Alignment of the LHCb Vertex Locator and Lifetime Measurements of Two-Body Hadronic Final States

Marco Gersabeck
Dipl. Phys.

Submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

Department of Physics and Astronomy
Faculty of Physical Sciences

September 2009

Abstract

Lifetime measurements offer excellent opportunities for precision tests of the Standard Model of Particle Physics as well as for discovery of effects involving particles beyond the Standard Model. This thesis presents a method for measurements of lifetimes and lifetime ratios and its application to two-body hadronic final states of heavy flavour decays at *LHCb*.

The *LHCb* experiment is designed to measure heavy flavour particle decays produced in proton-proton collisions at the *LHC*. Key to high quality vertexing is the spatial alignment of the Vertex Locator. The algorithms designed for this task, including a novel approach for the relative sensor alignment, are discussed in detail. Their performance is presented using test beam data as well as data using the first beam induced tracks from *LHC*. The precision of these algorithms is found to be of the order of $1-2$ μm.

A method for lifetime fitting using a Monte Carlo independent approach to determine a lifetime acceptance function on an event-by-event basis is presented. These acceptance functions are crucial to account for a bias caused by the trigger selection. The un-binned maximum likelihood fitter based on this method does not rely on a parametrised model for the lifetime distribution of combinatorial background.

The fit of the lifetime measured in $B_s^0 \to K^+K^-$ decays using a simulated data sample equivalent to an integrated luminosity of 0.1 fb^{-1} would yield $\tau(B_s^0 \to K^+K^-) = (1.498 \pm 0.030_{stat.} \pm 0.005_{syst.})$ ps with an average input lifetime of 1.500 ps. A competitive measurement of $\Delta\Gamma_s$ extracted from the $B_s^0 \to K^+K^-$ lifetime measurement would require a data set equivalent to about 0.7 fb^{-1} of luminosity.

With an integrated luminosity of only about 0.03 fb^{-1} it will be possible to make a competitive measurement of the D mixing parameter $y_{\mathcal{CP}}$. This uses a lifetime ratio measurement with prompt $D^0 \to h^+h'^-$ decays. A first event selection for prompt $D^0 \to h^+h'^-$ decays is presented. The major hurdle for this measurement is the contribution from secondary D decays. Possible solutions are discussed.

Contents

List of Figures

List of Tables

Preface

Flavour physics started with the discovery of the kaon in 1947. Since then it has led to many discoveries such as $\mathcal{CP}$ violation in the K^0, D^0, B_d^0, and B_s^0 systems. The high impact of these measurements is shown by the prediction of three quark families as a consequence of the observation of $\mathcal{CP}$ violation in the K^0 system at a time when only three quarks were known.

LHCb will open a new chapter in flavour physics with studies of unprecedented data sets of heavy flavour particle decays. Lifetime measurements at *LHCb* offer excellent opportunities for precision tests of the Standard Model of Particle Physics as well as for discovery of effects involving particles beyond the Standard Model. This thesis presents a method for measurements of lifetimes and lifetime ratios and its application to two-body hadronic final states of heavy flavour decays.

The first chapter of this thesis gives an introduction to the Standard Model of particle physics and the aspects of flavour physics relevant to this work. It focuses particularly on lifetime and lifetime ratio measurements. Lifetime ratio measurements like $\tau(B_s^0)/\tau(B_d^0)$ allow precision tests of Standard Model predictions by the Heavy Quark Expansion formalism. It is discussed how the lifetime measured in the decay $B_s^0 \to K^+K^-$ can be used for the extraction of $\Delta\Gamma_s$. To date, no precision measurement exists for $\Delta\Gamma_s$, which may reveal New Physics effects when measured in this channel. Lifetime measurements with D decays are discussed, which yield a measurement of the $\mathcal{CP}$ violation quantity $y_{\mathcal{CP}}$.

The *LHCb* experiment is designed to measure heavy flavour particle decays produced in proton-proton collisions at the *LHC*. Both the accelerator complex and the experiment are presented in detail in the second chapter. With its precision Vertex Locator and the two Ring Imaging Cherenkov detectors it is particularly suited for lifetime measurements involving hadronic final states. A particular focus is given to the Vertex Locator which plays a central role in this thesis.

Key to high quality vertexing is the spatial alignment of the Vertex Locator. The algorithms designed for this task are presented in the third chapter. Novel methods had to be exploited due to the design of the Vertex Locator with r and ϕ measuring silicon

strip sensors. The alignment of the Vertex Locator is split in three steps: the relative alignment of the two sensors on a module, the relative alignment of the modules in each half of the Vertex Locator, and the alignment of the two halves with respect to each other. Their performance is presented using test beam data as well as data from *LHC* injector commissioning test which resulted in secondary muons being recorded by *LHCb*. A study of the impact of misalignments of the tracking system is presented in section 5.4, which concludes that the remaining misalignments after application of the alignment algorithms will not have any deteriorating effects on lifetime measurements.

The fourth chapter presents a method designed for lifetime measurements using a Monte Carlo independent approach. It is discussed how the lifetime bias caused by the trigger selection can be accounted for using event-by-event acceptance functions. These are determined exploiting a data driven technique with an interface to the trigger software. The fitter based on this method uses a two-stage un-binned maximum likelihood fit. In the first stage the signal fractions are fitted and the second stage is the actual lifetime fit. The strength of the fitter is that it does not rely on a parametrised model for the lifetime distribution of combinatorial background.

The main physics aim of this thesis is the preparation of a measurement of $\Delta\Gamma_s$ with $B_s^0 \to K^+K^-$ decays. The extraction of $\Delta\Gamma_s$ from a lifetime measurement using $B_s^0 \to K^+K^-$ decays is discussed in the fifth chapter. The lifetime fitter has been tested extensively with toy MC simulation data. Its results for fits of the B_d^0 lifetime, the $B_s^0 \to K^+K^-$ lifetime, and of the B_s^0 to B_d^0 and Λ_b^0 to B_d^0 lifetime ratios are presented. Tests of potential systematic effects are discussed concluding in the potential sensitivity to $\Delta\Gamma_s$ of *LHCb*.

Due to the copious production of D mesons and the high branching ratio of $D^0 \to h^+h'^-$ decays, these are prime candidates for early measurements at *LHCb*. First studies of the measurement of $y_{\mathcal{CP}}$ from lifetime ratio measurements in $D^0 \to h^+h'^-$ decays are presented. A first event selection for prompt $D^0 \to h^+h'^-$ decays and an extension of the lifetime fitter for $B^0_{(s)} \to h^+h'^-$ decays to $D^0 \to h^+h'^-$ decays are demonstrated. The major hurdle for this measurement is the contribution from secondary D decays. Possible solutions are discussed.

The last chapter summarises the work of this thesis and provides an outlook to further measurements at *LHCb* and the longer term future of flavour physics.

Acknowledgement

To start with I would like to thank my supervisor Chris Parkes without whom this work would not have been possible. He spent a lot of time with never ending patience in discussions with me throughout the various stages of my thesis. During these, he prevented me from losing the path that led to the on-time completion of the thesis and taught me many invaluable lessons. At the same time, he always granted me the freedom to try new ideas which now form sections of this document. This mixture is what made this thesis a success and therefore: thanks a lot! The only lesson he failed to teach me is the skiing – thanks also for that.

The second big thank you to my second supervisor Paul Soler. He often managed to inject the right comment at the right time and was always available for advice. Special thanks for still making it to the airport on time. I would have never found Abingdon on my own.

Of course, the other members of the Glasgow *LHCb* group all contributed in one way or another to happy and successful three years of work with them. Thanks to Aldo and Franciole for the warm welcome to the group and for letting me participate in the epic story of the 'pink pig'. Special thanks to Franciole for having used the kernel method. Seb, thanks for introducing me to a world full of twisted coordinate systems. I owe Alison a big thank you for familiarising me with Italian physics and for agreeing on Dark Island for our first meeting. Thanks to Lars for joining me in the Gulag, trying to teach me to ski, and for showing a few beautiful hills in the highlands. Tomasz, thanks for the many nice wines we tried together. And after all you managed to show me your favourite restaurant on the last day of my thesis. Laurence, thanks for showing me that Irvine has much nicer guys than those often getting off there from the train to Prestwick. You are the only true matador in Glasgow. Thanks to Njaka for being a great flatmate and a very good friend! Also for teaching me so much about Malagasy culture. Eduardo, thank you very much for having had so much time for endless discussions and for always getting me to see my work from some distance. I am especially grateful that you were always available for coffee breaks even after beer o'clock as well as for many a good night out. Thanks to Vava for introducing me to

the business of lifetime fitting and for never getting tired of discussing it ever since. Silvia, an enormous portion of thanks for having taken over so quickly and for helping me out of that twisted world. Thanks as well for the green and friendly atmosphere in the office which made being there a pleasure. Thanks to Michael for the courage of taking over part of this work and for already having contributed a lot. Good luck to you!

It would be impossible to mention all the people I worked with at *CERN*. A big thank you to the whole *VELO* group for being such a great team to work with! Special thanks to Paula for always believing in me and for pushing the test beam analyses. Thanks to Jonas for teaching *LHCb* how to swim. Your comments often helped me make a leap forwards with my work. Also thanks to Jon for being a member of the swimming team. Thanks to Jacopo and Eduardo for the great team work which was a lot of fun!

Back at Glasgow, thanks a lot to Dima for many random things! Thanks also to Andrew and Kenny for keeping me alive at Oxford and for just being good mates. Thanks to Johanna for a lot of nice whiskies and parties. A big thank you to all PPE members whom I did not mention so far!

I also would like to thank SUPA, who gave me the financial support for this PhD. I also benefited a lot from the SUPA Graduate School lectures, which managed to give students access to the best lectures across Scotland.

Before the start of my PhD, I had the luck to get educated at the University of Mainz. Without the enormous amount I learned during my diploma thesis, I would not have been able to accomplish the work for my PhD thesis on this time scale. Thank you very much to my supervisors Konrad Kleinknecht and Rainer Wanke, as well as to all members of the NA48 group at Mainz. At this point I cannot value highly enough what I learned from Reinhard Simon and Anton Andronic during my visits to GSI. The time there is what got me hooked on to accelerator based physics and to working in international environments.

A few people dared to read this thesis before I wrote this page. They found many mistakes, typos, strange German phrasing, and a lot of other things that should not have been there. They were great and gave invaluable comments to help making sure that this thesis somehow makes sense. All remaining imperfections are likely to have been put in by myself after their reading, hence, they are not to blame. Thank you very much to Chris, Paul, Eva, Silvia, Eduardo, Vava, Simone, Volker, and Laurence.

Thanks also to two people who read the whole thesis with great care as well: my examiners Guennadi Borissov and Tony Doyle. They certainly made sure that there

are no loose ends and helped a lot in making this thesis more readable. Thanks a lot for the detailed and fruitful discussion.

Last but not least I owe a huge thank you to my family. Vielen Dank an meine Eltern, die mich haben ziehen lassen, damit ich meine Doktorarbeit im Ausland schreiben konnte. Ohne ihre Unterstützung wären weder mein Studium noch diese Arbeit möglich gewesen. Vielen Dank auch an meinen Bruder, der mir so viele interessante Orte gezeigt hat, an die mich die Physik noch nicht geführt hatte. Danke auch, dass du es so lange fernab der Zivilisation mit mir ausgehalten hast. Мерси много и на скъпа моя, Ева. Мерси за три хубави години с твоята любов и търпение. Много те обичам!

Declaration

I declare that except where explicit reference is made to the work of others, this dissertation is the result of my own work. This work has not been submitted for any other degree at the University of Glasgow or any other institution.

Marco Gersabeck

1 Theory of Standard Model and Flavour Physics

> Heutzutage haben Wissenschaftler mehr Fantasie als die Verfasser von Kriminalromanen.
>
> *Werner Heisenberg*

The start of the Large Hadron Collider (*LHC*) at the European Organization for Nuclear Research (*CERN*) will mark a new chapter in particle physics. With an energy that surpasses that of the previously most powerful accelerator, the TeVatron at Fermilab, by almost an order of magnitude, it will give access to a new range of physics. Key to the high precision measurements and discoveries of rare processes is its high luminosity.

This chapter will give the theoretical basis upon which the work in this thesis is based. After the introduction of the Standard Model of particle physics in section 1.1, section 1.2 will cover aspects of flavour physics relevant to this thesis. Section 1.3 will explain the concept of particle lifetimes and their measurements. The last two sections will specialise on lifetime measurements using the decay $B_s^0 \to K^+K^-$ and charm decays, respectively.

1.1 The Standard Model of Particle Physics

The fundamental building blocks of matter are fermions, particles of spin $\frac{1}{2}$. They are six quarks and six leptons together with their respective anti-particles. Both the quark and lepton sector are organised in three families of two particles each, where the particles of different families differ only in their mass. The quark sector consists of families of a quark with electrical charge $+\frac{2}{3}$ (up, charm, and top) and one with charge $-\frac{1}{3}$ (down, strange, and bottom). Quarks also have a colour charge that can take three

values and hence gives 18 different quark states. The lepton sector consists of particles of unit charge -1 (electron, muon, tau), and charge-less neutrinos (ν_e, ν_μ, and ν_τ).

Matter interacts via four forces: the strong, the electromagnetic, the weak, and the gravitational force. All but the latter can be described in the formalism of a quantum field theory (*QFT*) [1]. Compared to the others, the gravitational force is so weak that it can safely be neglected when dealing with quantum effects of elementary particles. The forces are mediated by bosons of spin 1. These are the photon for the electromagnetic force, the neutral Z^0 and charged $W^\pm$ for the weak force, and eight gluons for the strong force which carry a colour charge and an anti-colour charge.

The theory describing the interaction of elementary particles under the strong, electromagnetic, and weak force is the Standard Model (*SM*) [2, 3, 4, 5]. The following section will briefly introduce the principles of *QFT*. Section 1.1.2 will show the formalism for a *SM* with one generation of fermions. The full three generation *SM* will be explained in section 1.1.3.

1.1.1 Quantum Field Theory

The three forces that are of interest to particle physics – strong, electromagnetic, and weak – are described within a common theoretical framework of a quantum field theory. It is based on the concept of the Lagrange function, from which the equations of motion that describe the dynamics of a system can be obtained by evaluating the Euler-Lagrange equation.

In contrast to a classical Lagrange function, the coordinates of an N-point system are replaced by continuous fields and the Lagrange function becomes a Lagrange density. The quantisation replaces these fields by field operators. For a free scalar field ϕ this leads to

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2, \tag{1.1}$$

where m is the mass of the field carrying particle. Adding an interaction term that obeys the requirement of renormalisability of the theory leads to

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4. \tag{1.2}$$

For a Dirac field ψ the Lagrangian is given by

$$\mathcal{L} = \overline{\psi}(x)(i\gamma^\mu\partial_\mu - m)\psi(x). \tag{1.3}$$

Applying the Euler-Lagrange equation leads to the Dirac equation. For a free vector

field, A_μ, the Lagrangian is given by

$$\mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu}, \tag{1.4}$$

with $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$. If A_μ represents the electromagnetic field, this Lagrangian leads to the Maxwell equations.

1.1.2 The Standard Model for a Single Generation

The Lagrangian of a theory is the central piece in its description as outlined in the previous section. Many aspects of the *SM* can be discussed by only studying the Lagrangian for a single generation of fermions;

$$\mathcal{L}_{(SM,1)} = \mathcal{L}_{gauge\ bosons} + \mathcal{L}_{fermion\ masses} + \mathcal{L}_{fermion\ KT} + \mathcal{L}_{Higgs}. \tag{1.5}$$

In this case the Lagrangian can be split up into four parts as shown in equation 1.5. It contains:

- a term describing the kinetic terms for the gauge bosons (equation 1.6),
- one describing Yukawa couplings that lead to the fermion mass terms (equation 1.13),
- a third term describing the kinetic terms for fermions (equation 1.14),
- and finally the Higgs term, leading to mass terms for the gauge bosons (equation 1.20),

The content of these terms will be described in detail below.

In general, the particle content and the symmetries that a theory is to describe have to be predefined. Any $SU(n)$ group describes a field with $N = n^2 - 1$ gauge bosons. Experimental evidence, e.g. the existence of the Ω^- baryon or the relative fraction of $e^+e^- \to q\overline{q}$ with respect to $e^+e^- \to l^+l^-$, dictates that quarks have to have a quantum number which has three states: colour. The group to describe the quark fields is hence $SU(3)$. Therefore, the strong interaction is mediated by eight gluons. $U(1)_Y \times SU(2)_L$ describes the electroweak interaction with its four gauge bosons. Note that $U(1)_Y$ does not represent the electromagnetic interaction. Their connection is given in equation 1.18.

1.1.2.1 Kinetic Terms for the Gauge Bosons

The kinetic terms for the gauge bosons follow the usual structure of equation 1.4 and are given by

$$\mathcal{L}_{gauge\ bosons} = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} - \frac{1}{4}F^A_{\mu\nu}F^{A\mu\nu} + \mathcal{L}_{gauge\ fix} + \mathcal{L}_{FP\ ghosts}, \quad (1.6)$$

where the usual notation $X_{\mu\nu} = \partial_\mu X_\nu - \partial_\nu X_\mu$ is used.

B_μ is the $U(1)_Y$ field of hypercharge Y. F^a_μ with $a = 1..3$ represents the three fields of the $SU(2)$ group that together with B_μ combine to form the electroweak fields. F^A_μ with $A = 1..8$ denote the eight gluon fields of the $SU(3)$ group describing the strong interaction.

The final two terms allow for gauge fixing and so-called Fadeev-Popov ghosts. The latter are particles that appear inside additional loop processes which are introduced by the gauge fixing formalism.

1.1.2.2 Fermion Mass Terms

As the weak interaction is known to violate parity, the fermion content of the model is split into left-handed and right-handed components.

$$\psi = \psi_L + \psi_R = \frac{1-\gamma_5}{2}\psi + \frac{1+\gamma_5}{2}\psi, \quad (1.7)$$

where $\frac{1-\gamma_5}{2}$ and $\frac{1+\gamma_5}{2}$ are the projectors for the left-handed and right-handed components, respectively. For a one generation SM this gives two left-handed $SU(2)$ doublets and four right-handed $SU(2)$ singlets.

$$q_L \equiv \begin{pmatrix} u_L \\ d_L \end{pmatrix}; \quad l_L \equiv \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}; \quad u_R; \quad d_R; \quad e_R; \quad \nu_R \quad (1.8)$$

The right-handed neutrino will not be considered further as, in the approximation of massless neutrinos, it does not interact with any other field.

This leaves only mass terms of mixed chirality ($m\overline{\psi}\psi = m\overline{\psi_L}\psi_R + m\overline{\psi_R}\psi_L$), which are not invariant under $SU(2)_L$. A solution is the introduction of an $SU(2)$ doublet Φ. This is the Higgs field that leads to the Yukawa interaction terms

$$\mathcal{L}_{Yukawa} = -Y(f)\overline{f_L}^i\Phi_i f_R + h.c., \quad (1.9)$$

where $Y(f)$ denotes the Yukawa coupling constant for the fermion field f, and $h.c.$ stands for the hermitian conjugate of the first term.

The Higgs field introduces spontaneous symmetry breaking of the electroweak symmetry by having a potential with a minimum at $\Phi^*\Phi = \frac{1}{2}v^2$. This leads to the field

$$\Phi = \frac{e^{i(\omega_a T^a - \omega_3 Y)}}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix}, \tag{1.10}$$

with the generators of the electroweak groups, T^a and Y, and real parameters ω_i.

