \bar{f}. \quad (3.125)$$

An example is the case of semileptonic neutral meson decay $M^0 \rightarrow l^+ X$ and $\bar{M}^0 \rightarrow l^- X$. In this case, $|A_{l^+ X}| = |\bar{A}_{l^- X}|$ and $A_{l^- X} = \bar{A}_{l^+ X} = 0$ as it is the case in the Standard Model. In this semileptonic example, this is the only available source of **CP** violation. It is measured by looking at wrong signed decays induced by oscillations,

$$\mathcal{A}_{SC} = \frac{\frac{d\Gamma}{dt}(\bar{M}_p^0(t) \rightarrow l^+ X) - \frac{d\Gamma}{dt}(M_p^0(t) \rightarrow l^- X)}{\frac{d\Gamma}{dt}(\bar{M}_p^0(t) \rightarrow l^+ X) + \frac{d\Gamma}{dt}(M_p^0(t) \rightarrow l^- X)} = \frac{1 - \left| \frac{q}{p} \right|^4}{1 + \left| \frac{q}{p} \right|^4}, \quad (3.126)$$

where the p in M_p^0 and $\bar{M}_p^0$ indicates the state of the neutral meson at production.

3.6.3 CP Violation by Interference

This type of **CP** violation is produced by the interference between two decays, one with mixing $M^0 \rightarrow \bar{M}^0 \rightarrow f$ and the other without $M^0 \rightarrow f$, both decaying to the same final state, as shown in Figure 3.3. **CP** violation of this kind is defined by

$$\text{Im} \left(\frac{q \bar{A}_f}{p A_f} \right) \neq 0. \quad (3.127)$$

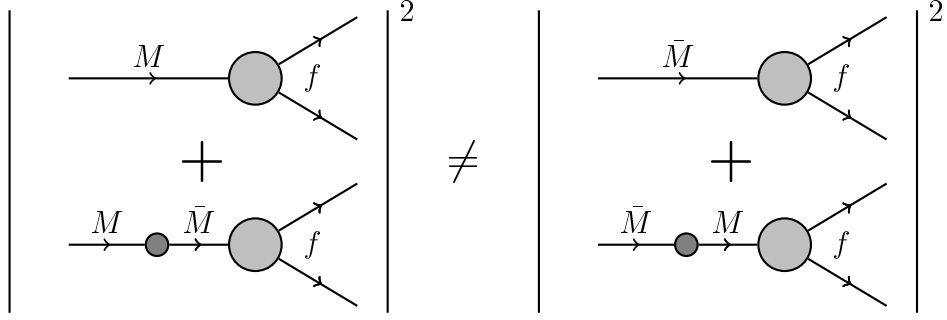


Figure 3.3: Illustration of **CP** violation by interference.

In the case of a neutral meson decaying to a final **CP** eigenstate $f = f_{CP}$, a **CP** asymmetry can be build,

$$\mathcal{A}_{f_{CP}}(t) = \frac{\frac{d\Gamma}{dt}(\bar{M}_p^0(t) \rightarrow f_{CP}) - \frac{d\Gamma}{dt}(M_p^0(t) \rightarrow f_{CP})}{\frac{d\Gamma}{dt}(\bar{M}_p^0(t) \rightarrow f_{CP}) + \frac{d\Gamma}{dt}(M_p^0(t) \rightarrow f_{CP})}. \quad (3.128)$$

One important remark about this kind of **CP** violation is that it can occur even in the absence of direct **CP** violation and violation in the mixing. This corresponds to a case where $|\bar{A}_f|^2 = |A_f|^2$ and $|\frac{q}{p}| = 1$, leading to the simple asymmetry term,

$$\mathcal{A}_{f_{CP}}(t) = \text{Im}\left(\frac{q\bar{A}_f}{pA_f}\right) \sin(x\Gamma t). \quad (3.129)$$

Chapter 4

Tracking and Alignment at LHCb

In this chapter, the tracking methods used in LHCb are presented along with the various track definitions. Also, the theoretical framework for alignment at LHCb is presented, especially for the Inner Tracker.

4.1 Track Reconstruction

At the LHCb experiment the pp collisions lead to high track multiplicities, with an average of 100 tracks in the VELO for typical $b\bar{b}$ events. In such a high track multiplicity environment, proper reconstruction of the tracks is a challenging task. This track reconstruction is performed in two steps. The first is pattern recognition, where hit clusters are attributed to tracks. The second step is the track fitting, where the parameters of the tracks are extracted through a fit of its trajectory. High precision track reconstruction is important for lifetime and mass measurements, as well as for B -meson flavour tagging.

4.1.1 Track Types

LHCb tracks are classified depending on which of the tracking stations they leave hits in. All LHCb tracks are included in one of the five following categories :

1. **VELO Tracks** have hits solely in the VELO. They may be tracks with large polar angle or going backward.

2. **Upstream Tracks** leave hits in the VELO and the TT. These tracks are bent out of the detector acceptance by the magnet due to their low momentum.
3. **Downstream Tracks** have their origin after the VELO and leave hits in the TT and Inner/Outer tracker stations.
4. **T Tracks** only leave hits within the Inner/Outer tracker stations. They can be used for the Inner/Outer tracker internal alignment and also for the RICH2 pattern recognition calibration.
5. **Long Tracks** traverse the entire tracking system, from the VELO to the Inner/Outer tracking stations. They are required to have a minimum number of hits in all the tracking stations.

4.1.2 Pattern Recognition

Pattern recognition is the reconstruction of track segments within the VELO or the T stations and then either the matching of this segment or their use as seed tracks, extending them to other tracking stations to reconstruct more complete tracks. Several algorithms are used in sequence to reconstruct the five types of tracks defined in Subsection 4.1.1. The different steps of the LHCb pattern recognition procedure are described in more detail below.

VELO Track Seeds

Clusters found in the VELO along straight lines are used to build a seed track [24, 25]. The use of straight lines is justified by the small magnitude of the magnetic field in the VELO region. The track seeds obtained here are then used as input for subsequent algorithms.

T Track Seeds

A track seed is built using clusters in the Inner Tracker and the Outer Tracker [26, 27]. Hits in x are taken in all three T-station and are fitted with a cubic curve (in fact the cubic term is very small, the effective fitted function is close to a parabola). Other x hits present in a window around the fitted curve are added. Finally stereo hits are added if available and compatible.

Forward Tracking

VELO track seeds are used as a starting point with one hit in the T stations to estimate the momentum of the candidate track. This candidate track trajectory is built from the VELO hits and the T hit. Other T stations hits are searched around the candidate's trajectory. If enough of these are found, and the track quality is high enough, the VELO track is promoted to a **Long** track. If the parameters of the track are such that it may have TT hits, these are searched for in a window and added to the tracks if available [28].

Track Matching

Matching between a set of VELO seed tracks to a set of T seed tracks is attempted for each possible pair of seed tracks. Both seed tracks are extrapolated to the bending plane of the magnet. Several parameters are then evaluated : the position in the bending plane, the variation of the slope, the number of compatible TT hits and others [29]. If the parameters are found to be compatible with a single track, then the two track seeds are merged and promoted to a **Long** track.

Upstream and Downstream Tracking

VELO seed tracks and T seed tracks are extrapolated to find compatible hits within the TT station. If this is the case, the VELO track seed is promoted to an **Upstream** track while a T seed track is promoted to a **Downstream** track.

VELO and T Ttracking

This algorithm uses leftover track seeds from the previous algorithms. The seed tracks are promoted to VELO tracks or T tracks depending on available hits. Due to the lack of information about the trajectory in the magnetic field region, these tracks have a poor momentum resolution.

Clone Killing

Clone tracks are tracks built with a fraction of identical hits. Clone killing is a step removing as many clone tracks as possible. This is the result of several algorithms attempting to reconstruct the same tracks. When such tracks are found, the track with the most hits is kept and the others are discarded. If two clone tracks have the same number of hits, the one with the best χ^2 is kept.

4.1.3 Track Fitting

The track fitting step start with the hits provided by the track finding step, and then fit a curve to these. The fitting procedure is based on a Kalman filter [30],[31]. This method starts building tracks from a small number of clusters and then adds hits one by one in a recursive way. It will return the track state vector parametrized as (x, y, t_x, t_y, κ) where $t_x = p_x/p_z$, $t_y = p_y/p_z$ and $\kappa = Q/p$. p is the momentum perpendicular to the magnetic field. Q corresponds to the charge of the track. In addition to the state vector, the filter returns the corresponding covariance matrix. At each iteration a hit from a new detection plane is added and the state vector and covariance matrix are updated.

The whole Kalman filter procedure is equivalent to a least χ^2 global fit, but presents several advantages compared to this latter method. The main advantages are that the pattern recognition and track fitting are in fact performed by one single algorithm and that the recursive nature of the procedure makes it faster than a global fit. Also, the Kalman procedure allows to account for track scattering and energy loss. The inclusion of different tracking technologies in the track fitting is straightforward with this method.

4.2 Alignment Strategy

In order to be able to achieve its physics goals, the positions of the parts of the LHCb detector have to be known to a very high precision. The reason for this is that any misalignment of the tracking system will notably lead to degradation of the momentum measurement and flight distance determination of the particles. This will in turn worsen the resolution on the measurement of the reconstructed particle masses and lifetimes. In addition, alignment is necessary to achieve the best possible tracking efficiency with the detector.

The objective of the procedure is to provide alignment of the different LHCb sub-detectors to a precision that is better than the intrinsic spatial resolution of these [32]. The subdetector alignment is first done internally only using the information from the subdetector. This internal alignment is possible for the VELO, IT and OT. In the TT case, an internal alignment is not possible due to the limited number of tracking planes. The different subdetectors are then aligned globally, using the tracking information from all the subdetectors. This strategy has been used during the start of the experiment and is further detailed in the following section.

4.2.1 VELO Internal Alignment

VELO alignment is extremely important as the detector is later used to fix the global coordinate system for the LHCb detector and also to determine the B -meson flight distance with respect to the primary vertices. The alignment procedure for the VELO is divided into three distinct phases : first the relative alignment of the r and ϕ sensors, then the alignment of the VELO modules contained by each VELO-half, and finally, the relative alignment of the two VELO halves.

4.2.2 IT and OT Internal and Relative Alignment

The tracker is formed by two detectors, the Inner and Outer Trackers. Both the Inner and Outer Tracker have 12 tracking planes, allowing to do an internal alignment in both cases. It is also possible to align both at the same time. Simultaneous alignment of the two Trackers is possible by using large angle tracks going through both or tracks crossing the small overlap region between the two.

The alignment of these Trackers is performed in several steps. First, the largest detector elements are aligned in order to resolve large misalignments. These large elements are the C-frames for the OT and the IT boxes. The OT C-frames are 12 support elements for OT modules, each covering the quarter of a station. Once this first step is performed, finer alignment is achieved by aligning the smaller detector elements down to the OT modules and the IT layers and ladders. Due to the tracker design, some detector elements are insensitive to some translations and/or rotations. An example is the x -layers of the IT that are completely insensitive to y translations and insensitive so small x -axis rotations. Also, if one apply a global x translation on the IT, without an

external reference point, it is not possible to solve it. This is an example of weak mode that can be encountered when aligning the tracking system.

Positions of small scale elements (OT modules, IT layers and ladders) of the Trackers where surveyed during the assembly of the tracker modular elements. These surveys provide information about large scale misalignments but are insufficient by themselves. In the IT, the ladder and layer positions where surveyed during the boxes assembly procedure. The precision of the survey for these elements is of $50\text{ }\mu\text{m}$ for the x -layers and x -ladders. For the stereo layers and ladders, it is not possible to survey them because they are masked by the x -ladders. However, as the stereo layers are mounted on the same support the x -layers, it is assumed that the x -layers corrections can also be applied to them. The position of the boxes was surveyed after installation in the LHCb cavern to a precision of $500\text{ }\mu\text{m}$. This survey was done with the IT stations away from the beampipe also called open position, and therefore does not account for possible variations due to the closing procedure. In the case of the OT, the modules and layers were surveyed along with a careful adjustment of the OT modules in the cavern to have them in their nominal positions.

4.2.3 VELO to T Station Relative Alignment

Once the VELO and the T stations are internally aligned, the next step is to align the T stations relative to the VELO. This part of the alignment uses **Long** tracks. At this step the VELO position in the LHCb frame is fixed to provide a constraint for the T stations positions. What is obtained at the end of this step is the relative alignment of the T stations with respect to the VELO.

4.2.4 TT Alignment from VELO to T Station Alignment and Long Tracks

Once the T stations and the VELO are aligned, the TT stations are aligned. The TT stations cannot be aligned internally because there are only four detection planes, making any standalone tracking unreliable. The alignment can be performed with **Long** tracks if the TT is aligned to both the T stations and the VELO. It is also possible to align it to the T stations using **Downstream** tracks or to the VELO using **Upstream** tracks. Once the TT stations are aligned with respect to the T stations and VELO,

the whole tracking system is aligned and the alignment of the other subdetectors can be started.

4.2.5 RICH1 and RICH2, Calorimeters and Muon Stations Alignment

All these detectors have a role in particle identification. Misalignment in these detectors will degrade their performances through wrong particle assignment. In the case of the RICH1 and RICH2 detectors, misalignment of the mirrors can change the measured Cherenkov angle and lead to lower particle identification performances. For the calorimeters, misalignment will degrade the reconstruction performance in decays dependent on information from these calorimeters. This leads to worse reconstructed mass resolutions. Finally, the muon stations need to be well aligned despite their rather low resolution. This is due to their important role in the Level 0 (L0) trigger by the information they provide by identifying muons and providing fast p_T measurement. The alignment precision for the calorimeters is approximately 0.5 mm and 1 mm for the muon system.

4.3 Alignment Theoretical Framework

All analytical linearized alignment paradigms imply the minimization of a χ^2 [33], built with track residuals and covariances, and end up with a set of linear equations in the alignment parameters. The size of this set of equations is simply the product of the number of aligned elements with the number of degrees of freedom aligned for. For example, aligning the 12 IT boxes for x and y translations (denoted Tx and Ty) and z -axis rotation (Rz) leads to a set of $12 \cdot 3 = 36$ linear equation,

$$\mathcal{M} \cdot \mathbf{a} = \mathbf{V}, \quad (4.1)$$

where $\mathcal{M}$ is the alignment inverse covariance matrix, $\mathbf{a}$ is the vector of alignment parameters and $\mathbf{V}$ the alignment vector.

If no regularization (which is further detailed in Subsection 4.3.1) procedure is applied, the solving will be stopped by the singularities of $\mathcal{M}$. These singularities originate

from movements and distortions to which we are insensitive, i.e. the problem (track reconstruction in particular) is invariant under certain combinations of the alignment parameters usually named "degenerate modes". Some of these modes are intuitively derivable, such as global translations and rotations of the whole system, longitudinal shearing along the track path; but other modes are less easy to find. The intuitive modes are exact singularities or "singular modes", i.e the corresponding eigenvalues are exactly null (in theory), while the remaining ones are pseudo-singularities (very small eigenvalues) or weak modes. In what follows, we will keep the expression "weak modes" for the whole set.

Obvious degeneracies, such as global translations, rotations and shearing, can be removed by applying canonical constraints on the χ^2 , i.e adding penalty terms with Lagrange multipliers. These simple modes can alternatively be removed using fixed aligned elements as references. For example, in the IT this may be done by fixing the positions of the first and last tracking planes, effectively closing down any global translations, rotations and shearings for all three coordinates. The problem with this method is that if the fixed elements are misaligned, their misalignment will be compensated by the other elements, leading to biases in the alignment solution. Remaining modes are to be treated in another way. The approach we follow here is a more systematical treatment using a spectral analysis. We first apply a diagonalization of $\mathcal{M}$:

$$\mathcal{M} = \mathcal{U}^T \mathcal{D} \mathcal{U}. \quad (4.2)$$

$\mathcal{U}$ is the matrix the columns of which are eigenvectors (or vectors of the diagonal basis) it is unitary, $\mathcal{U}^T \mathcal{U} = 1$, $\mathcal{U}^{-1} = \mathcal{U}^T$. $\mathcal{D}$ is the diagonal matrix containing the eigenvalues. With this, Equation 4.1 can be rewritten as

$$\mathcal{D} \mathcal{U} \mathbf{a} = \mathcal{D} \mathbf{a}' = \mathcal{U} \mathbf{V} = \mathbf{V}'. \quad (4.3)$$

All the primed vectors are vectors expressed in the diagonal basis. Since the inversion of $\mathcal{D}$ is trivial, the determination of $\mathcal{U}$ is the main effort in the solving. To build the solution, we proceed as follows:

$$\begin{aligned} \mathbf{V}' &= \sum_j \beta_j \mathbf{d}^j, \\ \mathbf{a} &= \sum_i \frac{\beta_i}{\lambda_i} \mathbf{d}^i, \end{aligned} \quad (4.4)$$

where $\mathbf{d}^k$ are the eigenbasis vectors, λ_l are the eigenvalues, and $\mathbf{V}'$ has been expressed in terms of the eigenbasis vectors, i.e eigenmodes.

By definition $\mathbf{a}$ is a linear combination of eigenmodes. We also see from its expression that modes corresponding to zero or very small eigenvalues are problematic.

4.3.1 Regularisation

In our way towards regularization and treating the weak modes problem, it is worth mentioning that the elements of the alignment matrix have the following features:

1. Large values make the determinant huge and thus the diagonalization/inversion unstable numerically.
2. Pronounced hierarchy between elements close to the diagonal and off-diagonal elements.

These problems are addressed with simple rescaling or more generally, preconditioning. Zeroth order rescaling consists in rewriting Equation 4.1 as $\frac{1}{\Lambda}\mathcal{M} \cdot \mathbf{a} = \mathcal{M}^\Lambda \cdot \mathbf{a} = \frac{1}{\Lambda}\mathbf{V} = \mathbf{V}^\Lambda$. A sound value for Λ is $\max_{i,j} |\mathcal{M}_{ij}|$. Then, $\mathcal{M}^\Lambda$ is diagonalized and the remainder of the procedure is applied.

As for preconditioning, the techniques are numerous [34] and we will mention the one applied in our case. The general rows and columns equilibration formula [35] is:

$$\mathcal{M}'_{ij} = r_i \cdot \mathcal{M}_{ij} \cdot c_j, \quad (4.5)$$

with $r_i = \frac{1}{\sqrt{\max_j |\mathcal{M}_{ij}|}}$ and $c_i = \frac{1}{\sqrt{\max_j |\mathcal{M}_{ji}|}}$. In our case, $\mathcal{M}$ is symmetric positive definite and dominated by its diagonal elements, $r_i = c_i = \frac{1}{\sqrt{\mathcal{M}_{ii}}}$. The consequence is that Equation 4.1 is rewritten as

$$\mathcal{C}^T \mathcal{M} \mathcal{C} \cdot (\mathcal{C})^{-1} \mathbf{a} = \mathcal{C}^T \mathbf{V}, \quad (4.6)$$

with $\mathcal{C} = \text{Diag}\{1/\sqrt{\mathcal{M}_{ii}}\}$.

This means that the weights are essentially the errors of the alignment parameters (if one neglects correlations). The scale of the eigenvalues of the equilibrated matrix is thus given by the covariances of the alignment parameters [36].

Singular modes of $\mathcal{M}'$ are related to singular modes of $\mathcal{M}$ through the $\mathcal{C}$ -transformation

but it is not the case of the remainder of the modes, and in particular the weak modes, since $\mathcal{C}$ is not orthogonal in general (as it is the case here).

4.3.2 Alignment Analysis Method

From what has been developed in Section 4.3, the following methodology is applied for the alignment studies, in both simulation and data :

1. Study of the eigenvalues spectrum.
2. Understand the nature of the weak modes associated to null/small eigenvalues.
3. Build a stable solution through the removal of the weak modes.

This methodology is first used in a simulation study and then applied to real data with the goal of achieving a precise alignment of the Inner Tracker.

Chapter 5

IT Alignment Simulation Studies

This chapter presents the studies performed on the Inner Tracker alignment using Monte Carlo data, with a focus on the identification of the weak modes of the Inner Tracker. The aim of the alignment studies performed on Monte Carlo simulation data was to learn to stabilize the alignment. The stabilization uses cuts on the eigenvalues and requires to understand the weak modes of the Inner Tracker.

5.1 Beam-Gas Simulation Studies

The Inner Tracker standalone alignment was first performed on beam-gas Monte Carlo samples. The data sets used for alignment contain 50k Long tracks. The goal was to classify and understand the IT weak modes at the boxes level. The next step was to use this understanding of the weak modes to stabilize the alignment by removing the appropriate number of eigenvectors from the alignment procedure. The method is then applied for the alignment of the IT elements on TED (beam dump) data acquired during LHC beam synchronization tests in Chapter 6.

5.2 Event and Track Selection

The simulations studies use long tracks sample only. The primary parameter used for the selection of alignment tracks is the track χ^2 , which measures the quality of the tracks.

In particular, this measure is sensitive to ghosts¹ and multiple scattering which tend to produce low quality tracks corresponding to high χ^2 values. It is also dependent on the detector misalignment, which tends to reduce the overall quality of tracks. The strategy used for the track selection is to use **Long** tracks and to apply a cut value on the track χ^2 where the cut evolves iteratively. The initial cut is rather loose, in order not to lose tracks due to χ^2 degradation from misalignment whilst still removing the worst of them. In addition, for matched tracks a cut is applied on the χ^2 of the **VELO** track segment to T stations track segment matching. These cuts are tightened at each iteration up to the fourth one. In the studies made on Monte Carlo, the χ^2 cut evolves as shown in Table 5.1.

	Track χ^2 cut value	Track matching χ^2 cut value
iteration 1	< 100	< 100
iteration 2	< 40	< 80
iteration 3	< 20	< 50
iteration 4+	< 10	< 30

Table 5.1: Situation of track χ^2 iterative cuts for MC.

5.2.1 Weak Modes Study Methodology

In order to understand the Inner Tracker standalone alignment, the weak modes arising in an IT box alignment with 3 degrees of freedom (x and y translation and z rotations) have been extracted and studied. This choice is driven by the high sensitivity of the IT to these three degrees of freedom. In principle it is possible to align the IT for degrees of freedom where the sensitivity is lower, but one would have to have access to a large set of data and a good pre-existing alignment. Since the goal is to prepare a first IT alignment on data, other degrees of freedom are not included in this alignment study.

The procedure consists of an alignment without any constraint where the eigenvectors are extracted after one iteration. To visualize the weak modes, the eigenvectors are multiplied by a magnification factor of 200 and applied to the geometry as an alignment condition database on top of the nominal geometry. This modified geometry is then visualized using the LHCb Panoramix visualisation package. In addition to this, the

¹Fake tracks from random hit combinations

spectrum of the eigenvalues is extracted. Weak modes will correspond to small eigenvalues of this eigenspectrum, as seen in Equation 4.4. The number of weak modes is proportional to the number of degrees of freedom aligned for. Figure 5.1 shows the eigenspectrum for an alignment on beam gas Monte Carlo simulation data. In this case the degrees of freedom are the x and y translations (Tx , Ty) along with the z -axis rotations (Rz). This spectrum has an internal structure with a group of 8 small eigenvalues followed by a step and progressively larger values. One important thing to remember when talking about removing eigenvectors is that while the ones corresponding to weak modes have the largest amplitudes, the next largest contribution is the solution of our alignment problem. This implies that removing too many eigenvalues will destroy large parts of the alignment solution. The spectrum structure in this case allow us to know that most of the detector element displacements are contained within the 8 first eigenvectors. One therefore knows that the number of eigenvalues to remove is lower than 8.

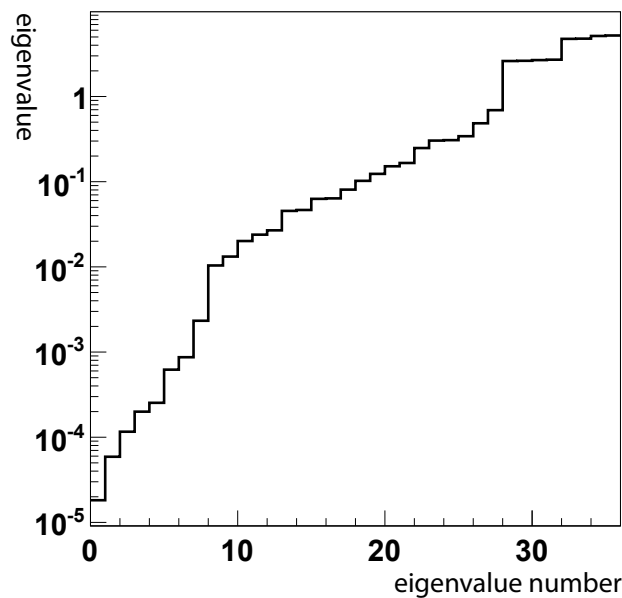


Figure 5.1: Eigenspectrum in the case of IT boxes alignment for Tx , Ty and Rz .

5.2.2 Weak Modes Analysis Results

Using the method described above, the IT weak modes at the boxes level have been identified. The result of this analysis is that in the case of an alignment with the 3 degrees of freedom Tx , Ty and Rz , there are up to 6 weak modes per free element in the IT. Free elements are uncorrelated sets of boxes in the IT. If the overlaps between the boxes are not used during the alignment, then the number of free elements is four. Each of these elements corresponds to a stack of 3 boxes along the z axis. If the overlaps are used, the number of free element goes down to one as the different boxes within a station are correlated. Moreover, when using **Long** tracks, the OT and VELO provide additional constraints. This further reduces the number of active weakmodes, as the OT and VELO positions are fixed and some **Long** tracks have hits both in the IT and OT. The weak modes have been found to be a combination of the transformations listed in Table 5.2 on the detector elements positions.

Degree of freedom	Description of the mode
Tx modes	global x translation x shearing along the z axis
Ty modes	global y translation y shearing along the z axis
Rz modes	global z axis rotation twisting along the z axis

Table 5.2: Weak modes in a Tx , Ty and Rz IT boxes alignment.

This result has been obtained by observing the different eigenvectors corresponding the smallest eigenvalues and decomposing them in terms of simple transformations of the IT aligned coordinates. The observed modes correspond to what would be naively expected.

As an example, a visualisation of the y shearing mode along the z axis is shown in Figure 5.2 while a global rotation around the z axis is made visible in Figure 5.3. A case where an eigenvector corresponds to the combination of a z axis twist for the Top and Bottom boxes and a combination of shearing in y along the z axis with a global z axis rotation for A-side and C-side boxes is shown in Figure 5.4. This latter example shows that a given eigenvector does not always correspond to a simple transformation

but may instead be a linear combination of simple transformations.

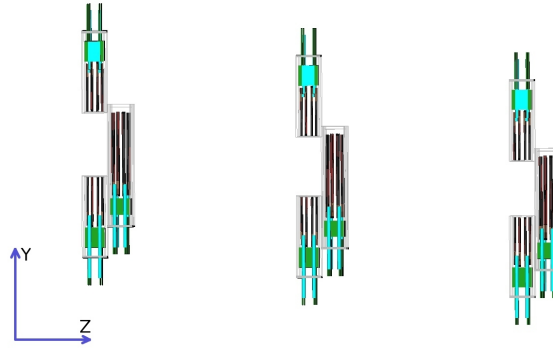


Figure 5.2: Visualisation of the y shearing along the z axis mode of IT boxes.

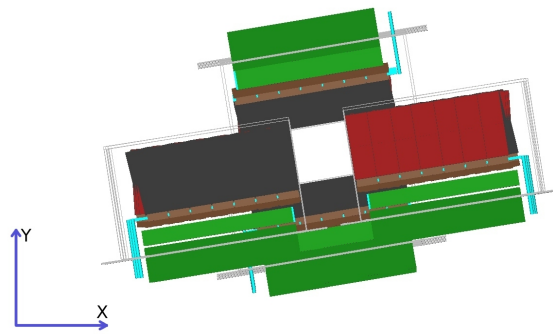


Figure 5.3: Visualisation of the global z axis rotation mode of IT boxes

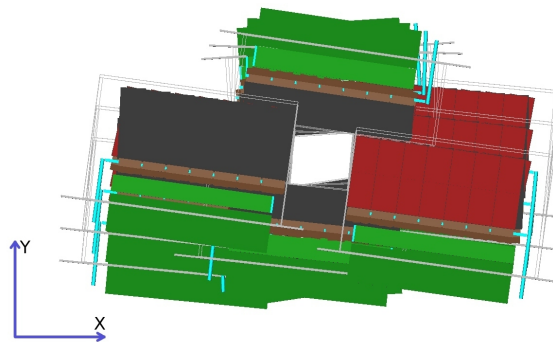


Figure 5.4: Composite mode of Top/Bottom z -axis twist with A-side/C-side z -rotation plus y shearing along z .

5.3 Alignment Stabilization using Eigenvalues

The knowledge of the IT box weak modes has then been applied to simulated test cases. In those, the goal is to solve alignment problems by stabilizing the system with cuts on the eigenvalues. The removal of the smallest eigenvalues is performed at the beginning of the alignment equation solving step of each iteration. The alignment test cases consist of random misalignments applied to each box for the T_x , T_y and R_z degrees of freedom. The random misalignments are sampled on a pdf built as the mean of four flat random variables with a sampling range of $[+1.0, -1.0]$ mm or $[+2.0, -2.0]$ mm. This probability density function (PDF) is an approximation of a Gaussian centered at zero with a variance σ of 0.5 mm or 1.0 mm and a cutoff at 2σ . The absence of tails prevents generating misalignments too large to be recovered.

The alignment tests on Monte Carlo simulated beam-gas events use **Long** tracks and are not standalone IT alignments. This alignment also make use of the boxes overalps. The other tracking detectors are fixed, and therefore act as constraints on the IT. These constraints will limit the amount of weak modes appearing in the alignment solution, meaning that not all of the six weak modes identified will be present. The correct number of eigenvalues removed is determined as the smallest possible amount leaving a stable alignment, as visible in Figures 5.5, 5.6 and 5.7. The effect of removing too few or too many eigenvalues is shown in Figures 5.8 and 5.9, respectively.

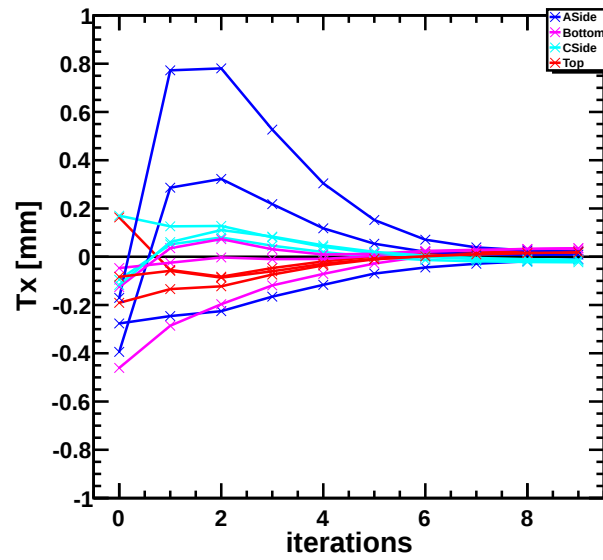


Figure 5.5: Alignment flow with iteration for x translations, with removal of the two smallest eigenvalues.

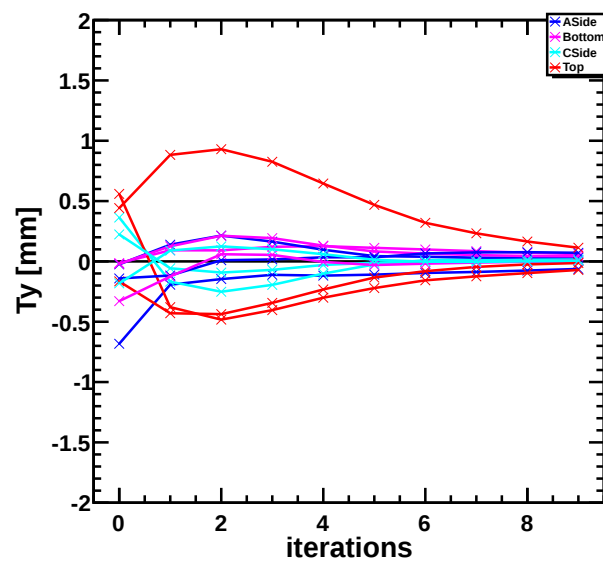


Figure 5.6: Alignment flow with iteration for y translations, with removal of the two smallest eigenvalues.

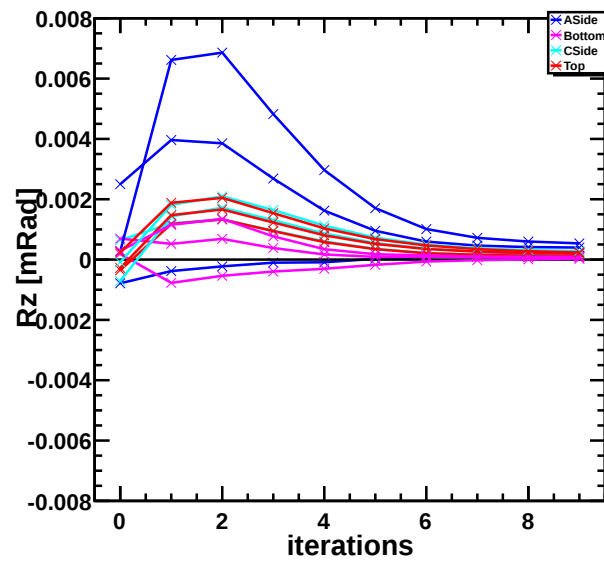


Figure 5.7: Alignment flow with iteration for z -axis rotations, with removal of the two smallest eigenvalues.

In Figures 5.5, 5.6 and 5.7, all the boxes converge as they all approximately return to a zero position corresponding to the nominal position. This effectively solves the initial misalignments applied to the boxes. The value of the starting misalignment is the starting value at iteration zero for all the plots. The choice of cutting only two eigenvalues is motivated by the stability of the alignment and the convergence to zero. In Figure 5.8, too few modes are removed and the y alignment converges to an incorrect solution, while in Figure 5.9, too many eigenvalues are removed, leading to poor precision on the y translation alignment of the IT boxes.

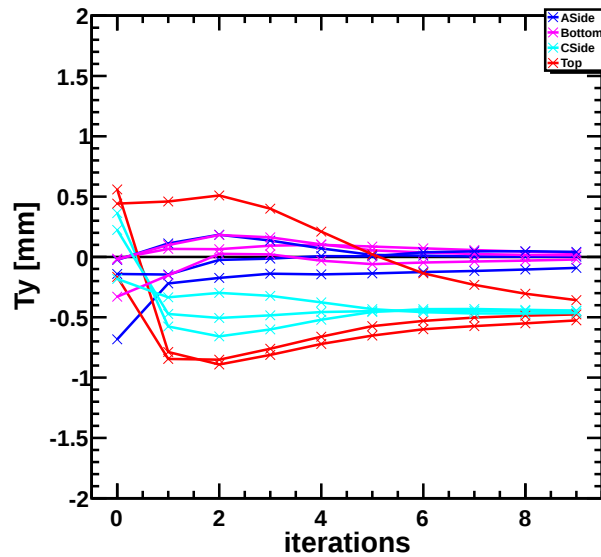


Figure 5.8: Alignment flow with iteration for y translations, with cut on only the smallest eigenvalue.

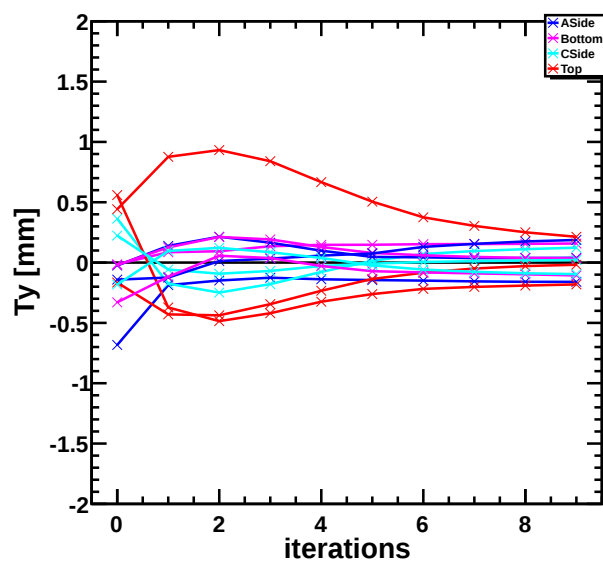


Figure 5.9: Alignment flow with iteration for y translations, with cut on the three smallest eigenvalues.

Chapter 6

Alignment of the Inner Tracker with Data

The goal of the study presented in this chapter is to provide a first alignment for the first collision data. This Inner Tracker alignment is performed with the so-called TED (beam dump) data collected by the LHCb detector. These data were produced by the LHC during beam synchronization tests in June and October 2009. The tests were conducted with 450 GeV proton beams from the SPS which were dumped on a beam stopper (TED.) The TED is situated 350 m downstream of the LHCb experiment, creating a shower of particles going through the detector. These events have a high track density of 20 times the nominal design, their particles have a moderately high momentum and were recorded with the LHCb magnet switched off. While the high density is detrimental to track reconstruction, the high momentum and magnet off simplify it, allowing to make use of the tracks for alignment studies, especially with cuts applied to remove the highest occupancy events.

To estimate the momentum of the tracks crossing the IT, TED Monte Carlo simulation data have been generated. The result of these simulations is that most particles going through the IT are muons with a momentum of $\mathcal{O}(10)$ GeV.

6.1 Alignment using the June 2009 TED Data

The data used for this alignment was recorded in June 2009 during TED events. The track and event selection will now be presented, followed by the alignment of the IT

boxes, layers and ladders. This alignment is performed without using the correlation terms due to the overlaps between the IT boxes during the solving, effectively aligning the twelve boxes as if they were completely independent.

6.1.1 TED Track Reconstruction

Due to the very high track density of the TED runs, the standard track reconstruction using the Kalman filter was found to be overwhelmed. An alternative and more robust track reconstruction algorithm called `ITGenericTracking` has been developed for this environment. This algorithm is only suitable in the absence of the magnetic field, since tracks are parametrized as straight lines. Because the algorithm execution time grows rapidly with the occupancy, excessively high track density events were discarded to limit the time needed for execution to acceptable levels. Steps before and after the track reconstruction are the same as in the standard alignment procedure.

The algorithm performance has been evaluated using TED Monte Carlo simulation data, and gave the following results :

1. For the top and bottom boxes of the IT, the tracking efficiency is 98% and the ghost rate is 0.8% for an average occupancy of 1.7%.
2. For the side boxes of the IT, the tracking efficiency is 81% and the ghost rate is 6.2% for an average occupancy of 4%.

6.1.2 Event and Track Selection

The alignment on the TED data is performed with data that is quite different from beam-gas Monte Carlo simulation data. The main differences are that the track density is much higher, track angles are small and tracks have relatively high momentum (6 GeV). The tracking used is also different, with the standard Kalman filter based tracking being replaced with simple straight line tracking more able to deal with the high track density. The iterative cut values applied to the χ^2 are tighter than in the simulation case. The values used in the TED case are listed in Table 6.1. In addition a cut on the total

IT occupancy is applied to avoid the highest occupancy events where the track density becomes too high for the `ITGenericTracking` algorithm. This rejection cut is placed

	χ^2 cut value
Iteration 1	< 20
Iteration 2	< 12
Iteration 3	< 8
Iteration 4+	< 6

Table 6.1: Track χ^2 cut used for alignment with June 2009 TED data.

at a number of IT hits of 2500 for the June 2009 TED data, giving a total statistics of 13k tracks. In addition to this, an isolation cut is applied on the tracks. This cut works by counting the number of hits that are not part of the considered track in a 20.0 mm cylinder around it. Tracks are rejected if more than six hits are found in the cylinder.

6.1.3 Alignment Sequence

The standalone IT alignment using TED data with the `ITGenericTracking` algorithm is performed in a top-down approach, aligning large detector elements first and then going to smaller ones by steps, fixing previously aligned larger elements. In this alignment procedure, overlaps between the elements of one given station of the IT are not taken into account. This leads to the effective alignment of 4 independent elements, each of which is a stack of three IT boxes along the z -axis.

The starting situation for this alignment is the alignment conditions obtained from an coarse alignment done with 2008 TED data. This 2008 TED alignment was performed only for top level elements, and only by looking at residual distributions to correct for the largest misalignment. The situation can be seen in Figures 6.1 and 6.2. The starting residual resolutions and biases are summarized in Table 6.2. In this table legend top and bottom boxes are denoted by T/B while A-side and C-side boxes are denoted by A/C.

The results presented in Figures 6.1, 6.2 and Table 6.2 are obtained as follows : for each ladder, the track hits position residual distribution is fitted with a Gaussian. Then, the means (bias) and widths (resolution) of the fits for each ladders are collected. The alignment results are then presented as the distribution of these values for all the IT ladders. The mean value of the bias distribution for each of the four groups of IT ladders is taken as the mean bias for the group's ladders, while the width is the uncertainty on

	Bias (μm)	Resolution (μm)
T/B X layers	-7 ± 48	95 ± 16
A/C X layers	4 ± 39	108 ± 18
T/B stereo layers	18 ± 117	139 ± 27
A/C stereo layers	8 ± 80	165 ± 36

Table 6.2: Situation before alignment with June 2009 TED data

the group's bias value. In a similar manner, for each of the four IT groups of ladders, the mean of the resolution distribution is taken as the group's mean resolution while the width is the uncertainty on the group's resolution.

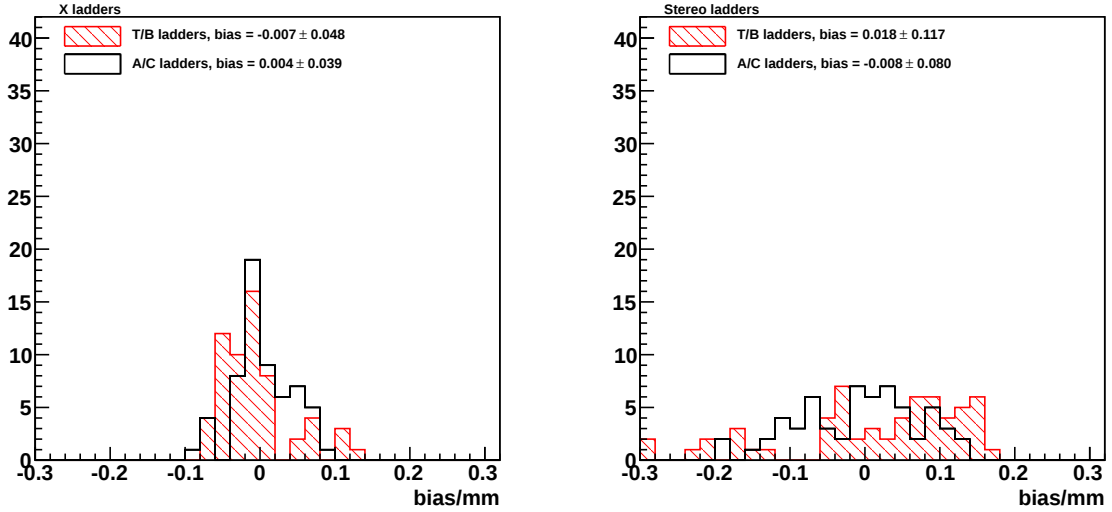


Figure 6.1: Ladders residual biases before box alignment on June TED data. Left plot is for X ladders and right plot is for stereo ladders.

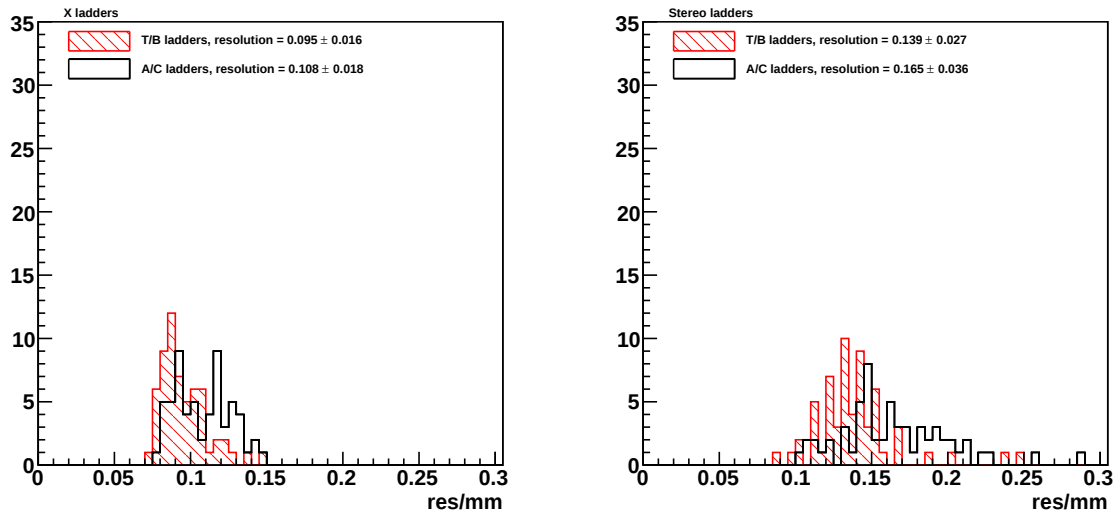


Figure 6.2: Ladders residual resolutions before box alignment on June TED data. Left plot is for X ladders and right plot is for stereo ladders.

Inner Tracker Box Alignment

The IT boxes are aligned for three degrees of freedom, the x and y translations and the z -axis rotations. This choice is dictated by the sensitivity of the IT to these degrees of freedom at the box level. While the sensitivity for x and y translations is clear, as the IT is designed to measure them, it is also sensitive to z -axis rotations. The sensitivity to the z -axis has been demonstrated in [37]. It can be explained by noticing that a z -axis rotation of a box is equivalent to the application of a sum of x and y translations on slices of the box. Sensitivity to other degrees of freedom also exists, but is a lot smaller.

Given the degrees of freedom considered, the weak modes for the IT boxes are x and y global translations and shearing, along with global z -axis rotations and z -axis twisting. These modes are suppressed by removing the corresponding eigenvectors from the alignment solution. This corresponds to removing the smallest six eigenvalues as explained in Section 4.3. One point to notice is that it is likely that some part of the solution is lost in this process. This can be understood as the IT only has three tracking stations, leading to a rather high probability that all the stations possess a misalignment in the same direction in x , y and z . There is no way of getting around this problem with an internal alignment, These parts of the solution can be retrieved when performing the alignment of IT relative to the VELO.

The alignment convergence for the IT boxes with June 2009 TED data is presented in Figures 6.3, 6.4 and 6.5. As visible in Figures 6.3, 6.4 and 6.5, the alignment parameters are stabilized in their evolutions after the second iteration. This shows that the alignment is stable. To measure the performance of this box alignment, the biases and resolutions after alignment are extracted, as shown in Figures 6.6 and 6.7 and summarized in Table 6.3.

	Bias (μm)	Resolution (μm)
T/B X layers	-5 ± 49	85 ± 11
A/C X layers	5 ± 38	91 ± 9
T/B stereo layers	14 ± 103	123 ± 20
A/C stereo layers	-11 ± 80	137 ± 23

Table 6.3: Situation after box alignment with June 2009 TED data.

As visible when comparing the status before box alignment summarized in Table 6.2

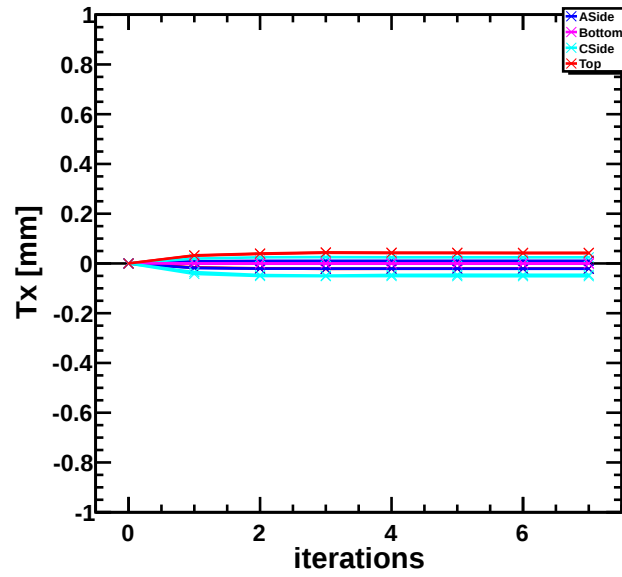


Figure 6.3: Flow of the boxes x translations degree of freedom with iterations.

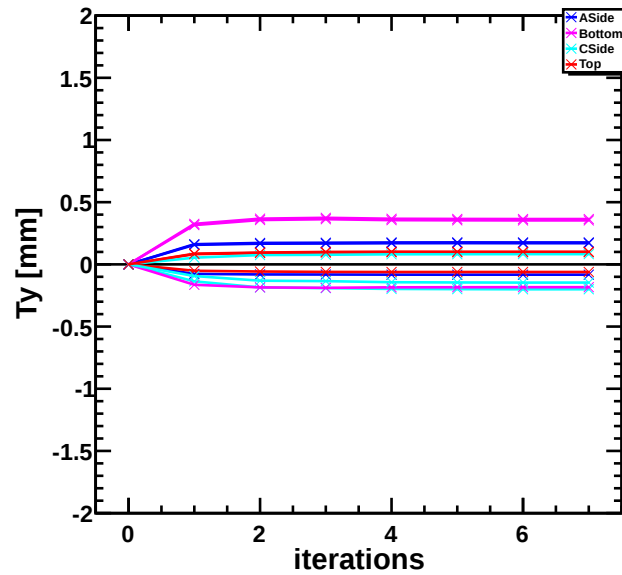


Figure 6.4: Flow of the boxes y translations degree of freedom with iterations.

with the result of the IT boxes alignment in Table 6.3, the mean IT resolution improves from $127 \pm 24 \mu\text{m}$ to $109 \pm 17 \mu\text{m}$.

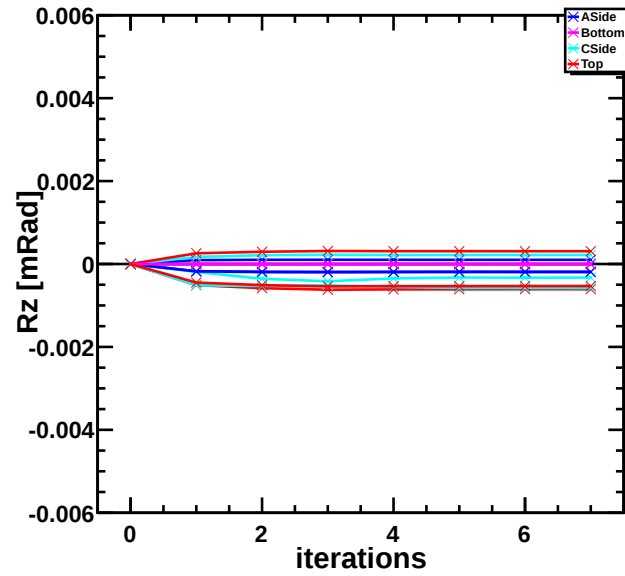


Figure 6.5: Flow of the boxes z rotations degree of freedom with iterations.

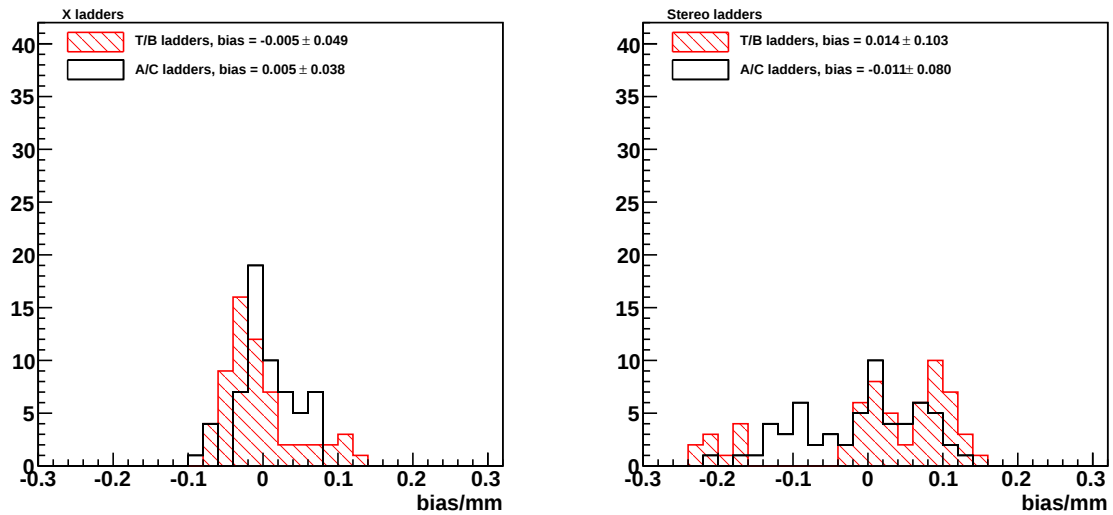


Figure 6.6: Ladder residual biases after box alignment. Left plot is for X ladders and right plot is for stereo ladders.

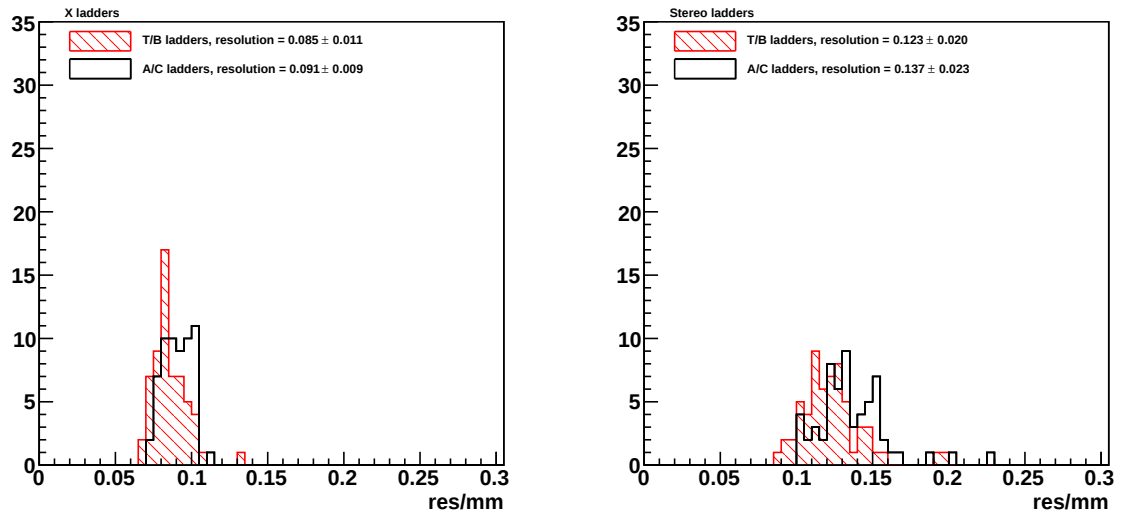


Figure 6.7: Ladder residual resolutions after box alignment. Left plot is for X ladders and right plot is for stereo ladders.

Inner Tracker Layers Alignment

After the alignment of the IT boxes, the layers are aligned using the alignment conditions from the boxes step as a starting point. For the layers, two degrees of freedom are aligned for : x translations and z -axis rotations. No alignment for the y translations is possible due to the lack of tracking information along this direction. This lack of information along the y axis reduces the alignment sensitivity to the z -axis rotations, but not enough to destabilize the alignment procedure.

Considering the degrees of freedom aligned for, there are five different weak modes to consider. While the first four are simple: global x translation, x shearing, z -axis global rotation and z -axis twist; the fifth one is non-trivial. It is a compensation between the pairs of X and stereo layers, where the two layers on each cooling rod move in opposite directions and create a displacement equivalent to a y translation combined with a x translation. The details of this weak mode can be found in Section 5.1.

Alignment convergence for the IT layers with June 2009 TED data is presented in Figures 6.8 and 6.9.

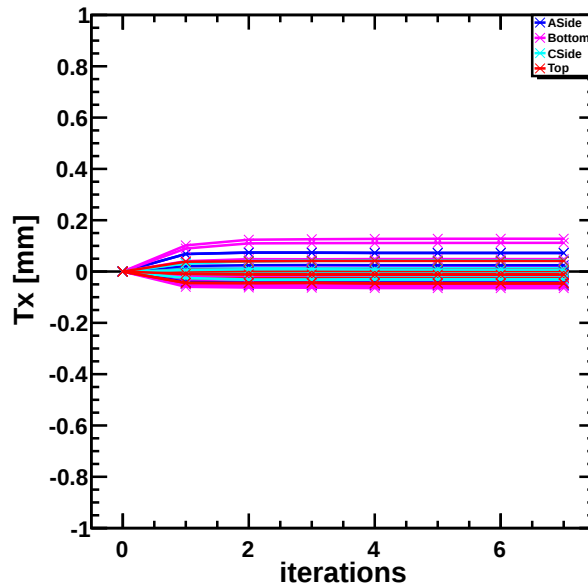


Figure 6.8: Flow of the layers x translations degree of freedom with iterations.

The alignment parameters are stabilized in their evolutions after the second iteration for x translations. However, as shown in Figure 6.9 the z -axis rotations need more iterations to stabilize due to the lower sensitivity to this degree of freedom. The performance

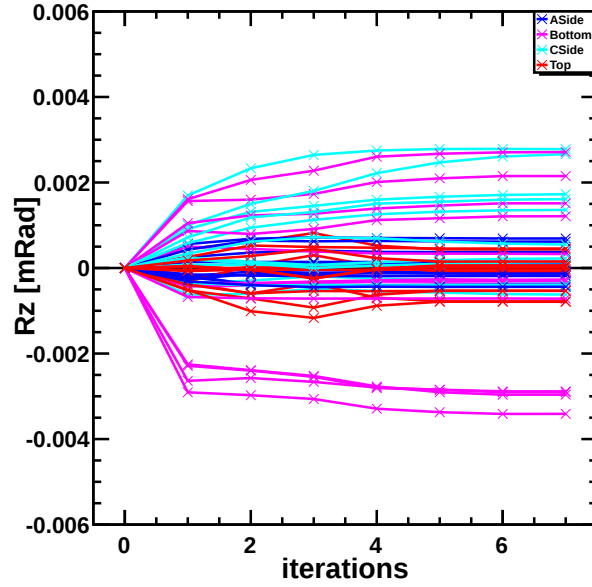


Figure 6.9: Flow of the layers x z -axis rotation degree of freedom with iterations.

of this layer alignment is measured by looking at the biases and resolutions. The results are presented in Figures 6.10 and 6.11 and summarized in Table 6.4.

	Bias (μm)	Resolution (μm)
T/B X layers	1 ± 15	83 ± 9
A/C X layers	1 ± 17	90 ± 8
T/B Stereo layers	1 ± 22	120 ± 18
A/C Stereo layers	-2 ± 37	134 ± 21

Table 6.4: Situation after layer alignment with June 2009 TED data.

If one compares the results obtained after box alignment summarized in Table 6.3 with the results after the layer alignment step summarized in Table 6.4 one observes a rather small improvement in the resolutions (from $109 \pm 16 \mu\text{m}$ to $107 \pm 14 \mu\text{m}$) but a much larger one in term of the biases distribution spread (from $1 \pm 68 \mu\text{m}$ to $0 \pm 23 \mu\text{m}$).

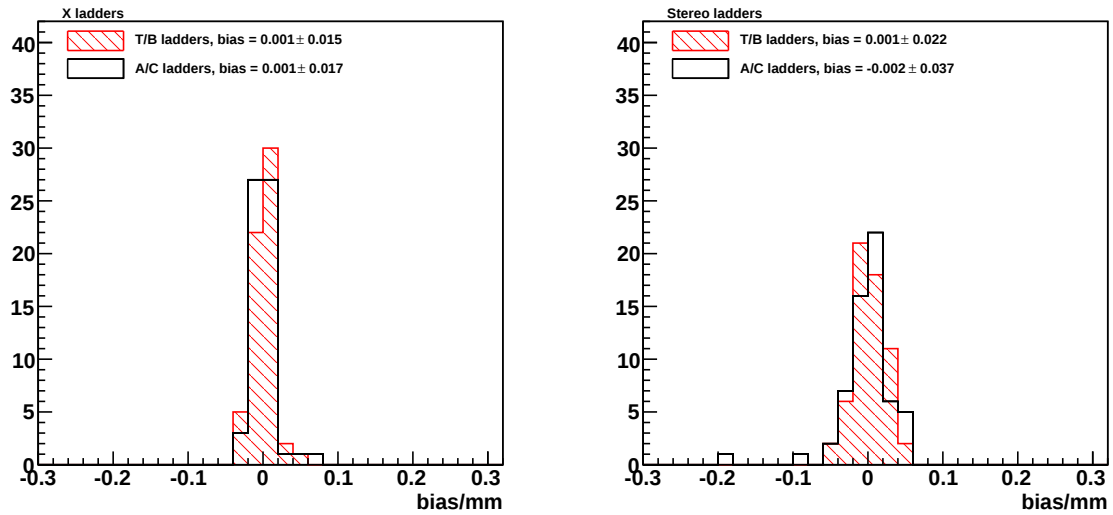


Figure 6.10: Ladder residual biases after layer alignment with June 2009 TED data.

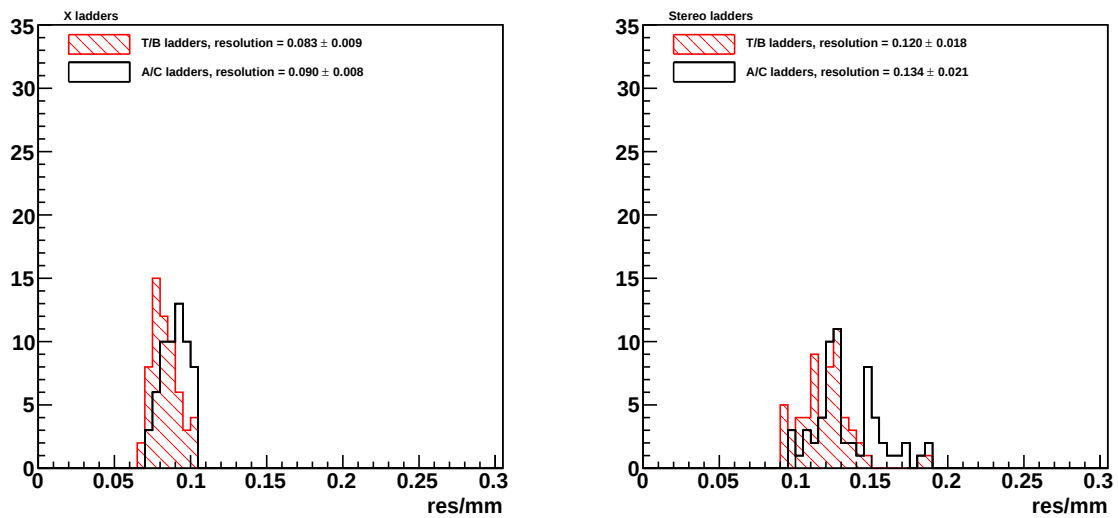


Figure 6.11: Ladder residual resolutions after layer alignment with June 2009 TED data.

Inner Tracker Ladder Alignment

The last step of the standalone Inner Tracker alignment is the ladder alignment. For reasons of stability, the outermost ladders in x are not aligned and are instead kept to

the survey values. For all the other ladders, the degrees of freedom that are aligned for are the x translations and z -axis rotations as for the layers. The motivations for this choice is the same as in the layers case.

In the ladder alignment, we have a total of eight weak modes to account for. Four of them are the x global translations and shearing along with the z -axis global rotation and twist. There are, however, four other modes that are more complicated and not clearly identified. At least one of them is a compensation mode similar to the one found in the layers case. The last three are modes of internal deformations of the ladder arrangement within the boxes and are of smaller amplitude than the first four. Therefore, the first 32 eigenvalues are removed. This number is the minimum to obtain convergence with the alignment.

	Bias (μm)	Resolution (μm)
T/B X layers	2 $\pm$ 10	80 $\pm$ 8
A/C X layers	1 $\pm$ 10	85 $\pm$ 8
T/B stereo layers	1 $\pm$ 4	113 $\pm$ 14
A/C stereo layers	0 $\pm$ 3	121 $\pm$ 13

Table 6.5: Situation after ladders alignment with June 2009 TED data.

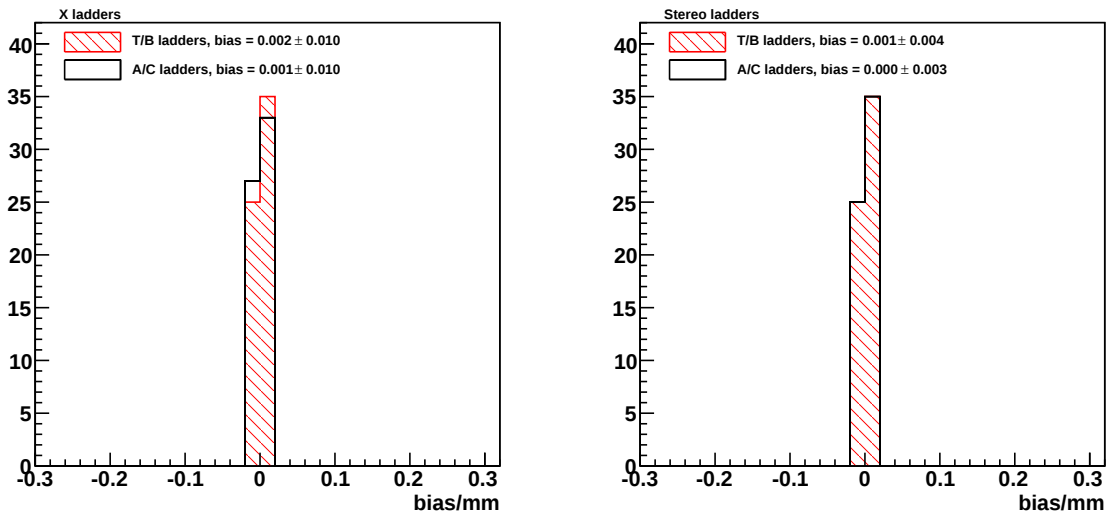


Figure 6.12: Ladder residual biases after ladder alignment.