Thus, the lepton part of the Lagrangian can be written as

$$\mathcal{L}_{Yukawa} = -\frac{Y(e)}{\sqrt{2}} \begin{pmatrix} \overline{\nu_L} & \overline{e_L} \end{pmatrix} \begin{pmatrix} 0 \\ v + H \end{pmatrix} e_R + h.c., \tag{1.11}$$

which leads to a mass term for the electron proportional to the Higgs field vacuum expectation value v

$$\frac{Y_e v}{\sqrt{2}}(\overline{e_L} e_R + \overline{e_R} e_L) = \underbrace{\frac{Y_e v}{\sqrt{2}}}_{\frac{1}{2}m_e} \overline{e}e. \tag{1.12}$$

The full Lagrangian describing the fermion mass terms can hence be written as

$$\mathcal{L}_{fermion\ masses} = -Y_e \overline{l_L}^i \Phi_i e_R - Y_d \overline{q_L}^i \Phi_i d_R - Y_u \epsilon_{ij} \overline{q_L}^i \Phi^{*j} u_R + h.c., \tag{1.13}$$

where ϵ_{ij} is a two-dimensional antisymmetric tensor that allows the generation of a mass for the up quark.

1.1.2.3 Fermion Kinetic Terms

The kinetic terms for fermions are constructed in the usual way (see equation 1.4)

$$\begin{aligned} \mathcal{L}_{fermion\ KT} &= i\overline{l_L}^T \gamma^\mu D_\mu l_L + i\overline{e_R}^T \gamma^\mu D_\mu e_R + i\overline{\nu_R}^T \gamma^\mu \partial_\mu \nu_R \\ &\quad + i\overline{q_L}^T \gamma^\mu D_\mu q_L + i\overline{d_R}^T \gamma^\mu D_\mu d_R + i\overline{u_R}^T \gamma^\mu \partial_\mu u_R, \end{aligned} \tag{1.14}$$

using the covariant derivatives that determine which fermions couple to which fields

$$\begin{aligned} D_\mu &= \partial_\mu + igT^a W_\mu^a + ig'Y(l_L)B_\mu && \text{for } l_L, \\ D_\mu &= \partial_\mu + ig'Y(e_R)B_\mu && \text{for } e_R, \\ D_\mu &= \partial_\mu + ig_s T_s^A G_\mu^A + igT^a W_\mu^a + ig'Y(q_L)B_\mu && \text{for } q_L, \\ D_\mu &= \partial_\mu + ig_s T_s^A G_\mu^A + ig'Y(d_R)B_\mu && \text{for } d_R. \end{aligned} \tag{1.15}$$

In these terms, the coupling constants for the various interactions appear: g for the weak $SU(2)$, g' for the hypercharge $U(1)$, and g_s for the strong $SU(3)$ coupling.

Writing this out for left-handed leptons yields

$$\mathcal{L}_{l_L\ interaction} = -\frac{g}{2} \begin{pmatrix} \overline{\nu_L} \\ \overline{e_L} \end{pmatrix}^T \gamma^\mu \left[\begin{pmatrix} W_\mu^3 & W_\mu^1 + iW_\mu^2 \\ W_\mu^1 - iW_\mu^2 & -W_\mu^3 \end{pmatrix} + \frac{g'}{g} B_\mu \right] \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}, \tag{1.16}$$

which can be written in terms of the gauge boson fields as

$$\mathcal{L}_{l_L\ interaction} = -\frac{g}{2}\begin{pmatrix}\overline{\nu_L}\\ \overline{e_L}\end{pmatrix}^T \gamma^\mu \begin{pmatrix}\frac{1}{\cos\theta_W}Z_\mu & \sqrt{2}W_\mu^-\\ \sqrt{2}W_\mu^+ & \cos\theta_W Z_\mu - 2\sin\theta_W A_\mu\end{pmatrix}\begin{pmatrix}\nu_L\\ e_L\end{pmatrix}, \tag{1.17}$$

using

$$\begin{pmatrix}Z_\mu\\ A_\mu\end{pmatrix} \equiv \begin{pmatrix}\cos\theta_W & -\sin\theta_W\\ \sin\theta_W & \cos\theta_W\end{pmatrix}\begin{pmatrix}W_\mu^3\\ B_\mu\end{pmatrix}; \qquad \tan\theta_W \equiv \frac{g'}{g}, \tag{1.18}$$

and

$$W^\pm \equiv \frac{1}{\sqrt{2}}(W^1 \mp iW^2). \tag{1.19}$$

This example shows how neutrinos interact only with each other through Z boson exchange (Z_μ), while electrons can also exchange photons (A_μ). It also introduces the fields of charged W bosons that provide the charged current couplings. For quarks, the coupling via gluons appears.

1.1.2.4 The Higgs Term

The fourth ingredient to the *SM* Lagrangian is the Higgs term which follows equation 1.2 with the potential with vacuum expectation value $v = \mu/\sqrt{\lambda}$.

$$\begin{aligned}\mathcal{L}_{Higgs} &= |D_\mu\Phi|^2 - \mu^2\Phi_i^*\Phi^i + \lambda(\Phi_i^*\Phi^i)^2\\ &= \frac{1}{2}(\partial_\mu H)^2 + \underbrace{\mu^2}_{\frac{1}{2}M_H^2}H^2 + \underbrace{\frac{g^2v^2}{4}}_{M_W^2}W^{+\mu}W_\mu^- + \underbrace{\frac{g^2v^2}{8\cos^2\theta_W}}_{\frac{1}{2}M_Z^2}Z_\mu Z^\mu\\ &\quad + interaction\ terms\end{aligned} \tag{1.20}$$

It immediately yields the mass terms for the massive gauge bosons. The Higgs is the only *SM* particle that has not yet been observed. It is one of the main goals of the *LHC* to reveal this last piece of the *SM* puzzle.

1.1.3 The Three Generation Standard Model

For the expansion to more than one fermion generation only the quark sector will be discussed. The existence of three generations in the lepton sector is well established through direct observation of the three charged leptons, and the measurement of the width of the Z boson which established the existence of three light neutrino families. In addition, the observation of neutrino mixing confirmed that neutrinos are not massless. However, a detailed discussion of the lepton sector is beyond the scope of the work presented here.

1.1.3.1 Two Quark Generations

The simpler case of two quark generations is discussed first to illustrate the major changes when going to more than one generation. The second quark generation consists of a left-handed quark doublet and two right-handed singlets, thus introducing the charm and strange quarks. In all but their mass and family-specific quantum numbers, charm and strange are identical to up and down, respectively.

The general form of Yukawa couplings of two quark generations, following equation 1.9, now becomes

$$\mathcal{L}_{Yukawa} = -[Y_d]_{ij}\overline{q_{Li}}\Phi_j d_R + -[Y_u]_{ij}\overline{q_{Li}}\epsilon_{jk}\Phi_k^* u_R + h.c., \tag{1.21}$$

which introduces the Yukawa couplings Y_q as matrices. Through these matrices quarks of different generations can interact.

As the size of the Yukawa coupling determines the quark masses, analogously to equation 1.12, these matrices have to be diagonal when acting on mass eigenstates. Since this is not generally the case, Cabibbo introduced the concept of a mixing matrix that rotates the mass eigenstates into the flavour eigenstates that couple to the gauge bosons [6]. Using the mixing matrix, V_C, the coupling of the mass eigenstate quarks to the W boson, following equation 1.14, can be written as

$$-\frac{g}{2\sqrt{2}}\begin{pmatrix}\overline{u}\\ \overline{c}\end{pmatrix}^T \gamma^\mu(1-\gamma^5)V_C\begin{pmatrix}d\\ s\end{pmatrix}W_\mu^- + h.c. \tag{1.22}$$

The most general observable mixing matrix is given by

$$V_C = \begin{pmatrix}\cos\theta_C & \sin\theta_C\\ -\sin\theta_C & \cos\theta_C\end{pmatrix}, \tag{1.23}$$

with the Cabibbo angle θ_C.

Writing out the expression for the coupling of d and s quarks to the Z boson the structure of the mixing matrix implies that there are no flavour changing neutral current processes at tree level. This consequence of the mixing structure of two quark generations is known as the GIM[1] mechanism [7]. The observation of the strangeness-violating decay $K_L^0 \to \mu^+\mu^-$ lead to the prediction of a charm quark with a mass different to the up quark. The mass difference is required since otherwise even higher order charged current amplitudes would cancel.

[1] Glashow, Iliopoulos, Maiani

1.1.3.2 Three Quark Generations

The generalisation to three quark generations reflects the same structure. The additional up-type quark is the top and the down-type quark is the bottom. The bottom quark is, among others, contained in B mesons, which are a key ingredient to this thesis. The connection between the mass and flavour eigenstates is given by

$$\begin{pmatrix} d \\ s \\ b \end{pmatrix}_{flavour} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{mass} , \tag{1.24}$$

with the 3×3 Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix [8].

The CKM matrix consists of 9 complex elements that can be expressed by 18 real numbers. 9 of these numbers can be removed taking into account the unitarity requirement. As five phases can be removed by global phase rotations of the quark fields, the number of free parameters for the CKM matrix is four. A general parametrisation is given by

$$\begin{aligned} V_{CKM} &= \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{13}s_{23} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{13}c_{23} \end{pmatrix}, \end{aligned} \tag{1.25}$$

with c_{ij} and s_{ij} representing $\cos\theta_{ij}$ and $\sin\theta_{ij}$, respectively. $\theta_{12} \equiv \theta_C$ is the Cabibbo angle mentioned above.

Another representation is the Wolfenstein representation which reflects more prominently the vastly different magnitudes of the individual elements.

$$V_{CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \tag{1.26}$$

with $\lambda \approx \sin\theta_C$, $A \approx 1$, $\rho - i\eta = e^{-i\delta_{13}}$.

This matrix is the basis for flavour physics, which describes processes involving quarks of different flavours. The following section will introduce the physics related to the CKM matrix in more detail and will show how $\mathcal{CP}$ violation arises from its structure.

1.2 Flavour Physics

Flavour physics started with the discovery of the kaon in 1947 [9]. Since then, it has studied the processes involving different quark families and led to the construction of the *CKM* matrix and subsequently to measurements of its parameters with increasing precision.

These measurements are made by observing the decay of hadrons, i.e. particles involving quarks. All stable particles have to be neutral with respect to the charge of the strong interaction, colour, due to colour confinement. This leads to the fact that hadronic matter consists of mesons, quark anti-quark pairs, and baryons, which contain three quarks. The quarks in mesons have the same absolute colour charge but with opposite signs to create a colour-neutral state. Baryons contain three quarks, each having a different colour charge, which also leads to a colour-neutral state.

The following sections will introduce the concept of mixing of neutral mesons as well as $\mathcal{CP}$ violation [10] and finally illustrate the connection of the *CKM* matrix with these phenomena.

1.2.1 Mixing of Neutral Mesons

The main topics of this thesis involve decays of neutral mesons. Therefore, this section will describe the concept of neutral meson mixing: a process by which a neutral meson changes into its anti-meson and vice versa. This process exists for the K^0, D^0, B_d^0, and B_s^0 mesons.

The mass eigenstates states are linear combinations of the particle and anti-particle states

$$\alpha|N^0\rangle + \beta|\overline{N^0}\rangle \equiv \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \tag{1.27}$$

with the time dependent Schrödinger equation

$$i\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = X\begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \tag{1.28}$$

The matrix X is given by $X = M - \frac{i}{2}\Gamma$, where M and Γ are Hermitian matrices. The elements of X are given by

$$\begin{aligned} X_{11} &= \langle N^0|\mathcal{H}|N^0\rangle, \quad X_{12} = \langle N^0|\mathcal{H}|\overline{N^0}\rangle, \\ X_{21} &= \langle \overline{N^0}|\mathcal{H}|N^0\rangle, \quad X_{22} = \langle \overline{N^0}|\mathcal{H}|\overline{N^0}\rangle, \end{aligned} \tag{1.29}$$

where $\mathcal{H}$ is the Hamiltonian describing the transition from a particle to an anti-particle state and vice versa. The eigenvalues of X are

$$\begin{aligned} \lambda_h &= m_h - \frac{i}{2}\Gamma_h, \\ \lambda_l &= m_l - \frac{i}{2}\Gamma_l, \end{aligned} \tag{1.30}$$

with the corresponding eigenstates

$$\begin{aligned} |N_h\rangle &= p|N^0\rangle - q|\overline{N^0}\rangle, \\ |N_l\rangle &= p|N^0\rangle + q|\overline{N^0}\rangle. \end{aligned} \tag{1.31}$$

These states are the physical states and have masses such that $\Delta m \equiv m_h - m_l > 0$. In addition one defines $\Delta\Gamma \equiv \Gamma_h - \Gamma_l$. Unitarity requires $|p|^2 + |q|^2 = 1$.

Using equation 1.30 it follows from $(\lambda_h - \lambda_l)^2$ that

$$(\Delta m)^2 - \frac{1}{4}(\Delta\Gamma)^2 = 4|m_{12}|^2 - |\Gamma_{12}|^2, \tag{1.32}$$

$$\Delta m \Delta\Gamma = 4\Re e(m_{12}\Gamma_{12}^*), \tag{1.33}$$

and furthermore

$$\frac{q}{p} = \pm\sqrt{\frac{m_{12}^* - \frac{i}{2}\Gamma_{12}^*}{m_{12} - \frac{i}{2}\Gamma_{12}}} = \pm\frac{2m_{12}^* - i\Gamma_{12}^*}{\Delta m - i\Delta\Gamma/2}, \tag{1.34}$$

where m_{ij} and Γ_{ij} are the i, j matrix elements of M and Γ, respectively, and where the sign $\pm$ corresponds to the choice between $\lambda_{h,l}$ (for a detailed discussion see chapter 7 in reference [11]).

1.2.1.1 Time Evolution of Mixing States

For the mass eigenstates the time evolution is given by

$$\begin{aligned} |N_h(t)\rangle &= e^{-im_h t - \frac{1}{2}\Gamma_h t}|N_h(0)\rangle, \\ |N_l(t)\rangle &= e^{-im_l t - \frac{1}{2}\Gamma_l t}|N_l(0)\rangle. \end{aligned} \tag{1.35}$$

Hence, the amplitude for an initially pure N^0 state is given by

$$\psi_N = \frac{1}{2}\left((e^{-im_h t - \frac{1}{2}\Gamma_h t} + e^{-im_l t - \frac{1}{2}\Gamma_l t})N^0 - \frac{q}{p}(e^{-im_h t - \frac{1}{2}\Gamma_h t} - e^{-im_l t - \frac{1}{2}\Gamma_l t})\overline{N^0}\right). \tag{1.36}$$

The probability for obtaining the state N^0 ($\overline{N^0}$) after time t when starting from an N^0 state is

$$\begin{aligned} P(N^0 \to N^0) &= \frac{1}{2}e^{-T}(\cosh yT + \cos xT), \\ P(N^0 \to \overline{N^0}) &= \frac{1}{2}\left|\frac{q}{p}\right|^2 e^{-T}(\cosh yT - \cos xT), \end{aligned} \tag{1.37}$$

	$K^0/\overline{K^0}$	$D^0/\overline{D^0}$	$B_d^0/\overline{B_d^0}$	$B_s^0/\overline{B_s^0}$
τ (ps)	89.58 ± 0.05, 51160 ± 200	0.4101 ± 0.0015	1.530 ± 0.009	1.470 ± 0.027
Γ (s^{-1})	5.59×10^9	2.4×10^{12}	6.5×10^{11}	6.8×10^{11}
x	0.946 ± 0.002	0.0097 ± 0.0028	0.776 ± 0.008	26.1 ± 0.5
y	-0.9965	0.0078 ± 0.0019	$\lvert y\rvert < 0.04$, 90% C.L.	$[0.09, -0.03]$, 95% C.L.

Table 1.1: *Parameters for decay and mixing of neutral mesons [12].*

where $T \equiv \Gamma t$, $x \equiv \frac{\Delta m}{\Gamma}$, and $y \equiv \frac{\Delta\Gamma}{2\Gamma}$ have been introduced. The case of starting with an $\overline{N^0}$ is identical apart from $\frac{q}{p}$ being replaced by its inverse.

The flavour asymmetry follows

$$A(T) = \frac{P(N^0 \to N^0) - P(N^0 \to \overline{N^0})}{P(N^0 \to N^0) + P(N^0 \to \overline{N^0})} = \frac{\cos xT + \delta \cosh yT}{\cosh yT + \delta \cos xT}, \tag{1.38}$$

with

$$\delta \equiv \langle N_h | N_l \rangle = |p|^2 - |q|^2 = \frac{1 - \left|\frac{q}{p}\right|^2}{1 + \left|\frac{q}{p}\right|^2}. \tag{1.39}$$

For the various systems of neutral mesons this asymmetry has drastically different time evolutions in reality due to large differences in the parameters involved. The current best measurements of these parameters are given in table 1.1.

1.2.2 CP Violation

Three discrete transformations play a central role in particle physics:

- $\mathcal{C}$: the charge transformation converts particles into anti-particles and vice versa,
- $\mathcal{P}$: the parity transformation reverses the spatial coordinates which leads to a change of chirality of particles,
- $\mathcal{T}$: the time transformation reverses time and hence changes sign of velocities, momenta, and angular momenta.

Following the $\mathcal{CPT}$ theorem [13, 14, 15], all realistic quantum field theories are invariant under the three combined transformations $\mathcal{CPT}$. As a consequence of the theorem, all particles and anti-particles must have the same masses, lifetimes, and magnetic dipole moments. This holds for all experimental tests conducted to date.

The weak interaction violates both $\mathcal{C}$ and $\mathcal{P}$ symmetries. This is clear as the former would transform a left-handed neutrino into a left-handed anti-neutrino, which does not exist (at least in the *SM*). Similarly, the right-handed neutrino is not found, which would be the result of a $\mathcal{P}$ transformation.

The combined transformation, $\mathcal{CP}$, however, initially appeared to be conserved. In 1964, Christenson, Cronin, Fitch, and Turlay discovered that K_L also decay into two pions [16]. This process is only allowed with the $\mathcal{CP}$ symmetry violated, as the K_L is an eigenstate of $\mathcal{CP}$ with eigenvalue -1 and a pion pair has an eigenvalue of $+1$. First evidence for $\mathcal{CP}$ violation in B mesons was found by the BaBar and Belle experiments in 2001 [17, 18].

Neutral mesons can be eigenstates of $\mathcal{CP}$. From

$$\begin{aligned} \mathcal{CP}|N^0\rangle &= e^{ia}|\overline{N^0}\rangle, \\ \mathcal{CP}|\overline{N^0}\rangle &= e^{-ia}|N^0\rangle, \end{aligned} \tag{1.40}$$

where the arbitrary phase a can be chosen to be 0, one obtains the $\mathcal{CP}$ eigenstates

$$\begin{aligned} \mathcal{CP}|N_+\rangle &= \frac{1}{2}\left(|N^0\rangle + |\overline{N^0}\rangle\right) \\ \mathcal{CP}|N_-\rangle &= \frac{1}{2}\left(|N^0\rangle - |\overline{N^0}\rangle\right), \end{aligned} \tag{1.41}$$

with the $\mathcal{CP}$ eigenvalues $+1$ and -1, respectively.

It follows from equation 1.31 that $\mathcal{CP}$ and mass eigenstates coincide if $|p| = |q|$. Furthermore, this implies that δ as defined in equation 1.39 vanishes. In this case, $P(N^0 \to \overline{N^0})$ and $P(\overline{N^0} \to N^0)$ are equal, or $\mathcal{CP}$ is conserved in the process of mixing. From equation 1.34 it is obvious that $\mathcal{CP}$ violation in mixing requires $\Im m\ m_{12} \neq 0$ or $\Im m\ \Gamma_{12} \neq 0$, in addition to a relative phase between m_{12} and Γ_{12}.