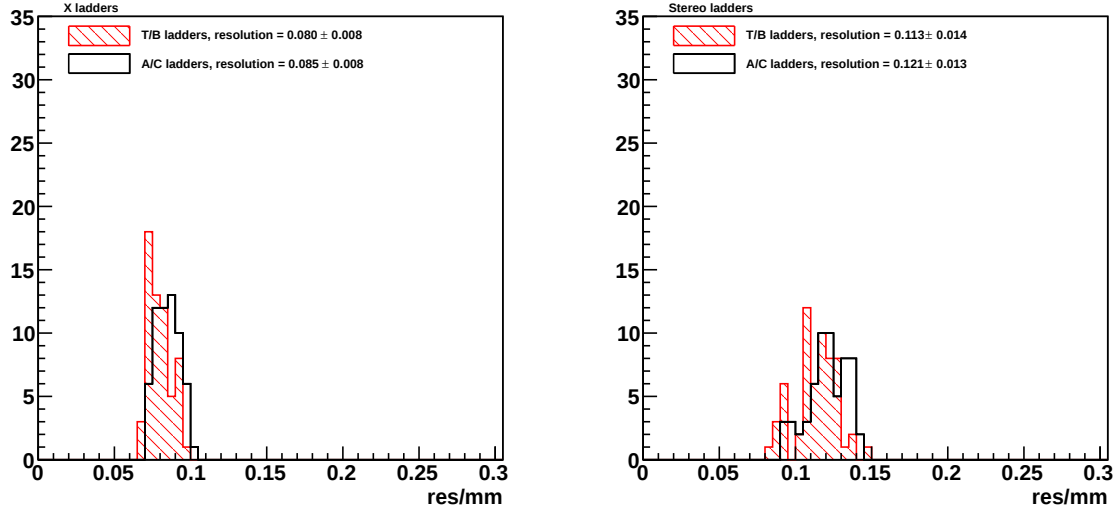


Figure 6.13: Ladder residuals resolution after ladder alignment.

The results of presented in Table 6.5 and Figures 6.12 and 6.13 are the final results obtained using the June 2009 TED data. One can notice the improvement over the previous alignment step as the resolutions improved from $107 \pm 14 \mu\text{m}$ to $100 \pm 11 \mu\text{m}$ and the biases distribution spread improved from $0 \pm 23 \mu\text{m}$ to $1 \pm 7 \mu\text{m}$.

The resolutions shown in Figure 6.13 are a quadratic sum of the IT intrinsic resolution of $57 \mu\text{m}$ and the alignment resolution. It is possible to retrieve the alignment part of the resolution to compare it the the intrinsic one. Both the total resolution and the extracted alignment only resolution are summarized in Table 6.6.

	Resolution (Total) μm	Resolution (Align.) μm
T/B X layers	80 $\pm$ 9	56 $\pm$ 9
A/C X layers	85 $\pm$ 8	63 $\pm$ 8
T/B stereo layers	113 $\pm$ 18	98 $\pm$ 18
A/C stereo layers	121 $\pm$ 21	107 $\pm$ 21

Table 6.6: Alignment contribution to resolution after final June 2009 TED data alignment.

6.2 Alignment using the October 2009 TED Data

In October 2009, several TED shots provided enough events to perform a new alignment of the IT. This alignment is used as a confirmation of the alignment performed in June 2009 and also to check if the opening and closing of the IT done between the two TED runs had introduced new misalignments. Finally, correlations coming from the overlap regions between IT boxes are now taken in account.

The alignment procedure is the same as the one used for June 2009 TED data. One important difference is the introduction of the correlations between boxes in the alignment. The inclusion of this feature reduces the number of cuts on the smallest eigenvalues, as it constrains the relative position of the boxes within an IT station.

The sizes of these overlaps between boxes change depending on the stations as the diameter of the beam pipe is not constant along the z -axis. The overlaps have the following sizes from the nominal geometry summarized in Table 6.7.

IT Station	$T \cap A$ [mm]	$T \cap C$ [mm]	$B \cap A$ [mm]	$B \cap C$ [mm]
Station 1	18.632	18.632	19.218	19.218
Station 2	11.632	11.632	12.228	12.228
Station 3	4.632	4.632	5.228	5.228

Table 6.7: Nominal overlap size in y for the IT sensors for each station, as determined from the nominal values.

6.2.1 Event and Track Selection

The χ^2 cuts values used in the October TED case are listed in Table 6.8. As for the analysis on June 2009 TED data, a total IT occupancy cut is applied to avoid the highest occupancy events and to keep a reasonable computing time with the `ITGenericTracking` algorithm. The rejection cut is increased to 4000 hits for the whole IT as it allows to increase the statistics to 11k tracks without increasing the CPU time. As for June data, the isolation cut is applied on the tracks. This cut works by counting the number of hits that are not part of the considered track in a 20.0 mm cylinder around it. If the algorithm finds more than 6 hits not part of the track within this window, the track is

	χ^2 cut value
Iteration 1	< 20
Iteration 2	< 12
Iteration 3	< 8
Iteration 4+	< 6

Table 6.8: Iterative cut values on the track χ^2 for October 2009 TED data.

discarded.

6.2.2 Alignment Sequence

The alignment sequence used for the October 2009 TED data is composed of two steps: First the IT boxes are aligned for the three Tx , Ty and Rz degrees of freedom, and then the Inner Tracker layers are aligned for the two Tx and Rz degrees of freedom. The IT ladders are not aligned during this alignment sequence as the boxes have not been opened.

Inner Tracker Box Alignment

The IT boxes are aligned for three degrees of freedom, namely x translations, y translations and z -axis rotations. Given the degrees of freedom considered, the weak modes for the IT boxes are x and y global translations and shearing, along with global z -axis rotations and z -axis twisting. These modes are the same as for the alignment on June 2009 TED data, and their corresponding eigenvectors are removed from the alignment solution.

The alignment convergence for the IT boxes with October 2009 TED data is presented in Figures 6.14, 6.15 and 6.16.

The observed movements of the boxes with respect to the June alignment results are due to two factors. The main one is the inclusion of the overlap constraint during the alignment procedure, which has a strong effect on the relative y position of the boxes in each station. The second effect is the opening of the IT between the June and October TED data taking, which introduced small misalignment in x .

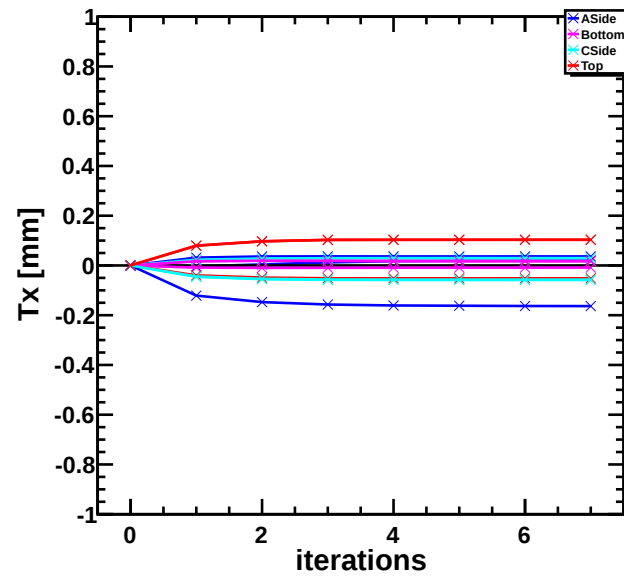


Figure 6.14: Evolution of the x translations of boxes with the number of iterations.

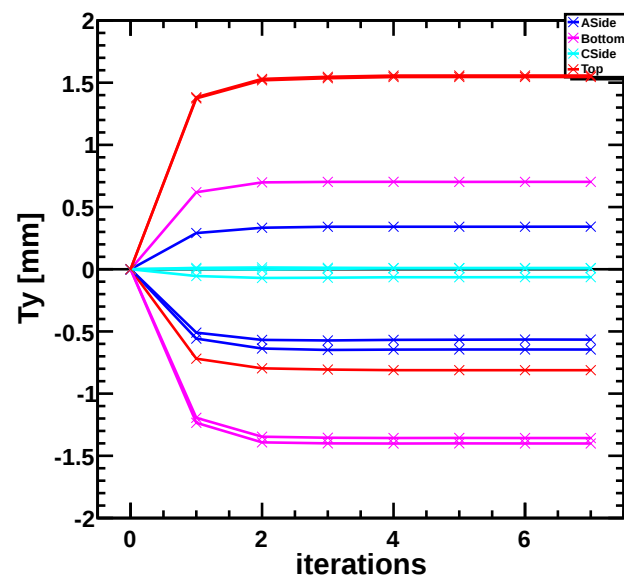


Figure 6.15: Evolution of the y translations of boxes with the number of iterations.

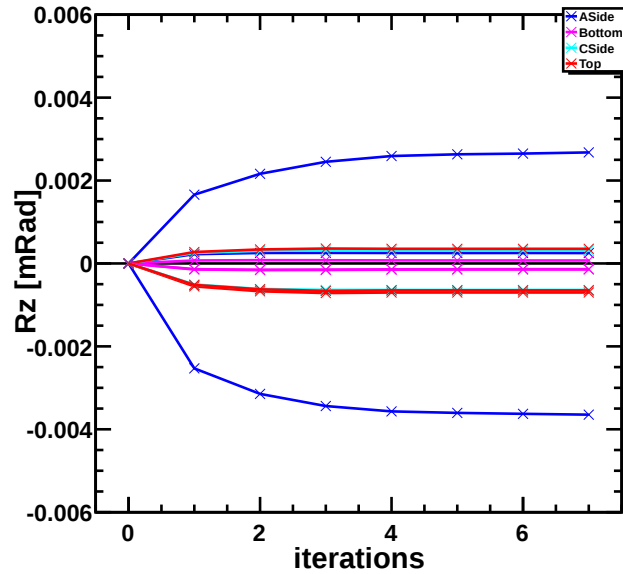


Figure 6.16: Evolution of the z rotations of boxes with the number of iterations.

The result of the box alignment using October TED data are presented in Table 6.9 and Figures 6.17 and 6.18. The resolutions obtained are close to the final ones obtained on June TED data, but the biases distribution spread has become wider. In both the resolutions and the biases, the difference with June results are concentrated on the T/B stereo layers. This degradations are recovered after the layer alignment step in Subsection 6.2.2.

	Bias (μm)		Resolution (μm)	
T/B X layers	-1	± 13	80	± 9
A/C X layers	2	± 8	86	± 6
T/B stereo layers	-1	± 36	120	± 16
A/C stereo layers	-2	± 12	123	± 11

Table 6.9: Situation after layer alignment with October 2009 TED data.

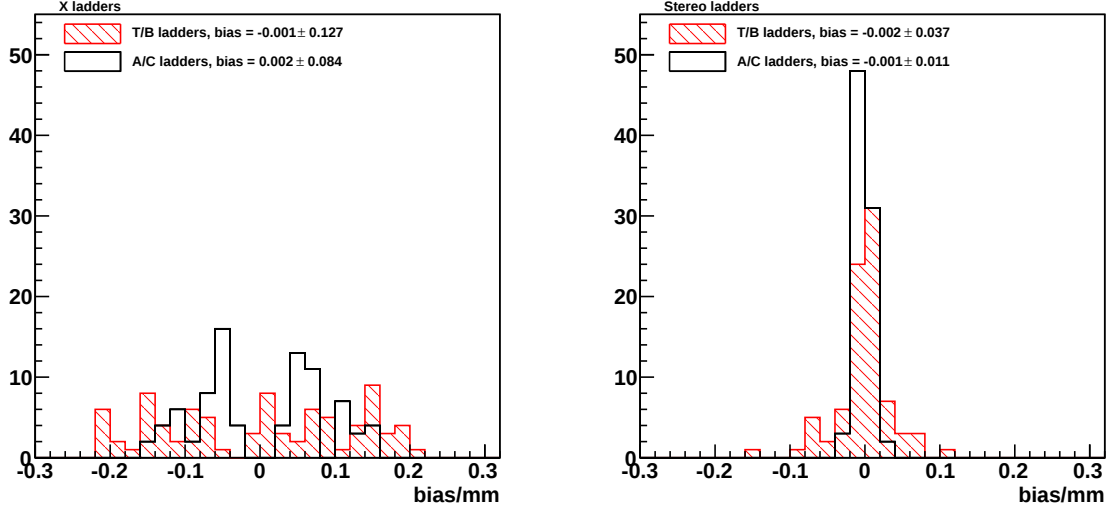


Figure 6.17: Ladder residual biases after box alignment

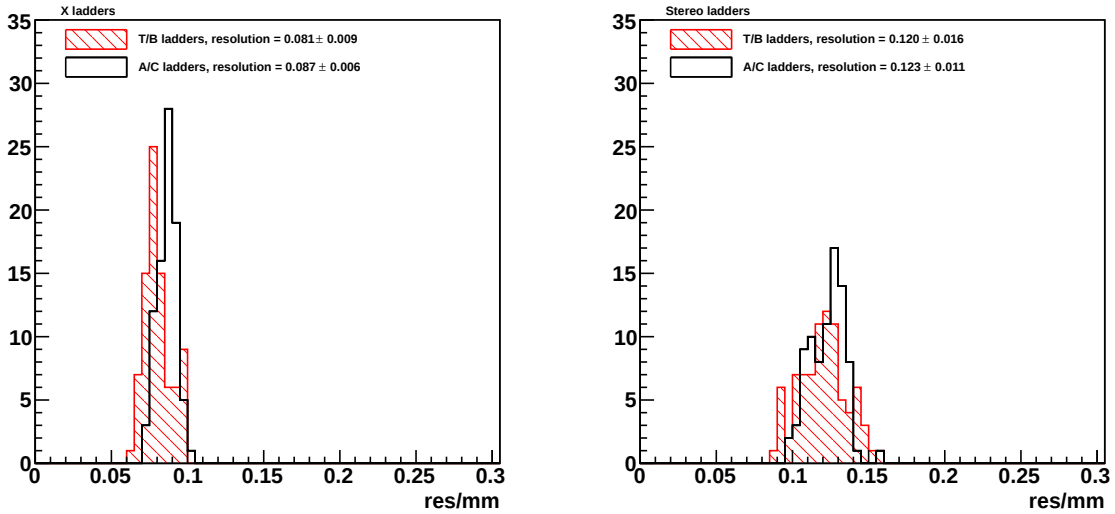


Figure 6.18: Ladder residual resolutions after box alignment

Inner Tracker Layer Alignment

Once the IT boxes have been aligned with October 2009 TED data, the layers are aligned. The starting point for the layer alignment is the set of alignment constants obtained from the boxes step. For the layers step, two degrees of freedom are aligned for : x translations and z -axis rotations. No alignment for the y translations is possible

due to the lack of tracking information along this direction. This lack of information along the y axis reduces the alignment sensitivity to the z -axis rotations, but not enough to destabilize the alignment procedure.

The weak modes present in the layers alignment are the same that are considered in the layer alignment on June 2009 TED data. Alignment convergence for the IT layers with October 2009 TED data is presented in Figures 6.19 and 6.20.

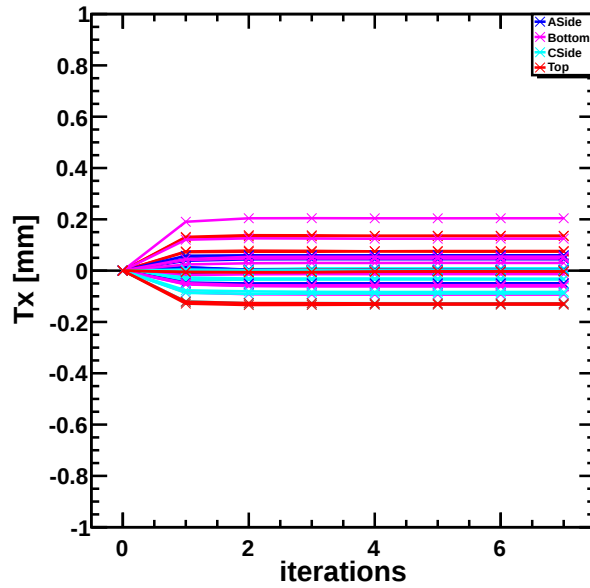


Figure 6.19: Evolution of the x translations of layers with iterations.

In Figure 6.14 the alignment parameters are stabilized in their evolutions after the second iterations for x translations. However, in Figure 6.15 the z -axis rotations need more iterations to stabilize due to the lower sensitivity to this degree of freedom. The performance of this layer alignment is measured by looking at the biases and resolutions. The results are presented in Figures 6.21, 6.22 and summarized in Table 6.10.

The alignment performance achieved using October TED data is similar to the one achieved with June TED data. One difference between these two alignments is that the IT ladders are not realigned on the October TED data. However, the results are similar in terms of resolution, which indicates that the ladders did not move significantly inside the boxes between the two TED data taking. This alignment also provided useful information about the actual size of the overlaps between the IT boxes, as overlap tracks were used to constrain the relative boxes position within the stations. The nominal overlaps

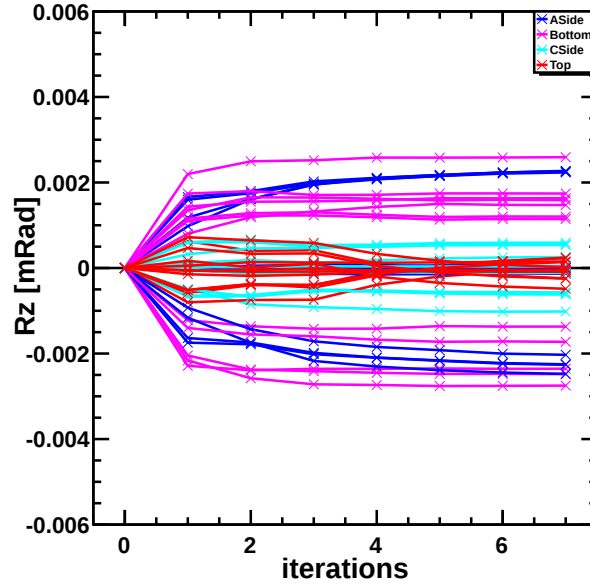


Figure 6.20: Evolution of the z -axis rotation of layers with iterations.

	Bias (μm)	Resolution (μm)
T/B X layers	0 $\pm$ 14	81 $\pm$ 7
A/C X layers	0 $\pm$ 11	87 $\pm$ 8
T/B stereo layers	0 $\pm$ 15	115 $\pm$ 12
A/C stereo layers	0 $\pm$ 12	126 $\pm$ 12

Table 6.10: Situation after layer alignment with October 2009 TED data.

size values are summarized in Table 6.7, while the values obtained after alignment on October TED data are summarized in Table 6.11. The deviations from the nominal values are quite large, with changes of $\mathcal{O}(5\text{mm})$ in some cases. In the third station the overlap between the top and C-side boxes is completely removed.

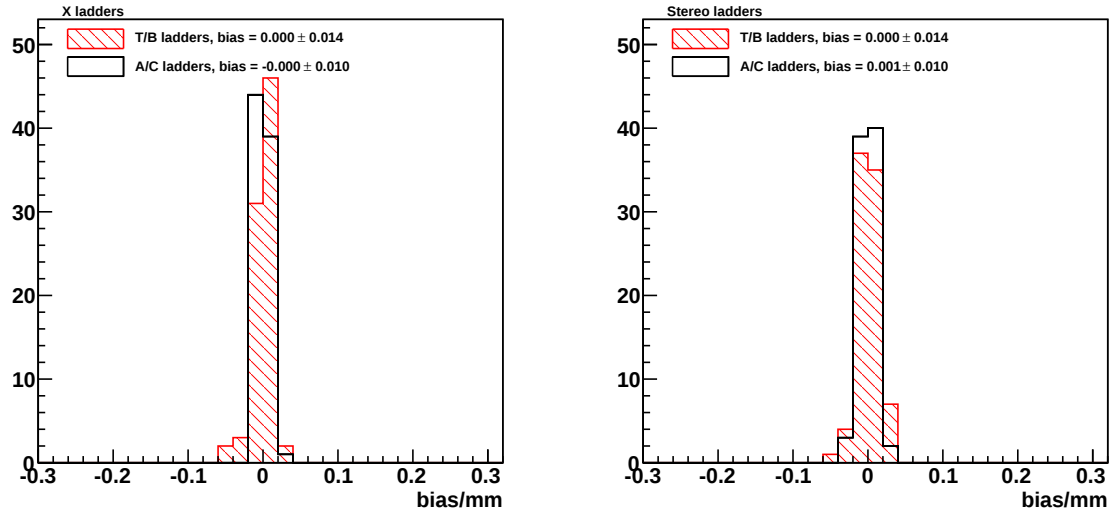


Figure 6.21: Ladder residual biases after layer alignment with October 2009 TED data.

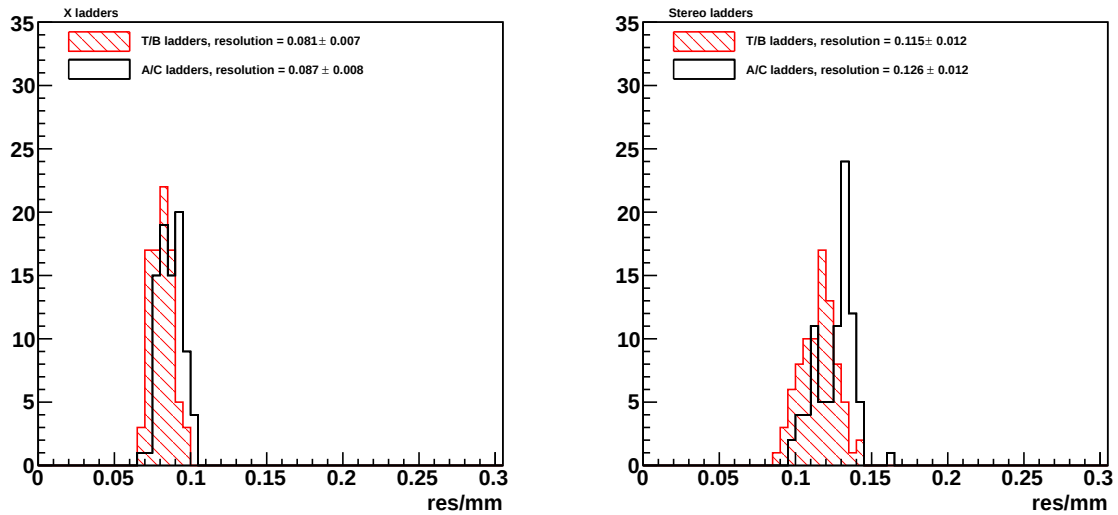


Figure 6.22: Ladder residual resolutions after layer alignment with October 2009 TED data.

IT Overlaps	$T \cap A$ [mm]	$T \cap C$ [mm]	$B \cap A$ [mm]	$B \cap C$ [mm]
Station 1	17.142	8.652	19.38	27.978
Station 2	12.252	9.012	14.088	17.328
Station 3	3.562	-2.578	8.568	17.708

Table 6.11: Measured overlap size in y for the IT sensors in X layers for each station, as determined from the alignment on october TED data.

6.3 Conclusion on Inner Tracker Alignment with TED Data

I have shown the alignment of the Inner Tracker using TED data and controlling the weak modes by removing the corresponding eigenvalues to be effective. The results obtained from the June and October data are compatible. In particular, the alignment on October data was only done for boxes and layers but not for ladders. Despite this, the ladder residual resolution is similar to the one achieved with June data, indicating that there has been no significant movement of ladders within the layers between the two series of TED shots.

The alignment with the October TED data has allowed to better determine the size of the overlaps between the IT sensors in neighbouring boxes. In particular it revealed that within the third IT station, where the overlaps are the smallest, some overlaps are non-existent or very small.

The final result is that the overall IT ladder residual resolution is $102 \pm 10 \mu\text{m}$, corresponding to a resolution due to alignment of $85 \pm 10 \mu\text{m}$, assuming an intrinsic resolution effect of $57 \mu\text{m}$. The bias is $0 \pm 13 \mu\text{m}$.

Chapter 7

Same Side Kaon Flavor Tagging Calibration with $D_s^+ \rightarrow \phi \pi^+$

7.1 Tagging Study Introduction

Flavour tagging is a set of methods used to determine the quantum state of the B^0 or B_s^0 and $\bar{B}^0$ or $\bar{B}_s^0$ mesons at their production. This information is essential for several **CP** violation measurements and also for the measurement of Δm_s in $B_s^0 \rightarrow D_s^- \pi^+$ [38]. A detailed description of LHCb tagging algorithms can be found in [39].

For any B decay a set of *taggers*, i.e. generic tagging methods, can be used. The so-called Opposite Side (OS) tagging exploits the presence of a muon, electron, kaon or secondary vertex in the event coming from the opposite (non-signal) B .

Conversely, the Same Side (SS) kaon tagger in the case of a B_s^0 (or pion tagger for a B^0 or B^+), exploits the presence of a particle from the fragmentation chain which is correlated in phase-space with the signal B . Details on OS and SS pion tagger optimization and calibration with 2010 data can be found in [40]. The Same Side Kaon (referred hereafter as SSK) tagger is of special interest in all analyses that need to tag the flavour of a B_s^0 meson at production. Two examples are the measurement of the B_s^0 mixing phase ϕ_s from $B_s^0 \rightarrow J/\psi \phi$ [41, 42], and the Δm_s measurement from $B_s^0 \rightarrow D_s^- \pi^+$ [38]. The optimizations and calibrations of the SS taggers are ideally performed using data with charged B mesons providing the true tag allowing a comparison with the tag provided by the considered tagger. Such decays used for calibration are called self-tagging modes.

With an integrated luminosity of 38 pb^{-1} collected in 2010 at $\sqrt{s}=7 \text{ TeV}$, the control channel of highest yield that permits a first tuning is $D_s^+ \rightarrow \phi\pi^+$, having 145k candidates compared to the 1.3k candidates of $B_s^0 \rightarrow D_s^- \pi^+$.

Both Monte Carlo (MC) and real data are used in order to motivate the tuning of the SSK tagger for B_s^0 decays using D_s data. The study includes the comparison of the fragmentation process in which the kaon is produced in both cases.

Once the tagger is optimized, the following step would be the calibration of the response of a Neural Net, originally trained on MC, that provides an estimate of the probability of the tag to be correct on an event-by-event basis.

This chapter is organised as follows. Section 7.2 explains how the SS kaons are selected. In Section 7.3, the control channels that might be used to evaluate the performance of this tagger are discussed, with a detailed comparison between $B_s^0 \rightarrow D_s^- \pi^+$ and $D_s^+ \rightarrow \phi\pi^+$ in Section 7.4. Section 7.5 presents the optimization of the SSK selection and Section 7.7 the calibration of the per-event mistag. Section 7.6 shows the preliminary performance in $B_s^0 \rightarrow D_s^- \pi^+$ after the optimization of the SSK tagging with $D_s^+ \rightarrow \phi\pi^+$ data. Finally, conclusions are given in Section 7.8.

7.2 Selection of Same Side Tagging Kaons

In the B_s^0 fragmentation process, a kaon or K^* might be produced, as depicted in Figure 7.1. The charge of this kaon (including kaons from K^* decay) tags the flavour of the B_s^0 produced. From reconstructed $B_s^0 \rightarrow D_s^- \pi^+$ MC events, 30% of events are found to have a fragmentation kaon. A common preselection of tagger candidates for both

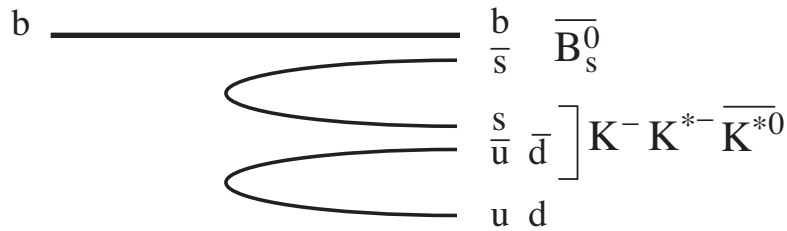


Figure 7.1: Illustration of the SS kaon production in the fragmentation of a b quark to a B_s^0 meson.

OS and SS is performed by requiring the tracks to be **Long** or **Upstream**¹, with no duplicates in the event; particles from the reconstructed B_s^0 decay are excluded and the cuts listed in Table 7.1 are applied. After this preselection, the average number of candidates per event is approximately 35. From those tagger candidates, the SSK selection

variable	value >
p	2 GeV
θ	12 mrad
$ \Delta\phi $	5 mrad
$IP/\sigma(PU)$	3

Table 7.1: Generic preselection cuts for taggers. $\Delta\phi$ refers to the difference in angle ϕ of the tagger candidate with respect to any daughter particle of the reconstructed B_s^0 meson. PU stands for PileUp vertex, any additional vertex in the event different from the primary vertex associated to the B_s^0 .

is performed by applying a set of cuts tuned for this particular tagger. These cuts are listed in Table 7.2. The cut values have been optimized using MC samples to maximize the effective tagging efficiency, defined as

$$\epsilon_{\text{eff}} = (1 - 2\omega)^2 \epsilon_{\text{tag}}, \quad (7.1)$$

where 7.1, ω , the so-called wrong tag fraction represent the fraction of wrongly tagged events, and ϵ_{tag} is the tagging efficiency or fraction of tagged events. The tuning of these tagger selection cuts from real data is presented in the following sections of this chapter.

On average, there are around 0.3 SSK candidates per event. In the current implementation of the algorithm, if there is more than one candidate in the event, the one with the highest p_T is retained. There is a clear dependence of ω on p_T , which will be discussed later.

The tagging performance on simulation data generated in 2010 (MC10) data with this selection is $\omega = 34.4 \pm 0.4 \%$, $\epsilon_{\text{tag}} = 20.30 \pm 0.17 \%$ and $\epsilon_{\text{eff}} = 1.98 \pm 0.10 \%$ without any trigger requirement. After applying L0 trigger conditions, the performance is improved: $\omega = 33.0 \pm 0.6 \%$, $\epsilon_{\text{tag}} = 23.17 \pm 0.25 \%$ and $\epsilon_{\text{eff}} = 2.68 \pm 0.19 \%$. This improvement is expected in hadronic modes due to the L0 hadron trigger, which

¹For a definition of the different types of tracks in LHCb, refer to Subsection 4.1.1.

Variable	Cut value	
	MC2010 ²	MC10 ³
$\Delta LL(K - \pi)$	> 1.0	> 3.0
$\Delta LL(K - p)$	> -1.0	> -2.0
p_T	$> 0.45 \text{ GeV}$	$> 0.55 \text{ GeV}$
p	$> 4 \text{ GeV}$	$> 2 \text{ GeV}$
Track χ^2/ndof	< 2	< 3
IP/ σ_{IP} (B PV ⁴)	< 3	< 3.5
$ \Delta\eta $	< 1.0	< 0.9
$ \Delta\phi $	< 1.1	< 1.0
$M(B + K) - M(B)$	$< 1.6 \text{ GeV}$	$< 2 \text{ GeV}$

Table 7.2: Selection cuts for SSK tagger optimized in MC 2010 ($\nu = 1$) and MC10 ($\nu = 2.5$), where ν is the number of interactions per crossing. $\Delta\phi$ and $\Delta\eta$ refer to the differences in angle ϕ and pseudo-rapidity η of the tagger candidate with respect to the reconstructed B_s^0 meson.

requires high E_T hadrons, and because of the correlation between the reconstructed B and the SSK tagger in term of kinematics.

Although 62% of events passing the SSK selection contain a fragmentation kaon, this fraction is reduced to 41% when considering events where the selected tagging particle is the actual fragmentation kaon.

Table 7.5 shows the detailed origin of the SSK tagging particle after selection. Note that in the case of a kaon coming from a K^* , the charge of the kaon is the correct one, thus providing a correct tag, while in the case of a kaon from a ϕ , the kaon selected can have the wrong charge.

In summary, around 38% of SSK are good fragmentation kaons, which means selected tagger kaon leading to the right tag. In the 62% of remaining cases, the selected tagger does not yield to a proper tag, but has instead a 50% chance of being wrong. In these cases, the selected tagger either comes from the PV (31%), as a decay product of a resonance (28%) or is a ghost (3%). To increase the tagger purity and to reduce the wrong tag fraction, cuts on the fragmentation kaons selection of Table 7.2 have to be optimized.

7.3 Same Side Kaon Control Channels

As already introduced, the optimization of the SSK selection, the calibration and measurement of the performance need to be performed using real data to avoid relying on MC predictions.

The main tagging control channels, where the tag of the B_s^0 meson is automatically known from the charge of the particles identified in the final state, are the semileptonic $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ decay and the $B_s^0 \rightarrow D_s^- \pi^+$ decay.

The $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ mode would be a good candidate to measure the ω of the SS kaon tagger, as its annual yield is estimated to be approximately 0.5 million events for an integrated luminosity of 1 fb^{-1} . However, due to the presence of the invisible neutrino, it is not straightforward to extract the mistag from the B_s^0 oscillations, as the reconstruction of the decay requires a good proper time resolution. Therefore, a *Double Tagging* method [43] is envisaged to obtain a global ω_{SameSide} without requiring the observation of the B_s^0 oscillations. However, with around 30k events from 2010 data, the sensitivity in ω is not good enough to be used in any mixing analysis. The number of expected $B_s^0 \rightarrow D_s^- \pi^+$ events is much lower, approximately 1.3k events. In order to perform the optimization of SSK with 2010 data, we require an alternative channel that provides a high yield. This channel is the $D_s^+ \rightarrow \phi \pi^+$ decay.

In this chapter, one concentrates on the use of $D_s^+ \rightarrow \phi \pi^+$ events from 2010 collision data in addition to MC for both $D_s^+ \rightarrow \phi \pi^+$ and $B_s^0 \rightarrow D_s^- \pi^+$ to assess the feasibility of the optimization with this channel. The Monte Carlo data was modified by a simple primary vertex smearing to better reproduce the features of real data. Studies with a further improved smeared MC to reproduce the characteristics of real data even better can be found in [44].

7.4 $D_s^+ \rightarrow \phi \pi^+$ as Control Channel for the SSK Tagger

$D_s^+ \rightarrow \phi \pi^+$ as a control channel presents the advantages, that : it is self-tagging and has large available statistics. The D_s meson, however, has a much lighter mass than the B_s^0 meson. This translates into different kinematics for the decay. It will be shown, however, that despite this difference the relevant tagging variables for particle track and the fragmentation process producing the tagging kaons tracks are very similar, leading

to the possibility of optimizing the tagging cuts using $D_s^+ \rightarrow \phi\pi^+$.

7.4.1 $D_s^+ \rightarrow \phi\pi^+$ Selection

In this study the $D_s^+ \rightarrow \phi\pi^+$ selection is based on the $B_s^0 \rightarrow D_s^- \pi^+$ selection used for the Δm_s measurement [38], i.e. the cuts used for the $D_s^+ \rightarrow \phi\pi^+$ selection are as similar as possible to the ones used in the $B_s^0 \rightarrow D_s^- \pi^+$ selection. For the MC10 simulation data the cuts are listed in Table 7.3.

Particle	Cut	Value
K from $\phi(1020)$	p_T	$> 300 \text{ MeV}$
	p	$> 2.0 \text{ GeV}$
	Track χ^2	< 5
	$\Delta LL(K - \pi)$	> -10
	Min IP- χ^2	> 9
	$M(K^+ + K^-)$	$< 1050 \text{ MeV}$
$\pi^\pm$ from $D_s^\pm$	p_T	$> 300 \text{ MeV}$
	p	$> 2.0 \text{ GeV}$
	Track χ^2	< 5
	$\Delta LL(K - \pi)$	< 10
	Min IP- χ^2	> 9
$D_s^\pm$	p_T	$> 2.0 \text{ GeV}$
	Vertex χ^2	< 12
	$ M(D_s^\pm) - M(D_{s,PDG}^\pm) $	$< 50 \text{ MeV}$
	$\cos \theta$ (dir. angle)	> 0.9999
	χ^2 separation from PV	> 64
$D_s^\pm$ Daughters	Max. Dist. of closest approach	$< 1.5 \text{ mm}$

Table 7.3: Selection cuts for $D_s^+ \rightarrow \phi\pi^+$ on MC10 Simulation Data.

When running on real data, the pre-selection line for prompt Charm is used, this is reconstructed with **Brunel** and stripped (pre-selection) with **DaVinci**. The selection cuts listed in Table 7.4 are then applied.

Particle	Cut	Value
K from $\phi(1020)$	p_T	$> 300 \text{ MeV}$
	p	$> 2.0 \text{ GeV}$
	Track χ^2	< 5
	$\Delta LL(K - \pi)$	> -10
	Min IP- χ^2	> 9
$\pi^\pm$ from $D_s^\pm$	p_T	$> 300 \text{ MeV}$
	p	$> 2.0 \text{ GeV}$
	Track χ^2	< 5
	$\Delta LL(K - \pi)$	< 10
	Min IP- χ^2	> 9
$D_s^\pm$	p_T	$> 2.0 \text{ GeV}$
	Vtx χ^2	< 12
	$ M(D_s^\pm) - M(D_{s,PDG}^\pm) $	$< 50 \text{ MeV}$
	$\cos \theta$ (dir. angle)	> 0.9999
	χ^2 separation from PV	> 100

Table 7.4: Selection cuts for $D_s^+ \rightarrow \phi \pi^+$ on real data.

7.4.2 Comparison of $D_s^+ \rightarrow \phi\pi^+$ vs $B_s^0 \rightarrow D_s^- \pi^+$ using MC10 Data

For the purposes of this study, we compare both the fragmentation and the tagging parameters of the $D_s^+ \rightarrow \phi\pi^+$ and $B_s^0 \rightarrow D_s^- \pi^+$ decay channels using simulated events. Both $D_s^+ \rightarrow \phi\pi^+$ and $B_s^0 \rightarrow D_s^- \pi^+$ MC10 data (mean $\nu = 2.5$ interactions per bunch crossing) are used for this study. The relevant fragmentation information is accessible from the MC-truth record. The SSK selection cuts presented in Table 7.2 are used.

Fractions	$B_s \rightarrow D_s \pi$	$D_s \rightarrow \phi\pi$
From Fragmentation	$41.02 \pm 0.34\%$	$44.59 \pm 0.42\%$
From String ⁵	$25.97 \pm 0.24\%$	$27.25 \pm 0.38\%$
From K^{*0} or higher ex ⁶ .	$8.19 \pm 0.15\%$	$8.30 \pm 0.24\%$
From K^{*+} or higher ex.	$4.16 \pm 0.11\%$	$4.45 \pm 0.18\%$
From $\phi(1020)$	$1.72 \pm 0.07\%$	$2.09 \pm 0.12\%$
From other	$0.99 \pm 0.05\%$	$2.49 \pm 0.13\%$
Not from fragmentation	$58.98 \pm 0.27\%$	$55.41 \pm 0.42\%$
K from PV	$22.23 \pm 0.23\%$	$18.89 \pm 0.33\%$
π from PV	$2.89 \pm 0.09\%$	$2.62 \pm 0.14\%$
p from PV	$5.50 \pm 0.12\%$	$5.45 \pm 0.19\%$
e from PV	$0.0 \pm 0.0\%$	$0.0 \pm 0.0\%$
K from X	$13.98 \pm 0.19\%$	$14.45 \pm 0.30\%$
π from X	$6.44 \pm 0.13\%$	$6.15 \pm 0.20\%$
p from X	$3.94 \pm 0.11\%$	$3.76 \pm 0.16\%$
e from X	$0.46 \pm 0.04\%$	$0.57 \pm 0.06\%$
Ghosts	$3.00 \pm 0.09\%$	$2.86 \pm 0.14\%$
'Good' kaon fraction	$38.32 \pm 0.27\%$	$40.01 \pm 0.51\%$

Table 7.5: Fraction of events corresponding to each different origin of the selected Same Side Kaon tagger. From MC10 $B_s^0 \rightarrow D_s^- \pi^+$ and $D_s^+ \rightarrow \phi\pi^+$ events. Tagging particles produced from resonances not directly related to the $b\bar{b}$ fragmentation.

Table 7.5 shows the origins of the selected SSK tagging candidate for both $B_s^0 \rightarrow D_s^- \pi^+$ and $D_s^+ \rightarrow \phi\pi^+$ MC10 data after selection. The fractions are broadly similar for the two decays in most categories, however there are two exceptions (red/light grey text):

prompt kaons⁷ and kaons from resonances in the fragmentation string which are not excited kaons or $\phi(1020)$ mesons.

One also has to divide the selected SSK tagger candidates into “good” and “bad” categories if we consider whether the candidate would lead to the correct tag decision. Only candidates from excited kaons in the fragmentation string or those produced within the fragmentation string itself fall into the “good” category, with a fraction of $38.32 \pm 0.27\%$ for $B_s^0 \rightarrow D_s^- \pi^+$ and $40.01 \pm 0.51\%$ for $D_s^+ \rightarrow \phi\pi^+$. All other candidates are either mis-identified particles, ghosts, or are from particles which produce two oppositely-charged kaons.

One also has to compare the distribution of the variables used for the tagging kaon selection in each dataset. In order to be able to optimise the SSK tagger for $B_s^0 \rightarrow D_s^- \pi^+$ decays using $D_s^+ \rightarrow \phi\pi^+$ data, it must be shown that the optimised tagging cuts will be similar for both the $D_s^+ \rightarrow \phi\pi^+$ and $B_s^0 \rightarrow D_s^- \pi^+$ channels. The selection variable distributions for both channels are shown in Figures 7.2 to 7.6.

Most of the variable distributions for $D_s^+ \rightarrow \phi\pi^+$ and $B_s^0 \rightarrow D_s^- \pi^+$ MC10 data are compatible within statistical errors. Exceptions to this are the p_T , $|\Delta\eta|$ and $|\Delta\phi|$ distributions where there are significant differences. From these observations, it can be expected that the cut values which optimize the tagging power for $D_s^+ \rightarrow \phi\pi^+$ and $B_s^0 \rightarrow D_s^- \pi^+$ should be identical for all variables but p_T , $|\Delta\eta|$ and $|\Delta\phi|$. For those later three variables, the optimisation is done on $B_s^0 \rightarrow D_s^- \pi^+$.

This hypothesis has to be put in a suitable form in order to be properly tested. The tagging power is given by :

$$\epsilon_{\text{eff}}(z) = \epsilon_{\text{tag}}(z) (1 - 2\omega(z))^2 \quad (7.2)$$

where z is the variable cut value and ϵ_{tag} is the tagging efficiency. If $\omega_{D_s}(z) = \omega_{B_s}(z) = \omega(z)$ and $\epsilon_{\text{tag}, D_s}(z) = \alpha \epsilon_{\text{tag}, B_s}(z) = \alpha \epsilon(z)$, with α a scaling factor. Then at the maximum of $\epsilon_{\text{eff}, B_s}$ we have:

$$\frac{d}{dz} \epsilon_{\text{eff}, B_s}(z) = \epsilon'(1 - 2\omega)^2 + \epsilon(-4\omega' + 8\omega'\omega) = 0. \quad (7.3)$$

⁷Kaons from the PV.

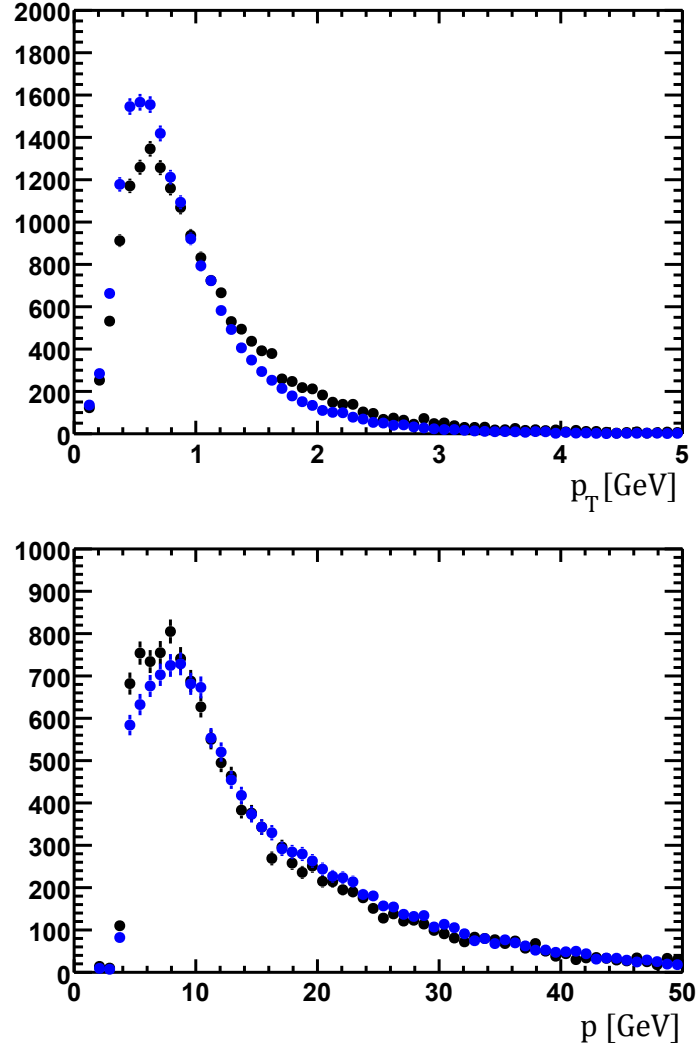


Figure 7.2: Distribution of p_T and p for D_s MC10 (black points) and B_s^0 MC10 (blue points).

Similarly, at the maximum of $\epsilon_{\text{eff}, D_s}$ we have:

$$\frac{d}{dz} \epsilon_{\text{eff}, D_s}(z) = \alpha \epsilon' (1 - 2\omega)^2 + \alpha \epsilon (-4\omega' + 8\omega' \omega). \quad (7.4)$$

From this it is clear that

$$\frac{d}{dz} \epsilon_{\text{eff}, D_s}(z) = \alpha \frac{d}{dz} \epsilon_{\text{eff}, B_s}(z), \quad (7.5)$$

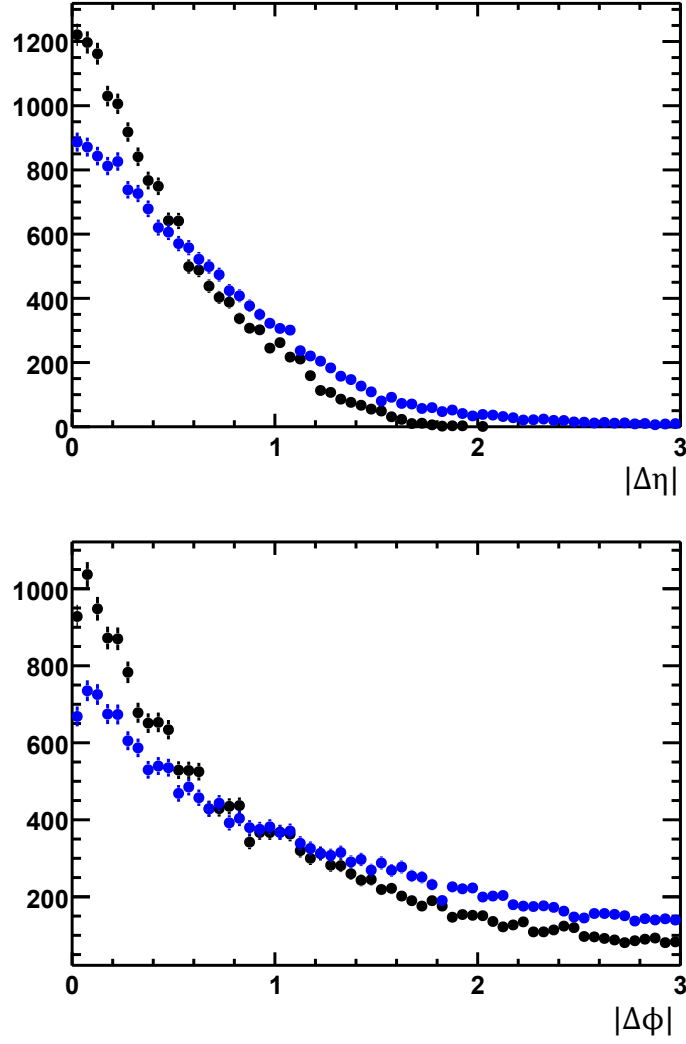


Figure 7.3: Distribution of $|\Delta\eta|$ and $|\Delta\phi|$ for D_s MC10 (black points) and B_s^0 MC10 (blue points).

which implies that if

$$\frac{d}{dz} \epsilon_{\text{eff}, B_s}(z) = 0, \quad (7.6)$$

then

$$\frac{d}{dz} \epsilon_{\text{eff}, D_s}(z) = 0. \quad (7.7)$$

Therefore, if $\epsilon_{\text{tag}}(z)$ and $\omega(z)$ follow the assumed hypothesis, then the maximum of $\epsilon_{\text{eff}}(z)$ will be at the same z for both D_s and B_s^0 .

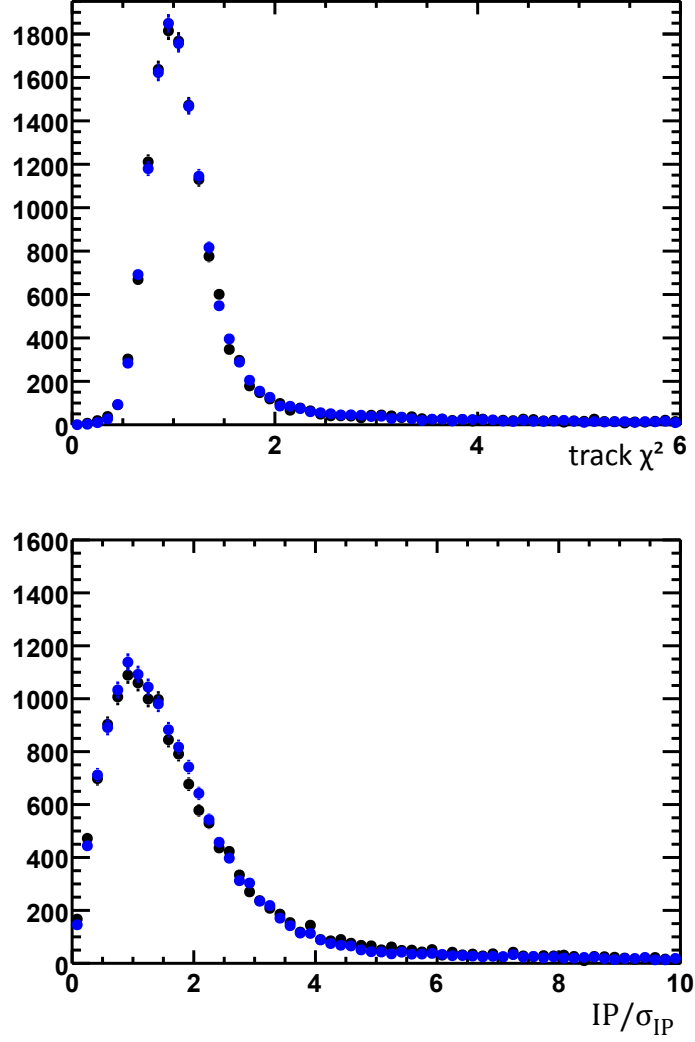


Figure 7.4: Distribution of χ^2 and IP/σ_{IP} for D_s MC10 (black points) and B_s^0 MC10 (blue points).

Figures 7.7 to 7.9 show the mistag rate $\omega(z)$ and the tagging efficiency $\epsilon(z)$ for several variables for both $D_s^+ \rightarrow \phi \pi^+$ and $B_s^0 \rightarrow D_s^- \pi^+$ MC10 data. In each efficiency plot, the $\epsilon_{\text{tag}, D_s}$ has been scaled by a factor $\alpha = 0.75$ obtained from a fit in order to make visible the point at which the relation $\epsilon_{\text{tag}, D_s}(z) = \alpha \epsilon_{\text{tag}, B_s}(z)$ is satisfied.

As visible in Figures 7.8 and 7.9, tagging variables with similar distributions have $\omega(z)$ and $\epsilon_{\text{tag}}(z)$ compatible with the hypothesis. For this reason, the maximum of the tagging power function will be the same for all of these variables. For the variables having $\omega(z)$ and $\epsilon_{\text{tag}}(z)$ incompatible with the hypothesis, (p_T , $|\Delta\eta|$ and $|\Delta\phi|$), as in Figure 7.7, the maximum of the tagging power function will be at different values of z

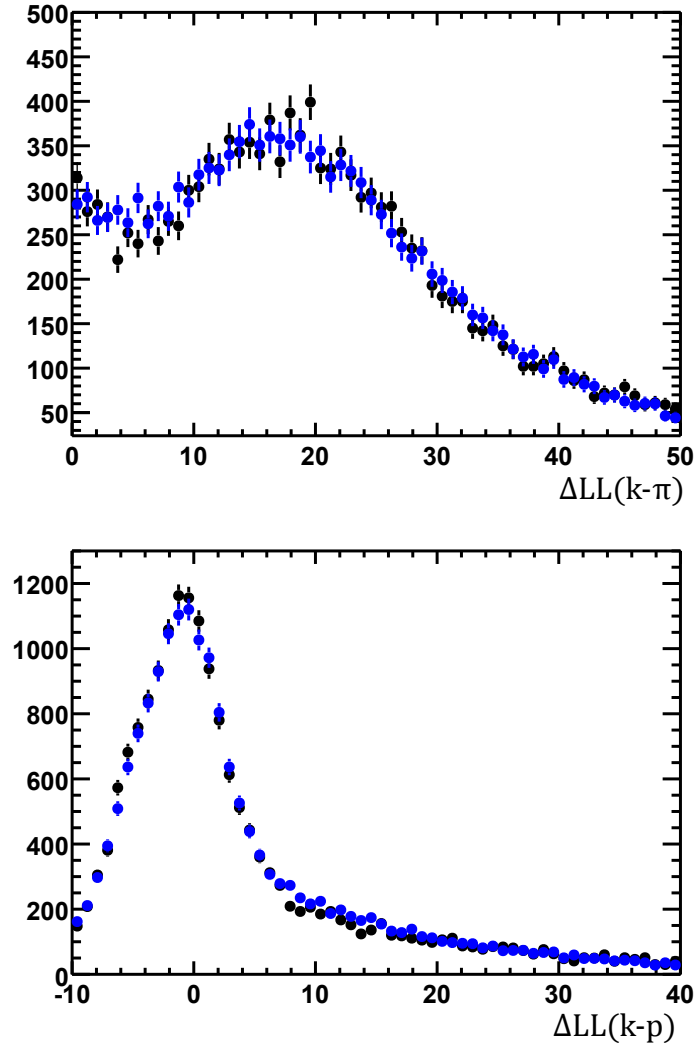


Figure 7.5: Distribution of $\Delta LL(K - \pi)$ and $\Delta LL(K - p)$ for D_s MC10 (black points) and B_s^0 MC10 (blue points).

for $D_s^+ \rightarrow \phi \pi^+$ and $B_s^0 \rightarrow D_s^- \pi^+$. Using $D_s^+ \rightarrow \phi \pi^+$ to optimize the $B_s^0 \rightarrow D_s^- \pi^+$ cuts will therefore be possible on the tagging variables except p_T , $|\Delta \eta|$ and $|\Delta \phi|$.

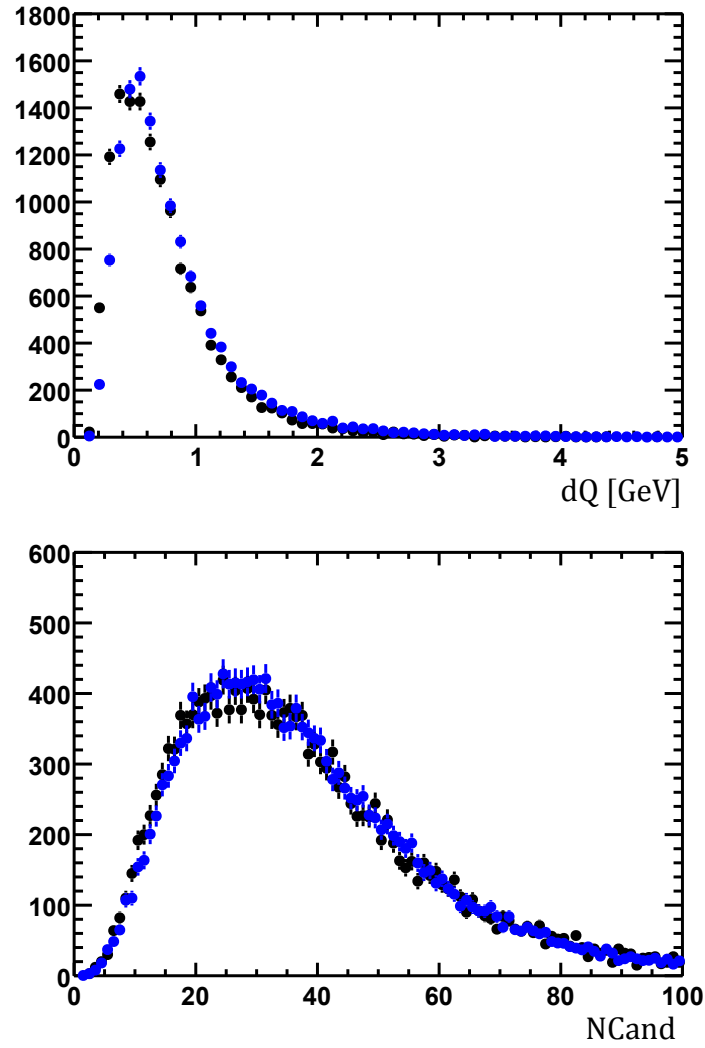


Figure 7.6: Distribution of $M(B + K) - M(B)$ (dQ) and $NCand$, the number of tagging candidates, for D_s MC10 (black points) and B_s^0 MC10 (blue points).

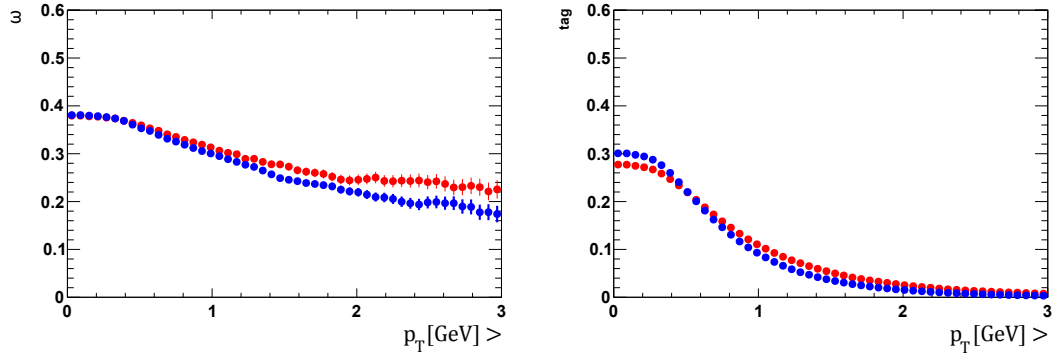


Figure 7.7: Hypothesis check on p_T with MC10 for D_s (red) and B_s^0 (blue), $\alpha = 0.75$ for ϵ_{tag, D_s} .

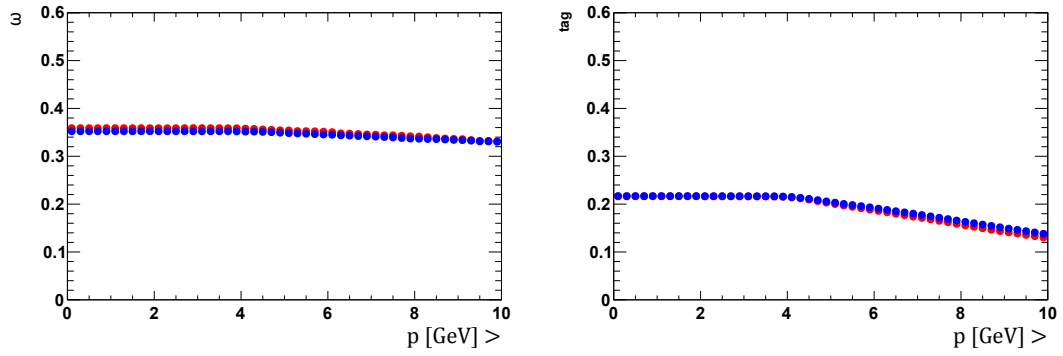


Figure 7.8: Hypothesis check on p with MC10 for D_s (red) and B_s^0 (blue), $\alpha = 0.75$ for ϵ_{tag, D_s} .

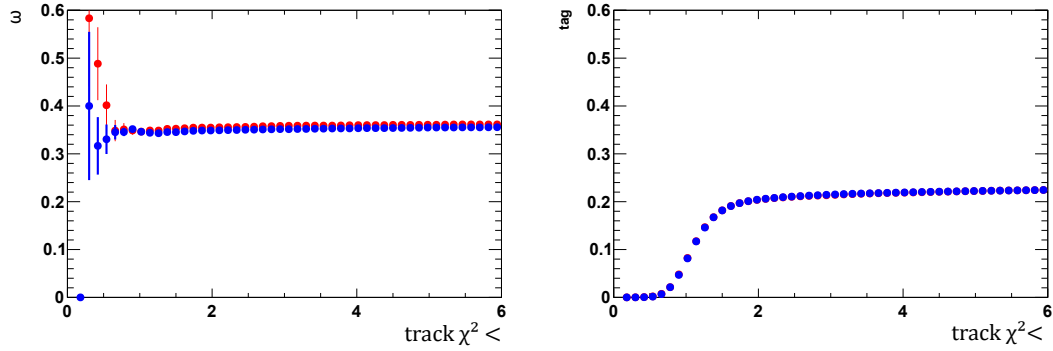


Figure 7.9: Hypothesis check on tagger track χ^2 with MC10 for D_s (red) and B_s^0 (blue), $\alpha = 0.75$ for ϵ_{tag, D_s} .

7.4.3 Comparison of $D_s^+ \rightarrow \phi\pi^+$ MC10 vs $D_s^+ \rightarrow \phi\pi^+$ Data

As for $D_s^+ \rightarrow \phi\pi^+$ a large statistics of real data is available, it is possible to make a comparison with MC10 simulation data. In this section, the tagging variables are compared, allowing to see the difference between MC10 simulation data and real data. The real data set contains 145k $D_s^+ \rightarrow \phi\pi^+$ candidates after selection shown in Table 7.4.

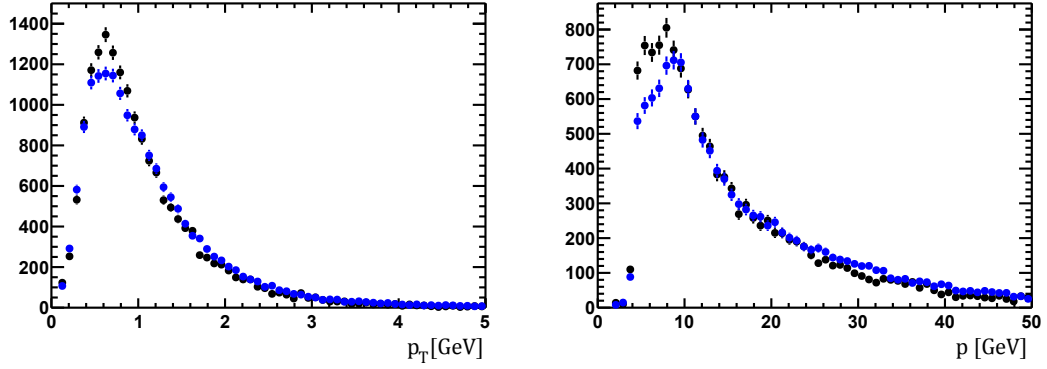


Figure 7.10: Distribution of p_T and p for D_s data (blue points) and MC10 (black points).

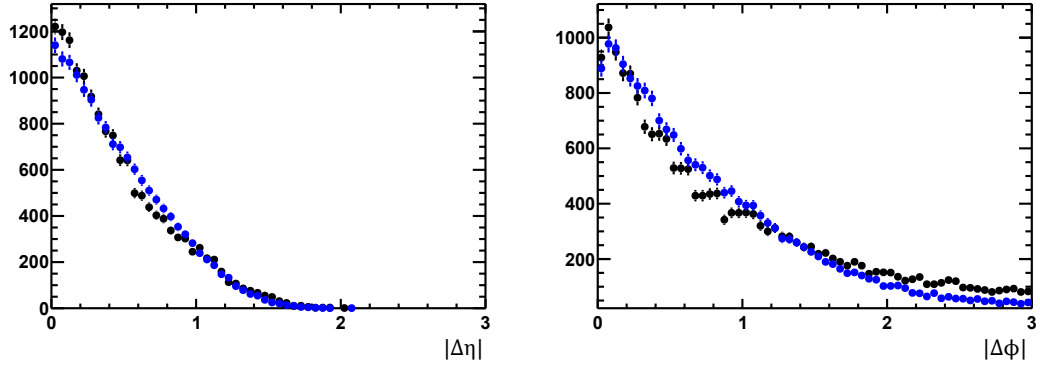


Figure 7.11: Distribution of $|\Delta\eta|$ and $|\Delta\phi|$ for D_s data (blue points) and MC10 (black points).

While the $|\Delta\eta|$ and $|\Delta\phi|$ distributions are similar for $D_s^+ \rightarrow \phi\pi^+$ data and MC10, the remaining distributions present slight differences. Because of these differences, it can be expected that the optimization of the cuts on the tagging variables for data will produce a different set of cuts with respect the one obtained from MC10 $D_s^+ \rightarrow \phi\pi^+$ simulation data.

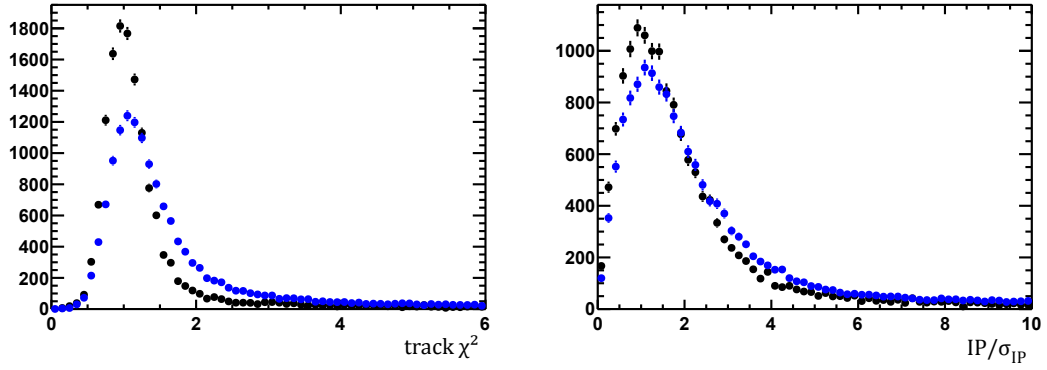


Figure 7.12: Distribution of χ^2 and IP/σ for D_s data (blue points) and MC10 (black points).