In addition to $\mathcal{CP}$ violation in mixing, also called indirect $\mathcal{CP}$ violation, there is also direct $\mathcal{CP}$ violation or $\mathcal{CP}$ violation in the decay. Introducing the decay amplitudes A ($\overline{A}$) of the B ($\overline{B}$) meson into the final state f or $\overline{f}$,

$$A_f \equiv \langle f|\mathcal{T}|N^0\rangle, \quad \overline{A}_f \equiv \langle f|\mathcal{T}|\overline{N^0}\rangle, \tag{1.42}$$

there is direct $\mathcal{CP}$ violation if

$$\begin{aligned} |A_f| &\neq |\overline{A}_{\overline{f}}|, \\ |A_{\overline{f}}| &\neq |\overline{A}_f|. \end{aligned} \tag{1.43}$$

For decays into a $\mathcal{CP}$ eigenstate this simplifies to $|A_f| \neq |\overline{A}_f|$. Evidence for direct $\mathcal{CP}$ violation in the neutral kaon system was first found by the NA31 experiment in

1988 [19] and finally established by NA48 [20] and KTeV [21] in 1999. In the B system direct $\mathcal{CP}$ was first observed by BaBar and Belle in the decay $B^0 \to K\pi$ in 2004 [22, 23].

A third possible form of $\mathcal{CP}$ violation arises from the interference of $\mathcal{CP}$ violation in mixing and decay. Analogously to equation 1.36, one obtains for the time-dependent decay amplitude of a pure N^0 ($\overline{N^0}$) state at $t = 0$ into a final state f

$$\begin{aligned} \langle f|\mathcal{H}|N^0(t)\rangle &= \frac{1}{2}e^{-imt}e^{-\Gamma t/2}\left[g_+A_f + \frac{q}{p}g_-\overline{A}_f\right], \\ \langle f|\mathcal{H}|\overline{N^0}(t)\rangle &= \frac{1}{2}e^{-imt}e^{-\Gamma t/2}\left[g_+\overline{A}_f + \frac{p}{q}g_-A_f\right], \end{aligned} \tag{1.44}$$

with

$$g_+ = e^{-(ix+y)T/2} + e^{(ix+y)T/2}, \quad g_- = e^{-(ix+y)T/2} - e^{(ix+y)T/2}. \tag{1.45}$$

Introducing

$$\lambda_f = \frac{q}{p}\frac{\overline{A}_f}{A_f}, \tag{1.46}$$

and approximating $|\frac{q}{p}| = 1$ it follows for the decay rate asymmetry that

$$\frac{\Gamma(N^0(t) \to f) - \Gamma(\overline{N^0}(t) \to f)}{\Gamma(N^0(t) \to f) + \Gamma(\overline{N^0}(t) \to f)} = \frac{(1 - |\lambda_f|^2)\cos(xT) + 2\Im m\ \lambda_f \sin(xT)}{(1 + |\lambda_f|^2)\cosh(yT) - 2\Re e\ \lambda_f \sinh(yT)}. \tag{1.47}$$

Introducing the observables

$$C^f = \frac{1 - |\lambda_f|^2}{1 + |\lambda_f|^2}, \quad S^f = \frac{2\Im m\ \lambda_f}{1 + |\lambda_f|^2}, \quad \mathcal{A}^f_{\Delta\Gamma} = \frac{2\Re e\ \lambda_f}{1 + |\lambda_f|^2}, \tag{1.48}$$

leads to

$$\frac{\Gamma(N^0(t) \to f) - \Gamma(\overline{N^0}(t) \to f)}{\Gamma(N^0(t) \to f) + \Gamma(\overline{N^0}(t) \to f)} = \frac{C^f\cos(xT) + S^f\sin(xT)}{\cosh(yT) - \mathcal{A}^f_{\Delta\Gamma}\sinh(yT)}. \tag{1.49}$$

For both neutral B mesons x is much larger than y. Therefore, it is of interest to use the approximation $y = 0$. In this case equation 1.47 simplifies to

$$\frac{\Gamma(B(t) \to f) - \Gamma(\overline{B}(t) \to f)}{\Gamma(B(t) \to f) + \Gamma(\overline{B}(t) \to f)} = C^f\cos(xT) + S^f\sin(xT). \tag{1.50}$$

1.2.3 The CKM Triangle

The *CKM* matrix has been introduced in equation 1.25. The *CKM* matrix has to be unitary under the assumption of the conservation of the total probability which is also a consequence of the $\mathcal{CPT}$ theorem. The unitarity requirement $V_{CKM}V^\dagger_{CKM} = \mathbb{1}$ leads

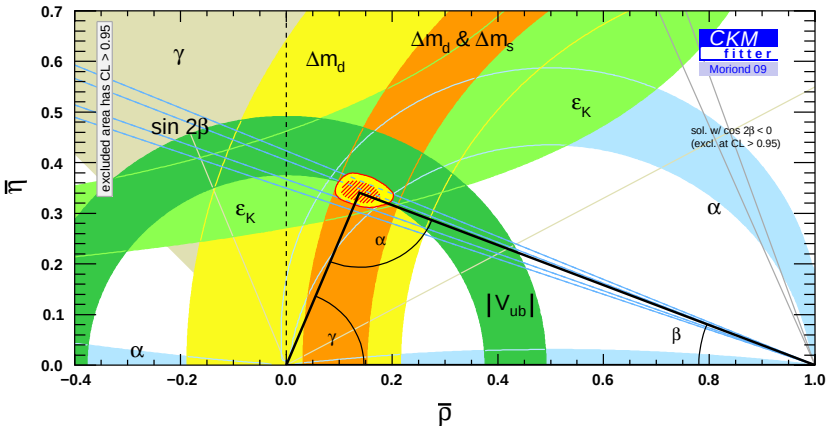


Figure 1.1: *CKM triangle showing the current experimental constraints. Reproduced from [24].*

to

$$\begin{aligned} \sum_{i=u,c,t} V_{ij}V_{ik}^* &= \delta_{jk}, \text{ or} \\ \sum_{i=d,s,b} V_{ji}V_{ki}^* &= \delta_{jk}, \end{aligned} \tag{1.51}$$

which, for the off-diagonal elements ($j \neq k$), are equations that define triangles in the complex plane. The only triangles with three sides of roughly equal length ($\mathcal{O}(\lambda^3)$) are given by

$$\begin{aligned} V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* &= 0, \\ V_{ud}V_{td}^* + V_{us}V_{ts}^* + V_{ub}V_{tb}^* &= 0, \end{aligned} \tag{1.52}$$

and normalising one side yields

$$\begin{aligned} \frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} + 1 + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} &= 0, \\ \frac{V_{ud}V_{td}^*}{V_{us}V_{ts}^*} + 1 + \frac{V_{ub}V_{tb}^*}{V_{us}V_{ts}^*} &= 0, \end{aligned} \tag{1.53}$$

Figure 1.1 shows the experimental constraints on the first of the two triangles. The apex of the triangle is over-constrained by the measurements. However, the constraint from the angle γ cannot compete with those from the other angles. The current best measurements are $(76 \pm 22_{stat.} \pm 5_{syst.} \pm 5_{Dalitz})^\circ$ from BaBar [25] and $(76^{+12}_{-13,stat.} \pm 4_{syst.} \pm 9_{model})^\circ$ from Belle [26]. It is one of the main goals of *LHCb* to reduce this uncertainty by about a factor of 10.

As γ is defined by

$$\gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right), \tag{1.54}$$

it is obvious that it is strongly related to the phase of V_{ub}, i.e. δ_{13} (see equation 1.25). This means that tree level $b \to u$ transitions are the most obvious way to measure γ. These occur in many decays of neutral B mesons into charged particles. The extraction of γ often suffers from manifold ambiguities when performed from B_d^0 decays alone. Therefore, the comparison of related B_d^0 and B_s^0 decays can greatly enhance the sensitivity to γ. One such concept has been introduced by Fleischer [27] and is pursued using channels like $B_{(s)}^0 \to D^{\pm}h^{\mp}$ [28] or $B_{(s)}^0 \to h^{\pm}h^{\mp}$ [29].

1.3 B Meson Lifetimes

B meson lifetime measurements form a sector of flavour physics which offers a number of interesting opportunities. This section outlines the possibilities of using lifetime measurements for precision *SM* tests as well as for probing *NP*. The following sections give examples of applications of lifetime measurements in the B_s^0 and the D^0 sector.

1.3.1 Heavy Quark Expansion

Lifetimes of B mesons can be calculated within the framework of heavy quark expansion (*HQE*, see [30] and references therein). In contrast to *QCD* which uses expansions in orders of the strong coupling constant, α_s, *HQE* expands in terms of Λ/m_b, i.e. a scale factor and the mass of the bottom quark.

The leading terms for the expansion of a decay rate are given by

$$\Gamma = \Gamma_0 + \frac{\Lambda^2}{m_b^2}\Gamma_2 + \frac{\Lambda^3}{m_b^3}\Gamma_3. \tag{1.55}$$

The first order term in Λ/m_b vanishes, the second order correction is due to the kinetic and chromomagnetic operator, and only the third order term depends on the spectator quark.

The weak influence of the spectator quark, which only enters at the level $\frac{\Lambda^3}{m_b^3}$, explains why the lifetime ratio of B_s^0 and B_d^0 is predicted as

$$\frac{\tau(B_s^0)}{\tau(B_d^0)} = 1.00 \pm 0.01. \tag{1.56}$$

Experimentally, slightly lower values are favoured following results from *CDF* and *D0*. The current experimental average is

$$\frac{\tau(B_s^0)}{\tau(B_d^0)} = 0.93 \pm 0.03, \tag{1.57}$$

when taking the ratio of the averaged individual lifetimes [12].

The theoretical precision of this ratio shows that precise measurements are an excellent test of *HQE*. Furthermore, it is important to either confirm the tension between theory and experiment or to resolve it by measurement with at least a percent level precision. A method for directly fitting lifetime ratios is presented in chapter 4.

1.3.2 B Meson Lifetime Parameters

The mass eigenstates of neutral mesons decay according to equation 1.35 with the decay rates Γ_h and Γ_l. These are given by

$$\Gamma_h = \Gamma + \frac{\Delta\Gamma}{2}; \quad \Gamma_l = \Gamma - \frac{\Delta\Gamma}{2}. \tag{1.58}$$

It is of interest to express Γ and $\Delta\Gamma$ in terms of the elements of the matrix X defined in equation 1.29. The matrix M is given by

$$\mathrm{m}_{ij} = m_0\delta_{ij} + \langle i|\mathcal{H}|j\rangle + \sum_n P\frac{\langle i|\mathcal{H}|n\rangle\langle n|\mathcal{H}|j\rangle}{\delta(m_0 - E_n)}, \tag{1.59}$$

where the sum goes over all internal virtual states, P denotes the projector for the principal part, and $|1\rangle$ and $|2\rangle$ represent the states $|N^0\rangle$ and $|\overline{N^0}\rangle$, respectively.

Similarly, the matrix Γ is given by

$$\Gamma_{ij} = 2\pi\sum_c \rho_c\langle i|\mathcal{H}|c\rangle\langle c|\mathcal{H}|j\rangle, \tag{1.60}$$

where the sum goes over the possible final states, and ρ_c is a phase space factor. Since Γ_{12} is limited to final states that are accessible to both meson and anti-meson it follows that $|\Gamma_{12}| < |\Gamma_{11}| = |\Gamma|$.

Experimental evidence (see table 1.1) shows that for B mesons it can be assumed that

$$|\Gamma_{12}| \ll |m_{12}|; \quad \Delta\Gamma \ll \Delta m. \tag{1.61}$$

Using this approximation and equations 1.32 and 1.33 it can be shown that

$$\Delta m = 2|m_{12}|\left[1 + \mathcal{O}\left(\left|\frac{\Gamma_{12}}{m_{12}}\right|^2\right)\right], \tag{1.62}$$

and

$$\Delta\Gamma = 2|\Gamma_{12}|\cos\phi\left[1+\mathcal{O}\left(\left|\frac{\Gamma_{12}}{m_{12}}\right|^2\right)\right], \tag{1.63}$$

where the mixing phase ϕ has been introduced as

$$\phi = \arg\left(\frac{m_{12}}{\Gamma_{12}}\right). \tag{1.64}$$

Further introducing

$$a = \Im m\frac{m_{12}}{\Gamma_{12}} = \left|\frac{m_{12}}{\Gamma_{12}}\right|\sin\phi, \tag{1.65}$$

it follows from using equations 1.62 and 1.63 in equation 1.34

$$\frac{q}{p} = \pm e^{-i\phi_m}\left[1-\frac{a}{2}\right]+\mathcal{O}\left(\left|\frac{\Gamma_{12}}{m_{12}}\right|^2\right), \tag{1.66}$$

where ϕ_m is the phase of m_{12}.

1.3.3 Time Evolution of Untagged B Mesons

Neglecting any production asymmetry in proton-proton collisions, the decay rate for untagged mesons can be written as

$$\Gamma(f,t) = \Gamma(N^0(t)\to f)+\Gamma(\overline{N^0}(t)\to f). \tag{1.67}$$

Neglecting the production asymmetry is a reasonable approximation as it is estimated to be about $1\%-2\%$ for B_d^0 mesons in *LHCb* [31]. However, for B_s^0 mesons it is much smaller.

Inverting 1.31 and using the time evolution given in 1.35 leads to

$$\Gamma(f,t) = \mathcal{N}_f\left[e^{-\Gamma_l t}|\langle f|B_l\rangle|^2+e^{-\Gamma_h t}|\langle f|B_h\rangle|^2\right]+\mathcal{O}(a), \tag{1.68}$$

where $\mathcal{N}_f$ is a normalisation factor. Further, using 1.31 and 1.46 this can be rewritten as

$$\Gamma(f,t) = \mathcal{N}_f\frac{|A_f|^2}{2}\left(1+|\lambda_f|^2\right)\left[(1-\mathcal{A}^f_{\Delta\Gamma})e^{-\Gamma_l t}+(1+\mathcal{A}^f_{\Delta\Gamma})e^{-\Gamma_h t}\right]+\mathcal{O}(a). \tag{1.69}$$

For practical reasons it is interesting to consider the case of fitting an untagged time distribution to a single exponential with decay rate $\hat{\Gamma}_f$. Writing

$$\Gamma(f,t) = Ae^{-\Gamma_l t}+Be^{-\Gamma_h t}, \tag{1.70}$$

the fit result would be

$$\hat{\Gamma}_f = \frac{A/\Gamma_l+B/\Gamma_h}{A/\Gamma_l^2+B/\Gamma_h^2}. \tag{1.71}$$

Expanding to second order in $\Delta\Gamma$ and reintroducing the observables leads to

$$\hat{\Gamma}_f = \Gamma - \mathcal{A}^f_{\Delta\Gamma}\frac{\Delta\Gamma}{2} - \frac{1-(\mathcal{A}^f_{\Delta\Gamma})^2}{2}\frac{(\Delta\Gamma)^2}{\Gamma} + \mathcal{O}\left(\frac{(\Delta\Gamma)^3}{\Gamma^2}\right). \tag{1.72}$$

For $\mathcal{CP}$ eigenstates $\mathcal{A}^f_{\Delta\Gamma} = \pm 1$ and one measures $\Gamma \mp \frac{\Delta\Gamma}{2}$. In flavour specific modes $\mathcal{A}^f_{\Delta\Gamma} = 0$ leads to a measurement of $\Gamma - \frac{(\Delta\Gamma)^2}{2\Gamma}$.

1.3.4 New Physics Influence on B Lifetime Measurements

New Physics (*NP*) particles can only influence *SM* processes at the level of loop diagrams. Hence, it can be assumed for B mesons that Γ_{12} is not influenced by *NP* as it is dominated by tree level transitions. The *NP* influence on m_{12} is generally parametrised by an additional complex factor Δ^{NP}

$$m_{12} = m_{12}^{SM}\Delta^{NP} = |m_{12}^{SM}||\Delta^{NP}|e^{i(\phi_m+\phi^{NP})}. \tag{1.73}$$

Examples for *NP* scenarios that can lead to sizeable *NP* phases are models with a fourth generation of quarks [32] or the Littlest Higgs model with T-parity [33]. Using equations 1.62 to 1.66, this leads to

$$\begin{aligned}
\Delta m &= 2|m_{12}^{SM}||\Delta^{NP}|, \\
\Delta\Gamma &= 2|\Gamma_{12}|\cos(\phi^{SM}+\phi^{NP}), \\
a &= \left|\frac{m_{12}^{SM}}{\Gamma_{12}}\right||\Delta^{NP}|\sin(\phi^{SM}+\phi^{NP}), \\
\frac{q}{p} &= \pm e^{-i(\phi_m^{SM}+\phi^{NP})}\left[1-\frac{a}{2}\right] + \mathcal{O}\left(\left|\frac{\Gamma_{12}}{\mathrm{m}_{12}}\right|^2\right).
\end{aligned} \tag{1.74}$$

For $\mathcal{A}^f_{\Delta\Gamma}$ this leads to

$$\mathcal{A}^f_{\Delta\Gamma} = \pm\frac{2|\lambda_f|}{1+|\lambda_f|^2}\cos(\phi_m^{SM}+\phi^{NP}), \tag{1.75}$$

when neglecting the phase of $\frac{\overline{A}_f}{A_f}$ in equation 1.46. For equation 1.72 this results in two cosine terms entering through $\mathcal{A}^f_{\Delta\Gamma}$ and $\Delta\Gamma$. Thus, *NP* effects will decrease the lifetime difference between a $\mathcal{CP}$ averaged and a $\mathcal{CP}$ specific mode.

1.4 The Decay $B_s^0 \to K^+K^-$

The decay $B_s^0 \to K^+K^-$ is one of a group of decay modes which is usually labelled as $B^0_{(s)} \to h^+h'^-$. Their decay processes are governed by the diagrams shown in figure

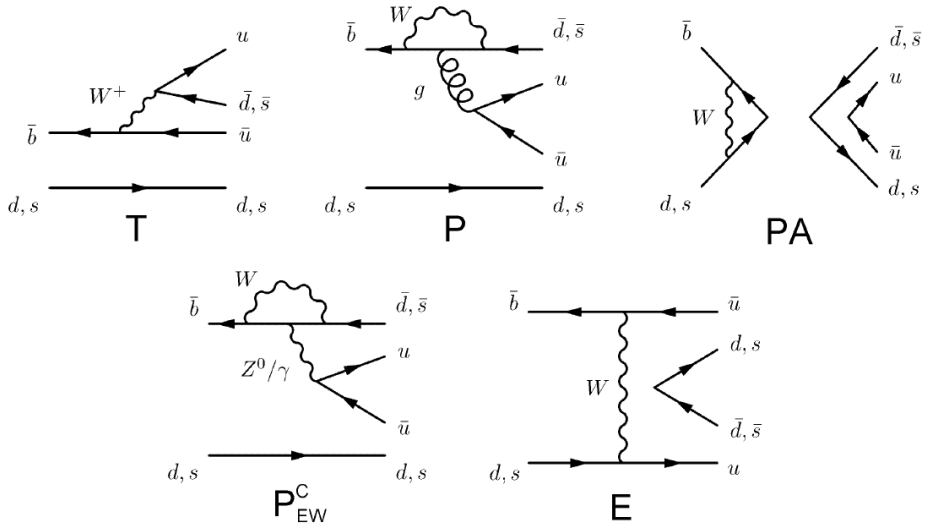


Figure 1.2: *Feynman graphs of processes contributing to the $B^0_{(s)} \to h^+h'^-$ decays. The diagrams are tree (T), gluonic penguin (P), penguin annihilation (PA), colour-suppressed electroweak penguin (P^C_{EW}), and W exchange (E).*

1.2. As the tree transition is proportional to V_{ub} and V_{us}, and hence to $\approx \lambda^4$, the $B^0_s \to K^+K^-$mode is dominated by penguin diagrams. Thus, *NP* particles entering the various loop processes can significantly influence the $B^0_s \to K^+K^-$ decay.