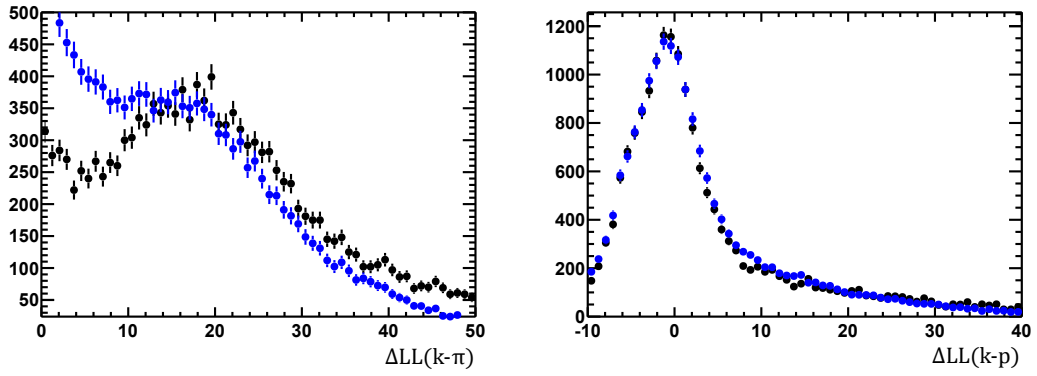


Figure 7.13: Distribution of $\Delta LL(K-\pi)$ and $\Delta LL(K-p)$ for D_s data (blue points) and MC10 (black points).

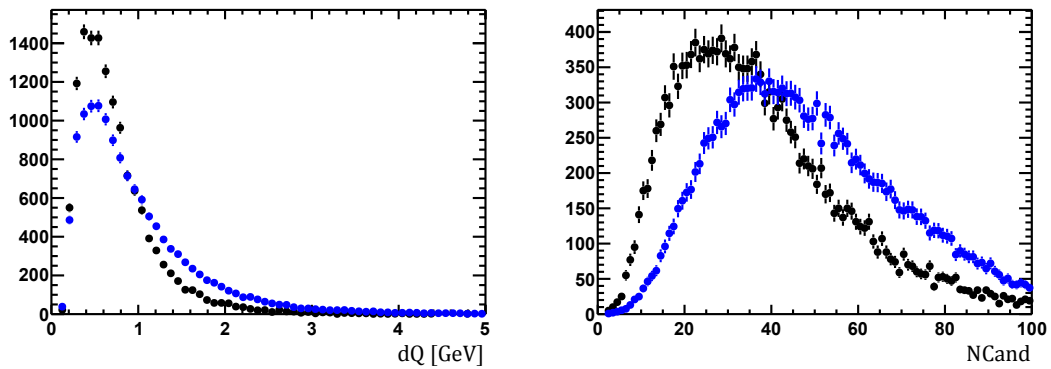


Figure 7.14: Distribution of $M(B+K) - M(B)$ and $NCand$, the number of tagging candidates for D_s data (blue points) and MC10 (black points).

7.5 Optimization of the SSK Selection Cuts

The procedure used here to optimize the selection cuts of SS kaons consists of a discrete scanning of each cut variable, keeping the others fixed, and finding the value of the variable that maximizes the effective tagging efficiency (tagging power), ϵ_{eff} . Figure 7.15 shows two examples of ϵ_{eff} as function of the value of the cut on the SSK p_T and $|\Delta\phi|$.

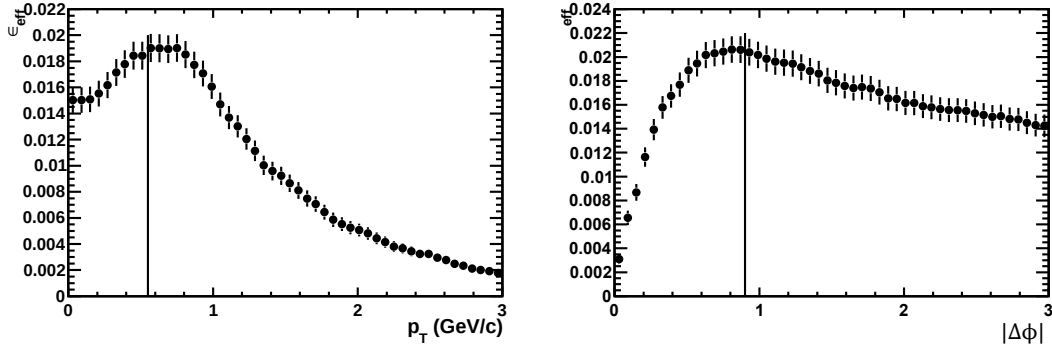


Figure 7.15: Tagging effective efficiency as a function of the cut in SSK p_T , requiring the p_T to be above the cut value (left) and $|\Delta\phi|$, requiring this variable to be below the cut value (right).

The procedure is performed for each of the tagging variables, leading to an initial set of optimized cuts. The procedure is then repeated with the cuts of the previous iteration as the starting point until the set of cuts stabilizes. The optimization has been performed on $D_s^+ \rightarrow \phi\pi^+$ data, $D_s^+ \rightarrow \phi\pi^+$ MC10 and $B_s^0 \rightarrow D_s^- \pi^+$ MC10. The results are shown in Table 7.6. The tagging cut values after optimization are nearly identical between $D_s^+ \rightarrow \phi\pi^+$ MC10 and $B_s^0 \rightarrow D_s^- \pi^+$ MC10 with the exception of the $|\Delta\eta|$, $|\Delta\phi|$ and p_T cut values.

In the case of $D_s^+ \rightarrow \phi\pi^+$ optimization on data, a correction has been applied to remove the contribution from background events to the mistag and the efficiency. This is done because the background raises the mistag rate. The procedure for this is the following :

- One $\phi\pi^+$ invariant mass histogram is filled for all events.
- A second histogram is filled for correctly-tagged (correct) events.
- A third one is filled for wrongly-tagged (wrong) events.

On these mass plots, a linear (background) plus double gaussian (signal) fit is performed to extract the correct $D_s^+ \rightarrow \phi\pi^+$ yields as shown in Figure 7.16. These yields are used to extract the tagging efficiency and mistag rate. This allows the background corrected effective efficiency (tagging power), ϵ_{eff} to be computed. During the tagging optimization where the tagging cuts are scanned, the correction procedure is applied at each step of the scan.

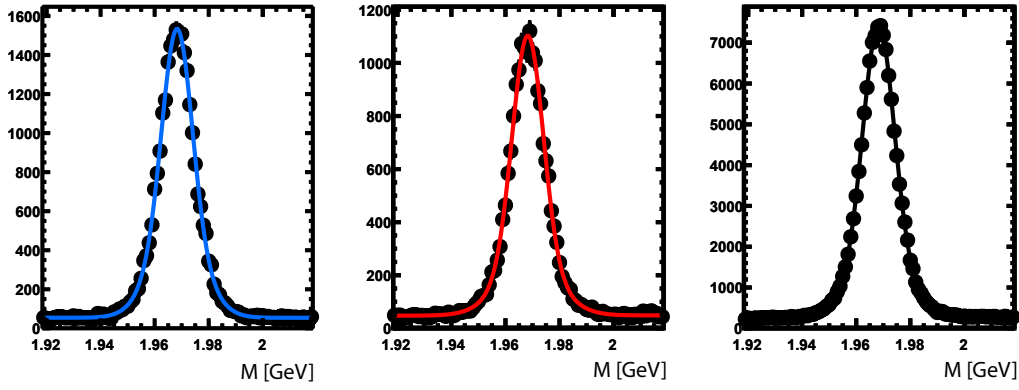


Figure 7.16: Left : $(\phi\pi^+)$ invariant mass for correctly tagged events. Center : $(\phi\pi^+)$ invariant mass for wrongly tagged events. Right : $(\phi\pi^+)$ invariant mass for all events.

The optimization results shown in Table 7.6 in the data D_s column includes the

Cuts	MC10 Default	MC10 B_s^0 Opt.	MC10 D_s Opt.	Data D_s Opt.
$p_T >$	0.55	0.50	0.55	0.55
$p >$	2.0	2.0	2.0	2.3
$ \Delta\eta <$	0.90	0.90	0.60	0.55
$ \Delta\phi <$	1.00	0.96	0.85	0.8
$IP/\sigma <$	3.5	4.0	3.8	3.5
$M(B + K) - M(B) <$	2.0	2.0	1.9	1.9
track χ^2/ndof	3.0	3.0	2.5	4.0
$\Delta LL(K - \pi) >$	3.0	4.0	4.0	3.2
$\Delta LL(K - p) >$	-2.0	-3.5	-3.2	-2.0

Table 7.6: Results of tagging cut optimization on $D_s^+ \rightarrow \phi\pi^+$ data and MC10 and $B_s^0 \rightarrow D_s^- \pi^+$ MC10.

	Mistag ω	Tagging Efficiency ϵ_{tag}	Tagging Power ϵ_{eff}
Cuts Optimized on MC10 B_s^0 data			
MC10 D_s	$36.49 \pm 0.39 \%$	$33.35 \pm 0.22 \%$	$2.44 \pm 0.14 \%$
MC10 B_s^0	$35.97 \pm 0.25 \%$	$25.35 \pm 0.11 \%$	$1.99 \pm 0.07 \%$
MC10 D_s , L0 req	$36.87 \pm 0.64 \%$	$36.66 \pm 0.39 \%$	$2.53 \pm 0.25 \%$
MC10 B_s^0 , L0 req	$35.36 \pm 0.33 \%$	$28.85 \pm 0.17 \%$	$2.47 \pm 0.11 \%$
Data D_s	$41.35 \pm 0.24 \%$	$35.75 \pm 0.14 \%$	$1.07 \pm 0.06 \%$
Cuts Optimized on MC10 D_s data			
MC10 D_s	$32.89 \pm 0.44 \%$	$24.14 \pm 0.20 \%$	$2.83 \pm 0.15 \%$
MC10 B_s^0	$33.28 \pm 0.30 \%$	$16.44 \pm 0.10 \%$	$1.84 \pm 0.07 \%$
MC10 D_s , L0 req	$33.71 \pm 0.73 \%$	$27.47 \pm 0.36 \%$	$2.92 \pm 0.26 \%$
MC10 B_s^0 , L0 req	$32.54 \pm 0.36 \%$	$19.44 \pm 0.15 \%$	$2.37 \pm 0.11 \%$
Data D_s	$38.75 \pm 0.28 \%$	$24.66 \pm 0.13 \%$	$1.25 \pm 0.06 \%$
Cuts Optimized on Real D_s data			
MC10 D_s	$32.31 \pm 0.48 \%$	$20.61 \pm 0.19 \%$	$2.58 \pm 0.14 \%$
MC10 B_s^0	$32.46 \pm 0.33 \%$	$13.76 \pm 0.09 \%$	$1.69 \pm 0.06 \%$
MC10 D_s , L0 req	$32.93 \pm 0.78 \%$	$23.91 \pm 0.34 \%$	$2.79 \pm 0.26 \%$
MC10 B_s^0 , L0 req	$31.38 \pm 0.43 \%$	$16.49 \pm 0.14 \%$	$2.29 \pm 0.11 \%$
Data D_s	$37.36 \pm 0.30 \%$	$21.73 \pm 0.12 \%$	$1.39 \pm 0.07 \%$

Table 7.7: Results of tagging cut optimization on $D_s^+ \rightarrow \phi\pi^+$ data and MC10 and $B_s^0 \rightarrow D_s^- \pi^+$ MC10. L0 req stands for requiring MC10 events to pass the L0 trigger.

background correction. As expected, these optimal cuts are different from the ones obtained from the $D_s^+ \rightarrow \phi\pi^+$ MC10 data, as the tagging variable distributions are not the same.

Given that $D_s^+ \rightarrow \phi\pi^+$ and $B_s^0 \rightarrow D_s^- \pi^+$ MC10 data share similar tagging variable distributions (with the exception of $|\Delta\eta|$, $|\Delta\phi|$ and p_T) and have similar optimization results, if the correspondence between the two is assumed to hold for tagging on real data, then the results of the cut optimization on $D_s^+ \rightarrow \phi\pi^+$ data can be used for the $B_s^0 \rightarrow D_s^- \pi^+$ tagging on data.

The tagging variables that are not suitable for this are $|\Delta\eta|$, $|\Delta\phi|$ and p_T . For $|\Delta\eta|$ and $|\Delta\phi|$ this is due to the difference in the kinematics between the D_s and B_s^0 decays,

since the first one contains a charm quark which is much lighter than the bottom quark in the B_s^0 . For the p_T variable, despite the different distributions in $B_s^0 \rightarrow D_s^- \pi^+$ and $D_s^+ \rightarrow \phi\pi^+$, we end up with nearly the same optimized value. This is probably accidental and not linked to any properties of the mistag or tagging efficiency as a function of this variable; for this reason the p_T variable should also be optimized on $B_s^0 \rightarrow D_s^- \pi^+$ real data. In summary, the cuts that will be applied for tagging B_s^0 in real data are the ones optimized on the D_s MC10 data, except for variables $|\Delta\eta|$, $|\Delta\phi|$ and p_T which are optimized on B_s^0 MC10 data.

The results obtained from applying the sets of cuts obtained on $B_s^0 \rightarrow D_s^- \pi^+$ MC10, $D_s^+ \rightarrow \phi\pi^+$ MC10 and $D_s^+ \rightarrow \phi\pi^+$ real data are summarized in Table 7.7. While the tagging power on MC10 simulation data is close to 2.5 % after L0 for both $B_s^0 \rightarrow D_s^- \pi^+$ and $D_s^+ \rightarrow \phi\pi^+$, it is much lower on real $D_s^+ \rightarrow \phi\pi^+$ data with a value of 1.4 %, even with cuts optimized on real data. This suggests that the tagging power on real data for the Same Side Kaon tagger is lower than on MC10 simulation data both in $B_s^0 \rightarrow D_s^- \pi^+$ and $D_s^+ \rightarrow \phi\pi^+$. The results obtained with real B_s^0 data are presented in the following Section 7.6.

7.6 Performance on $B_s^0 \rightarrow D_s^- \pi^+$ Data

After the optimization performed using $D_s^+ \rightarrow \phi\pi^+$ data, a preliminary estimation of the performance of the SSK tagger in the $B_s^0 \rightarrow D_s^- \pi^+$ decay mode can be made. For a flavour specific decay like $B_s^0 \rightarrow D_s^- \pi^+$, where no CP violation is present, the time dependent asymmetry has an amplitude which is the combination of the time resolution and the dilution factor $D = (1 - 2\omega)$ from which the mistag fraction can be extracted. The complete time dependant asymmetry expression is

$$A^{\text{mix}}(t) = \frac{N^{\text{unmixed}}(t) - N^{\text{mixed}}(t)}{N^{\text{unmixed}}(t) + N^{\text{mixed}}(t)} = (1 - 2\omega) \times e^{-\Delta m_s^2 \sigma_t^2 / 2} \times \cos(\Delta m_s t). \quad (7.8)$$

In Equation 7.8 the $e^{-\Delta m_s^2 \sigma_t^2 / 2}$ is the amplitude factor from the finite time resolution and $1 - 2\omega$ is the amplitude factor from tagging.

Figure 7.17 shows the oscillations seen in the B_s^0 system using only the SSK tagger. This is obtained from 1020 SSK tagged events from 2010 and 2011 data, where 586 ± 31 events correspond to B_s^0 signal events.

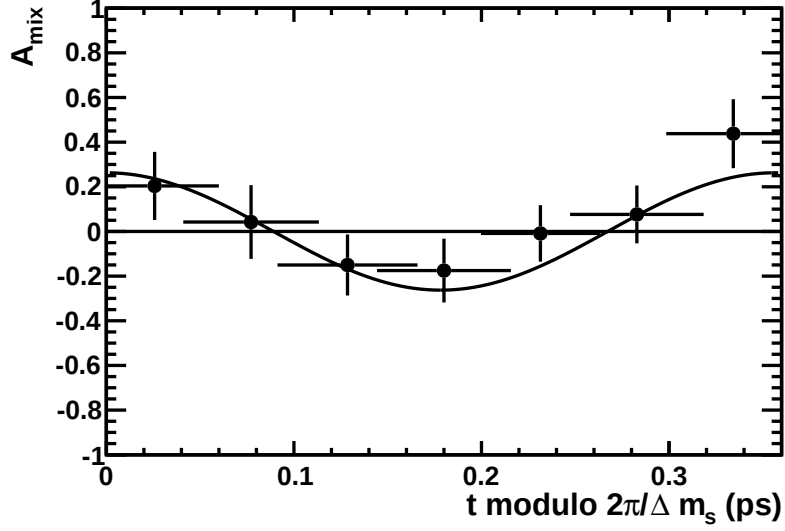


Figure 7.17: Mixing asymmetry in $B_s^0 \rightarrow D_s^- \pi^+$ as function of time modulo $2\pi/\Delta m_s$.

Background is taken into account using *sWeights* obtained from a mass fit. The amplitude of the asymmetry is higher with D_s optimized cuts⁸ than using the cuts optimized with B_s^0 MC which proves the validity of this work. The tagging efficiency for this sample is $16.0 \pm 0.6\%$ and the wrong tag fraction obtained from the asymmetry fit, $\omega = 33 \pm 5\%$. These values can be compared to the ones obtained with the set of cuts entirely optimized on B_s^0 MC10 data and used for the same time dependent asymmetry fit on a larger real data set with the same selection applied. In this latter case, the tagging efficiency is $27.3 \pm 0.3\%$ and the wrong tag fraction obtained is $\omega = 41 \pm 2\%$. This shows that using the cut from optimization on $D_s^+ \rightarrow \phi \pi^+$ 2010 data leads to an enhanced tagging power and a much reduced mistag rate, as presented in Table 7.8.

7.7 Calibration of the Mistag Estimate

The final decision of the flavour of the signal at production is taken according to the charge of the tagging particles from the different taggers available. To improve the global performance of the tagging, the sample is subdivided into a number of *tagging*

⁸Default cuts in the Flavour Tagging algorithms slightly differ from the ones presented in this chapter. Differences between D_s and B_s are taken into account by relaxing for instance the $\Delta\eta$ and $\Delta\phi$ cuts according to the B_s^0 MC optimization.

Tag parameters	D_s based cuts	B_s^0 MC10 based cuts
ϵ_{tag}	$16.0 \pm 0.6 \%$	$27.3 \pm 0.3 \%$
ω	$33 \pm 5 \%$	$41 \pm 2 \%$
ϵ_{eff}	$1.88 \pm 1.1 \%$	$0.87 \pm 0.04 \%$

Table 7.8: Results obtained with cuts optimized on D_s data and B_s^0 MC10.

categories. This splitting can be made following different criteria. One possibility is to sort the events based on the type of *tagger* (*PID approach*). Another possibility is to consider a number of kinematic variables for each tagger and the event itself and use them as the input to a Neural Net (NNet) that estimates a probability of the tagger decision to be correct. These probabilities are then combined and treated as independent (*Neural Net approach*).

The combined probability of OS taggers is calibrated to remove correlations among the Secondary Vertex Charge tagger and the other OS taggers [40, 45]. The Secondary Vertex Charge tagger is a tagger which collects all the particle around the OS B meson and measure the total charge to provide a tag. Afterward, the signal samples are sorted into 5 categories of decreasing mistag probability. This way to proceed gives the advantage of a higher effective tagging efficiency.

The quality of the tag may depend on the event multiplicity, the $B_s^0 p_T$ spectrum and properties of the tagger. Four examples of the dependence of the SSK mistag on some of these variables are shown in Figure 7.18.

In the case of the SSK, the input to the Neural Net, which is a Multi Layer Perceptron, are : the number of tagger candidates (proportional to the number of tracks in the event), p_T of the B, p of the SSK, p_T of SSK, IP/σ with respect to the PV, $|\Delta\phi|$, $|\Delta\eta|$, $M(B+K)-M(B)$, the number of PV and the number of SSK candidates.

The probability per tagger and per event needs to be calibrated from data. The NNet output, here called η , trained using B_s^0 MC data, must be translated into the mistag probability, ω . To first approximation the dependence of this probability on η has been parametrized as $\omega = p_0 + p_1 \cdot (\eta - \bar{\eta})$, where $\bar{\eta}$ is the average value. The parameters p_0 and p_1 are fitted in the control channels.

Left Fig. 7.19 shows the distribution of the per-event η and the right figure its relation with ω for SSK of a $B_s^0 \rightarrow D_s^- \pi^+$ MC10 sample. The linear fit applied in the latter plot

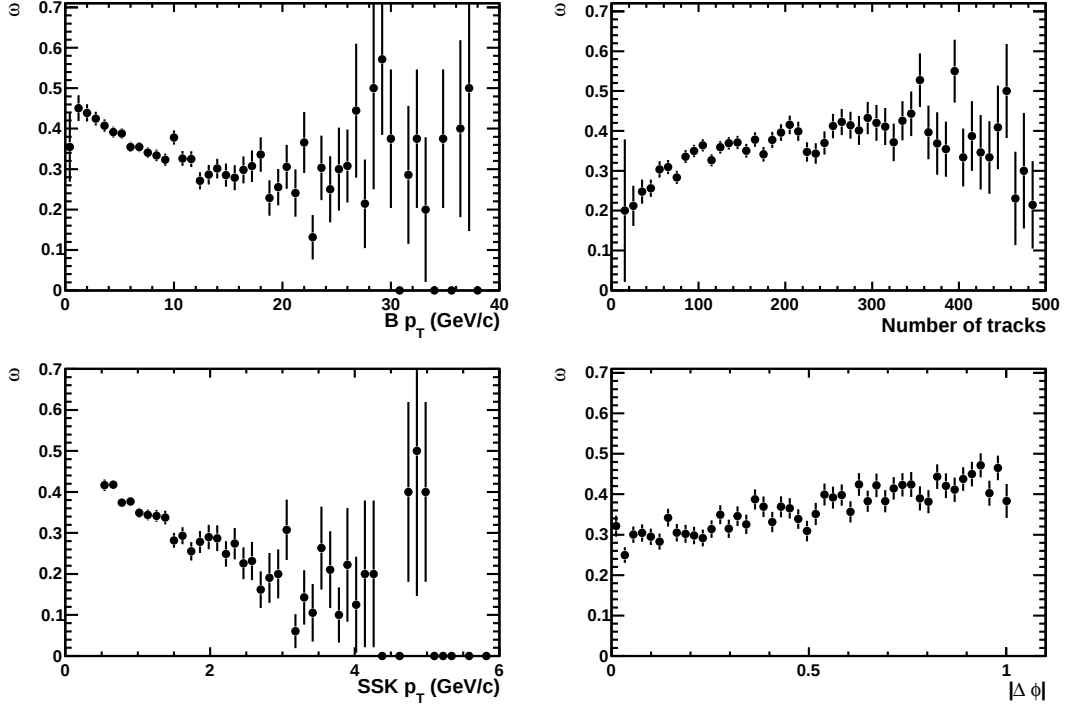


Figure 7.18: Wrong tag fraction of SSK as a function of: p_T of the reconstructed B, number of reconstructed tracks in the event, p_T of the kaon tagger and $|\Delta\phi|$ between the kaon and the B. From MC10 $B_s^0 \rightarrow D_s^- \pi^+$ events.

corresponds to the formula given in the paragraph above. The parameters p_0 and p_1 obtained are close to $\bar{\eta}$ but differ on the slope (1.0 is expected). The cause of this difference is that the MC data used in the training was 2010 MC data, while the fit has been done with the more recent MC10 data.

Using gathered informations, with the MC we can check the response of the NNet when the selected SSK is or is not a kaon from the B_s^0 fragmentation process. These distributions are given in Fig. 7.20. Note that high values of η correspond to high probabilities of wrong tagging decisions, so a cut on this probability to reject the worst candidates would benefit the final tagging performance. Such a cut should be applied on the calibrated ω .

Taggers with estimated probabilities of mistag above 0.5 should be rejected, as in a combination with OS taggers, this could bias the decision towards the opposite tag. In Figure 7.21 is shown ϵ_{eff} as a function of a cut in the mistag probability. As expected, a cut on this variable increases the tagging power. In order not to significantly reduce the tagging efficiency, ϵ_{tag} , a cut at in the 0.45 range is applied to avoid events with

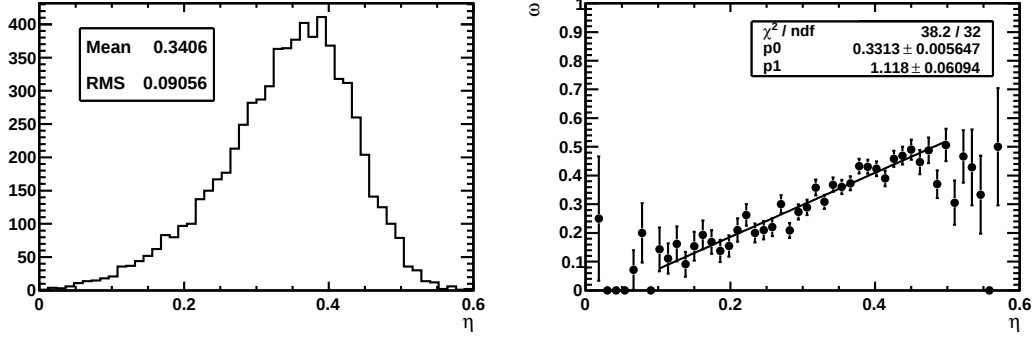


Figure 7.19: Distribution of the probability of the tag to be wrong, η , obtained from a Neural Net for MC10 SSK (left). True mistag fraction, ω , as a function of the estimated mistag, η , showing the results of a linear fit $\omega = p_0 + p_1 \cdot (\eta - \bar{\eta})$ (right)

mistag estimate close to the 0.5 limit. This choice is reasonable, since the impact on the tagging power of rejected events with $0.45 < \eta < 0.50$ is minimal.

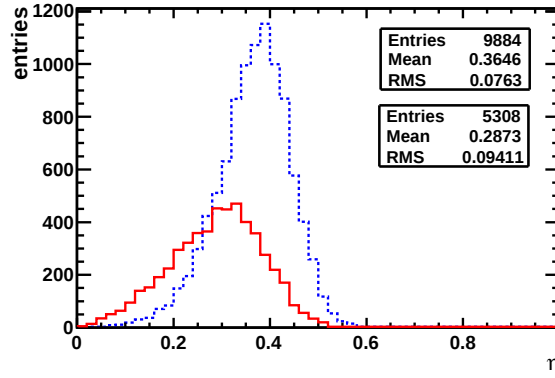


Figure 7.20: Distribution of η for true fragmentation kaons (red points) and non-fragmentation kaons (blue dashed points). From MC10 $B_s^0 \rightarrow D_s^- \pi^+$ SSK.

7.7.1 Using $D_s^+ \rightarrow \phi \pi^+$ with NNet for SSK Tagger Calibration

Given the opportunity to optimize the tagging cuts for $B_s^0 \rightarrow D_s^- \pi^+$ using $D_s^+ \rightarrow \phi \pi^+$, it is legitimate to investigate the NNet response with $D_s^+ \rightarrow \phi \pi^+$ data for the Same Side Kaon. If the response is similar, one may use the $D_s^+ \rightarrow \phi \pi^+$ real data to perform a calibration of the NNet instead of relying on $B_s^0 \rightarrow D_s^- \pi^+$ MC simulation data. It is clearly visible that the results of the fits in Fig. 7.22 barely overlap within errors, with

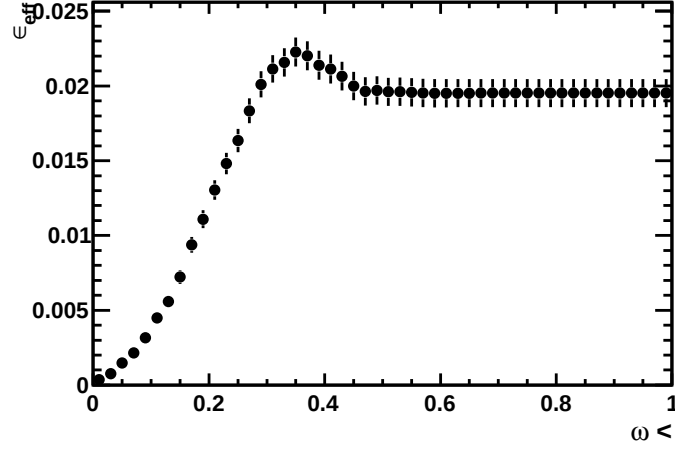


Figure 7.21: Effective efficiency as a function of the cut applied on the mistag probability for MC10 $B_s^0 \rightarrow D_s^- \pi^+$ SSK.

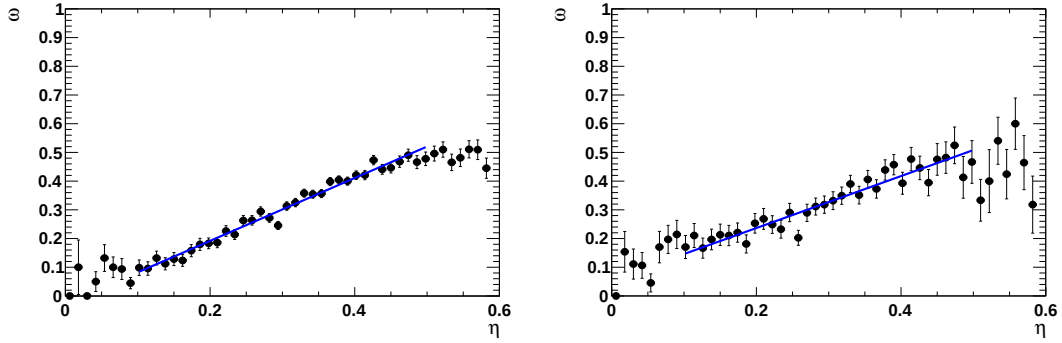


Figure 7.22: Calibration Fit for the NNet using MC10 B_s^0 (left) with a slope of 1.10 ± 0.03 and MC10 D_s (right) with a slope of 0.91 ± 0.06 .

a returned slope of 1.10 ± 0.03 for MC10 B_s^0 and 0.91 ± 0.06 for MC10 D_s . There are several elements explaining the difference. The first is that these plots use the estimation of the mistag given by the neural network. This network uses the same variables used for the tagging decision exception made of the p_T of the B_s^0 or D_s which is an additional variable. This additional variable is very different, as shown in Figure 7.23. Moreover some of the other variables are not compatible, in particular $|\Delta\eta|$, $|\Delta\phi|$ and the p_T of the kaon, as shown in Subsection 7.4.2.

The second element is the fact that the correlations in pairs of variables are different for MC10 B_s^0 and MC10 D_s . An example of this might be observed in Figure 7.24. As neural networks are very sensitive to correlations between variables, differences in

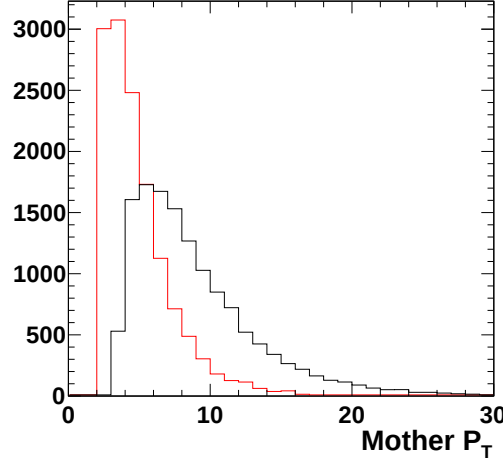


Figure 7.23: p_T of the B_s^0 meson (black) and D_s meson (red).

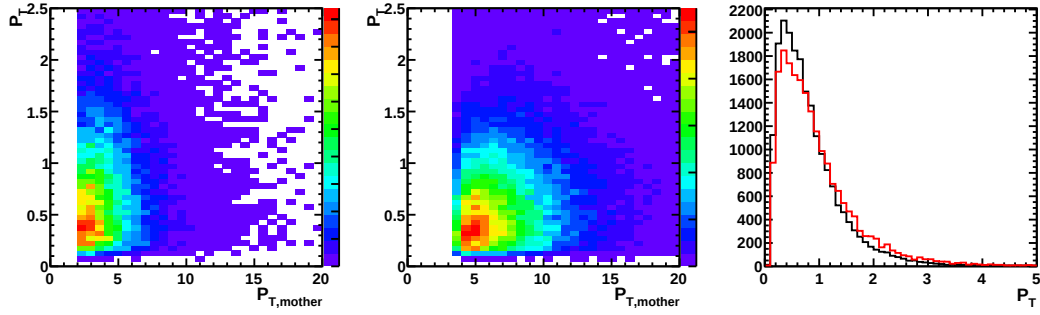


Figure 7.24: p_T of the kaon in MC10 B_s^0 (black) and D_s (red) versus coproduced mother p_T (right plot). The left plot stands for MC10 D_s , the middle one is for MC10 B_s^0 .

these will lead to different outputs, even in the case where the individual variables have compatible distributions as in Figures 7.25, 7.26 and 7.27.

For these reasons, the neural network calibration for $B_s^0 \rightarrow D_s^- \pi^+$ cannot be done using $D_s^+ \rightarrow \phi \pi^+$ data.

7.8 Conclusions

A first optimization of the selection cuts of the Same Side Kaon tagger has been performed using $D_s^+ \rightarrow \phi \pi^+$ 2010 data. The oscillations seen in $B_s^0 \rightarrow D_s^- \pi^+$ using this op-

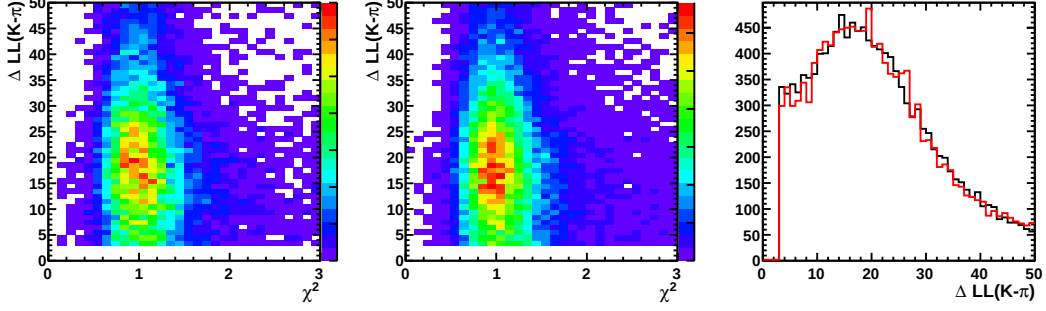


Figure 7.25: $\Delta LL(K - \pi)$ of the kaon in MC10 B_s^0 (black) and D_s (red) versus kaon track χ^2 (right plot). The left plot stands for MC10 D_s , the middle one is for MC10 B_s^0 .

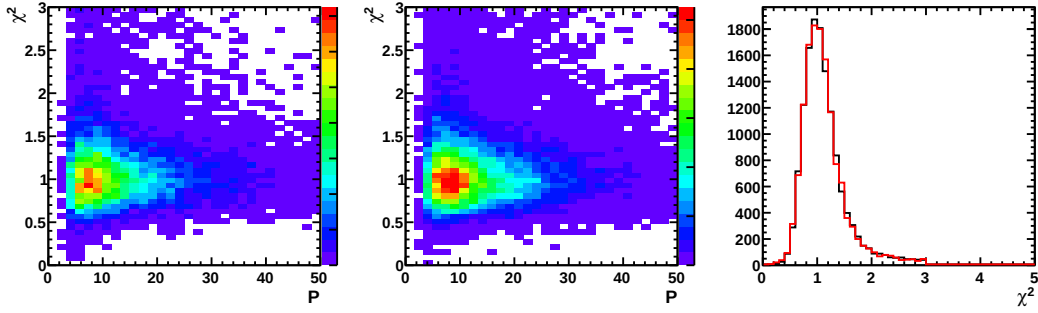


Figure 7.26: kaon track χ^2 in MC10 B_s^0 (black) and in D_s (red) versus kaon momentum p (right plot). The left plot stands for MC10 D_s , the middle one is for MC10 B_s^0 .

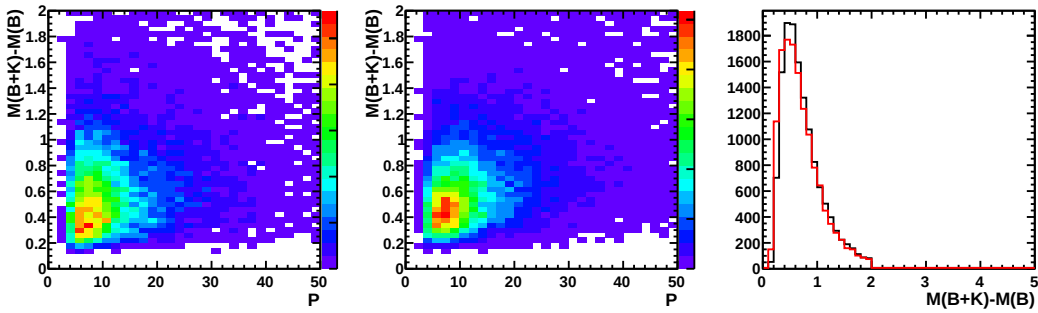


Figure 7.27: $M(B + K) - M(B)$ of the kaon in MC10 B_s^0 (black) and D_s (red) versus kaon momentum p (right plot). The left plot stands for MC10 D_s , the middle one is for MC10 B_s^0 .

timization are promising. Nevertheless, from comparison using Monte Carlo, we expect some differences between B_s^0 and D_s modes. Thus there are some variables whose cut value should be tuned directly from a B_s^0 control channel. For other variables, we expect no big differences in the final set of cuts that will be obtained optimizing with $B_s^0 \rightarrow D_s^- \pi^+$ data. In short, using the $D_s^+ \rightarrow \phi \pi^+$ decay to optimize the Same Side Kaon tagger has proved to be effective, leading to an improvement in tagging power and mistag rate. As shown in this study, the limit of this method is that it is not possible to calibrate the neural network used for mistag estimation for the $B_s^0 \rightarrow D_s^- \pi^+$ decay with $D_s^+ \rightarrow \phi \pi^+$ 2010 data.

Chapter 8

B_s^0 Lifetime Measurement from $B_s^0 \rightarrow D_s^- \pi^+$

This chapter presents work on the measurement of the B_s^0 meson lifetime using the $B_s^0 \rightarrow D_s^- \pi^+$ with $D_s^- \rightarrow K^- K^+ \pi^-$ decay channel. A second chapter presents the lifetime measurement of the B^0 meson using the $B^0 \rightarrow D^- \pi^+$ with $D^- \rightarrow K^- K^+ \pi^-$. All data used for these measurements has been recorded in 2011. The sequence of steps required to perform the measurement of the B_s^0 lifetime are listed below:

- Event reconstruction and selection, with the steps of the selection being the trigger, the stripping (offline pre-selection) and final selection. These steps form the content of Section 8.1.
- Studies on Monte Carlo simulation data are presented in Section 8.2. In particular, the determination of the mass shape parameters from MC11 data. A mass and decay time fit is performed to determine which parameters need to be fixed to MC-determined values and which should be left free. The goal is to have the largest possible number of free parameters whilst keeping the fitter stable.
- In Section 8.3 the acceptance function parameter values and their associated uncertainties are determined from MC11 pure signal data.
- The treatment of the signal decay time and the effect of the Heavy and Light propagation rates of the B_s^0 is presented in Section 8.4
- The Toy MC studies to determine systematics and investigate possible biases in the

Decay Mode	HLT1 triggers	HLT2 triggers
$B^0 \rightarrow DX$	Hlt1TrackAllL0	Hlt2Topo2BodyBBDT Hlt2Topo3BodyBBDT Hlt2Topo4BodyBBDT

Table 8.1: The triggers required for both the $B_s^0 \rightarrow D_s^- \pi^+$ and $B^0 \rightarrow D^- \pi^+$ modes.

measured decay time and decay time uncertainty form the content of Section 8.5. Toy MC studies are fits on data sets generated using the PDF model used by the fit where one parameter is varied. The variation applied corresponds to a change of 1σ on the parameters determined on MC11. The sets of Toy used for systematics contain 1000 Toys of 20k events. A set of 2000 Toys of 20k events with no parameter variation is used for fitter biases investigation.

- Final results with both statistical and systematic uncertainties are obtained in Section 8.6.

8.1 $B_s^0 \rightarrow D_s^- \pi^+$ Event Reconstruction and Selection

The common selection of the $B_s^0 \rightarrow D_s^- \pi^+$ and $B^0 \rightarrow D^- \pi^+$ decay modes whose lifetimes are measured in this thesis is described below.

8.1.1 Trigger Selection

The triggers required for the various modes are listed in Tab. 8.1. A set of explicit HLT1 and HLT2 triggers is required. The three HLT2 triggers used are the 2,3 and 4 body topological triggers. Each of these make use of a set of cuts applied through a binary boosted decision tree. These lines do not make use of any particle identification (PID) information.

8.1.2 Offline Pre-selection

A first offline pre-selection, called a stripping line, is applied. The candidates are built using data passing the B2DX stripping selection. This selection incorporates stripping lines which selects generic $B_q \rightarrow D_q h$ and $D_q \rightarrow h h h$ decays where h represents any hadron and q stands for a d or s quark. For this reason, no RICH information is used to form the candidates.

8.1.3 $B_q \rightarrow D_q X$ Offline Selection

The offline selection of the is $B_q \rightarrow D_q X$ with $q = d, s$ is common for the $B_s^0 \rightarrow D_s^- \pi^+$ and $B^0 \rightarrow D^- \pi^+$ modes and is designed to limit systematic uncertainties, since enough signal events exist for the $B_s^0 \rightarrow D_s^- \pi^+$ modes to comfortably surpass existing measurements. To this end, the following cuts are used to reject specific backgrounds or minimize specific systematics:

- All final state particles are required to have a minimum impact parameter χ^2 with respect to the primary vertex greater than 9, to ensure that they could not have formed part of the primary vertex (and hence could not bias the PV position);
- The D_q is required to have a flight distance χ^2 greater than 4 with respect to the B_s^0 vertex to suppress charmless backgrounds;
- The D_q is required to have an impact parameter χ^2 with respect to the primary vertex greater than 4 in order to suppress prompt charm backgrounds.
- The D_q signal mass window are chosen to be between 1848 and 1890 for the $D^\pm$ and between 1948 and 1990 for the D_s .
- The bachelor track is required not to be incompatible with a muon hypothesis to simplify the modeling of the lower mass sideband of the B^0 or B_s^0 .

Next, the D_q sample is cleaned up using PID and Dalitz cuts. The decay of interest, $D_s^- \rightarrow K^- K^+ \pi^-$ proceeds through two prominent resonances, the ϕ and K^* , as well as having a non-resonant component. The PID cuts applied are an OR of three selections, one for each of the resonances and one for the non-resonant, and are listed in Table 8.2. In all cases a Λ_c^+ veto is applied: either the kaon with the same charge as the pion has to be $DLL(K-p) > 5$ or the mass of the D_q is under the Λ_c^+ mass hypothesis and is more

Decay Mode	K^+	K^-	π^-	Mass Window
$\phi\pi$	> -5	> -5	< 5	$ m_{KK} - 1020 < 10$
K^*K	> 5	> -5	< 5	$ m_{K^-\pi^+} - 892 < 50$
non-resonant	> 5	> 0	< 5	N/A

Table 8.2: The DLL cuts for $D_s^- \rightarrow K^- K^+ \pi^-$.

than 25 MeV away from 2285 MeV. The motivation for this veto is the large amount of $\Lambda_b \rightarrow \Lambda_c^+ \pi^-$ decays.

Finally, a simple decision tree based on the OR of some very hard cuts is used to suppress combinatoric backgrounds. The event is kept if all of the following conditions is satisfied:

- The cosine of the angle between the B_q momentum vector and direction of flight is greater than 0.9999975, which represent a rather tight cut.
- The sum of the transverse momenta of the final state particles is greater than 15 GeV.
- Both the bachelor track and D_q have a minimum impact parameter χ^2 greater than 225.
- All final state tracks have a minimum impact parameter χ^2 greater than 100.
- The B_q mass is refitted with the D_q mass as a constraint.

If multiple candidates remain in an event, one is chosen at random for the final fit. This occurs in fewer than 2% of the cases.

Bachelor Pion Particle Identification Cut

The last step in the selection process is to impose a cut on the particle identified as the bachelor pion in the $B_s^0 \rightarrow D_s^- \pi^+$ decay or $B^0 \rightarrow D^- \pi^+$. Each final state particle has an associated probability to be of a certain species. This probability is based on the output of a neural net trained on MC10 data [46]. Candidates are rejected where the particle identified as the bachelor pion has a calculated probability to be a kaon greater than

Channel	Cut Efficiency
$B_s^0 \rightarrow D_s^- \pi^+$	81.8%
$B_s^0 \rightarrow D_s^- K^+$	9.10%
$\Lambda_b \rightarrow \Lambda_c^+ \pi^-$	82.1%

Table 8.3: Efficiency of a cut at 0.2 on the probability of the candidate's bachelor pion to be a kaon for the signal and both peaking background channels.

0.2. The efficiency of this cut for the signal and both peaking backgrounds is presented in Table 8.3.

8.2 Studies on Monte Carlo Simulated Events

The final fit of the B_s^0 decay time distribution uses information from the prediction based on MC distributions for the candidates invariant mass, decay time uncertainty, and decay time acceptance. The fitter is described in Section 8.2.1 and the samples used to provide constraints are discussed in Section 8.2.2. The shapes of the invariant mass distributions are determined in Section 8.2.3.

8.2.1 Fitter Description

The B_s^0 lifetime measurement is performed using a two dimensional maximum likelihood fit. The fitted observables are the reconstructed invariant mass and decay time. The decay time uncertainty is introduced in the fit as a conditional parameter on a per-event basis. The two-dimensional nature of the fit is needed to separate the multiple background contributions from the signal, in particular the peaking background from the Λ_b decay. The fitter observables and parameters are stored in separate xml configuration files, allowing for simple configuration of starting values and fit ranges.

The mass model probability density functions (PDF) used in the fit fall into two categories : histogram-based PDFs using the MC mass distributions of the partially reconstructed backgrounds, and PDFs whose shapes were determined from fits to the peaking background MC mass distributions. Most of the histogram-based PDFs are

ignored in the final fit due to their mass range being below the fit mass window, with the exception of the $B_s^0 \rightarrow D_s^{*-} \pi^+$ partially reconstructed physics background. The lifetime fit range is from 0.35 to 14 ps, the lower bound being used to avoid a combinatorial background peaking at small reconstructed decay times.

8.2.2 Monte Carlo Event Samples

The initial characterisation of the background and signal distributions was performed using 2011 Monte Carlo datasets (MC11). The datasets were generated using the conditions under which the bulk of the data collected during 2011 was recorded. Each dataset is formed from two subsamples corresponding to the two magnet polarities, in all cases the subsamples contain near-equal amounts of events.

For each dataset, the Monte Carlo truth information of events passing the stripping selection (pre-selection) and offline selection is checked and candidates formed from ghost tracks (tracks formed from random hits), track clones or tracks originating from the primary vertex as well as spillover events are discarded. The number of candidates present at each stage of this process and the overall candidate efficiency without the kaon PID cut on the bachelor is presented in Table 8.4.

Name	Generated	After stripping	After offline selection	Efficiency/ 10^{-2}
$B_s^0 \rightarrow D_s^- \pi^+$	1012495	46781	17003	1.68
$B_s^0 \rightarrow D_s^{*-} \pi^+$	524098	20667	7645	1.49
$B_s^0 \rightarrow D_s^+ \rho^-$	1953391	28964	11107	0.57
$B_s^0 \rightarrow D_s^{*+} \rho^-$	1019191	3971	1399	0.14
$B^0 \rightarrow D^- \pi^+$	1014995	9040	7	0.00
$B_s^0 \rightarrow D_s^- K^+$	1887293	85620	31950	1.69
$\Lambda_b \rightarrow \Lambda_c^+ \pi^-$	2033496	49970	64	0.00
$\Lambda_b \rightarrow \Lambda_c^+ \pi^-$ (no veto)	2033496	49970	2437	0.12

Table 8.4: Selection efficiency for signal and backgrounds, without kaon PID cut on the bachelor.

In Table 8.4, the Generated column gives the number of generated events for the considered decay. The Candidates column gives the number of events after pre-selection, the Selected column gives the numbers of events after selection. The Efficiency col-

umn provides the ratio between the number of candidates and the number of selected events. With the exception of B_s⁰ → D_s⁻π⁺, all the samples correspond to potential backgrounds. The backgrounds enters in two broad categories, the partially reconstructed physics backgrounds and the resonant backgrounds. The partially reconstructed physics backgrounds in this table are B_s⁰ → D_s^{*-}π⁺, B_s⁰ → D_s⁺ρ⁻ and B_s⁰ → D_s^{*+}ρ⁻. The resonant backgrounds are B_s⁰ → D_s⁻K⁺, B⁰ → D⁻π⁺ and Λ_b → Λ_c⁺π⁻. In the Λ_b → Λ_c⁺π⁻ case, numbers with and without Λ_c⁺ veto are presented as the later will be used to determine the Λ_b → Λ_c⁺π⁻ background mass shape.

8.2.3 Mass Distributions

The mass models used for each channel are presented. Functional forms are used for the peaking distributions and for the combinatorial background. In all other cases, a histogram of the MC mass distribution is used.

B_s⁰ → D_s⁻π⁺ Signal Mass Model

The invariant mass distribution of the B_s⁰ → D_s⁻π⁺ signal and conjugate decay is modelled using the sum of a Crystal ball with a Gaussian, where the Crystal ball distribution shares the same mean as the additional gaussian distribution. The Crystal ball distribution describes a central gaussian distribution with an exponential distribution on one side which conforms to a radiative tail. The exponential part of the Crystal ball distribution helps fit the low mass tail. A fit of this model to the Monte Carlo signal data is shown in Figure 8.1, the parameters which define the distribution are given in Table 8.5.

Peaking Background Model

The only two significant peaking backgrounds in the mass region are the Λ_b → Λ_c⁺π⁻ and B_s⁰ → D_s⁻K⁺ decays. For the Λ_b → Λ_c⁺π⁻ background, a proton is misidentified as a kaon, causing the reconstructed mass of the Λ_b mistaken as a B_s⁰ to lie in the upper region of the B_s⁰ mass peak, as shown in Figure 8.2. The model used for the Λ_b → Λ_c⁺π⁻ background is a Crystal ball whose tail is used to describe the low mass tail of the background.

It should be noted that since a Λ_c⁺ veto is implemented in the selection, the number

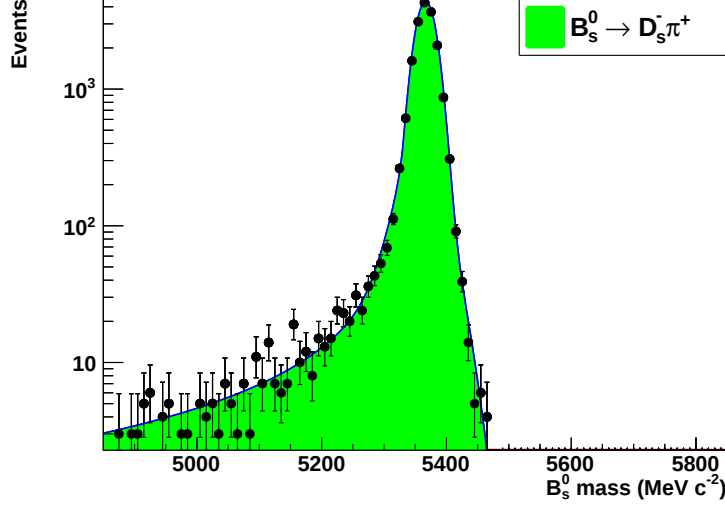


Figure 8.1: Fit of the mass distribution for $B_s^0 \rightarrow D_s^- \pi^+$ signal on Monte Carlo pure signal data.

Parameter	Fit Value	Fit Uncertainty
Centre of Distribution	$5.3669 \cdot 10^3$	$1.0 \cdot 10^{-1}$
Width of CB Component	$1.513 \cdot 10^1$	$2.32 \cdot 10^0$
Exponent of CB Tail	$1.396 \cdot 10^{-1}$	$1.12 \cdot 10^{-2}$
CB Switch to Exponential	$3.059 \cdot 10^0$	$1.95 \cdot 10^{-1}$
Width of Gaussian Component	$3.611 \cdot 10^1$	$2.32 \cdot 10^0$
Fraction of CB Component	$9.403 \cdot 10^{-1}$	$1.12 \cdot 10^{-2}$

Table 8.5: The parameters which define the $B_s^0 \rightarrow D_s^- \pi^+$ invariant mass distribution.

of MC candidates passing the selection is very low. In order to fit the Λ_c^+ invariant mass distribution, the veto was relaxed. The results of the fit with the Crystal Ball model is summarized in Table 8.6. It is possible that the veto distort slightly the mass shape, but in the limit given by the low number of MC11 $\Lambda_b \rightarrow \Lambda_c^+ \pi^-$ passing the selection with the veto, no effect has been observed.

The model used for the $B_s^0 \rightarrow D_s^- K^+$ background is also a Crystal ball, whose tail is used to describe the low mass tail of the background.

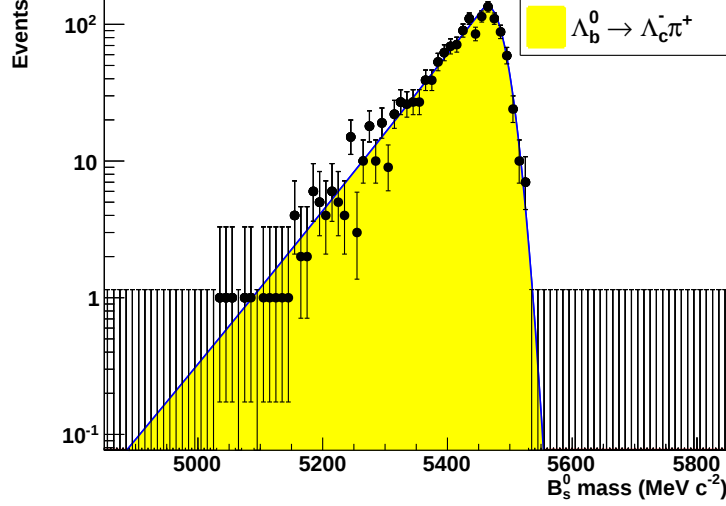


Figure 8.2: Fit of the mass resolution model for $\Lambda_b \rightarrow \Lambda_c^+ \pi^-$ background on Monte Carlo pure $\Lambda_b \rightarrow \Lambda_c^+ \pi^-$ reconstructed as signal.

Parameter	Fit Value	Fit Uncertainty
Centre of Distribution	$5.4640 \cdot 10^3$	$2.3 \cdot 10^0$
Width of CB	$2.294 \cdot 10^1$	$1.49 \cdot 10^0$
CB Tail Exponent	$1.159 \cdot 10^2$	$1.0 \cdot 10^{-1}$
CB Switch to Exponential	$3.083 \cdot 10^{-1}$	$2.47 \cdot 10^{-2}$

Table 8.6: The parameters which define the $\Lambda_b \rightarrow \Lambda_c^+ \pi^-$ invariant mass distribution.

It should be noted that since a cut on the PID is implemented in the selection, the number of $B_s^0 \rightarrow D_s^- K^+$ MC candidates passing the selection is very low. In order to properly model the mass distribution, the number of events is increased by removing the bachelor PID cut. This bias the mass distribution slightly, however no better option was available. The results of the fit with the Crystal ball model is summarized in Table 8.7.

Due to the low number of expected events in this channel relative to the signal ($\mathcal{O}(1\%)$ from Tables 8.3 and 8.4), coupled with the fact that this PDF distribution overlaps significantly with both the signal and the $B_s^0 \rightarrow D_s^{*-} \pi^+$ background, the fit is performed without using contribution from $B_s^0 \rightarrow D_s^- K^+$ model. The effect of this decision is accounted for in the systematic studies discussed in Section 8.5.

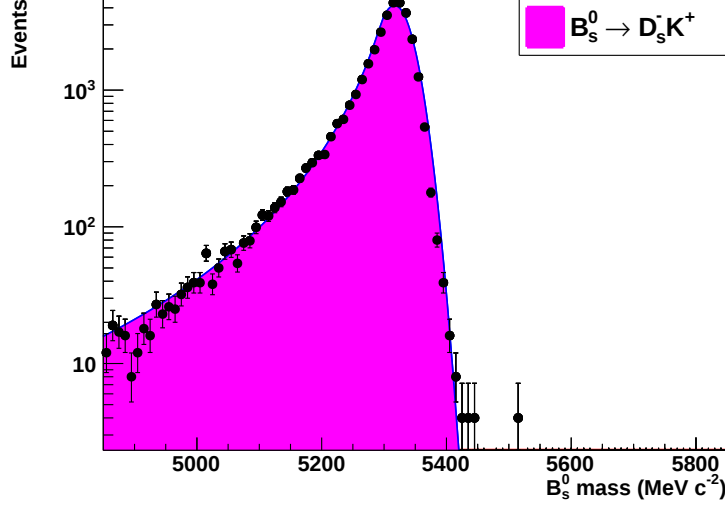


Figure 8.3: Fit of the mass resolution model for $B_s^0 \rightarrow D_s^- K^+$ background on Monte Carlo pure $B_s^0 \rightarrow D_s^- K^+$ reconstructed as signal.

Description	Fit Value	Fit Uncertainty
Centre of Distribution	$5.3166 \cdot 10^3$	$3.0 \cdot 10^{-1}$
Width of CB	$2.635 \cdot 10^1$	$1.9 \cdot 10^{-1}$
Exponent of Tail	$2.871 \cdot 10^0$	$8.9 \cdot 10^{-2}$
Switch to Exponential	$8.71 \cdot 10^{-1}$	$1.6 \cdot 10^{-2}$

Table 8.7: The parameters which define the $B_s^0 \rightarrow D_s^- K^+$ invariant mass distribution.

Partially Reconstructed Physics Backgrounds Mass Models

The missing mass backgrounds are situated below the B_s^0 mass peak and have no clear peaking structure. The concerned backgrounds are the following : $B_s^0 \rightarrow D_s^{*-} \pi^+$, $B_s^0 \rightarrow D_s^{*-} K^+$, $B_s^0 \rightarrow D_s^+ \rho^-$ and $B_s^0 \rightarrow D_s^{*+} \rho^-$. All these backgrounds pass the selection by being reconstructed while missing on one particle, resulting in missing mass with respect to a real signal event. These backgrounds are modelled using PDFs formed from histograms obtained from MC11 data. The mass distributions of these three backgrounds are presented in Figure 8.4.

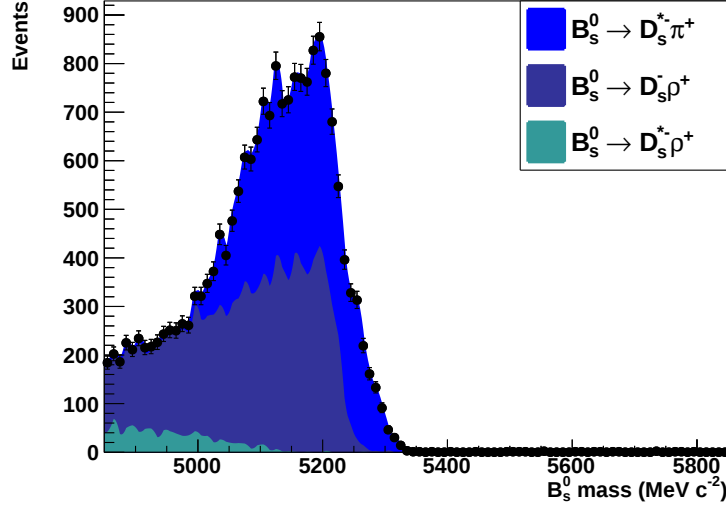


Figure 8.4: The mass model for missing mass backgrounds on pure Monte Carlo missing mass background data. Each PDF is constructed from the relevant selected Monte Carlo mass distribution.

Combinatorial Background

Combinatorial background candidates are random combinations of particles which are not produced in a decay of the same particle. When the candidates are formed, a cut is applied both on the D_s invariant mass and the B_s^0 invariant mass. For this reason, there is no available sideband data that can be used to determine the combinatorial background. However, fits to an older dataset of B_s^0 cocktail events show that the distribution is well modelled by an exponential with negative exponent in the region above the Λ_b peak. This background accounts for most events having a mass above the B_s^0 and Λ_b mass peaks.

8.3 Decay Time Distributions and Acceptance

In order to correctly fit the decay time of the B_s^0 meson, it is important to take into account all the effects which can alter the decay time distribution. To do this we introduce an acceptance function; the value of which gives the probability of a decay with a certain decay time to be ultimately selected. The acceptance function therefore represents the average combination of all decay time-biasing effects for all events. From a physical

point of view the acceptance corresponds to the fact that the detector can have a different probability of reconstructing an event, depending on its decay time. This effect has its source in the selection and reconstruction of events. Acceptance from selection can come from cuts present in the trigger, pre-selection or the selection. Acceptance from the reconstruction has its source in the detector itself. The methods used to obtain the decay time acceptance are discussed in Section 8.3.1.

In addition to the decay time acceptance, each candidate is also supplied with an estimate of the resolution achieved on the measured decay time. This information is used to convolve the basic decay time distribution with a Gaussian resolution on an event-by-event basis. The per-event decay time resolution is discussed in Section 8.3.2.

8.3.1 Decay Time Acceptance

It is clear from the observed decay time distribution that the sample suffers from decay time-biasing effects. Since the aim of this analysis is to measure the B_s^0 lifetime, it is important to account for such effects.

The combination of all decay time biasing effects can be modelled as the probability to observe a candidate with a given measured decay time. The shape of the decay time acceptance function can be determined from MC data if we assume that the average bias in the reconstructed time is close to zero. The residual distribution of the reconstructed decay time, $t_{\text{reco}} - t_{\text{true}}$ is shown in Figure 8.5. A double Gaussian fit to the observed weighted mean distribution has a mean of $(-7.28 \pm 2.88) \cdot 10^{-04}$. As visible in Figure 8.5 the mean is compatible with zero, which means that it is reasonable to use the MC data to determine the acceptance.

The following process is used to set the shape of the decay time acceptance. A lifetime fit is performed on pure signal MC11 data, where the B_s^0 lifetime is fixed to generation value, and the parameters of the acceptance function defined by Equation 8.1 are let free. The parameters obtained that way are then used to fix the acceptance function used in the fitter,

$$\text{acc}(t_{\text{reco}}) = \text{Max}(0, (1 - tS_2)(1 - e^{-S_1(t-t_P)})). \quad (8.1)$$

In Equation 8.1 the S_2 parameter correspond to the slope of the linear component of the acceptance, having an effect mainly at long lifetimes. The t_P parameter give the

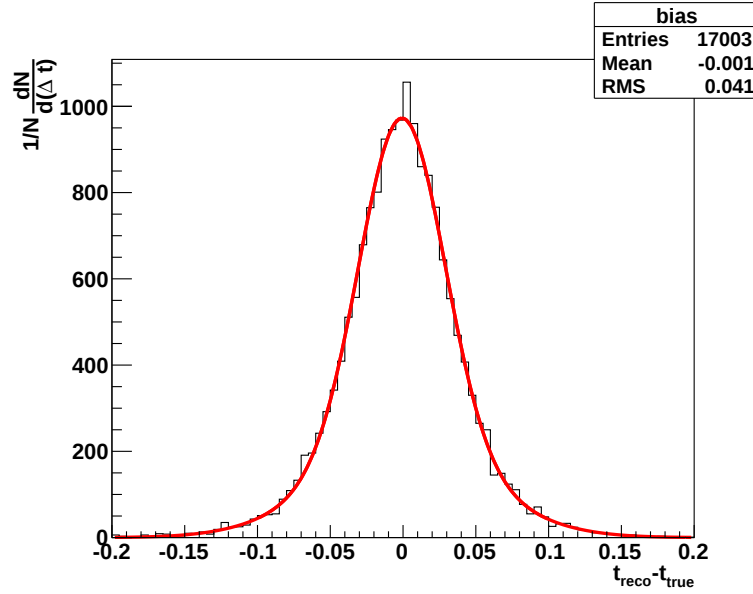


Figure 8.5: Residual distribution of the reconstructed decay time, i.e. $t_{\text{reco}} - t_{\text{true}}$

turning point where the function takes a non zero value while the S_2 parameter the the exponential slope controlling the rise of the function toward 1.

The corresponding fit is shown in Fig. 8.6. It should be noted that in the above procedure, a mix of the reconstructed and generated decay times was used. The acceptance function parameters obtained from the fit are summarised in Table 8.11.

Parameters	Value	Uncertainty
t_P	0.1901	0.0207
S_1	1.3972	0.1336
S_2	0.0242	0.0086

Table 8.8: Acceptance function parameters as determined from a fit to the MC11 signal data.

We use the signal acceptance on all channels but the combinatorial background. Any bias to the lifetime of the background channels introduced are inconsequential since they will be compensated for in the value of the fitted background lifetime.

The decay time acceptance is obtained from Monte Carlo simulation data. The fitter

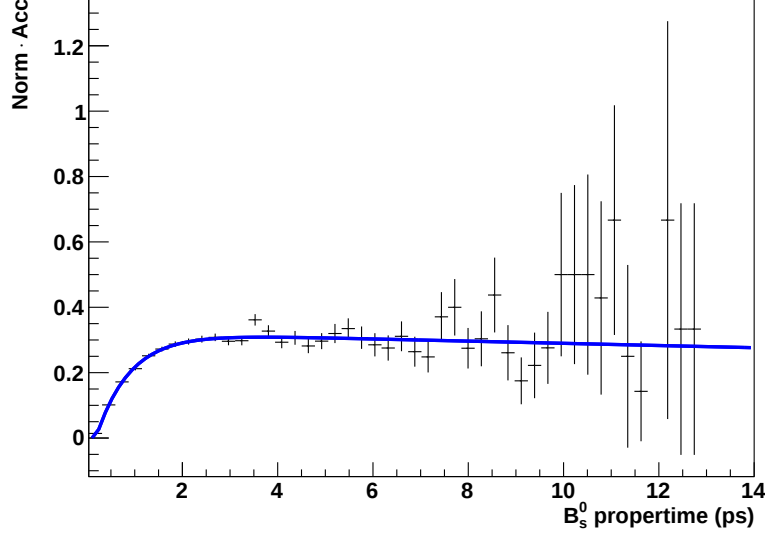


Figure 8.6: Acceptance function fit on pure MC11 signal data.

makes use of one acceptance function for all the background contribution and for the signal. The only exception is the combinatorial background where no acceptance is applied. MC11 samples of each backgrounds have been generated for this purpose, as well as a sample for signal. The acceptance function is extracted by a fit where the B_s^0 lifetime is fixed at generation value while the acceptance parameters are free parameters of the fit. The obtained acceptance is a function of the reconstructed decay time and not the true decay time. It is also possible to use an acceptance in terms of the true decay time, but this is very CPU intensive and slows the fitting process a lot. The difference in CPU time comes from the integration performed when applying the decay time uncertainty distributions as conditional PDF for the 2D PDF model. When the acceptance is a function of the measured decay time, it is applied after the 2D PDF model is set to be conditional to the decay time uncertainty, keeping the acceptance function out of the integration performed at this step. If the acceptance as function of the true decay time is used, it is then applied before the 2D PDF model is set to be conditional on the decay time uncertainty PDF's, and the integration at this step is not analytical anymore and therefore consume a lot of CPU time. For this reason, the choice is to use the acceptance in terms of the reconstructed decay time.