The K^+K^- final state is $\mathcal{CP}$ even, however, the interference of the tree diagram with the loop diagrams leads to a small $\mathcal{CP}$ violating effect. Therefore, already within the *SM* there is a $\mathcal{CP}$ violating phase

$$A_f = \langle K^+K^-|B^0_s\rangle \propto e^{i\sigma}; \quad \overline{A}_f = \langle K^+K^-|\overline{B^0_s}\rangle \propto e^{-i\sigma}. \tag{1.76}$$

1.4.1 Lifetime Measurements with $B^0_s \to K^+K^-$

The previous section showed the impact of $\mathcal{A}^f_{\Delta\Gamma}$ on lifetime measurements. For a decay into a $\mathcal{CP}$ even final state, it can also be written as

$$\mathcal{A}^f_{\Delta\Gamma} = -\frac{2|\lambda_f|}{1+|\lambda_f|^2}\cos(\arg\lambda_f), \tag{1.77}$$

highlighting the importance of the phase of λ_f. This phase consists of two components, the phase of q/p and that of $\overline{A}_f/A_f$. In the *SM* the former is approximately equal to $\phi^{SM}_m \approx \phi^{SM} \approx -2\beta_s = (-0.04 \pm 0.01)$ rad [30]. As described above (see equation 1.76), the latter adds a $\mathcal{CP}$ violating phase 2σ.

This leads to

$$\mathcal{A}^f_{\Delta\Gamma} = -\frac{2|\lambda_f|}{1+|\lambda_f|^2}\cos(2\beta_s + 2\sigma). \tag{1.78}$$

Within the *SM*, one obtains the constraints [34]

$$-1 \lesssim \mathcal{A}_{\Delta\Gamma}(B_s^0 \to K^+K^-) \lesssim -0.95, \tag{1.79}$$

which means that $B_s^0 \to K^+K^-$is at least 95% $\mathcal{CP}$ even.

NP particles can contribute to the mixing phase as described in section 1.3.4. In addition, such processes may contribute a $\mathcal{CP}$ violating phase σ^{NP}. Neglecting the small *SM* phase $2\beta_s$, combining equation 1.72 with equations 1.74 and 1.75 and assuming $|\lambda_f| = 1$ leads to

$$\begin{aligned}\hat{\Gamma}(B_s^0 \to K^+K^-) &= \Gamma_s \\ &\quad + \cos(-\phi^{NP} + 2\sigma^{SM} + 2\sigma^{NP})\frac{\Delta\Gamma_s^{SM}}{2}\cos(\phi^{NP}) \\ &\quad - \sin^2(-\phi^{NP} + 2\sigma^{SM} + 2\sigma^{NP})\frac{(\Delta\Gamma_s^{SM})^2}{2\Gamma_s}\cos^2(\phi^{NP}) \\ &\quad + \mathcal{O}\left(\frac{(\Delta\Gamma_s^{SM})^3}{\Gamma_s^2}\right).\end{aligned} \tag{1.80}$$

1.4.2 Interpretation of Lifetime Measured with $B_s^0 \to K^+K^-$

One of the exponentials in equation 1.69 is strongly suppressed in the case of a decay into a $\mathcal{CP}$ eigenstate. Hence, it is not possible to directly measure both Γ_s and $\Delta\Gamma_s$ with a lifetime measurement using $B_s^0 \to K^+K^-$. Taking an independent, precise measurement of Γ_s gives access to the second term in equation 1.80. Such a measurement can be made using high statistics flavour specific modes like $B_s^0 \to D_s^{\pm}\pi^{\mp}$. The third term in this equation can be neglected due to the $\frac{\Delta\Gamma_s^{SM}}{\Gamma_s}$ ($\lesssim 0.18$ see table 1.1) suppression.

The interpretation of a precise measurement of the second term in equation 1.80 is still ambiguous as the phases as well as $\Delta\Gamma_s^{SM}$ have not yet been measured to high precision. A full interpretation of the result is made possible through the comparison with measurements from other modes.

1.4.2.1 Comparison with $B_s^0 \to J/\psi\phi$

The decay $B_s^0 \to J/\psi\phi$ is dominated by a tree level $b \to c\bar{c}s$ transition. Therefore, *NP* effects can only enter through B_s^0 mixing. The orbital angular momentum (L) of the two decay products is of importance, as this decay has a final state with two vector mesons. The final state $(J/\psi\phi)_{L=0}$ is $\mathcal{CP}$ even and due to the absence of a significant $\mathcal{CP}$ phase in the *SM* one measures

$$\hat{\Gamma}(B_s^0 \to (J/\psi\phi)_{L=0}) = \Gamma_s + \frac{\Delta\Gamma_s^{SM}}{2}\cos^2(\phi^{NP})$$

$$
\begin{aligned}
&- \sin^2(\phi^{NP}) \frac{(\Delta\Gamma_s^{SM})^2}{2\Gamma_s} \cos^2(\phi^{NP}) \\
&+ \mathcal{O}\left(\frac{(\Delta\Gamma_s^{SM})^3}{\Gamma_s^2}\right). \qquad (1.81)
\end{aligned}
$$

or

$$
\begin{aligned}
\hat{\Gamma}(B_s^0 \to (J/\psi\phi)_{L=0}) &= \Gamma_s + \frac{\Delta\Gamma_s^{SM}}{2} \cos^2(\phi^{NP}) \left(1 - \sin^2(\phi^{NP}) \frac{\Delta\Gamma_s^{SM}}{\Gamma_s}\right) \\
&\quad + \mathcal{O}\left(\frac{(\Delta\Gamma_s^{SM})^3}{\Gamma_s^2}\right). \qquad (1.82)
\end{aligned}
$$

A comparison with $B_s^0 \to K^+K^-$ allows to constrain the phase $\sigma^{SM} + \sigma^{NP}$. It should be noted that the decay $B_s^0 \to J/\psi\phi$ allows the extraction of $\Delta\Gamma_s$ without comparison to other decay modes when measuring the $\mathcal{CP}$ odd $L = 2$ final state. This analysis, however, is significantly more challenging than the measurement with $B_s^0 \to K^+K^-$as it requires a lifetime measurement in combination with an angular analysis.

1.4.2.2 Comparison with $B_s^0 \to \phi\phi$

The decay $B_s^0 \to \phi\phi$ cannot occur at tree level. Thus, it has no $\mathcal{CP}$ phase in the *SM*. Similarly to $B_s^0 \to J/\psi\phi$, it requires an angular analysis to disentangle the $\mathcal{CP}$ even $L = 0$ state. This is even more challenging than for $B_s^0 \to J/\psi\phi$ as the branching ratio is only $\sim 10^{-5}$ compared to $\sim 10^{-3}$ for $B_s^0 \to J/\psi\phi$.

Like for $B_s^0 \to K^+K^-$, *NP* particles can introduce a phase σ^{NP}, responsible for $\mathcal{CP}$ violation in the decay, in addition to a possible $\mathcal{CP}$ violating *NP* mixing phase ϕ^{NP}. This leads to a measurement of

$$
\begin{aligned}
\hat{\Gamma}(B_s^0 \to (\phi\phi)_{L=0}) &= \Gamma_s \\
&\quad + \cos(\phi^{NP} + 2\sigma^{NP}) \frac{\Delta\Gamma_s^{SM}}{2} \cos(\phi^{NP}) \\
&\quad - \sin^2(\phi^{NP} + 2\sigma^{NP}) \frac{(\Delta\Gamma_s^{SM})^2}{2\Gamma_s} \cos^2(\phi^{NP}) \\
&\quad + \mathcal{O}\left(\frac{(\Delta\Gamma_s^{SM})^3}{\Gamma_s^2}\right). \qquad (1.83)
\end{aligned}
$$

This measurement will, however, only be able to serve as a cross check to $B_s^0 \to K^+K^-$, due to its branching ratio of $B_s^0 \to \phi\phi$ being lower than that of $B_s^0 \to K^+K^-$ by about a factor 2 and due to the requirement of an angular analysis.

1.5 Charm Lifetime Measurements

Lifetime measurements in the charm sector are of high interest at *LHCb* due to the abundant production of charm mesons. Section 5.3 gives more details on the potential for charm measurements at *LHCb*.

The approximation made in equation 1.61 for B mesons is not valid for charm mesons, as x_D and y_D are of the same order. The charm mixing parameters are very small, as shown in table 1.1,

$$x_D, y_D = \mathcal{O}(1\%) \ll 1. \tag{1.84}$$

Therefore, it is useful to expand the decay rates up to leading order in x_D and y_D [35]. For the $\mathcal{CP}$ even final state K^+K^-, squaring equation 1.44 and expanding leads to

$$\begin{aligned}
\Gamma(D^0(t) \to K^+K^-) &= e^{-\Gamma_D t}|A_{KK}|^2 \left\{1 + [\Re e(\lambda_{KK})y_D - \Im m(\lambda_{KK})x_D]\Gamma_D t\right\}, \\
\Gamma(\overline{D^0}(t) \to K^+K^-) &= e^{-\Gamma_D t}|\overline{A}_{KK}|^2 \left\{1 + [\Re e(\lambda_{KK}^{-1})y_D - \Im m(\lambda_{KK}^{-1})x_D]\Gamma_D t\right\},
\end{aligned} \tag{1.85}$$

where λ_{KK} follows the definition given in equation 1.46. For the non $\mathcal{CP}$ eigenstate $K^-\pi^+$ the decay rate is given by

$$\begin{aligned}
\Gamma(D^0(t) \to K^-\pi^+) &= e^{-\Gamma_D t}|A_{K\pi}|^2, \\
\Gamma(\overline{D^0}(t) \to K^+\pi^-) &= e^{-\Gamma_D t}|\overline{A}_{K\pi}|^2.
\end{aligned} \tag{1.86}$$

Experimentally, one obtains a contribution from the doubly Cabibbo suppressed modes $\overline{D^0} \to K^-\pi^+$ and $D^0 \to K^+\pi^-$. As their contribution would not significantly change the measured lifetime it is henceforth ignored.

Under the assumption of no significant direct $\mathcal{CP}$ violation the rates mentioned above can be written as

$$\begin{aligned}
\Gamma(D^0(t) \to K^+K^-) &= e^{-\Gamma_D t}|A_{KK}|^2 \left\{1 + |\lambda_{KK}|[y_D\cos(\phi) - x_D\sin(\phi)]\Gamma_D t\right\}, \\
\Gamma(\overline{D^0}(t) \to K^+K^-) &= e^{-\Gamma_D t}|A_{KK}|^2 \left\{1 + |\lambda_{KK}|^{-1}[y_D\cos(\phi) + x_D\sin(\phi)]\Gamma_D t\right\}, \\
\Gamma(D^0(t) \to K^-\pi^+) &= \Gamma(\overline{D^0}(t) \to K^+\pi^-) = e^{-\Gamma_D t}|A_{K\pi}|^2,
\end{aligned} \tag{1.87}$$

where ϕ is the phase of λ_{KK}. It should be noted that there is neither experimental evidence nor theoretical motivation for significant direct $\mathcal{CP}$ violation. Fitting with single exponentials would yield measurements of

$$\begin{aligned}
\hat{\Gamma}(D^0 \to K^+K^-) &= \Gamma_D \left\{1 + |\lambda_{KK}|[y_D\cos(\phi) - x_D\sin(\phi)]\right\}, \\
\hat{\Gamma}(\overline{D^0} \to K^+K^-) &= \Gamma_D \left\{1 + |\lambda_{KK}|^{-1}[y_D\cos(\phi) + x_D\sin(\phi)]\right\}, \\
\hat{\Gamma}(D^0 \to K^-\pi^+) &= \hat{\Gamma}(\overline{D^0}(t) \to K^+\pi^-) = \Gamma_D.
\end{aligned} \tag{1.88}$$

Using the production asymmetry

$$A_{prod} = \frac{N(D^0) - N(\overline{D^0})}{N(D^0) + N(\overline{D^0})}, \tag{1.89}$$

one can define the quantity

$$\begin{aligned} y_{\mathcal{CP}} &\equiv \frac{\hat{\Gamma}(D \to K^+K^-)}{\hat{\Gamma}(D^0 \to K^-\pi^+)} - 1 \\ &= y_D \cos(\phi) \left[\frac{1}{2}(|\lambda_{KK}| + |\lambda_{KK}|^{-1}) + \frac{A_{prod}}{2}(|\lambda_{KK}| - |\lambda_{KK}|^{-1})\right] \\ &\quad -x_D \sin(\phi) \left[\frac{1}{2}(|\lambda_{KK}| - |\lambda_{KK}|^{-1}) + \frac{A_{prod}}{2}(|\lambda_{KK}| + |\lambda_{KK}|^{-1})\right], \end{aligned} \tag{1.90}$$

where $\hat{\Gamma}(D \to K^+K^-)$ refers to the measurement of untagged $D \to K^+K^-$ decays. Approximating $|\lambda_{KK}|^{\pm 1} = 1 \pm \frac{A_m}{2}$, with the small parameter A_m, leads to

$$y_{\mathcal{CP}} = y_D \cos(\phi) - x_D \sin(\phi) \left(\frac{A_m}{2} + A_{prod}\right). \tag{1.91}$$

The case $\phi = 0$ results in the equality of $y_{\mathcal{CP}}$ and the mixing parameter y_D. In case measurements of the two parameters differ, this would be a sign of $\mathcal{CP}$ violation in the interference of mixing and decay. The situation $y_{\mathcal{CP}} \neq y_D$ can furthermore be used to determine $y_D \cos(\phi)$ and $x_D \sin(\phi)$ separately if a tagged measurement is used to determine both $\hat{\Gamma}(D^0 \to K^+K^-)$ and $\hat{\Gamma}(\overline{D^0} \to K^+K^-)$.

The lack of a precision measurement of $y_{\mathcal{CP}}$ and the abundance of D mesons produced at *LHCb* makes these measurements very interesting. A strategy for these analyses is presented in section 5.3.2.

1.6 Summary

This chapter outlined the Standard Model of particle physics which represents the current knowledge about particle interactions involving the electromagnetic, the weak, and the strong force. An overview of flavour physics, which builds the basis of the *LHCb* physics programme, was given.

The second part of the chapter described the potential of lifetime measurements. Measurements in $B^0_{(s)} \to h^+h'^-$ decays, particularly in $B^0_s \to K^+K^-$, offer possibilities for both *SM* tests via lifetime ratios and for sensitivities to *NP* through $\Delta\Gamma_s$ measurements. Finally, charm lifetime measurements allow measurements of the mixing parameter y and even sensitivity to $\mathcal{CP}$ violation in the charm sector via measurements of y_{CP}.

2 The LHCb experiment

> Wir dürfen und wir können nicht aus der Technik aussteigen, sondern wir müssen es besser lernen, sie mit ihrer eigenen Hilfe zu beherrschen.
>
> *Richard von Weizsäcker*

In 1993 three collaborations proposed experiments for heavy flavour physics at the Large Hadron Collider (*LHC*) at *CERN*, a proton-proton collider that would follow the Large Electron Positron collider (*LEP*) in its tunnel. The *LHB* collaboration envisaged a fixed target experiment with an extracted *LHC* beam [36]. The *COBEX*[1] experiment was intended to have a forward spectrometer to detect B decays from collisions of the two *LHC* beams [37]. The third proposal from the *GAJET*[2] collaboration planned the study of B physics with collisions originating from a gas jet target [38].

After having been asked to develop a common proposal for a collider based experiment with a forward spectrometer by the *LHC* experiments Committee (*LHCC*), the three collaborations merged and produced the *LHC-B* Letter Of Intent in 1995 [39]. Three years later the Technical Proposal of the slightly renamed *LHCb* collaboration followed [40]. After a review process, a revised version was produced [41] that reduced the overall material budget and at the same time optimised the trigger efficiency.

In order to set the scene for describing the *LHCb* experiment in more detail, this chapter will start with an overview of the *CERN* accelerator complex. This is followed by a description of each of the components of the *LHCb* detector, its data acquisition system, trigger and software.

2.1 The Accelerator Complex at CERN

2.1.1 From Linac to SPS

CERN started as a high energy physics laboratory in 1954. One of the first accelerators

[1] Collider Beauty Experiment
[2] Gas Jet

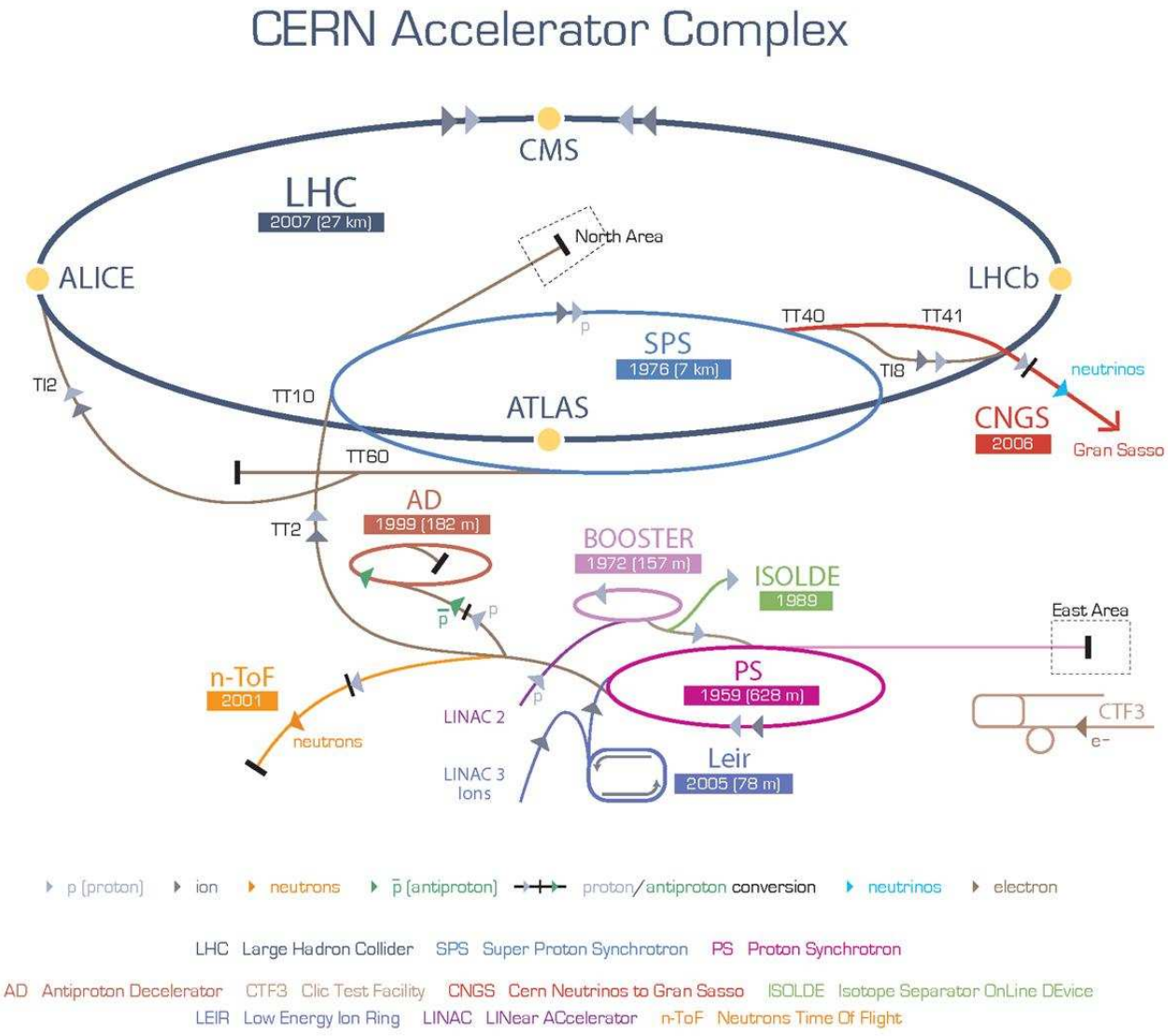


Figure 2.1: *Overview of the CERN accelerator complex. Reproduced from [42].*

to be constructed at *CERN* was the Proton Synchrotron (*PS*) which started operation in 1959. It is now a part of the accelerator chain preceding the *LHC* [43, 44, 45]. Two types of particles, protons and heavy ions (lead nuclei), can be accelerated and collided in the *LHC*. As *LHCb* is designed to acquire data from proton-proton collisions, only their path will be described.

All protons originate from a small bottle of hydrogen gas. Through a high electric field and so-called stripping foils the hydrogen molecules are ionised and thus serve as the proton source. Their first stage of acceleration is a linear accelerator called *LINAC2* (see Fig. 2.1). A linear accelerator is used at the start of the accelerator chain as the protons quickly gain in speed.

After having been accelerated from thermal energies to a kinetic energy of 50 MeV, the protons are injected into the Proton Synchrotron Booster (*PSB*). It is the first circular accelerator that the protons pass through. The *PSB* consists of four rings sitting on top of each other and used in parallel in order to increase the possible proton intensity. At a kinetic energy of 1.4 GeV the protons are injected into the aforementioned Proton Synchrotron (*PS*). At an energy of 25 GeV they are transferred into the Super Proton Synchrotron (*SPS*) which forms the last acceleration stage before the *LHC*. The 7 km circumference *SPS* accelerates the protons up to an energy of 450 GeV. At this energy they are injected via two transfer lines in opposite directions into the two beam pipes of the *LHC*. One of the transfer lines ends close to *LHCb*, as discussed further in section 3.4.

2.1.2 The LHC

When arriving at the *LHC* the protons come in bunches of up to 1.1×10^{11} particles. Under nominal conditions each *LHC* beam will have 2808 of these bunches which means that their separation in the beam pipe is only about 8 m or 25 ns. The particles are accumulated with increasing bunch charges in the various acceleration stages to achieve these high intensities. This starts with two extractions from the *PSB*, which are separated by 1.2 s and repeated every 3.6 s. The period for accumulating and accelerating particles in the *SPS* is 21.6 s. To fill the whole *LHC* then takes just under 10 minutes.

The beams need to pass through separate beam pipes with opposite magnetic fields, as the *LHC* is a proton-proton collider. The only way to achieve the field strength necessary to keep TeV protons on a 27 km ring under the given spatial constraints is with superconducting magnets. After acceleration, the protons will finally have an energy of 7 TeV which requires a peak magnetic field of 8.33 T.

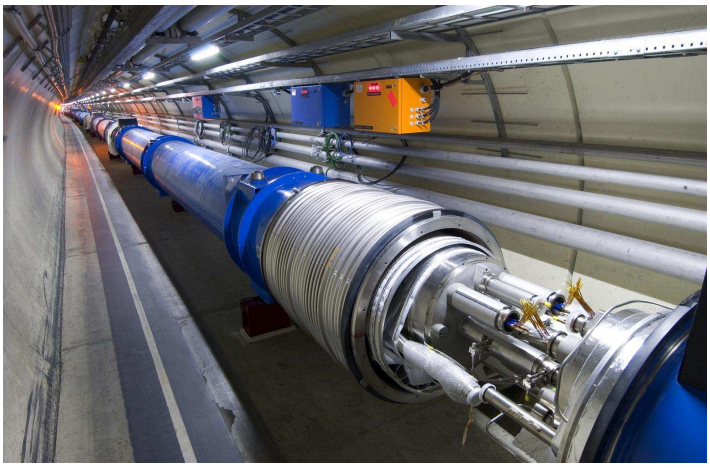


Figure 2.2: *Photo of the open interconnection between two LHC dipoles showing the beam pipes, the connection of the superconducting cables and the tubes for the liquid helium supply. Reproduced from [46].*

The protons are kept on a circular trajectory by 1232 dipole magnets (see Fig. 2.2). 392 quadrupole magnets are used to ensure a well focused beam. In addition, about 8000 smaller magnets are used to fine tune the beams. This is necessary to complete the beam optics of the main magnets and to compensate for imperfections in the construction and performance of the accelerator. The main superconducting magnets are cooled using a total of 120 tons of superfluid helium.

The *LHC* is split in eight sectors, each of which consists of repeated sections of dipole and quadrupole magnets (see Fig. 2.3). Interaction points (*IP*s) are located in the centre of the respective octants shown in figure 2.3. Only at four of the eight *IP*s are actual collision points at which experiments are located: *ATLAS* is located at *IP* 1, *ALICE* at *IP* 2, *CMS* is at *IP* 5, and finally *LHCb* is at *IP* 8.

The other four *IP*s are used to host beam related instrumentation. Beam cleaning facilities are located at *IP* 3 and *IP* 7. At *IP* 3 the so-called momentum cleaning removes particles with large ($> 6\sigma$) longitudinal oscillation amplitudes and at *IP* 7 particles with large transversal oscillation amplitudes are removed. The main beam dump for both beams sits at *IP* 6. The protons are accelerated using high frequency cavities at *IP* 4. These are superconducting cavities operating at a frequency of 400.8 MHz. With an energy gain per turn of 485 keV it takes the *LHC* 20 minutes to ramp up to nominal beam energy of 7 TeV, where the speed is limited by the ramping speed of the magnets.

Figure 2.4 shows how *CERN* and its accelerator complex is located between Geneva in the east and the Jura mountains in the west. Following a stratum of rock rising towards the mountains, the *LHC* tunnel has an inclination of 3.6 mrad. The *LHCb*

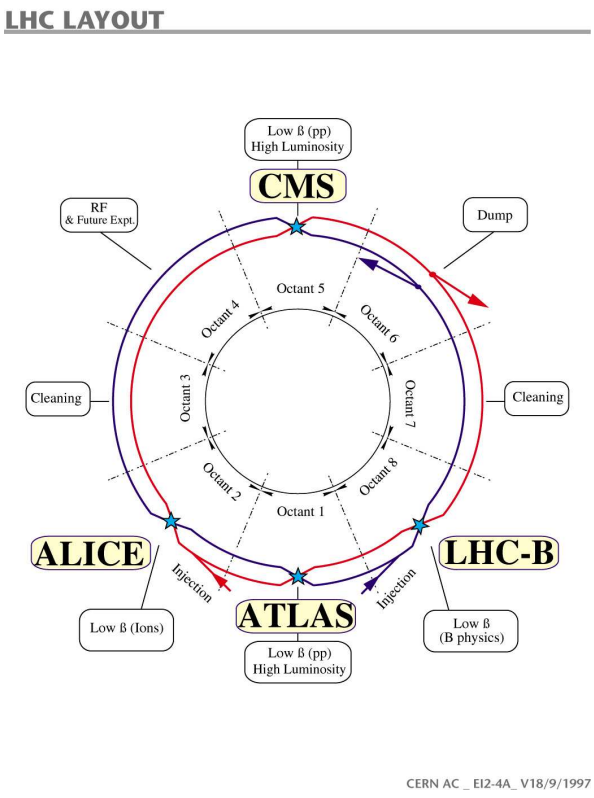


Figure 2.3: *Schematic view of the LHC. The four interaction points where the experiments are located and the regions for beam cleaning and the beam dump are indicated. Reproduced from [47].*

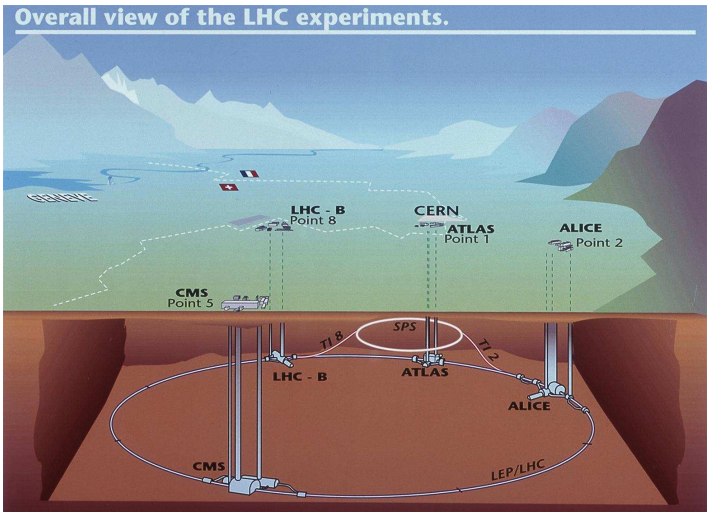


Figure 2.4: *View of the location of the LHC in the Geneva region between the Jura on the right and the Alps on the left. Reproduced from [48].*

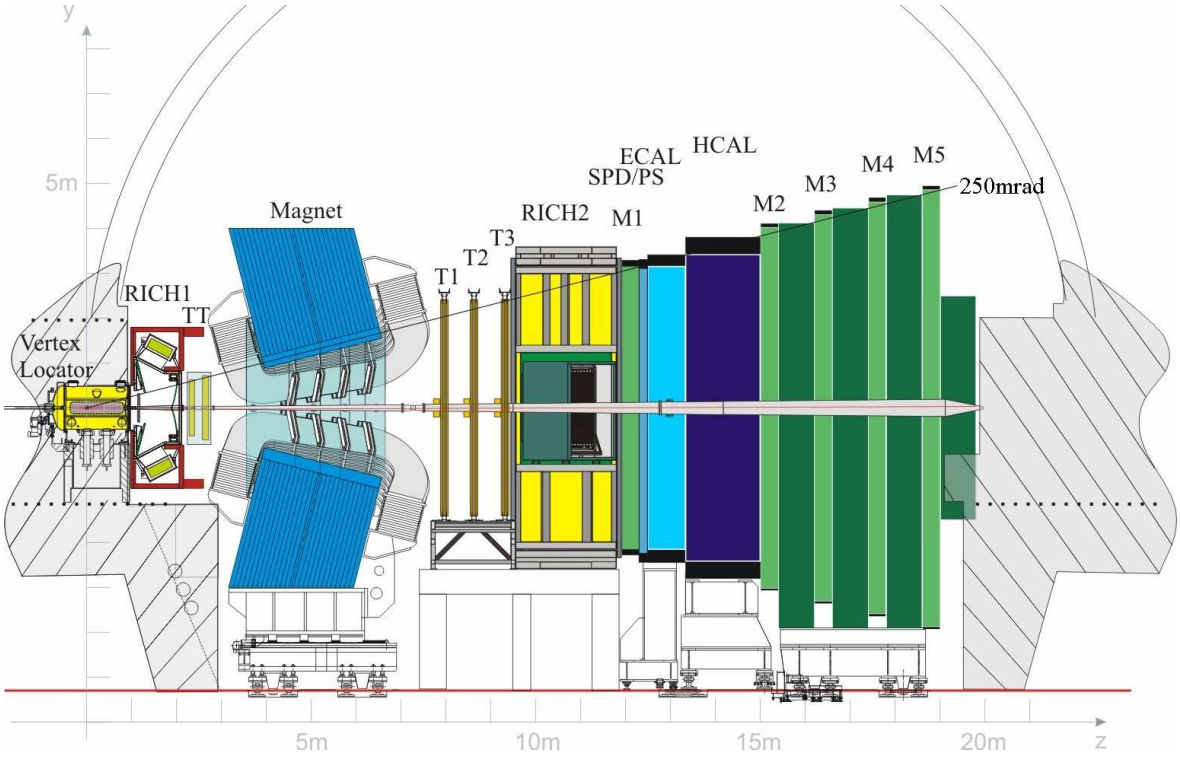


Figure 2.5: *Schematic drawing of the LHCb detector as seen from inside the LHC ring. Reproduced from [49].*

experiment is located at the French-Swiss border close to Geneva airport in the east of the ring.

2.2 The LHCb Detector

The *LHCb* experiment has been conceived as a so-called forward spectrometer – a detector that covers only a comparatively small region of the full solid angle around the beam line. The nominal interaction point of the two proton beams has been moved from the centre to one side of the cavern to maximise the space for a single arm spectrometer. As shown in figure 2.5 the Vertex Locator of the *LHCb* experiment (*VELO*) is located at one side of the experimental hall while the other sub-detectors fill the 20 m up to the other end. All quantities given subsequently in this section are quoted from reference [49] unless otherwise stated.

B mesons are produced highly forward boosted such that they are nearly collinear with one of the two beams. Both B and $\overline{B}$ mesons fly in the same direction (see figure 2.6). Hence, *LHCb* can detect roughly 34% of the B mesons produced while covering only about 2% of the full solid angle.

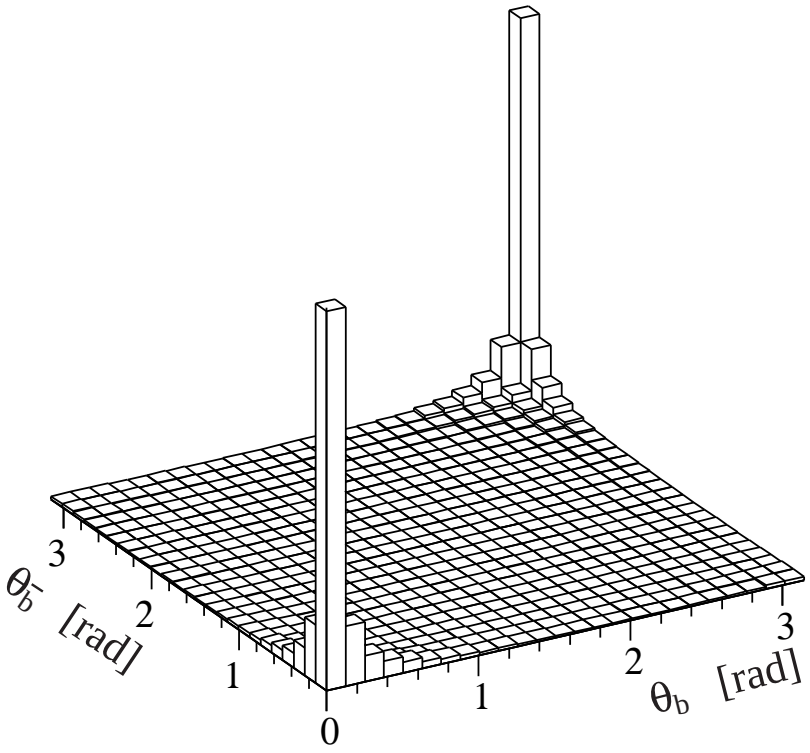


Figure 2.6: *Angular distribution of the production of B and $\overline{B}$ mesons with respect to the beam axis at the LHC. Reproduced from [40].*

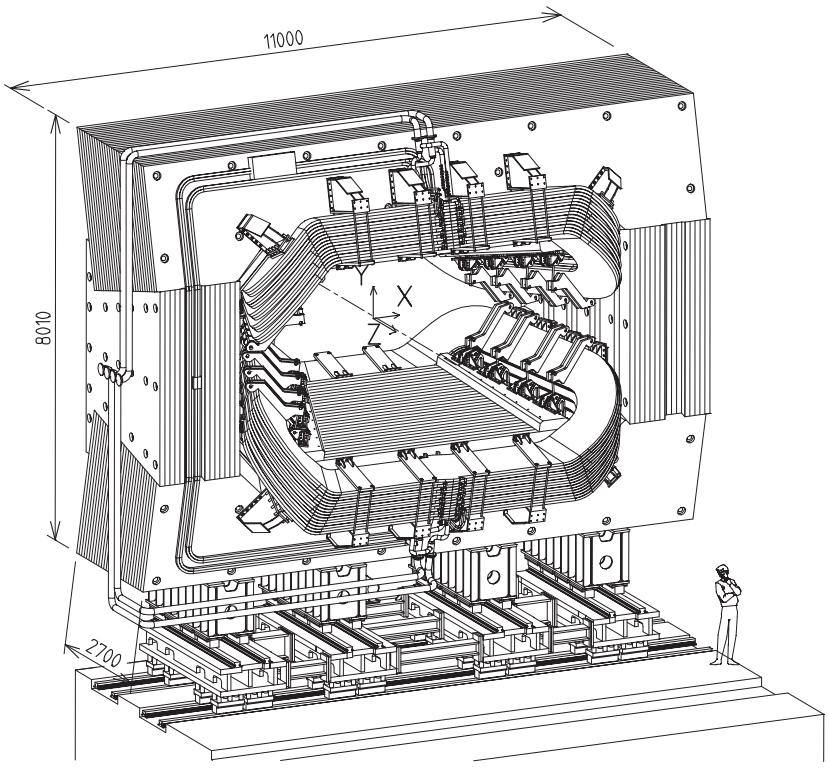


Figure 2.7: *Schematic drawing of the LHCb dipole magnet. Reproduced from [49].*

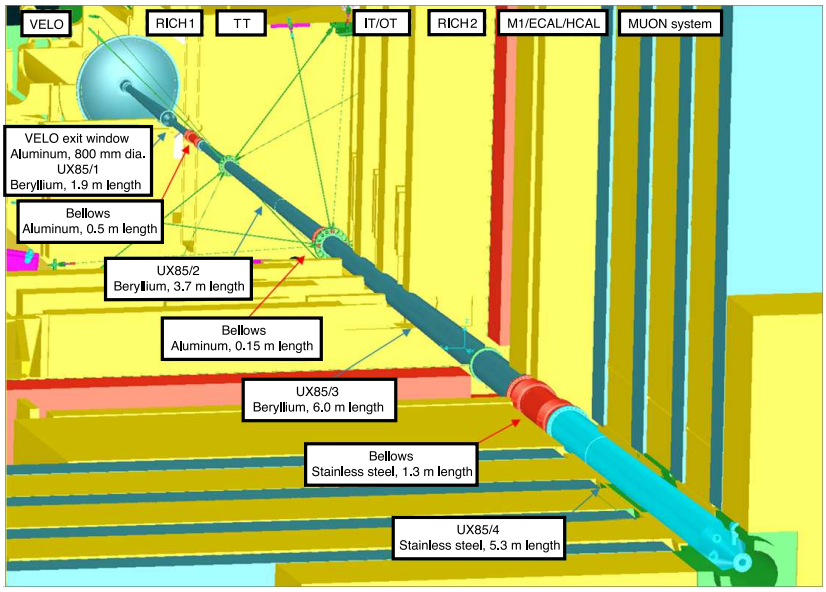


Figure 2.8: *Schematic drawing of the beam pipe inside the LHCb detector. Reproduced from [49].*

The *LHCb* tracking system consists of *VELO* (see Sec. 2.2.1), the two silicon trackers Tracker Turicensis[3] (*TT*) and Inner Tracker (*IT*) (see Sec. 2.2.2), Outer Tracker (*OT*) (see Sec. 2.2.3), and the magnet. A water-cooled dipole magnet is between *TT* and *IT* to allow the momentum of charged particles to be measured (see Fig. 2.7). Its magnetic field reaches a peak of just over 1 T and has a total bending power of about 4 Tm over a length of 10 m, roughly corresponding to a momentum kick in the horizontal plane of 1.2 GeV/c.

The *LHCb* coordinate system is defined as a right-handed Cartesian system with the z axis along the nominal direction of beam 1 (towards *IP* 1), the y axis roughly upwards[4] and the x axis towards the outside of the *LHC* ring. For reasons of structural stability not all sub-detectors have been constructed according to this coordinate system. All detectors downstream of the magnet have their centre along the nominal beam line but their vertical axes coincide with the gravitational axis.

The particle identification system consists of two Ring Imaging CHerenkov (*RICH*) detectors (see Sec. 2.2.5), one between *VELO* and *TT* and the other between the tracking stations and the calorimeters (see Sec. 2.2.6), which, together with the muon identification system (see Sec. 2.2.7), complete the system.

The *VELO* is located in the beam vacuum (see below). An exit window of 800 mm

[3] formerly known as Trigger Tracker

[4] As the beam line is inclined by 3.6 mrad with respect to the horizontal plane, the y axis does not coincide with the gravitational axis.

Figure 2.9: *Drawing of the closed VELO (left) and photo of modules of a fully assembled VELO half (right). The different sensor colours in the drawing indicate R and Φ sensors, respectively. Reproduced from [50, 51]*

diameter connects the *VELO* to a beam pipe (see Fig. 2.8). The first 12 m of this conical beam pipe are made of Beryllium to reduce the amount of material traversed by highly forward boosted particles. The final 7 m of the beam pipe inside the calorimeters and the muon stations consist of stainless steel.