The treatment of the acceptance in the fitter can be improved by replacing the global acceptance function by an acceptance on a per-event basis. This is possible by using the vertex swimming method [47, 48] and will be used in the future. For each event,

the primary vertex position is moved incrementally, at each step the trigger decision is recalculated. This leads to a multistep function for each event, where the event is accepted for a decay time within certain ranges and rejected elsewhere. For each event the integration range is modified dynamically to match the decay time regions where the considered event is within the acceptance and reject the ranges of decay time where the event to be out of acceptance.

If one makes the summation of the multistep function recovered for each event, a global acceptance function due to the trigger would be obtained. The swimming method also reduces the dependency on Monte Carlo simulated data, as it is performed on the data. It does not completely remove the need of an acceptance function determined on simulated data. The swimming method resolve acceptance caused by selection cuts, but not the acceptance due to track reconstruction as it does not rerun it at each iteration.

8.3.2 Per-Event Decay Time Uncertainty

In this analysis, the decay time uncertainty is treated on a per-event basis. This is done by using the decay time uncertainty of each event as the width of the decay time resolution model. The decay time resolution model used is a single Gaussian.

To properly use a per-event decay time uncertainty in a conditional fit, it is necessary to take into account the distribution of the conditional parameter in the signal and the backgrounds. This has to be done for all event classes, even if the conditional parameter is only included for one class of event [49]. If this is not done, a bias is introduced in the fit. To illustrate this problem, the case of a simple fit with only one signal, one background and only the decay time as free parameter is be used. If one assume a fit where one maximize the probability of measuring the observed data with our simple model for signal and background, this probability is described by the likelihood function,

$$L = \prod_i^i f \cdot p(\tau_i | S) + (1 - f) \cdot p(\tau_i | B). \quad (8.2)$$

If one use a fixed resolution, the one may write

$$L = \prod_i^i f \cdot p(\tau_i | S, \sigma) + (1 - f) \cdot p(\tau_i | B, \sigma). \quad (8.3)$$

Now, in order to introduce a per-event resolution, in our case on the decay time τ , the

naive and wrong way of doing it is

$$L = \prod_i f \cdot p(\tau_i | S, \sigma_i) + (1 - f) \cdot p(\tau_i | B, \sigma_i). \quad (8.4)$$

This is wrong because in the per-event case, σ_i is a measurement. The measured quantity is not τ_i anymore, but $\vec{x}_i = (\tau_i, \sigma_i)$. This means that one need to determine the form of $p(\vec{x}_i | A) = p((\tau_i, \sigma_i) | A)$. The expression for this term is

$$p(\tau_i \cap \sigma_i | A) = p(\tau_i | \sigma_i \cap A) \cdot p(\sigma_i | A). \quad (8.5)$$

Using the result of Equation 8.5, the correct form for the Likelihood function is

$$L = \prod_i f \cdot p(\tau_i | \sigma_i \cap S) \cdot p(\sigma_i | S) + (1 - f) \cdot p(\tau_i | \sigma_i \cap B) \cdot p(\sigma_i | B). \quad (8.6)$$

It is now visible from Equation 8.6 that one need to take the conditional parameter σ_i distribution into account both for the signal and background.

The decay time uncertainty is applied to the decay time PDF by convolving an exponential with a Gaussian of width equal to $\tau_{\text{err}} \cdot \sigma$, where τ_{err} is the decay time uncertainty and σ is a scaling parameter. The scaling parameter σ is a free parameter of the fit and allows to account for an eventual difference between the τ_{err} distribution in Monte Carlo data and real data.

It can be observed that most of the event classes have a near identical τ_{err} distribution, as shown in Figure 8.7. The Monte Carlo data sample of combinatorial background not large enough to accurately determine the decay time uncertainty distribution. Instead, the decay time uncertainty distribution is fixed to that of the data in the high mass region above all peaking backgrounds.

In principle, the τ_{err} distributions for the different classes of events can be extracted from data using S-Plots[50]. In this analysis, the distributions are taken from Monte Carlo as they fit the real data very well as shown in Figure 8.8.

Given the very similar distributions for the signal and the peaking backgrounds, this analysis uses two decay time uncertainty distributions : one for the signal and resonant backgrounds and another for the combinatorial background. These PDFs are introduced as conditional PDF for the 2D mass and decay time PDF, leading to a PDF of the type

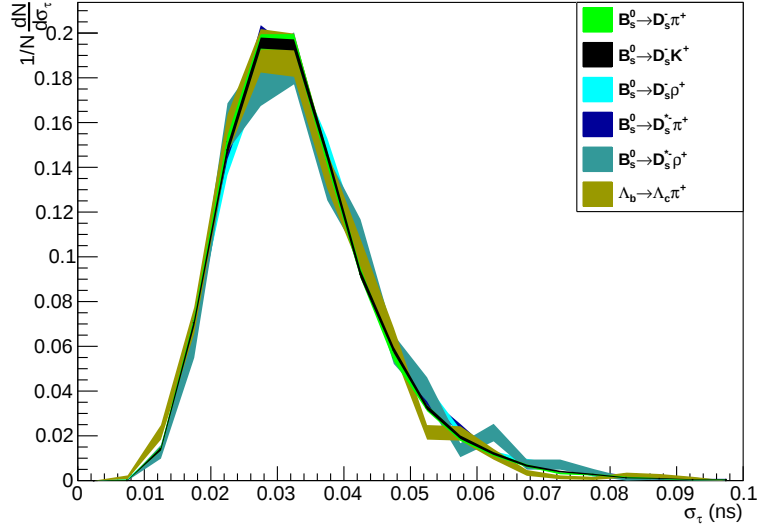


Figure 8.7: τ_{err} distributions for all classes of events but combinatorial background.

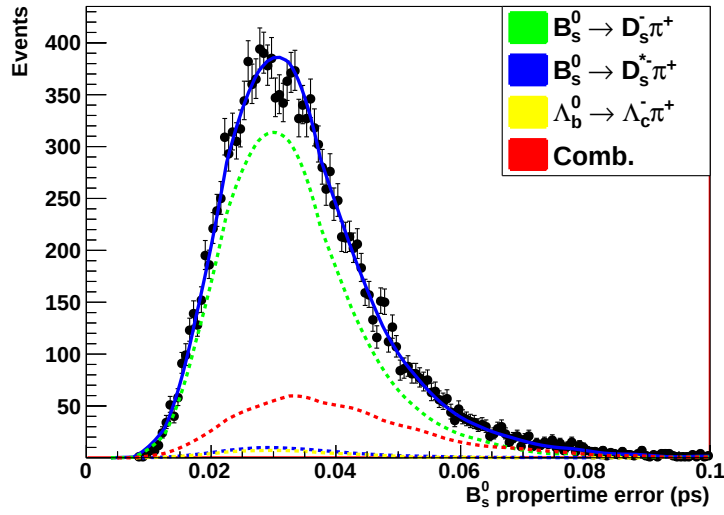


Figure 8.8: τ_{err} distributions compared to data.

shown in Equation 8.7 for each class of events.

$$(Mass \times \tau | \tau_{\text{Err}}). \quad (8.7)$$

This procedure allows to correctly take into account the dependence of the decay time

resolution with the decay time itself by assigning a decay time uncertainty corresponding to the reconstructed decay time of each event.

8.4 Heavy and Light States

As discussed in Section 3.6 and in more depth in [23], the neutral meson system can oscillate between their $|B_q\rangle$ and $|\bar{B}_q\rangle$ flavour states, that is $\langle B_q | \bar{B}_q \rangle \neq 0$. This oscillation is due to the weak interaction between the component quarks, and has the consequence that the general state of the B meson system is a superposition of the two states $|B_q\rangle$ and $|\bar{B}_q\rangle$,

$$a|B_q\rangle + b|\bar{B}_q\rangle. \quad (8.8)$$

As shown in Section 3.6, the evolution of the state shown in Equation 8.8 is described by the Schrödinger equation. With the assumption that B_q and $\bar{B}_q$ have identical masses and decay widths (from **CPT** conservation), the mass eigenstates of the B meson system,

$$|B_{H,L}\rangle = p|B_q\rangle \pm q|\bar{B}_q\rangle, \quad (8.9)$$

and have well-defined decay times. In the fit model, the acceptance function determined in 8.3.1 is not uniform in decay time and can therefore alter the ratio of these two states in the selected candidates if $\Delta\Gamma_s \neq 0$. To avoid this, a sum of the decay time distributions of $B_{s,L}$ and $B_{s,H}$ is used as the signal decay time distribution, both for fitting and when determining the acceptance function parameters.

The B meson, in our case B_s^0 , will propagate as a mixture of these Heavy and Light states. Because the B_s^0 mesons are produced through the strong interaction, they are created in their flavour eigenstates, which means an equal fraction of $B_{s,L}$ and $B_{s,H}$ at production. The evolution of this mixture is described by:

$$\frac{1}{2}|B_{s,L}(t)\rangle + \frac{1}{2}|B_{s,H}(t)\rangle = e^{-im_L + \Gamma_L/2t}|B_{s,L}\rangle + e^{-im_H + \Gamma_H/2t}|B_{s,H}\rangle. \quad (8.10)$$

The mass and decay width differences are defined as

$$\Delta m_s = m_{s,H} - m_{s,L} \quad \Delta \Gamma_s = \Gamma_{s,L} - \Gamma_{s,H}. \quad (8.11)$$

The difference in lifetime used in the fitting procedure and acceptance determination is $\Delta \Gamma_s / \Gamma_s = 0.092 \pm 0.055$ from [1]. Equation 8.10 can be rewritten using 8.11 and $\frac{\Gamma_L + \Gamma_H}{2} = \bar{\Gamma}_s$

$$\begin{aligned} \frac{1}{2} |B_{s,L}(t)\rangle + \frac{1}{2} |B_{s,H}(t)\rangle &= \frac{1}{2} e^{-im_L t} e^{-\frac{\bar{\Gamma}_s}{2} \left(1 - \frac{\Delta \Gamma_s}{\bar{\Gamma}_s}\right) t} |B_{s,L}\rangle \\ &+ \frac{1}{2} e^{-im_H t} e^{-\frac{\bar{\Gamma}_s}{2} \left(1 + \frac{\Delta \Gamma_s}{\bar{\Gamma}_s}\right) t} |B_{s,H}\rangle \end{aligned} \quad (8.12)$$

This permits the measurement of the average lifetime $\tau_{B_s} = 1/\bar{\Gamma}$.

8.5 Systematic Uncertainties and Fitter Bias

This section presents the different sources of systematic uncertainties considered and their estimated values. The systematic uncertainties from the fitting model are presented in subsection 8.5.1 and the acceptance corrections in subsection 8.5.2. A summary of all the systematic uncertainties is given in subsection 8.5.5. The presence of a possible bias in the lifetime measurement from the fitter is investigated in 8.5.6 using MC Toy experiments.

8.5.1 Systematic Uncertainties due to the Mass Models

The systematic uncertainties due to the mass models are estimated for both the signal mass model and for the Λ_b background. The systematic uncertainty estimation is performed for each parameter of the mass models that has been determined using Monte Carlo simulation data. The relevant parameters in the two models considered are listed in Table 8.9.

The signal mass model is the sum of a Gaussian and a Crystal ball, where the gaussian part is used to pick up a small upper mass tail. The parameters of the gaussian and tail of the crystal ball are determined from pure signal Monte Carlo simulation data. The Λ_b mass model is a single crystall ball, where all parameters but the mass are determined

Model	Parameters
B_s^0 Signal	Exponent of CB Tail
	Fraction of the CB
	Width of the Gaussian
Λ_b Background	Width of CB
	Exponent of CB Tail
	Switch to Exponential of CB

Table 8.9: Fixed parameters of the mass models.

on Monte Carlo simulation data.

Two sets of Toy Monte Carlo experiments are performed for each of the parameters listed in Table 8.9. Each parameter is varied over a symmetric range centered on the parameter fit value and has a width of the parameter uncertainty as determined in the original fit. Each set contains a total of 1000 experiments, where each experiment has a total of 20k events, corresponding approximately to the statistics available in data. In each of the 1000 experiments, the generated model is based on the fit PDF where parameters determined on MC11 data are changed by 1σ deviation. The same fitting algorithm used for the B_s^0 lifetime extraction is used to fit these generated data sets. The means of the B_s^0 lifetime measured with the respective $+1\sigma$ and -1σ deviation of the studied parameter are measured. These two values provide the upper and lower bounds of the systematic uncertainty on the B_s^0 lifetime for the considered parameter. The results obtained from the Toy MC experiments are shown in Table 8.10.

Model	Parameters	Lower Sys. [ps]	Upper Sys. [ps]
Signal	Tail of CB	0.0007	0.0004
	Fraction of CB	0.0005	0.0004
	Width of Gaus.	0.0009	0.0005
Λ_b	Width of CB	0.0010	0.0000
	Tail of CB	0.0002	0.0011
	Alpha of CB	0.0001	0.0003

Table 8.10: Systematic uncertainties due to fixed parameters of the mass models.

Given the number of Toy experiments generated, the observed systematic uncertainty for these parameters are consistent with zero. We conservatively assume the largest observed deviation to be the systematic uncertainty.

8.5.2 Systematic Uncertainty from Acceptance Model

The acceptance model is given by

$$\text{Max}(0, (1 - tS_2)(1 - e^{-S_1(t-t_P)})), \quad (8.13)$$

where t_P , the turning point, represents the time after which the acceptance function has a non-zero value, S_1 determines the gradient of the initial turn on and S_2 dictates the slope of the long lifetime acceptance. The results of the Toys are summarised in Table 8.11.

The value of the three parameters t_P , S_1 and S_2 were determined using MC11 simulation data for pure signal as described in 8.3.1. The procedure used to determine the systematic uncertainties from these parameters is identical to the one used for the mass models in Subsection 8.5.1. The systematic uncertainties on the lifetime obtained with this method for t_P , S_1 and S_2 are presented in Table 8.11.

Model	Parameter	Lower sys. [ps]	Upper sys. [ps]
Acceptance	t_P	0.0124	0.0125
	S_1	0.0266	0.0315
	S_2	0.0249	0.0236

Table 8.11: Systematic uncertainties due to acceptance function.

8.5.3 Systematic Uncertainty from $\Delta\Gamma_s/\Gamma_s$

The systematic uncertainty caused by the uncertainty on the $\Delta\Gamma_s/\Gamma_s$ taken from the Particle Data Group [1] is treated using the same method as the parameters studied in Subsections 8.5.1 and 8.5.2. As in these previous cases, a total of 1000 Toy MC experiments of 20k events were generated. The systematic uncertainty due to $\Delta\Gamma_s/\Gamma_s$ is given in Table 8.12.

Model	Lower Sys. [ps]	Upper Sys. [ps]
$\Delta\Gamma_s/\Gamma_s$	0.0009	0.0001

Table 8.12: Systematic uncertainties due to the $\Delta\Gamma_s$ parameter.

8.5.4 Other Systematic Uncertainties

There are two sources of systematic uncertainties which remain to be discussed : the momentum scale and the decay length scale. Both values can be found in Table 8.13 and are taken from [51].

The momentum scale is calibrated with a precision of $0.05 \pm 0.02\%$ [52]. The proper time expression is given by

$$t = \frac{d m}{p}. \quad (8.14)$$

In this equation d represents the flight distance, m the reconstructed mass and p the momentum of the candidate. Since the reconstructed mass m depends on the momentum, the momentum scale contribution to the systematic uncertainty is small as the uncertainty on p partially cancels in the ratio $\frac{m}{p}$.

The decay length scale is a representation of the uncertainty on the measurement of the B_s^0 flight distance along the z axis. It mainly depends on the alignment of the detector, the number of tracks in the decay and the number of VELO modules containing a hit. A more complete explanation can be found in [51], where the relative uncertainty on the lifetime is calculated to be 0.0026%.

The last systematic uncertainty is due to the presence of $B_s^0 \rightarrow D_s^- K^+$ decays which is not modelled. This uncertainty is estimated by adding a PDF component corresponding to the fitted Monte Carlo $B_s^0 \rightarrow D_s^- K^+$ distributions with an amplitude fixed to the number of expected events. This number is found to be 200 expected events, and is determined by combining branching fraction information, the selection efficiency for the $B_s^0 \rightarrow D_s^- K^+$ decay and the mass window of the fit. The lifetime of this PDF is the same one used for the signal PDF. The systematic uncertainty is taken as the difference in measured lifetime between the fit with and without this component. This uncertainty has been estimated to be ± 0.00035 [ps].

8.5.5 Summary of the Systematic Uncertainties

The systematic uncertainties determined in Subsections 8.5.1, 8.5.2, 8.5.3 and 8.5.4 are summarized here in Table 8.13 and are then combined taking into account the correlations between parameters.

Parameter	Lower sys. [ps]	Upper sys. [ps]
Sig. CB tail	0.0005	0.0004
Sig. CB frac.	0.0007	0.0004
Sig Gaus. width	0.0009	0.0006
Total sig. model	± 0.00123	
Λ_b CB width	0.0001	0.0003
Λ_b CB tail	0.0002	0.0011
Λ_b CB alpha	0.0010	0.0000
Total Λ_b model	± 0.00123	
t_P	0.012	0.013
S_1	0.027	0.032
S_2	0.025	0.024
Total acc. model	0.039	0.041
$\Delta\Gamma/\Gamma$	0.0009	0.0001
Momentum scale	0.0008	0.0008
Decay length scale	0.0000	0.0000
Expected $B_s^0 \rightarrow D_s^- K^+$	0.0004	0.0004
Total other	± 0.00084	
Total systematic uncertainty	0.0387	0.0414

Table 8.13: Summary of all systematic uncertainties.

Correlations between parameters are accounted for by applying a transition matrix that diagonalizes the correlation matrix to the uncertainty vector. The details of the method used can be found in the statistics appendix of the Particle Data Group booklet [1]. The final systematic uncertainty is calculated by summing the magnitude of the uncertainties in the eigenbase in quadrature.

8.5.6 Fitter Bias

A possible bias in the fitter is investigated with Toy Monte Carlo experiments. The parameters of the generation model are fixed either to the values obtained from the fit performed on real data or on values obtained from MC11 simulation data if this is not possible. A total of 4000 Toy experiments of 20k events have been generated and fitted. A Gaussian is then fitted to the pull distribution. The measured values of the pull width and mean from this set of Toy MC experiments are shown in Table 8.14.

Parameter	Fit value	Uncertainty
Pull Mean	-0.036	0.017
Pull Width	1.026	0.014

Table 8.14: Pull width and mean from 4k Toy MC Experiments.

The results presented in Table 8.14 are compatible with the absence of bias due to the fitter, both for the measured lifetime and the lifetime resolution.

8.6 Results of the B_s^0 Lifetime Measurement in the $B_s^0 \rightarrow D_s^- \pi^+$ Channel

As visible in the mass and decay time fits in Figures 8.9, 8.10 and 8.11, the fit describes the data well and there are no unexpected deviations from the model. Combining the fit results with the systematic uncertainties studies, the B_s^0 lifetime $\tau_{B_s^0}$ is measured to be,

$$\tau_{B_s^0} = 1.515 \pm 0.015(\text{stat})_{-0.039}^{+0.041}(\text{sys.}) [\text{ps}]$$

This result has a comparable precision with the combined result quoted by the Particle Data Group [1] measured using $B_s^0 \rightarrow D_s X$ decays

$$\tau_{B_s^0} = 1.425 \pm 0.041 \text{ ps}$$

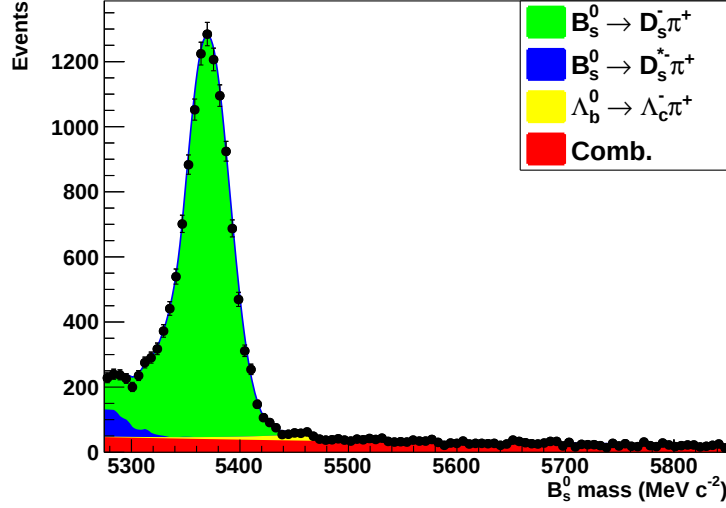


Figure 8.9: B_s^0 measured mass distribution projection of the 2D lifetime fit on real data, where the contribution from each component is displayed.

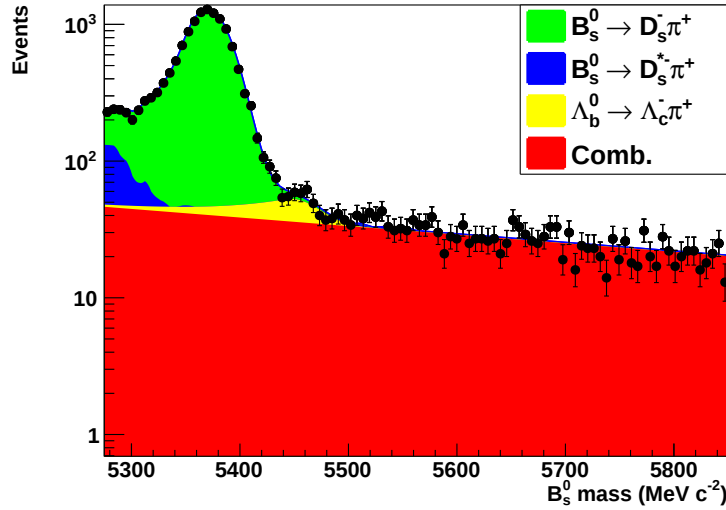


Figure 8.10: Log scale B_s^0 measured mass distribution projection of the 2D lifetime fit on real data. Contributions from individual channels are shown.

The main contribution to the overall resolution is the systematic uncertainty from the acceptance function and presents the limit to the precision of this method of measurement. The possibility exists to further improve this result through the use of per-event acceptance estimates.

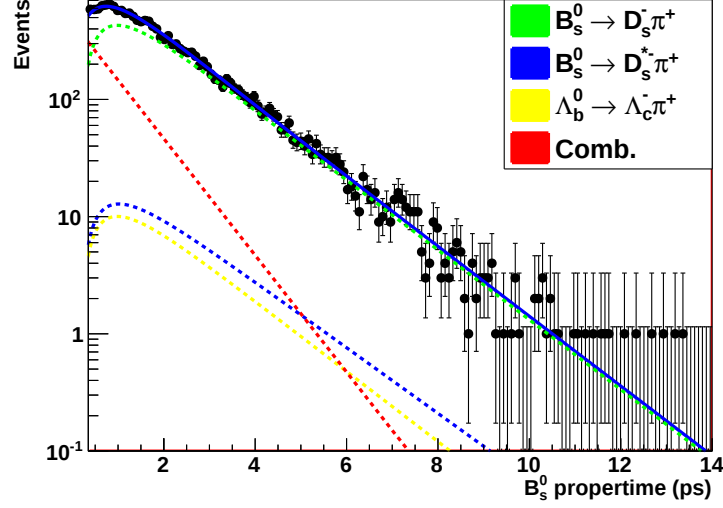


Figure 8.11: B_s^0 lifetime distribution projection of the 2D lifetime fit on real data. The shape and amplitude of the individual components is shown.

Parameter	Fit value	Statistical Uncertainty
B_s^0 Lifetime	$1.515 \cdot 10^0$	$0.015 \cdot 10^0$
σ_t Scale for B_s^0	$7.737 \cdot 10^{-1}$	$4.929 \cdot 10^{-1}$
B_s^0 Mass	$5.370 \cdot 10^3$	$0.000 \cdot 10^3$
B_s^0 CB Width	$1.970 \cdot 10^1$	$0.033 \cdot 10^1$
B_s^0 CB Switch to Exp.	$1.117 \cdot 10^0$	$0.039 \cdot 10^0$
Nbr. of $B_s^0 \rightarrow D_s^- \pi^+$	$1.245 \cdot 10^4$	$0.014 \cdot 10^4$
Λ_b Lifetime	$1.514 \cdot 10^0$	$0.183 \cdot 10^0$
Λ_b Mass	$5.448 \cdot 10^3$	$0.007 \cdot 10^3$
Nbr. of $\Lambda_b \rightarrow \Lambda_c^+ \pi^-$	$2.915 \cdot 10^2$	$0.585 \cdot 10^2$
$B_s^0 \rightarrow D_s^{*-} \pi^+$ Bkg. Lifetime	$1.623 \cdot 10^0$	$0.141 \cdot 10^0$
Nbr. of $B_s^0 \rightarrow D_s^{*-} \pi^+$	$3.918 \cdot 10^2$	$0.564 \cdot 10^2$
Comb. Bkg. Lifetime	$8.720 \cdot 10^{-1}$	$0.222 \cdot 10^{-1}$
σ_t Scale for Comb.	$7.228 \cdot 10^{-1}$	$3.111 \cdot 10^{-1}$
Comb. Bkg. Mass Slope	$-1.446 \cdot 10^{-3}$	$0.183 \cdot 10^{-3}$
Nbr. of Comb. Bkg.	$3.144 \cdot 10^3$	$0.097 \cdot 10^3$

Table 8.15: Final fit parameters and resolutions.

Chapter 9

B^0 Lifetime Measurement from $B^0 \rightarrow D^- \pi^+$

This chapter presents the measurement of the B^0 meson lifetime using the $B^0 \rightarrow D^- \pi^+$ with $D^+ \rightarrow K^- K^+ \pi^+$ decay channel. This mode is Cabibbo suppressed, but has the advantage of sharing the same selection used for the B_s^0 lifetime measurement. As for this previous lifetime measurement, all data used for this measurements has been recorded in 2011. The sequence of steps required to perform the measurement of the B^0 lifetime are listed below:

- Event reconstruction and selection, with the steps of the selection being the trigger, the stripping (offline pre-selection) and final selection. Section 9.1 will remind that these are identical to what is done in the B_s^0 measurement and may be found in Section 8.1 of Chapter 8.
- Studies on Monte Carlo simulation data are presented in Section 9.2. In particular, the determination of the mass shape parameters from MC11 data. The mass shapes for signal and backgrounds are determined along with the number of parameters to be fixed to attain the convergence of the fit.
- In Section 9.3 the acceptance function parameter values and their associated uncertainties are determined from MC11 pure signal data.
- Toy MC studies used to determine systematics in the measured decay time and decay time uncertainty form the body of Section 9.4. As for the B_s^0 measurement, the sets of Toy used for systematics contain 1000 Toys of 20k events.

- Final results with both statistical and systematic uncertainties are given in Section 9.5.

9.1 $B^0 \rightarrow D^- \pi^+$ Event Reconstruction and Selection

As explained in Section 8.1, the event reconstruction and selection is shared by the B_s^0 and B^0 measurements, and therefore they will not be described again. They can be found in Section 8.1 of Chapter 8.

9.2 Studies on Monte Carlo Simulated Events

The B^0 lifetime fit uses information from the MC distribution of the candidate invariant mass, decay time, decay time uncertainty and decay time acceptance. As the fitter has been described in the previous chapter, the focus will be on the shape of invariant mass used for the signal and backgrounds.

9.2.1 Monte Carlo Event Samples

Signal and background characterisation is done with MC11 datasets. As for the B_s^0 lifetime study, the Monte Carlo truth information of events passing the stripping selection (pre-selection) and offline selection is checked. Candidates formed from ghost tracks (tracks formed from random hits), track clones or tracks originating from the primary vertex as well as spillover events are discarded.

The initial characterisation of the background and signal distributions was performed using 2011 Monte Carlo datasets (MC11). The datasets were generated using the conditions under which the bulk of the data collected during 2011 was recorded. Each dataset is formed from two subsamples corresponding to the two magnet polarities, in all cases the subsamples contain near-equal amounts of events.

The Monte Carlo samples used are the divided in three categories, partially reconstructed physics backgrounds, resonant backgrounds and signal. The partially reconstructed physics backgrounds are $B^0 \rightarrow \rho^- \pi^+$ and $B^0 \rightarrow D^{*-} \pi^+$. The potential resonant

backgrounds are B_s⁰ → D_s⁻π⁺, B_s⁰ → D⁻π⁺, B⁰ → D⁻K⁺ and Λ_b → Λ_c⁺π⁻. The signal used is B⁰ → D⁻π⁺.

After running the selection on Monte Carlo sample and considering the branching fraction, it appeared that only a small subset of these decays were actually relevant backgrounds. The first reason is that the lower mass boundary for the fit is 5200 MeV, which is high enough to exclude both the B⁰ → ρ⁻π⁺ and B⁰ → D^{*-}π⁺ partially reconstructed backgrounds from the fitting range. The second reason is that in the case of the resonant backgrounds B_s⁰ → D_s⁻π⁺, B_s⁰ → D⁻π⁺ and Λ_b → Λ_c⁺π⁻, the combination of the selection efficiency and branching fraction information lead to a negligible amount of backgrounds with less than 10 events expected for each. The only remaining potential background is the B⁰ → D⁻K⁺ decay with the kaon mistaken as a pion.

9.2.2 Mass Distributions

The mass models used for each relevant channel are now presented. Functional forms are used for the peaking distributions and for the combinatorial background. In all other cases, a histogram of the MC mass distribution is used.

B⁰ → D⁻π⁺ Signal Mass Model

The invariant mass distribution of the B⁰ → D⁻π⁺ signal and conjugate decay is modelled using the sum of a Crystal ball with a Gaussian. The Crystal ball distribution shares the same mean as the Gaussian distribution. The exponential part of the Crystal ball distribution helps fit the low mass tail. A fit of this model to the Monte Carlo signal data is shown in Figure 9.1, the parameters which define the distribution are given in Table 9.1.

Peaking Background Model

The only significant peaking backgrounds in the mass region is the B⁰ → D⁻K⁺ decays. The model used for the B⁰ → D⁻K⁺ background is also a Crystal ball with a Gaussian, where the Crystal ball tail is used to describe the low mass tail of the background.

It should be noted that since a cut on the PID is implemented in the selection, the

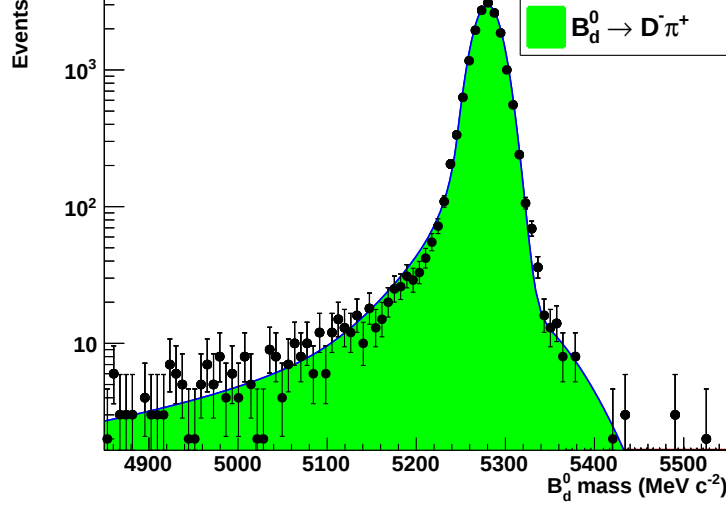


Figure 9.1: Fit of the mass resolution for $B^0 \rightarrow D^- \pi^+$ signal on Monte Carlo pure signal data.

Parameter	Fit Value	Fit Uncertainty
Centre of Distribution	$5.27992 \cdot 10^3$	$1.4 \cdot 10^{-1}$
Width of CB Component	$1.558 \cdot 10^1$	$1.30 \cdot 10^0$
Exponent of CB Tail	$1.255 \cdot 10^{-1}$	$7.47 \cdot 10^{-2}$
CB Switch to Exponential	$2.119 \cdot 10^0$	$5.19 \cdot 10^{-1}$
Width of Gaussian Component	$6.727 \cdot 10^1$	$4.36 \cdot 10^0$
Fraction of CB Component	$9.707 \cdot 10^{-1}$	$3.8 \cdot 10^{-3}$

Table 9.1: The parameters which define the $B^0 \rightarrow D^- \pi^+$ invariant mass distribution.

number of $B_s^0 \rightarrow D_s^- K^+$ MC candidates passing the selection is very low. In order to properly model the mass distribution, the number of events is increased by removing the bachelor PID cut. The result of the fit with the Crystal ball plus Gaussian model is summarized in Table 9.2.