2.2.1 The Vertex Locator

The *VELO* is a very special and challenging sub-detector. It has the highest position resolution of any device in all four main *LHC* experiments. Of their sub-detectors it is the closest to the *LHC* beam. It is operated inside the *LHC* vacuum. Hence, its design has required a number of novel ideas.

2.2.1.1 The VELO System

Given the goal to measure tracks of small angles with respect to the beam line a series of circular disks is the most natural choice from a geometrical point of view. This has been adopted in the design of the *VELO* which consists of two halves with 42 half circular silicon sensors each (see Fig. 2.9). The split into two halves allows the retraction of each half away from the beam line. This is necessary in order to protect the detector during injection of the beam into the *LHC*, acceleration to nominal momentum, the beam dump, or other unstable beam conditions.

Each *VELO* half contains 21 modules, each of which has two sensors. One sensor on each module measures the radial coordinate (R sensor), the other one measures the angular coordinate around the beam line (Φ sensor). In addition to the 84 *VELO*

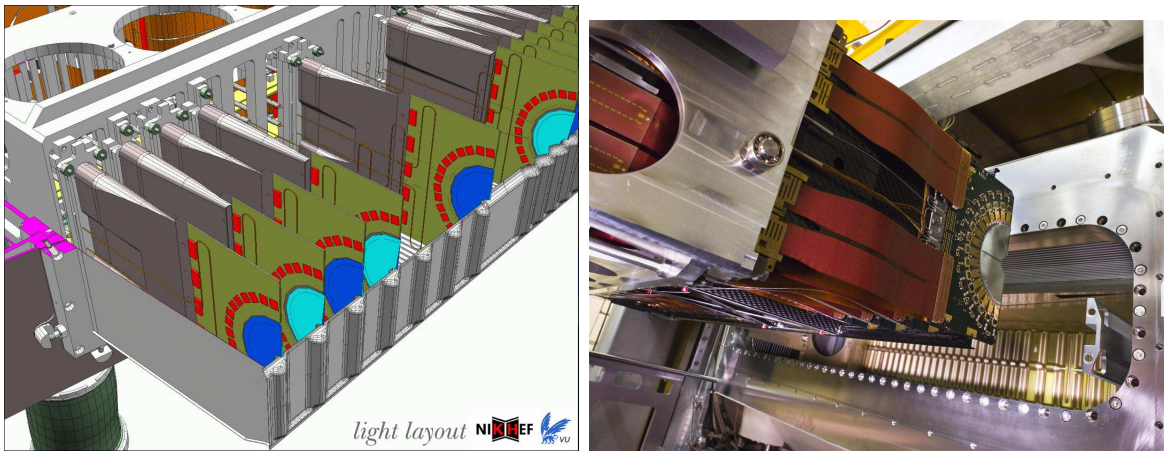


Figure 2.10: *Drawing of a VELO half with the RF-box shown in the lower half and photo of its insertion into the vacuum vessel. Reproduced from [50, 52]*

sensors, two R sensors placed at the upstream end of either $VELO$ half form the so-called pile-up stations. They can be used in the trigger to suppress events with very high track multiplicity or events with more than one hard interaction.

The $VELO$ halves are placed inside so-called RF-boxes (see Fig. 2.10). This is to protect the $VELO$ sensors from the RF field caused by the beam and to suppress the beam's wake field which can influence subsequent bunches. The side of the boxes facing the beam is made of 300 μm thin Aluminium foil to reduce the amount of their material. This requires the $VELO$ sensors to be operated in a secondary vacuum with a differential pressure with respect to the beam vacuum of less than 10^{-3} mbar.

As shown in figures 2.9 and 2.10 the front side of the RF-boxes has an undulating shape. This allows the sensors from both halves to slightly overlap such that they create a gap-less acceptance in ϕ.

The motion system that retracts and reinserts the $VELO$ halves is able to move each half independently away from the beam by 30 mm in the horizontal plane. The system can adjust the 'closed' position to anywhere within 5 mm of the nominal beam position both horizontally and vertically, as the beam position may change from fill to fill.

The motion system has been designed to have a reproducibility of about 3 μm for a certain reference position. The position of each of the halves with respect to their reference position is read out by resolvers with a precision of 10 μm. The use of the motion system information in the reconstruction of $VELO$ hits is explained further in section 3.1.3.1. A relative alignment of the positions of the two halves is essential, as the motion system can only determine the position of each half separately. A solution for this task is presented in section 3.2.4.

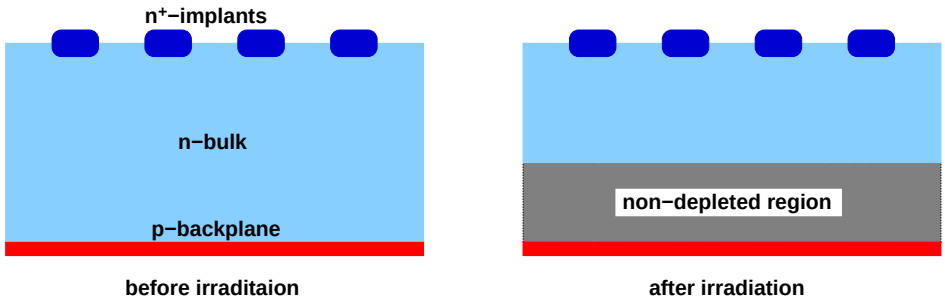


Figure 2.11: *Cross-section showing the effect of irradiation on n^+-on-n sensors.*

2.2.1.2 The VELO sensors

The *VELO* sensors [53] are made of n-doped silicon wafers with n^+ implants that form the strips. The backplane consists of a pattern of p-doped guard rings.

For a non-irradiated sensor the n-doped bulk is depleted by applying a bias voltage. The size of the signal created by a particle that passes through the sensor is proportional to the size of the depletion region. The voltage necessary for full depletion changes with irradiation. Thus, with a given maximal voltage the sensors can no longer be fully depleted after a certain radiation dose. The advantage of n^+-on-n sensors is that their non-depleted region is at the backplane (see Fig. 2.11). Hence, the (smaller) signal is still collected by the implants without additional diffusion that would be caused by a non-depleted region near the implants.

For very high radiation doses the bulk of the sensors undergoes type inversion, i.e. the sensors are then n^+-on-p sensors. In this case the non-depleted region would be located at the implants, thus generating a larger spread of the signal with lower intensity on the individual strips.

It is crucial to permanently monitor the performance of the sensors[5], as the changes with irradiation are continuous. A thinner depleted region or a larger charge diffusion can both influence the position resolution as they change the charge distribution among neighbouring strips. A method to measure the sensor resolution is introduced in sections 3.3.3.3 and 3.4.3 together with results from data taken with non-irradiated *VELO* sensors.

All sensors have a cut-out up to a radius of 7 mm to form a hole that allows the beams to pass through the *VELO*, as shown in figure 2.12. The active area of both R and Φ sensors extends from a radius of 8.2 mm to a radius of 41.9 mm. The 2048 strips on the R sensors form arcs of concentric circles. The R sensors are separated in four

[5]The *VELO* sensors have been qualified to withstand a dose of 1.3×10^{14} n_{eq}/cm^2 roughly equivalent to three full years under nominal running conditions.

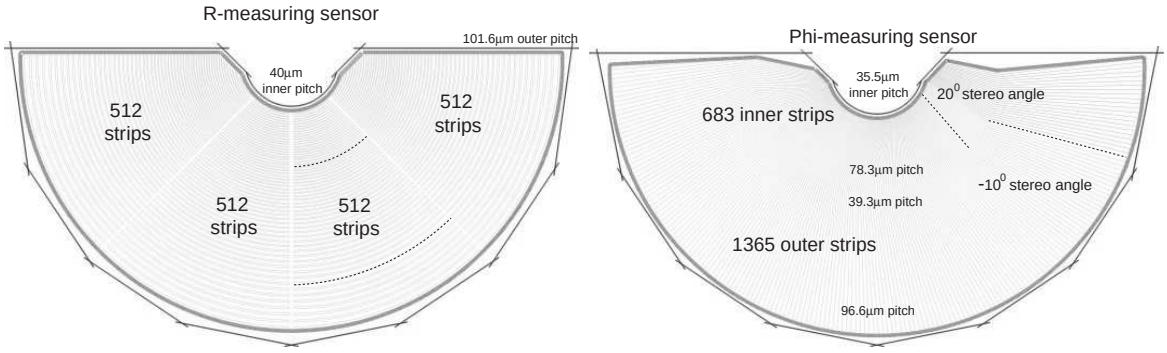


Figure 2.12: *Layout of the VELO R and Φ sensors. Reproduced from [41]*

sectors such that each strip covers approximately 45° in ϕ. The strip pitch[6] increases linearly with r from 40.0 μm to 101.6 μm.

The strips on the Φ sensors are separated into inner and outer regions with the boundary at a radius of 17.25 mm. The boundary serves several purposes. By reducing the overall strip length the occupancy per strip is reduced, which helps the pattern recognition in the track finding algorithm. The shorter strip length also reduces the strip capacity and, hence, the noise. It also allows the strips in the outer region to start at a smaller pitch which improves the overall resolution. The 683 inner strips have a pitch ranging from 35.5 μm to 78.3 μm, while the 1365 outer strips have a pitch between 39.3 μm and 96.6 μm.

The strips on the Φ sensors are oriented in the radial direction to first order. As described on page 78, they are lying on tangents to a circle of 2.8 mm (3.1 mm) radius for the inner (outer) region. Thus, they have a stereo angle with a radial line of $-20°$ (10.35°) at their respective innermost radius. The Φ sensors on subsequent modules are rotated by 180° such that their strips have opposite stereo angles to avoid a four-fold ambiguity in the pattern recognition of two tracks.

2.2.1.3 The VELO readout chain

The signals from the strips are routed via a second metal layer to pitch adaptors which are placed around the outer edge of the sensors. From there, the signals enter one of 16 Beetle front-end chips, i.e. each chip processes the signals of 128 strips.

For R sensors each sector of 512 strips is read out by four Beetle chips. The innermost 128 strips are always read out in the opposite order than the outermost 384. The readout order is reversed for neighbouring sectors. In the case of asymmetric cross-talk

[6]The strip pitch is defined as the distance between the midpoints of the centre of the current strip to the centres of the neighbouring strips.

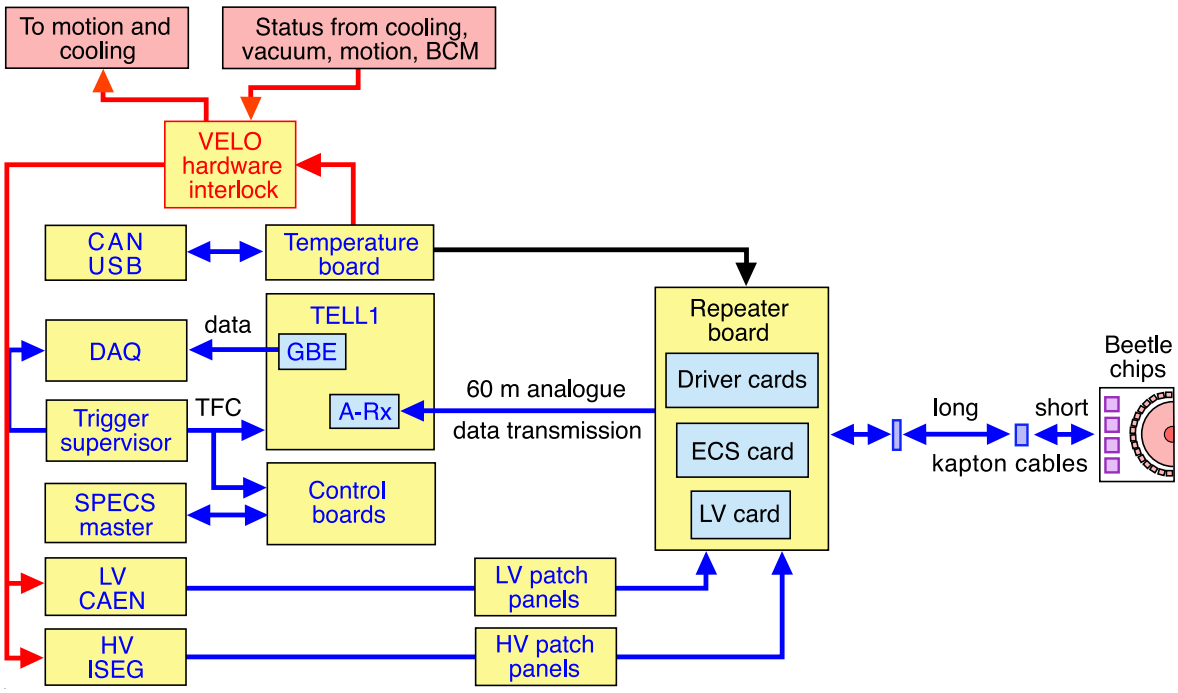


Figure 2.13: *Overview of the VELO readout chain. Reproduced from [49].*

this pattern can result in non-trivial biases in the reconstructed hit positions. This effect has been observed and successfully corrected for in test beam data as explained in section 3.3.3.1.

For Φ sensors the readout order is more complicated. The signals from inner strips are routed on top of outer strips to the pitch adaptors. Despite there being twice as many outer strips than inner strips, the readout pattern does not follow a modulo three pattern but rather a modulo 12 pattern. Due to this readout order, cross-talk effects similar to those for R sensors are not expected for Φ sensors.

The full *VELO* readout chain is shown in figure 2.13. Starting with the Beetle front-end chips shown on the right, the signals are routed out of the vacuum vessel into repeater boards via two sets of kapton cables. From the repeater boards the signals go via 60 m analogue cables to a radiation free area where they are fed into the so-called *TELL1* boards [54].

The *TELL1* boards perform a range of signal processing steps:

- **Pedestal subtraction**: The analogue signals all have a common offset of around 512 ADC counts[7]. The exact value of this offset can either be measured prior to data taking or followed with a running average.

[7] The total available range is 1024 ADC counts. 1 ADC count is equivalent to ≈ 450 e^-, hence a minimum ionising particle produces a signal around 50 ADC counts.

- **Beetle cross-talk correction**: This correction removes cross-talk from the Beetle chip header bits in the raw data.
- **FIR filter**: This finite impulse response filter removes cross-talk effects by applying corrections to three to five neighbouring channels.
- **MCMS**: The mean common mode suppression counteracts shifts in the Beetle baseline caused by very large signals.
- **Reordering**: To simplify the reconstruction the channels are reordered to reflect the strip numbering on the sensors, i.e. physically neighbouring strips have consecutive numbering after reordering.
- **LCMS**: The linear common mode suppression removes sensor noise common to strips of the same region, hence its place after the reordering.
- **Zero suppression**: The final processing step forms clusters by applying several thresholds to the ADC values of the individual strips.

The output of the *TELL1* processing is a set of clusters that can subsequently be combined to form tracks by software pattern recognition algorithms.

A software package called `VETRA` [55] has been developed for the analysis of raw *VELO* data. Most notably, it includes a bit-perfect emulation of the processing algorithms of the *TELL1* boards which allows detailed debugging, analysis, and tuning of the *VELO* readout.

2.2.1.4 The VELO assembly and performance

The individual *VELO* modules were assembled at Liverpool. Several optical survey measurements were made to ensure the precise placement of the sensors on the module. The final *VELO* halves were assembled at *CERN*, followed by a survey of the whole system. Each module underwent a so-called burn-in procedure upon arrival at *CERN* [56] to ensure stability in the long term operation. In this procedure, the modules went through several temperature cycles in a vacuum chamber before and after which the functionality and performance of the sensors was assessed. The measurements done during the burn-in process serve as a reference for regular checks during normal operation.

The sensor resolution has been measured with the final readout chain in a test beam to assess the performance of the *VELO*. A detailed description of the apparatus and

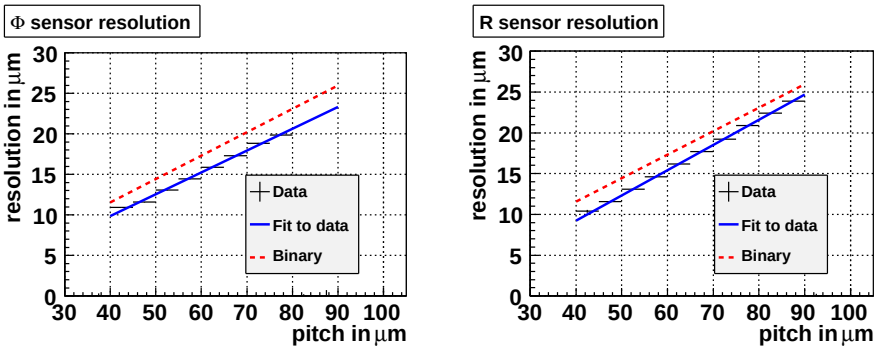


Figure 2.14: *VELO sensor resolution as measured with data from the 2006 test beam. The dashed line indicates the resolution for a binary readout.*

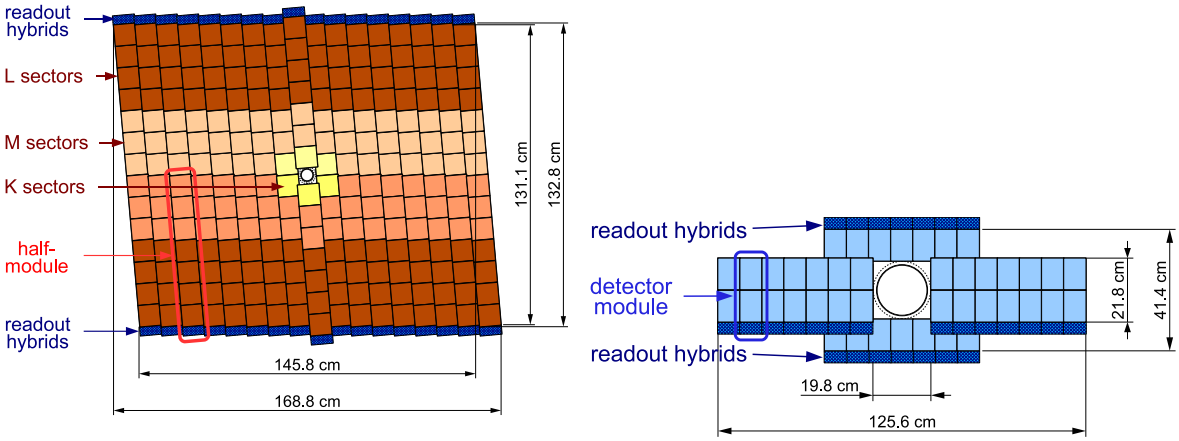


Figure 2.15: *Schematic view of a TT v-layer (left) and an IT x-layer (right). Reproduced from [49].*

the results is given in section 3.3. As shown in figure 2.14 the measured resolution for both R and Φ sensors is well below the binary resolution that would be expected for a binary readout scheme. For a pitch of 40 μm both types of sensors reach a resolution below 10 μm.

2.2.2 The Silicon Trackers

The Silicon Trackers (ST) comprise the other two tracking detectors, apart from the $VELO$, which use silicon strip technology. The Tracker Turicensis (TT) is located before the magnet and the Inner Tracker (IT) is positioned behind the magnet.

The TT (see Fig. 2.15) consists of one station comprising four layers of silicon strip detectors which are arranged in modules. The first and last layer of the station have their silicon strips oriented vertically, i.e. they measure x and are hence called x-layers. The middle two layers, called u- and v-layer, have their strips oriented at an angle of

$\pm 5^\circ$ with respect to the vertical, respectively. This x-u-v-x scheme is repeated for the IT.

Each layer consists of 17 modules, which are split into upper and lower half-modules with 7 sensors each. The front-end readout electronics sit at the outer end of the half-modules to where the signals are transported from the sensors by kapton cables. The readout of the sensors is such that only the sensors around the beam pipe are read out individually (K sectors). The innermost three sensors (or two around the beam pipe) are read out such that strips of the same vertical co-ordinate are combined (M sectors). The outermost four sensors on each half-module (L sectors) are read out together.

The sensors are made of 500 μm thick p-on-n silicon. Having a size of 9.64×9.44 cm^2 they consist of 512 strips at a pitch of 183 μm. These sensors are designed to have a single hit resolution in x of about 50 μm. Through a staggering of the modules a gap-less acceptance is achieved.

In contrast to the TT, the IT does not cover the full acceptance. Its total active area of about 4 m^2 is meant to cover the high multiplicity region around the beam pipe, while the rest of the acceptance is covered by the OT (see below). The IT consists of three stations that all comprise four layers with the same x-u-v-x scheme as the TT. Each station is split into four boxes, one above, one below and one on either side of the beam pipe (see Fig. 2.15).

The top and bottom boxes each have 7 one-sensor modules, while the side boxes have the same number of two-sensor modules. The sensors are again made of p-on-n silicon with a thickness of 320 μm for the one-sensor modules and 410 μm for the two-sensor modules. Each sensor has a size of 7.6×11 cm^2 and holds 384 readout strips with a pitch of 198 μm. The IT sensors are designed to have a single hit resolution of roughly 50 μm.

2.2.3 The Outer Tracker

The Outer Tracker (OT) like the IT consists of three stations of x-u-v-x layers (see Fig. 2.16). The individual layers consist of modules that each contain two staggered layers of straw tubes. The single straws are up to 4.85 m long and are placed with a pitch of 5.25 mm. They operate with a mixture of Ar (70%) and CO_2 (30%) gas that has been chosen for its fast drift time (< 50 ns). As this is still slower than the time between two LHC bunch crossings (25 ns) the drift time is digitised every 25 ns and upon a positive trigger decision is read out for three consecutive bunch crossings. The active area of nearly 29 m^2 covers an acceptance of 300 mrad in the horizontal magnet

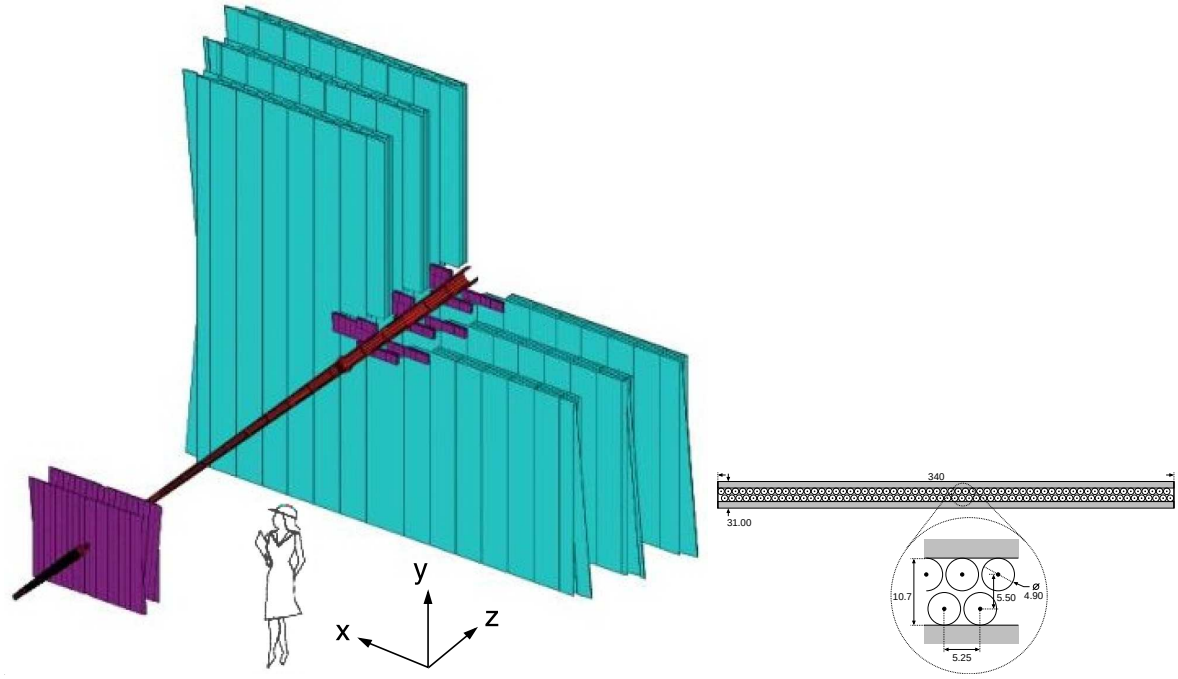


Figure 2.16: *Layout of the tracking system (left) and cross-section of an OT module (right). Reproduced from [49].*

bending plane and of 250 mrad in the vertical plane. Using the drift time information a single hit resolution in x of 200 μm can be achieved.

2.2.4 Performance of the Tracking System

A high performance of the tracking system is mandatory to perform heavy flavour physics measurements. A good mass resolution allows a tight selection and, hence, a good suppression of background events. Aiming for a mass resolution of 10 MeV for the decay $B_s^0 \to D_s^- \pi^+$ translates into a required momentum resolution of $\delta p/p \approx 0.4\%$. This in turn needs both precise knowledge of the magnetic field throughout the detector and excellent position resolution, particularly in the bending plane.

A second important point is the track reconstruction efficiency. An efficiency of 95% per track would translate in an overall reconstruction efficiency of 81% for a decay with four tracks like $B_s^0 \to D_s^- \pi^+$.

However, not all tracks can be fully reconstructed as they escape the acceptance of the detector. Figure 2.17 shows the y-component of the magnetic field on top and an illustration of the various track categories below. For physics measurements, the tracks of highest interest are the Long tracks which pass through all tracking stations. Tracks from decays of longer lived particles like K_S or Λ can escape the $VELO$ and would then

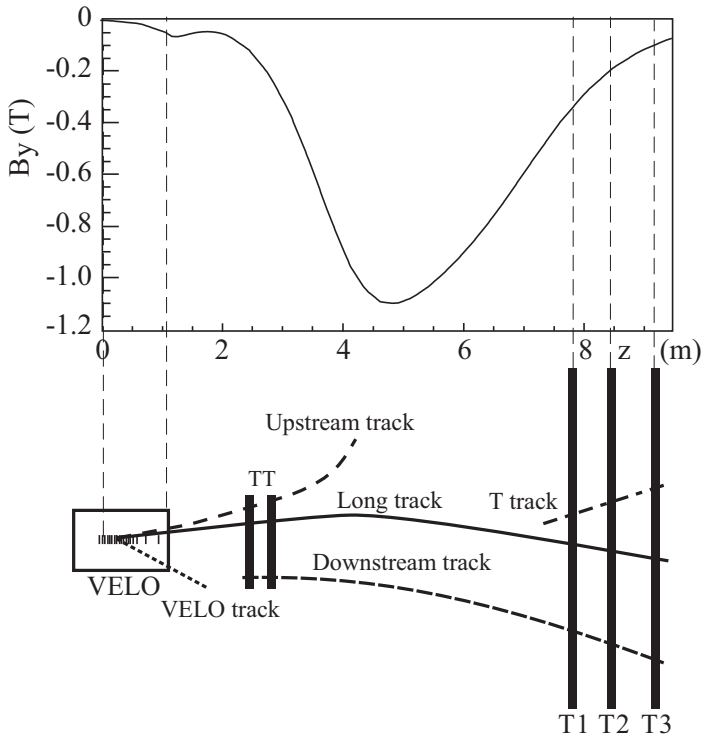


Figure 2.17: *Illustration of the different track types and of the y-component of the magnetic field. Reproduced from [49].*

only be reconstructed as Downstream tracks. Particles of very low momentum can be deflected by the magnet such that they are reconstructed only as Upstream tracks. These can for instance occur in the decay chain $D^{*+} \to D^0\pi^+$ (see Sec. 5.3.1). *VELO* tracks add to the precision of the reconstruction of vertices from primary interactions. Pure *T* tracks are of no interest to physics measurements.

According to simulation studies, the track reconstruction efficiency for Long tracks with a momentum above 10 GeV/c is 94%. The momentum resolution varies from $\delta p/p = 0.35\%$ to $\delta p/p = 0.55\%$ depending on the track momentum, as shown in figure 2.18. Another important quantity is the impact parameter (*IP*), i.e. the shortest distance of an extrapolated track to the primary vertex. The *IP* is a powerful variable to distinguish B decays with their finite lifetime from prompt decays. Its resolution is determined to be $\sigma_{IP} = 14\ \mu\text{m} + 35\ \mu\text{m}/p_T$.

2.2.5 The RICH Detectors

LHCb has two Ring Imaging Cherenkov[8] (*RICH*) detectors, one located immediately

[8]Cherenkov is chosen as the transcription of the original Cyrillic Черенков, as it best suggests the correct pronunciation as an English word.

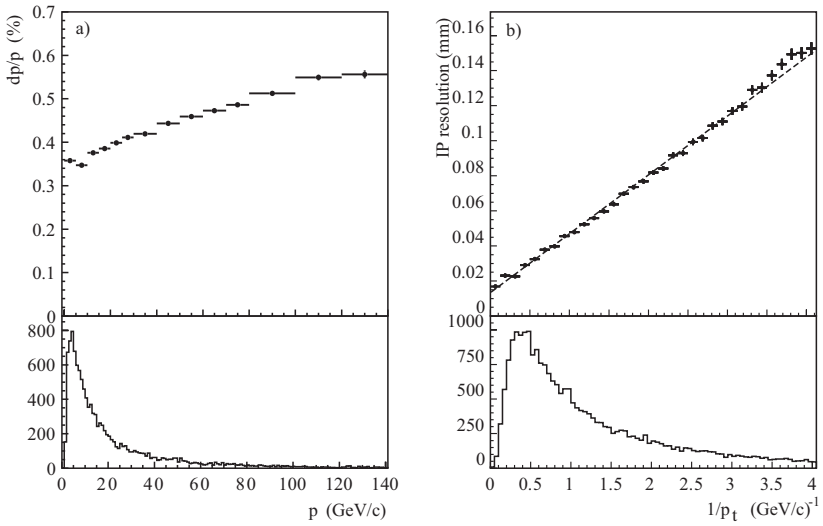


Figure 2.18: *Resolution of momentum (a) and impact parameter (b). Reproduced from [49].*

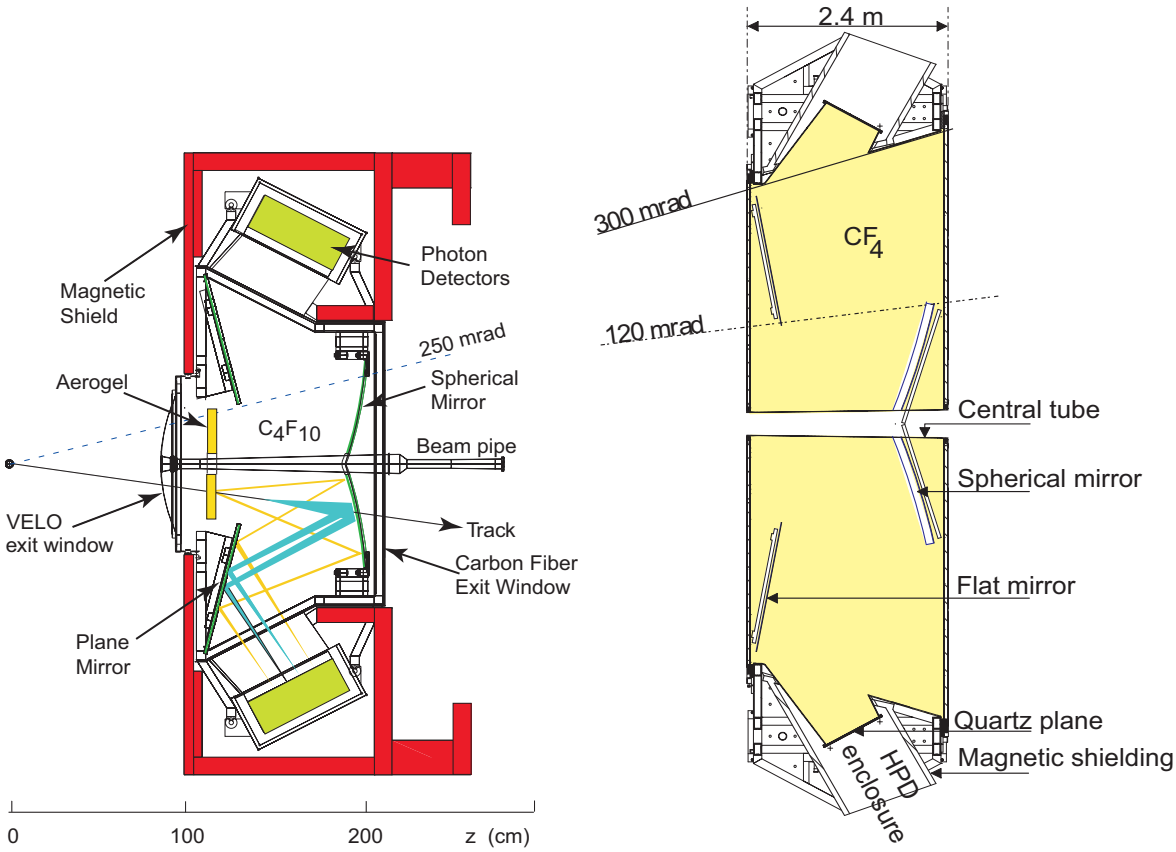


Figure 2.19: *Schematic view of RICH1 (left) and RICH2 (right). Reproduced from [49].*

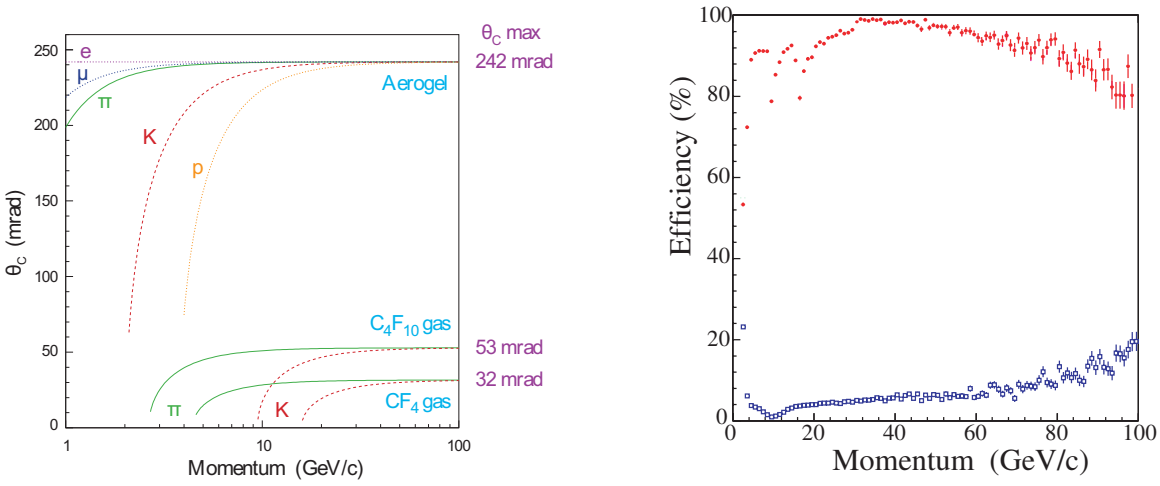


Figure 2.20: *Cherenkov angle as function of momentum for different particles and different radiators (left). Kaon identification efficiency (red) and pion misidentification rate (blue) as function of momentum (right). Reproduced from [49].*

after the *VELO*, and the second following the *IT* and *OT*. As shown in figure 2.19 both operate with two mirror planes to guide the Cherenkov light to the photon detectors which are placed outside the acceptance.

RICH1 has two radiator materials: blocks of 5 cm thick aerogel, and C_4F_{10} gas. This allows the identification of particles in a momentum range of 1 GeV/c to 60 GeV/c as shown in figure 2.20. *RICH2* operates with CF_4 gas and has a good separation power for momenta between 15 GeV/c to above 100 GeV/c. Between 2 GeV/c and 100 GeV/c track momentum the average kaon identification efficiency is 95% with a pion misidentification rate of 5%.

The photon detection in both *RICH* detectors uses Hybrid Photon Detectors (*HPD*). Photoelectrons from the cathode are accelerated and focused onto a silicon pixel sensor. The granularity on the cathode plane of 2.5×2.5 mm^2 allows a Cherenkov angle resolution of 1.6 mrad.

2.2.6 The Calorimeters

The Calorimeters in *LHCb* follow the typical scheme of an electromagnetic calorimeter (*ECAL*) followed by a hadronic calorimeter (*HCAL*). The readout is organised in cells of increasing size with increasing distance to the beam axis. The cells have a projective geometry through the various calorimeter stations.

A Scintillating Pad Detector (*SPD*) has been placed in front of the *ECAL* to identify

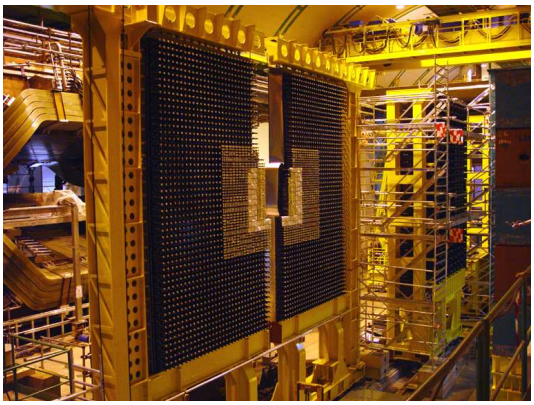
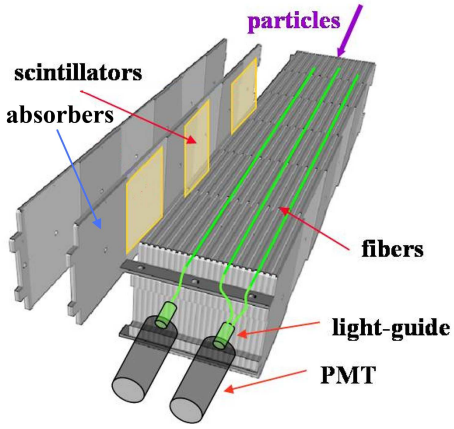
(a)  (b) 

Figure 2.21: *Module of the electromagnetic calorimeter (a) and complete system (b). Reproduced from [49].*



(a) (b) 

Figure 2.22: *Module of the hadronic calorimeter (a) and complete system (b). Reproduced from [49].*

electrons. It is designed to register the signals of charged particles with scintillator pads read out by wavelength shifting fibres into multi anode photo multiplier tubes. Separated by a 15 mm layer of lead, a second almost identical layer of scintillator pads forms the Pre-Shower detector (*PS*). The shower fragments detected in the *PS* together with the *SPD* information are used to quickly identify electrons with large transverse energy at the first trigger level, while suppressing individual photons or those from π^0 decays.

The main part of the *ECAL* (see Fig. 2.21) consists of a shashlik structure of 66 layers of 2 mm lead and 4 mm scintillator tiles. The scintillators are read out by wavelength shifting fibres into photomultiplier tubes. With a total depth equivalent to 25 electromagnetic radiation lengths, the *ECAL* is designed to have a resolution of $\sigma_E/E = 10\%/\sqrt{E} \oplus 1\%$ (E in GeV) which has been confirmed by test beam results.