Due to the low number of expected events in this channel relative to the signal ($\mathcal{O}(1\%)$ from Tables 8.3 and 9.3), coupled with the fact that this PDF distribution overlaps significantly with the signal both in decay time and mass, the fit is performed without using contribution from $B^0 \rightarrow D^- K^+$ model. The effect of this decision is accounted for

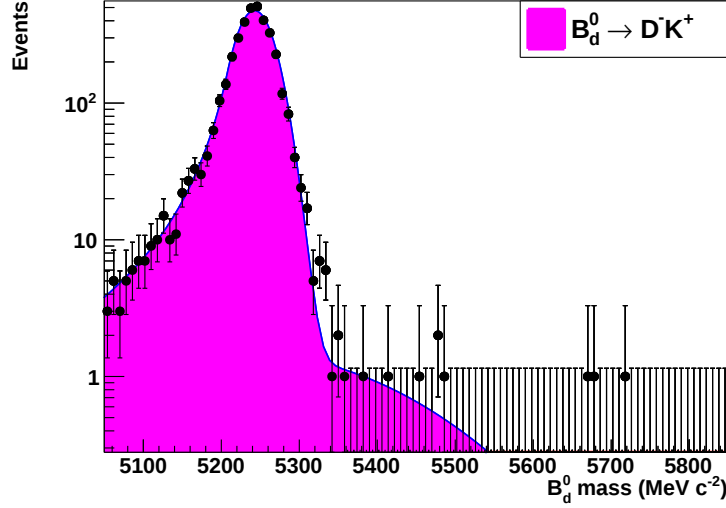


Figure 9.2: Fit of the mass resolution model for B⁰ → D⁻K⁺ background on Monte Carlo pure B⁰ → D⁻K⁺ background data.

Description	Fit Value	Fit Uncertainty
Centre of Distribution	$5.24253 \cdot 10^3$	$4.9 \cdot 10^{-1}$
Width of CB Component	$2.338 \cdot 10^1$	$0.46 \cdot 10^0$
Exponent of CB Tail	$2.525 \cdot 10^0$	$4.91 \cdot 10^{-1}$
CB Switch to Exponential	$1.443 \cdot 10^0$	$8.69 \cdot 10^{-2}$
Width of Gaussian Component	$1.660 \cdot 10^2$	$2.41 \cdot 10^1$
Fraction of CB Component	$9.828 \cdot 10^{-1}$	$5.0 \cdot 10^{-3}$

Table 9.2: The parameters which define the B⁰ → D⁻K⁺ invariant mass distribution.

in the systematic studies discussed in Section 9.4.

Combinatorial Background

The model used for the combinatorial background in B⁰ lifetime measurement is the same used in the B_s⁰ measurement, that is simple exponential with negative exponent. This background accounts for all events having a mass above the B⁰ mass peak.

Name	Generated	Candidates	Selected	Efficiency/10 ⁻²
B ⁰ → D ⁻ π ⁺	1014995	45507	17489	1.723
B ⁰ → D ^{*-} π ⁺	202798	8767	3410	1.681
B ⁰ → D ⁻ K ⁺	2005992	87179	3928	0.196
B ⁰ → ρ ⁻ π ⁺	206799	2618	1037	0.501
B _s ⁰ → D _s ⁻ π ⁺	1012495	8456	68	0.007
B _s ⁰ → D _s ⁻ K ⁺	1887293	15672	16	0.001
Λ _b → Λ _c ⁺ π ⁻	2033496	26008	100	0.005
Λ _b → Λ _c ⁺ π ⁻ (no veto)	2033496	26008	1227	0.060

Table 9.3: Selection efficiency for signal and backgrounds, with kaon PID cut.

9.3 Decay Time Distributions and Acceptance

In order to correctly fit the decay time distribution of the B⁰ meson, it is necessary to introduce the decay time acceptance function in the measurement. The function used is has the same form as the one used in the B_s⁰ lifetime measurement, but the parameters values are different. The concept of acceptance and the chosen method to determine its shape and parameters have been explained in the B_s⁰ lifetime measurement Chapter 8, and therefore they will not be presented here. The result obtained for the acceptance parameters values are contained in Subsection 9.3.1.

In both lifetime measurements, each candidate is also supplied with an estimate of the resolution achieved on the measured decay time. This information is used to carry the basic decay time distribution on an event-by-event basis. This per-event decay time resolution in the B⁰ is discussed in Section 9.3.2.

9.3.1 Decay Time Acceptance

The acceptance function used for this B⁰ lifetime measurement is identical to the one used and defined in the B_s⁰ measurement procedure,

$$\text{acc}(t_{\text{reco}}) = \text{Max}(0, (1 - tS_2)(1 - e^{-S_1(t-t_P)})). \quad (9.1)$$

In Equation 9.1 the S_2 parameter correspond to the long lifetime acceptance component, and is linear slope. The t_p parameter give the turning point where the function takes a non zero value, and the S_1 parameter is the exponential slope controlling the rise of the function to 1. The acceptance parameters are obtained from a fit on Monte Carlo MC11 B⁰ → D⁻π⁺ signal data with the lifetime fixed at generation value. The results of this fit are summarised in Table 9.4.

Parameters	Value	Uncertainty
t_P	0.1921	0.0231
S_1	1.9928	0.2262
S_2	0.0238	0.0103

Table 9.4: Acceptance function parameters as determined from a fit to the MC11 signal data.

The signal acceptance is applied to all channels, except the combinatorial background. Any biases to the lifetime of the background channels introduced are compensated by a change in the value of the considered background lifetime. The decay time acceptance parameters obtained for the B⁰ → D⁻π⁺ signal are different to the ones obtained on the B_s⁰ → D_s⁻π⁺ signal. The exception is the turning point which are fully compatible. These differences in the slope parameter values imply that it will not be possible to cancel the systematics linked to the acceptance slopes with a lifetime ratio.

9.3.2 Per-Event Decay Time Uncertainty

The treatment of the per-event decay time uncertainty in the B⁰ lifetime measurement is identical to what is done for the B_s⁰ lifetime. In summary, the decay time uncertainty is introduced in a per-event basis. The decay time resolution model used is a single Gaussian. As using the decay time uncertainty in a per-event basis makes it a conditional variable for the fit, it is necessary to account for the distribution of this conditional parameter for the signal and background. Further details and justifications can be found in Subsection 8.3.2, which is based on the paper [49].

In this analysis, the decay time uncertainty distributions are taken from Monte Carlo as they reproduce the real data very well as shown in Figure 9.3.

Given the very similar distributions for the signal and the peaking backgrounds, this

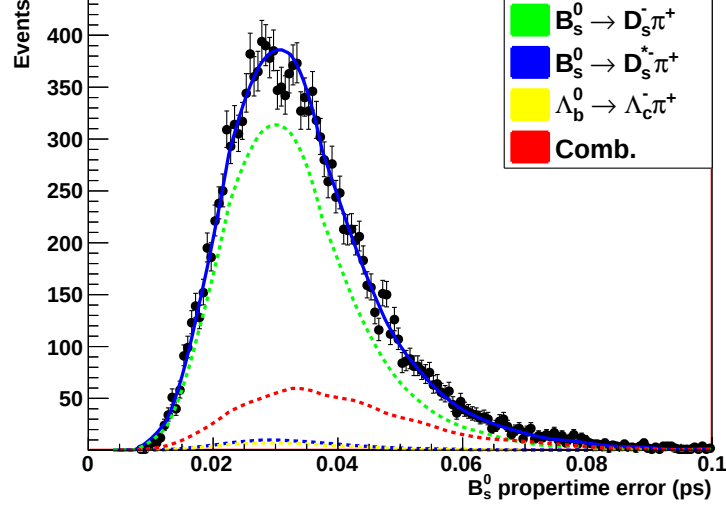


Figure 9.3: τ_{err} distributions compared to data.

analysis uses two decay time uncertainty distributions : one for the signal and another for the combinatorial background. These PDFs are conditional PDF for the 2D mass and decay time PDF. This leads to a PDF of the type shown in Equation 9.2 for each class of events.

$$(Mass \times \tau | \tau_{\text{Err}}). \quad (9.2)$$

9.4 Systematic Uncertainties and Fitter Bias

This section contains the different sources of systematic uncertainties considered and their estimated values. Systematic uncertainties from the fitting model are presented in Subsection 9.4.1 and the acceptance corrections in Subsection 9.4.2. The complete summary of the systematic uncertainties from the different sources is given in subsection 9.4.4.

9.4.1 Systematic Uncertainties due to the Mass Models

The systematic uncertainties are estimated for the signal mass model. The systematic uncertainty estimation is performed for each parameter of the signal mass model that

has been determined using Monte Carlo simulation data. The relevant parameters are listed in Table 9.5.

Model	Parameters
B_s^0 Signal	Exponent of CB tail
	Fraction of the CB
	Width of the gaussian

Table 9.5: Fixed parameters of the mass models.

The signal mass model is the sum of a Crystall ball and Gaussian. The gaussian part is used to pick up a small upper mass tail. The parameters of the Gaussian and tail of the Crystal ball are determined from pure signal Monte Carlo simulation data.

Two sets of Toy Monte Carlo experiments are performed for each of the parameters listed in Table 9.5. Each Toy MC set contains a total of 1000 experiments, where each experiment has a total of 20k events, which correspondd approximately to the statistics available in data.

In each of the 1000 experiments, the generated model is based on the fit PDF where parameters determined on MC11 data are changed by 1σ deviation. The same fitting algorithm used for the B_s^0 lifetime extraction is used to fit these generated data sets. The means of the B^0 lifetime measured with the respective $+1\sigma$ and -1σ deviation of the studied parameter are measured. These two values provide the upper and lower bounds of the systematic uncertainty on the B^0 lifetime for the considered parameter. The results obtained from the Toy MC experiments are shown in Table 9.6.

Model	Parameters	Lower sys. [ps]	Upper sys. [ps]
Signal	Sig. CB tail	0.0003	0.0008
	Sig. CB alpha	0.0010	0.0011
	Sig. CB frac.	0.0000	0.0006
	Sig Gaus. width	0.0010	0.0003

Table 9.6: Systematic uncertainties due to fixed parameters of the mass models.

In the limit of the the precision available from the number of Toy MC experiments,

these parameters are consistent with zero. The largest observed deviation is therefore taken as the systematic uncertainty to be conservative.

9.4.2 Systematic Uncertainty from Acceptance Model

The acceptance model is given by

$$\text{Max}(0, (1 - tS_2)(1 - e^{-S_1(t-t_P)})). \quad (9.3)$$

The turning point t_P , represents the time after which the acceptance function takes a non-zero value. The S_1 slope determines the gradient of the initial turn on and S_2 is the slope of the long lifetime acceptance. The results of the Toys are summarised in Table 9.7.

The value of the three t_P , S_1 and S_2 parameters were determined using MC11 data for pure signal as described in 9.3.1. The procedure used to determine the systematic uncertainties from these parameters is identical to the one used for the mass models in Subsection 9.4.1. The systematic uncertainties on the lifetime obtained with this method for t_P , S_1 and S_2 are presented in Table 9.7.

Model	Parameter	Lower sys. [ps]	Upper sys. [ps]
Acceptance	t_P	0.0111	0.0125
	S_1	0.0247	0.0299
	S_2	0.0255	0.0057

Table 9.7: Systematic uncertainties due to acceptance function.

9.4.3 Other Systematic Uncertainties

As in B_s^0 lifetime measurement, the momentum scale and the decay length scale are both source of systematic uncertainties. Both values can be found in Table 9.8 and are taken from [51].

The momentum scale has been calibrated to a precision of $0.05 \pm 0.02\%$ [52]. The

proper time expression is given by

$$t = \frac{d m}{p}. \quad (9.4)$$

In this equation d represents the flight distance, m the reconstructed mass and p the momentum of the candidate. As the reconstructed mass m depends on the momentum, the momentum scale contribution to the systematic uncertainty is small as the uncertainty on p partially cancels in the ratio $\frac{m}{p}$.

The decay length scale is a representation of the uncertainty on the measurement of the B_s⁰ flight distance along the z axis. It is described in details in [51]. The relative uncertainty on the lifetime is calculated to be 0.0026%.

The last systematic uncertainty is due to the presence of B⁰ → D⁻K⁺ decays which is not included in the fit. This uncertainty is estimated by adding a PDF component corresponding to the fitted Monte Carlo B⁰ → D⁻K⁺ distributions. The amplitude is fixed to the number of expected events, which is found to be 70 expected events, and is determined by combining branching fraction information, the selection efficiency for the B⁰ → D⁻K⁺ decay and the mass window of the fit. This PDF shape the lifetime parameter with the signal PDF. The systematic uncertainty is taken as the difference in measured lifetime between the fit with and without this component. It has been estimated to be ± 0.00014 [ps].

9.4.4 Summary of the Systematic Uncertainties

The systematic uncertainties determined in Subsections 9.4.1, 9.4.2 and 9.4.3 are summarized here in Table 9.8 and are then combined taking into account the correlations between parameters.

Correlations between parameters are accounted for by applying a transition matrix that diagonalizes the correlation matrix to the uncertainty vector. The method is detailed in the statistics appendix of the Particle Data Group booklet [1]. The final systematic uncertainty is calculated by summing the magnitude of the uncertainties in the eigenbase in quadrature.

Parameter	Lower sys. [ps]	Upper sys. [ps]
Sig. CB tail	0.0003	0.0008
Sig. CB alpha	0.0010	0.0011
Sig. CB frac.	0.0000	0.0006
Sig Gaus. Width	0.0010	0.0003
Total sig. model	± 0.00178	
t_P	0.0111	0.0125
S_1	0.0247	0.0299
S_2	0.0255	0.0057
Total acc. Model	0.0372	0.0329
Momentum scale	0.0008	0.0008
Decay length scale	0.0000	0.0000
Expected B ⁰ → D ⁻ K ⁺	0.0001	0.0001
Total other	± 0.0008	
Total systematic error	0.0372	0.0330

Table 9.8: Summary of all systematic uncertainties.

9.5 Results of the B⁰ Lifetime Measurement in the B⁰ → D⁻π⁺ Channel

As visible in the mass and decay time fits in Figures 9.4, 9.5 and 9.6, the fit describes the data well and there are no unexpected deviations from the model. Combining the fit results with the systematic uncertainties studies, the B⁰ lifetime τ_{B^0} is measured to be,

$$\tau_{B^0} = 1.513 \pm 0.020(\text{stat})_{-0.037}^{+0.033}(\text{sys}) [\text{ps}]$$

The main contribution to the overall resolution is the systematic uncertainty from the acceptance function and presents the limit to the precision of this method of measurement. The possibility exists to further improve this result through the use of per-event acceptance estimates. Also, the decay mode used for this measurement is Cabibbo suppressed, and the available statistics may be greatly increased by using D⁺ → K⁻π⁺π⁺

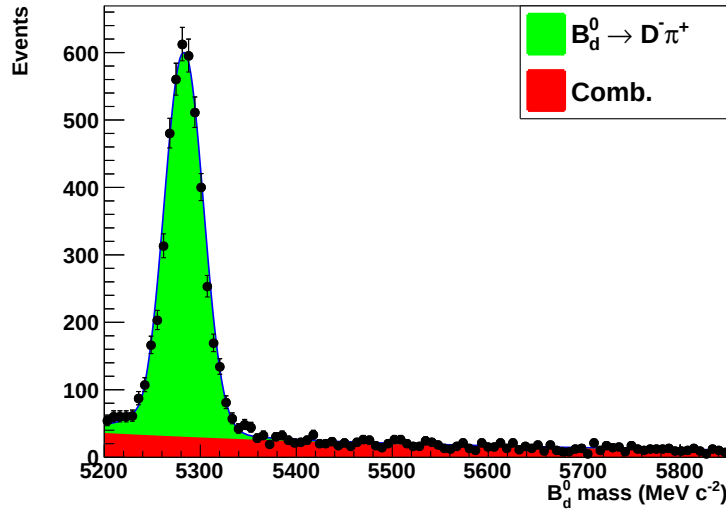


Figure 9.4: B^0 measured mass distribution projection of the 2D lifetime fit on real data, where the contribution from each component is displayed.

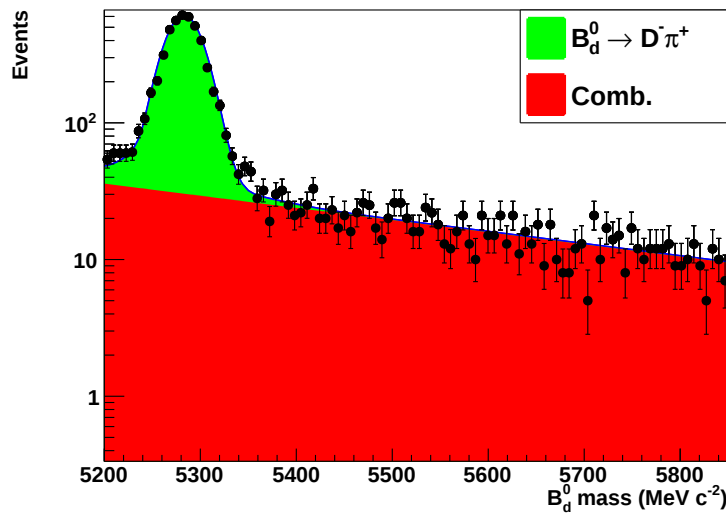


Figure 9.5: Log scale B^0 measured mass distribution projection of the 2D lifetime fit on real data. Contributions from individual channels are shown.

instead of $D^+ \rightarrow K^- K^+ \pi^+$, but the the B_s^0 and B^0 lifetime measurements would then use different selections.

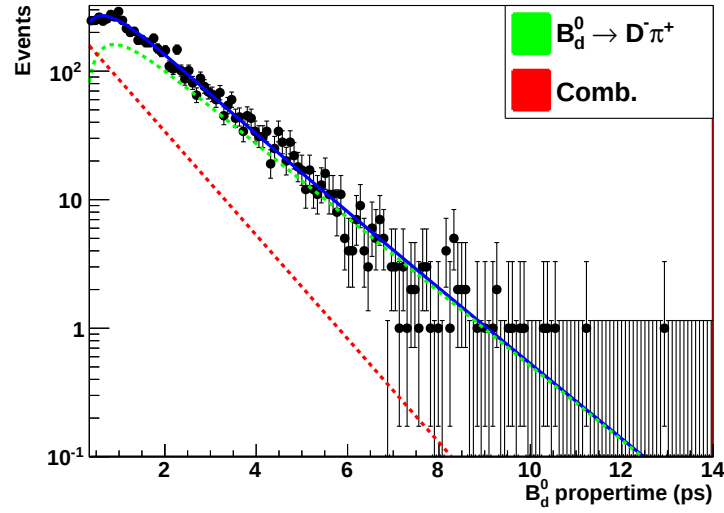


Figure 9.6: B^0 lifetime distribution projection of the 2D lifetime fit on real data. The shape and amplitude of the individual components is shown.

Parameter	Fit value	Statistical Uncertainty
B^0 Lifetime	$1.513 \cdot 10^0$	$0.020 \cdot 10^0$
σ_t Scale for B^0	$1.037 \cdot 10^0$	$0.337 \cdot 10^0$
B^0 Mass	$5.283 \cdot 10^3$	$0.000 \cdot 10^3$
B^0 CB Width	$1.958 \cdot 10^1$	$0.034 \cdot 10^1$
Nbr. of $B^0 \rightarrow D^- \pi^+$	$4.436 \cdot 10^3$	$0.076 \cdot 10^3$
Comb. Bkg. Lifetime	$1.077 \cdot 10^0$	$0.030 \cdot 10^0$
σ_t Scale for Comb.	$1.280 \cdot 10^0$	$0.584 \cdot 10^0$
Comb. Bkg. Mass Slope	$-2.049 \cdot 10^{-3}$	$0.159 \cdot 10^{-3}$
Nbr. of Comb. Bkg.	$1.985 \cdot 10^3$	$0.058 \cdot 10^3$

Table 9.9: Final fit parameters and resolutions.

Chapter 10

Conclusions

In order to be able to achieve its physics goals, the positions of the parts of the LHCb detector have to be known to a very high precision. Misalignments of the tracking system degrade the momentum measurement and flight distance determination of particles. Such quantities are vital for accurate lifetime and mass measurements.

A first standalone alignment of the Inner Tracker with no fixed elements was performed on first data. A method to stabilize the alignment of the Inner Tracker which does not require any elements to be fixed was described. The alignment obtained using this method was presented.

The final alignment results are that an overall Inner Tracker ladder residual resolution of $102 \pm 10 \mu\text{m}$ could be obtained, corresponding to a resolution due to alignment of $85 \pm 10 \mu\text{m}$, assuming an intrinsic resolution of $57 \mu\text{m}$, with a bias of $0 \pm 13 \mu\text{m}$.

A first optimization of the selection cuts of the Same Side Kaon tagger has been performed using $D_s^+ \rightarrow \phi\pi^+$ 2010 data. A selection for the Same Side Kaon tagger was developed using a combination of Monte Carlo and $D_s^+ \rightarrow \phi\pi^+$ data. The $D_s^+ \rightarrow \phi\pi^+$ channel was shown to be suitable for this purpose in the absence of a large B_s^0 calibration sample. Compared to the initial kaon selection, tuned on Monte Carlo data, an improvement both in tagging power and mistag rate is observed.

Finally, a method to measure the average B_s^0 and B^0 lifetimes was developed using a global decay time acceptance function and per-event decay time uncertainties. The precision of the fit results is found to be dominated by the systematic uncertainty due to the precision to which the acceptance function can be modelled. The average B_s^0 lifetime

is measured to be

$$\tau_{B_s^0} = 1.515 \pm 0.015(\text{stat})_{-0.039}^{+0.041}(\text{sys.}) [\text{ps}]$$

The result has an accuracy comparable to the world average provided by the Particle Data Group [1] using $B_s^0 \rightarrow D_s X$ decays

$$\tau_{B_s^0} = 1.425 \pm 0.041 \text{ ps}$$

For the average B^0 lifetime is measured to be

$$\tau_{B_d^0} = 1.513 \pm 0.020(\text{stat})_{-0.037}^{+0.033}(\text{sys}) [\text{ps}]$$

For both results, it is expected that the precision of this measurement can be improved by using per-event acceptance estimates in place of a global function.

Appendix A

CV

Vincent Fave

Bvd Henri-Plumhof 8, 1800 Vevey
 (+41) 78 652 60 56 or (+41) 21 921 86 62
 vincent.fave@gmail.com

Key Strengths

Physics PhD with good communication and teamwork skills,
 several years of work in international environment.

Education

Sep. 2007 – Apr. 2012	PhD in Particle Physics, EPFL Member of the LHCb experiment at CERN
Sep. 2006 – Sep. 2007	Master in Physics, 2nd year, Tsinghua University of Beijing Mandarin courses, Tsinghua intermediate level 1 exam passed. MSc. in Physics, final grade 5.31/6
Sep. 2005 – Sep. 2006	Master in Physics, 1st year, EPFL Specialization : quantum field physics, nuclear physics and graph theory.
Sep. 2002 – Sep. 2005	Bachelor in Physics, EPFL Specialization : quantum physics and mathematical physics.
Sep. 1999 – Sep. 2002	Maturité Fédérale, Gymnase de Burier

Professional Experience

Jul. 2005 – Aug. 2006	State tax office (Vaud) Summer job, auxiliary employee. Computation of legacy taxes, in particular high fortune cases.
Jan. 2004 – Jan. 2005	EPFL Assistant for 1 st year C++ programming courses.

PhD Projects

Jul. 2009 – Feb. 2012	B_s^0 and B^0 particle lifetime measurements Development of the measurement of the mean lifetimes of the B_s^0 and B^0 mesons.
May. 2009 – Aug. 2011	B_s^0 quantum state tagging by co-produced $K^\pm$ particles Lead the development of an innovative calibration method. Performances increased by 70% with new method.
Sep. 2007 – Dec. 2010	LHCb, Inner Tracker alignment strategy Inner Tracker elements position determined to $82\mu m$ with first data. Final result corresponds to a 38% improvement in alignment resolution.
Apr. 2010	Vietnam, Collaboration with Hanoi University Deployment of LHCb physics software and instruction of the users.
Sep. 2007 – Nov. 2008	LHCb, Inner Tracker cooling control system Development of the system, inclusion in the global control system.

Languages

French (native), English (fluent), Mandarin (social), German (basic)

Software

Matlab, Statistical Analysis (ROOT), Industrial Control (PVSS)
 programming in C, C++, python and perl.

Interests

Sep. 2006 – Oct. 2011	Chinese Language, culture and cooking.
Mar. 2001 – Oct. 2011	Organization of boardgames and role playing sessions.

References

- [1] K. Nakamura, W.-M. Yao, C. Amsler, D. Asner, and B. et al., “Review of Particle Physics,” *Journal of Physics G*, vol. 37, pp. 1+, 2010.
- [2] W. Kienzle, M. Oriunno, A. L. Perrot, S. Weisz, M. Bozzo, A. Buzzo, M. Macri, A. Santroni, G. Settle, M. Buénerd, F. Malek, Y. Muraki, K. Kasahara, G. Sanguinetti, G. Matthiae, P. Privitera, V. Verzi, A. Faus-Golfe, J. Velasco, and S. Torii, “TOTEM, Total Cross Section, Elastic Scattering and Diffraction Dissociation at the LHC Technical Proposal,” Tech. Rep. CERN-LHCC-99-007. LHCC-P-5, CERN, Geneva, Mar 1999.
- [3] O. Adriani, L. Bonechi, M. Bongi, R. D’Alessandro, D. A. Faus, M. Haguenauer, Y. Itow, K. Kasahara, K. Masuda, Y. Matsubara, H. Menjo, Y. Muraki, P. Papini, T. Sako, T. Tamura, S. Torii, A. Tricomi, W. C. Turner, J. Velasco, and K. Yoshida, *LHCf experiment Technical Design Report*. Technical Design Report LHCf, Geneva: CERN, 2006.
- [4] *ALICE Technical Proposal for a Large Ion Collider Experiment at the CERN LHC*. LHC Tech. Proposal, Geneva: CERN, 1995.
- [5] A. A. Alves *et al.*, “The LHCb Detector at the LHC,” *JINST*, vol. 3, p. S08005, 2008.
- [6] G. Aad *et al.*, “The ATLAS Experiment at the CERN Large Hadron Collider,” *JINST*, vol. 3, p. S08003, 2008.
- [7] R. Adolphi *et al.*, “The CMS experiment at the CERN LHC,” *JINST*, vol. 3, p. S08004, 2008.
- [8] “Roadmap for selected key measurements of lhcb,” Feb 2010. LHCb-PUB-2009-029.

- [9] T. Sjöstrand, “High Energy physics Event Generation with PYTHIA 6.1.” Computer Physics Communication 135 (2001) 238.
- [10] “LHCb Reoptimized Detector, Design and Performance,” Sep 2003. CERN/LHCC 2003-030.
- [11] T. L. Collaboration, “The RICH Technical Design Report,” 2000. CERN/LHCC 2000-037.
- [12] T. L. Collaboration, “The Trigger System Technical Design Report,” 2003. CERN/LHCC 2003-031.
- [13] T. L. Collaboration, “The LHCb Simulation Program.”
- [14] “The EvtGen Package.”
- [15] “GEANT4, a Toolkit for the Simulation of the Passage of Particles through Matter.”
- [16] “The Moore Project.”
- [17] “The Brunel Project.”
- [18] “ROOT, a data analysis framework.”
- [19] “The DaVinci Project.”
- [20] M. E. Peskin and D. V. Schroeder, *An Introduction To Quantum Field Theory (Frontiers in Physics)*. Westview Press, 1995.
- [21] M. Maggiore, *A modern introduction to quantum field theory*. Oxford master series in physics, Oxford University Press, 2005.
- [22] L. Wolfenstein, “Parametrization of the Kobayashi-Maskawa Matrix,” *Phys. Rev. Lett.*, vol. 51, pp. 1945–1947, Nov 1983.
- [23] I. I. Y. Bigi and A. I. Sanda, *CP violation*. Cambridge monographs on particle physics, nuclear physics, and cosmology, Cambridge: Cambridge Univ. Press, 2000.
- [24] D. Hutchcroft, “VELO Pattern Recognition,” Tech. Rep. LHCb-2007-013. CERN-LHCb-2007-013, CERN, Geneva, Mar 2007.
- [25] T. Lastoviicka, “Generic VELO Pattern Recognition,” Tech. Rep. LHCb-2007-002. CERN-LHCb-2007-002, CERN, Geneva, Feb 2008.

-
- [26] O. Callot and M. Schiller, “PatSeeding: A Standalone Track Reconstruction Algorithm,” Tech. Rep. LHCb-2008-042. CERN-LHCb-2008-042, CERN, Geneva, Aug 2008.
 - [27] J. Albrecht and M. Schiller, “Fast T-Station Seeding for HLT1,” Tech. Rep. LHCb-2009-012. CERN-LHCb-2009-012, CERN, Geneva, Jan 2009.
 - [28] O. Callot, “Online Pattern Recognition,” Tech. Rep. LHCb-2004-094. CERN-LHCb-2004-094, CERN, Geneva, Oct 2004.
 - [29] M. Needham, “Performance of the track matching,” Tech. Rep. LHCb-2007-129. CERN-LHCb-2007-129, CERN, Geneva, Oct 2007.
 - [30] G. Gracia, M. Merk, W. Rückstuhl, and R. Van der Eijk, “Track reconstruction for LHCb,” Tech. Rep. LHCb-98-045, CERN, Geneva, Feb 1998.
 - [31] R. E. Kalman, “A new approach to linear filtering and prediction problems,” *Journal Of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.
 - [32] LHCb Collaboration, “LHCb Inner Tracker TDR,” Tech. Rep. CERN-LHCC-2002-029, CERN, 2002.
 - [33] V. Blobel and C. Kleinwort, “Conference on Advanced Statistical Techniques in Particle Physics,” Tech. Rep. CERN-ATL-INDET-PUB-2005-002, CERN, Geneva, March 2002.
 - [34] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins University Press, third ed., 1996.
 - [35] K. Chen, *Matrix preconditioning techniques and applications*. Cambridge: Cambridge University Press, 2005.
 - [36] A. Bocci and W. Hulsbergen, “TRT Alignment For SR1 Cosmics and Beyond,” Tech. Rep. ATL-INDET-PUB-2007-009. ATL-COM-INDET-2007-011. CERN-ATL-COM-INDET-2007-011, CERN, Geneva, Jun 2007.
 - [37] L. Nicolas, “Alignment of the LHCb Tracking Stations and Selection of X(3872) and Z(4430) in pp Collisions at 14 TeV,” Jun 2009.
 - [38] “Measurement of Δm_s in the decay $B_s^0 \rightarrow D_s^-(K^+K^-\pi^-)(3)\pi$,” Apr 2011. LHCb-ANA-2011-005.

- [39] M. Calvi, O. Leroy, and M. Musy, “Flavour Tagging Algorithms and Performances in LHCb,” Tech. Rep. LHCb-2007-058. CERN-LHCb-2007-058, CERN, Geneva, May 2007.
- [40] “Optimization and calibration of the LHCb flavour tagging performance using 2010 data,” Mar 2011. LHCb-ANA-2011-003.
- [41] J. A. et Al., “Road map for the measurement of mixing induced CP violation in $B_s^0 \rightarrow J/\psi\phi$ at LHCb,” Feb 2009.
- [42] “Tagged time-dependent angular analysis of $B_s^0 \rightarrow J/\psi\phi$ decays with the 2010 LHCb data,” May 2011. LHCb-ANA-2011-006.
- [43] O. Awunor and U. Egede, “Using Double Tagging to measure the Performance of the Same Side Kaon Tagger in Data,” Tech. Rep. LHCb-2007-127. CERN-LHCb-2007-127, CERN, Geneva, Oct 2007.
- [44] G. Krocker, “Presentation at Delta ms and Delta md mixing meeting,” Feb 2010.
- [45] M. Calvo, M. Grabalosa, and M. Musy, “Combination of same-side with opposite-side flavour tagging,” Tech. Rep. LHCb-PUB-2009-027. CERN-LHCb-PUB-2009-027, CERN, Geneva, Feb 2010.
- [46] C. Jones, “Charged Particle ID.” Private communication (2012), 2012.
- [47] M. Gersabeck, V. V. Gligorov, J. Imong, and J. Rademacker, “A Monte Carlo simulation free method of measuring lifetimes using event-by-event acceptance functions at LHCb,” Tech. Rep. LHCb-PUB-2009-022. CERN-LHCb-PUB-2009-022, CERN, Geneva, Nov 2009.
- [48] V. V. Gligorov and J. Rademacker, “Monte Carlo Independent Lifetime Fitting at LHCb in Lifetime Biased Channels,” Tech. Rep. LHCb-2007-053. CERN-LHCb-2007-053, CERN, Geneva, Jun 2007.
- [49] G. Punzi, “Comments on Likelihood fits with variable resolution,” 2004.
- [50] M. Pivk and F. R. Le Diberder, “splot: a statistical tool to unfold data distributions,” Tech. Rep. physics/0402083. LAL-2004-07, Paris 11. Lab. Accél. Linéaire, Orsay, Feb 2004.
- [51] “ b -hadron lifetime measurements with exclusive $b \rightarrow J/\psi X$ decays reconstructed in the 2010 data,” Mar 2011. LHCb-ANA-2011-001.

-
- [52] M. Needham, “Momentum scale calibration with first LHCb data,” Tech. Rep. LHCb-INT-2011-012. CERN-LHCb-INT-2011-012, CERN, Geneva, Apr 2011.