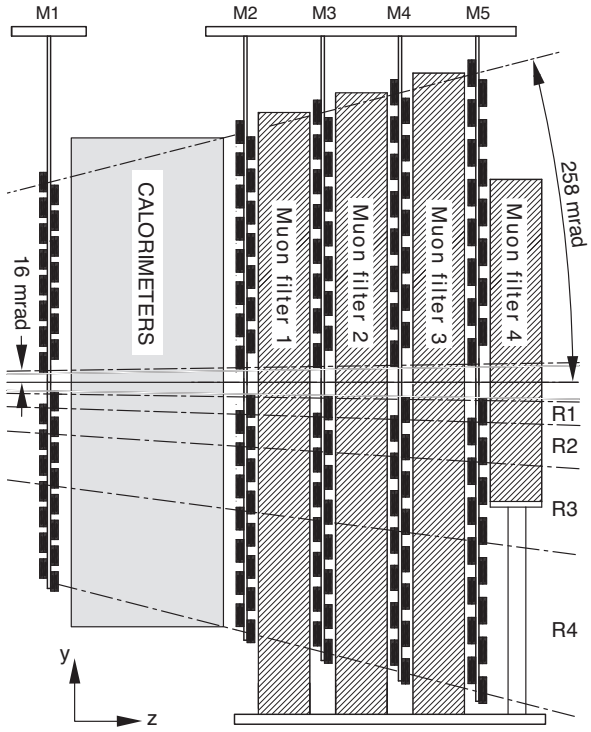


Figure 2.23: *Schematic side view of the five muon stations. Reproduced from [49].*

The *HCAL* also uses tiles of an absorber material, iron in this case, and scintillator tiles. As shown in figure 2.22 the tiles are assembled parallel to the beam axis. There are three scintillator tiles in the longitudinal direction, these are read out in the same way as for the *ECAL*. The total thickness of the material is equivalent to 5.6 nuclear interaction lengths (λ_I) while the material in the *ECAL* adds another 1.2 λ_I. The resolution has been determined to be $\sigma_E/E = (69 \pm 5)\%/\sqrt{E} \oplus (9 \pm 2)\%$ (E in GeV).

2.2.7 The Muon Detector

The muon detector plays a vital role in a large number of *LHCb* analyses, such as the benchmark decays $B_s^0 \to J/\psi(\mu^+\mu^-)\phi(K^+K^-)$, $B_s^0 \to \mu^+\mu^-$, or $B_s^0 \to K^{*0}\mu^+\mu^-$. In addition, muons are used for flavour tagging, i.e. for determining whether a reconstructed B or $\overline{B}$ meson was produced as a B or $\overline{B}$, via detection of a muon originating from a semileptonic B decay. It is therefore essential to have a very efficient muon identification system.

The muon detector consists of five stations, one before the calorimeters, and four thereafter (see Fig. 2.23). The first station is to support the fast matching of tracks to muon hits in the first trigger level. The latter four stations are each separated by iron absorber walls of 80 cm thickness. Including the calorimeters, this amounts to a total

of 20 nuclear interaction lengths.

Like the calorimeters, the muon stations use increasing sizes of their readout cells with increasing distance from the beam axis. All but the innermost region of the first station use Multi-Wire Proportional Chambers. Due to the higher particle rate triple-GEM (Gas Electron Multiplier) chambers will be used in the first muon station.

A simulation study of $B_d^0 \to J/\psi K_S^0$ decays estimates the muon identification efficiency to $\varepsilon(\mu \to \mu) = 94\%$ with a pion misidentification rate of $\varepsilon(\pi \to \mu) = 3\%$, which is flat for particles with a momentum above 10 GeV/c. These numbers can be significantly improved by also using the information from the *RICH* and calorimeter systems to achieve $\varepsilon(\mu \to \mu) = 93\%$ and $\varepsilon(\pi \to \mu) = 1\%$ for all particles with a momentum above 3 GeV/c.

2.3 The Data Analysis Infrastructure

2.3.1 The Trigger System

The trigger system plays a crucial role in the *LHCb* data taking. It is responsible for efficiently selecting events of interest while suppressing the unwanted background to as low a level as possible.

As the detailed implementation of the *LHCb* trigger keeps being adapted to the changing start-up scenarios of the *LHC*, this section will introduce the basic constraints and goals of the trigger system, and will outline the current implementation and its evolution as it is foreseen at this stage.

LHCb operates at a lower design luminosity (2×10^{32} cm^{-2} s^{-1}) than the one envisaged for *ATLAS* and *CMS*. This is achieved by having a larger beam diameter and by having an effective bunch crossing rate of 30 MHz. Only every third bunch crossing produces a visible event in the detector. These 10 MHz of visible interactions have to be reduced to 2 kHz of events that will be written to tape.

To achieve this, the trigger is split in three levels (see Fig. 2.24). The first level, *L0*, is a hardware trigger whose decision is based on the calorimeters and the muon stations, which are read out at 40 MHz at a coarser granularity than for off-line analysis. A *L0* decision can be triggered by one or two muons, one or two hadrons, an electron, or photons.

After a reduction to a rate of 1 MHz the full detector information is read out and processed in a software trigger stage called *HLT1*. The principle of this trigger stage is to confirm the *L0* decision using the full detector information. For example a *L0*

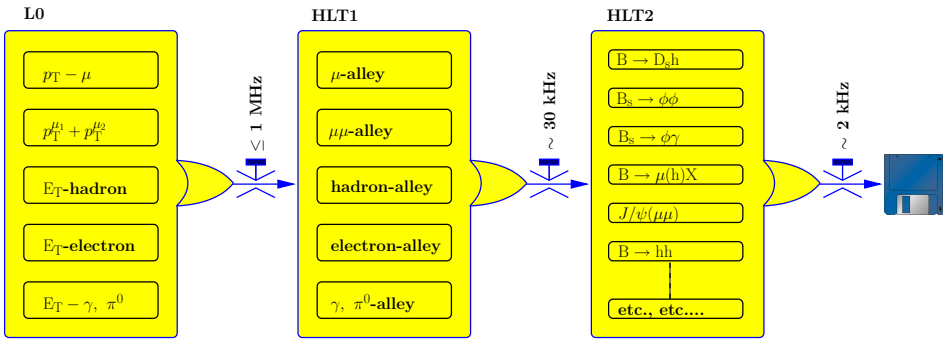


Figure 2.24: *Overview of the LHCb trigger system. Reproduced from [49].*

hadron trigger has to be confirmed through the so-called hadron alley in *HLT1*. A more detailed description of the hadron alley in the context of lifetime measurements will be given in section 4.2.

After a reduction of the rate to about 30 kHz a second software trigger, *HLT2*, will perform the final selection to reach the desired output rate. Eventually, this stage is foreseen to have a large number of exclusive selections to select individual channel (groups) of interest. However, at the beginning a simpler strategy is currently being anticipated which exploits a few very inclusive selections. The motivation for this approach is its simplicity during the early learning period of the experiment.

2.3.2 The LHCb Computing Resources

The bulk of the *LHCb* computing resources is dedicated to the so-called online farm providing the computing power needed for running the *HLT*. This farm comprises up to 2200 units of multi-core CPUs.

For the reconstruction and analysis stages (see below) *LHCb* will use the Grid. The first processing of raw data and the subsequent storage will be performed at the Tier-0 centre at *CERN*. All further processing and analysis is foreseen to be performed at Tier-1 centres. Tier-2 centres will be used to produce Monte Carlo (MC) simulation data.

2.3.3 The LHCb Software

The *LHCb* software is based on the `Gaudi` framework [57]. The main projects, representing the various computing tasks, are:

- `Gauss` provides the event generation for MC simulation. The processing is split in two parts, a generator phase to create the initial decays, for example using

`EvtGen` [58] for B-decays, and a simulation phase which tracks the various particles through the detector based on `GEANT4` [59, 60].

- `Boole` performs the digitisation step for simulated events, i.e. transforms the energy deposits calculated by `Gauss` into detector signals.

- `Brunel` is responsible for the event reconstruction which is common to real and simulated data. At this stage pattern recognition algorithms combine detector hits to tracks or calorimeter clusters. Also, a first particle identification is performed. This is the stage where detector calibration information is used to account for deviations from the ideal detector design, e.g. misalignments whose treatment will be explained in the following chapter.

- `DaVinci` is used for all physics analysis purposes. It combines reconstructed particles to their mother particles and applies selections to distinguish signal events from background. It also computes various observables that serve as input to the final analyses. This is discussed in further detail in chapter 4 for the lifetime fits.

All *LHCb* software is designed for being run under both Linux and Windows systems. It can be installed on local desktop machines and is installed on the Grid to ensure a high performance of all stages of data processing.

3 Alignment

> Das Bild hängt schief.
>
> *Loriot*

The goal of operating the *VELO* (see Sec. 2.2.1) as the detector at the *LHC* with the best position resolution is a great challenge. In order to achieve such a resolution, the position of the individual sensors must be known to high precision. Any misalignment beyond the level of few microns will significantly worsen the resolution. Misalignments are defined as translations or rotations of an object with respect to its nominal position.

This chapter describes the alignment procedure, starting with an overview of the *LHCb* alignment strategy in section 3.1. The main part of the chapter, section 3.2, explains the different methods used in the *VELO* software alignment. Finally, sections 3.3 and 3.4 demonstrate the performance of the *VELO* software alignment with data from a test beam and from an *LHC* synchronisation test, respectively.

3.1 The LHCb Alignment Strategy

For any detector the alignment effort is structured in various steps. Key to high precision operation of any detector is a precise assembly. This does not have to be at the same level of precision as the desired resolution but should be within the same order of magnitude. In most cases displacements can be corrected by software during the reconstruction stage (see below). Only if a track passes close to the edge of a sensor displacements can lead to the particle missing the sensor and the signal not being recorded.

For the *VELO* particular care was taken with the placement of the R sensors. They will be used for a first, fast track fit in r-z coordinates in the trigger. In an ideal geometry hits on the circular strips of the R sensors are points in the r-z coordinate plane. However, a displacement of these sensors leads to the strips no longer lying on concentric circles. Hence, sizeable displacements (several 10 μm) significantly deteriorate a simple r-z tracking.

3.1.1 The Conditions Database

Misalignments of individual detector elements can be taken into account at reconstruction level when combining individual hits to form tracks. This is done by applying corrections to the hit positions according to the knowledge of the displacement of the detector element they belong to. In general, such a displacement can be described by six parameters, three translations and three rotations.

In *LHCb*, these parameters, also called alignment constants are stored in the conditions database (`CondDB`). The `CondDB` contains `XML` files to store the alignment constants for various levels in the detector hierarchy. In case of the *VELO*, one file contains the alignment information for individual sensors, a second contains the constants for modules, and a third holds the information for the two *VELO* halves and for the full system.

When applying a correction for the displacement of a particular detector element its own alignment constants have to be combined with those of its parent to obtain the correct position in the global frame. This is done automatically by the classes which provide the information about the hit positions.

3.1.2 Optical Survey Measurements

During assembly and installation every detector underwent frequent metrology measurements to ensure that the quality margins were kept. For the *VELO* these started with optical surveys at the various institutes where the individual parts were assembled.

In this process, the module base plate and the vacuum vessel were surveyed by photogrammetry at Nikhef [61] while the individual sensors and, later on, the fully assembled modules were measured with Coordinate Measurement Machines (CMM) and Smartscope systems at Liverpool [62, 63]. After arrival at *CERN* and assembly of all modules on their base plate, the fully equipped *VELO* halves underwent a CMM metrology. Also after assembly at *CERN*, the vacuum vessel was measured in another photogrammetry procedure.

Measurements from metrologies at *CERN* (for the vessel and the module positions) and at Liverpool (for the relative sensor position on each module) have been used to generate a first set of alignment constants which describe the best knowledge of the initial position of all detector elements [64]. These constants have been stored in the `CondDB`. The precision of these measurements has already been tested with data as discussed in sections 3.3 and 3.4.

3.1.3 Detector Alignment During Operation

At the time of operation external survey measurements are no longer possible due to several reasons, such as the high radiation environment and the inaccessibility of certain parts of the detector. Hence, two possibilities are left to study misalignment during operation of the detector. Hardware based alignment systems or track based systems using the software reconstruction.

3.1.3.1 Hardware Based Alignment Systems and the VELO Motion System

Hardware based alignment systems are devices specially installed to monitor the detector position in situ. Examples are laser systems where the laser photons replace particles from an interaction and are detected with the active parts of the detector, or optical systems attached to the detector frame that are able to detect deformations.

A laser alignment system is used in the *RICH* detectors where the beams are directed from definite positions via the mirrors onto the *HPD*s. Any movement in the mirrors can thus be traced and corrected for. The *OT* uses a RASNIK system, which is based on projecting an image through a lens onto a CCD camera.

In the *VELO*, the motion system that measures the position of each of the two halves is a hardware based alignment system. The horizontal position of each half and their common vertical position is measured by resolvers to an accuracy of 10 μm (see page 53). These measurements are stored in the online conditions database and they are combined with the alignment constants stored in `CondDB` at the time of reconstruction. Further hardware based alignment systems, e.g. to monitor the position of the individual modules, have not been considered as they require an increase of material in the active area.

3.1.3.2 Track Based Alignment Systems

The second option for alignment strategies during operation are track based systems using the software reconstruction. As these are in principle available for all sub-detectors due to their close connection to the event reconstruction such approaches are widely used – not only in *LHCb*.

Various algorithms have been developed to align tracking detectors [65]. Their common base is the use of so-called track residuals (see Fig. 3.1). A residual measures the distance of a single measurement to a track. This can be the distance of closest approach in 3D-space or the distance measured in the plane defined by the detector which the measurement is associated to (the latter is shown in Fig. 3.1). One distinguishes

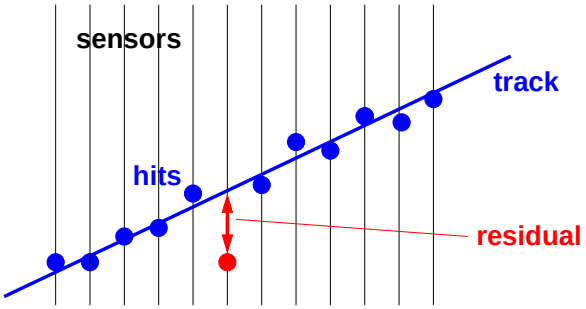


Figure 3.1: *Illustration of the principle of track residuals. The diagram shows a track producing hits on several sensors. The residual is given as the distance between a hit and the track in the sensor plane of the hit.*

between biased residuals, when the measurement in question is used in the track fit and, hence, causes a bias, and unbiased residuals, when the track is fitted without using the measurement in question.

It is important for the understanding of alignment algorithms to distinguish the terms misalignment and residual. Misalignments are physical displacements that are, to first order, constant over a data taking period. Residuals vary for each measurement due to the single hit resolution of the detector. For a perfectly aligned device the average of the residuals on a given detector should be zero. Hence, misalignments only affect the average of the residuals.

The goal of every alignment algorithm is to derive alignment constants. These are constants that describe the position of a detector element with respect to its nominal position. Hence, they can be expressed as a translation vector, a rotation defined by three angles and, optionally, a pivot point. The storage and handling of these constants in *LHCb* is described in the following section.

The computation of alignment constants from residuals can follow many methods. The underlying idea is to solve a minimisation problem that yields the optimal set of alignment constants given the measured residuals. In general, two basic principles exist: closed form or global alignment techniques, or open form, iterative techniques.

3.1.3.3 Iterative Alignment Algorithms

Within the iterative techniques two approaches are distinguished: solving the alignment problem inside the track fit; and fitting descriptions of the shape of residual distributions as a function of the alignment parameters. The former is often used with a Kalman filter track fit and sequentially updates the alignment parameters as part of the track fit. In practice, this means that the set of alignment parameters evolves with

every track being processed. This approach is able to account for correlations between the individual alignment parameters. However, this is at the expense of high computing requirements. It has to be iterated so that eventually all tracks are fitted with the same set of alignment parameters once the problem has converged. An advantage of using a Kalman filter track fit in an alignment algorithm is that this fit is able to account for multiple scattering effects and that it can apply energy loss corrections.

The second iterative approach uses the knowledge of the shape of residual distributions as a function of one or more coordinates. This technique has been used to align the SLD vertex detector [65]. The residual distributions are evaluated using the full data set and thereafter fitted to a model depending on the alignment constants. Thus, all tracks are always fitted using the same set of alignment constants. However, properly taking into account the correlations between the alignment constants can be challenging in this method. Therefore, this approach also requires several iterations. This approach has been adopted for the *VELO* sensor alignment (see Sec. 3.2.1).

3.1.3.4 Global Alignment Algorithms

A closed form alignment algorithm solves a global minimisation problem, i.e. one that determines the solution for all alignment constants at the same time. The most widely used approach is the `MILLEPEDE` algorithm developed by Volker Blobel for the *H1* collaboration [65]. This technique is used in the *VELO* alignment and is described in detail in the following section. The principle is to express the dependency of the measurements on all track parameters and all alignment constants as a large set of linear equations that can be solved by a single matrix inversion. The requirement of linearity prevents the use of a Kalman filter track fit as the track model, and instead a linear track model is used.

Driven by work for the *ST* and *OT*, a closed form alignment algorithm based on the Kalman filter track fit has also been developed in *LHCb* [66]. Similarly to the `MILLEPEDE` approach, it populates a large matrix to describe the connection between track parameters and alignment constants, rather than updating the alignment constants on a track by track basis. This algorithm requires only a small number of iterations to converge, as correlations between track parameters are taken into account.

3.1.3.5 The Structure of Alignment Systems

A decision to take for all alignment problems of large tracking detectors is whether to attempt a full alignment of the whole detector at once, or whether to first align

sub-detectors individually and later treat them as rigid bodies when combining them to achieve an overall alignment. In most cases the systems are sufficiently decoupled that the latter approach is used.

A benefit of introducing a hierarchical structure into the alignment problem is that it reduces the requirement on computing resources. On the other hand, sometimes it is the sub-structure of a detector which dictates that a particular alignment algorithm is not applicable to the system with the finest granularity. This is the case for the *VELO* where the linearity requirement of the `MILLEPEDE` algorithm means that the smallest alignable entity is a module rather than a sensor as described below.

A problem for any alignment approach, but particularly for those with a hierarchical structure are weak modes. These are systematic deformations of the detector that leave the quantity to be minimised (usually a χ^2) invariant. These deformations have to be constrained by either forcing certain parts of the detector to remain unchanged or by constraining the affected set of alignment constants as a whole. Examples are given below for the *VELO* module alignment.

For *LHCb* the alignment follows a hierarchical structure. As explained below, the *VELO* alignment is done in three stages of different granularity. Other alignment algorithms have been developed to align *ST* and *OT* based on both the `MILLEPEDE` approach and the closed form Kalman approach. Furthermore, the Kalman approach has been extended to be able to align the *VELO* and muon sub-detectors. An alignment algorithm has also been developed to align the *RICH* mirrors, following a similar approach to the one used in the *VELO* sensor alignment.

A global alignment strategy that aligns all sub-detectors simultaneously is currently being developed. One approach exploits the Kalman filter alignment. As this requires reasonably precise starting conditions it will only be run after the sub-detectors have been aligned individually. One major advantage of aligning the full system is the reduced sensitivity to weak modes within the individual sub-detectors (see section 3.2.3).

3.2 The VELO Software Alignment Method

Two aspects are central to understanding the choice of strategy for the *VELO* alignment. First, the *VELO* alignment should be able to give a fast response to whether or not the alignment constants have changed[1] after re-insertion of the two halves. The risk of movements of individual parts and, hence, sudden changes in the alignment

[1] As explained above, the position of the two *VELO* halves as measured by the resolvers is treated separately and is not part of the alignment constants.

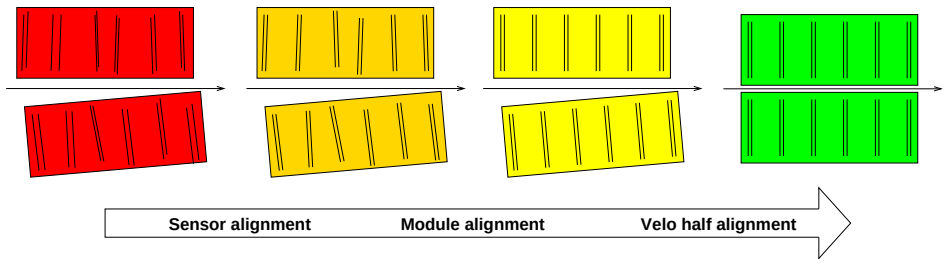


Figure 3.2: *Structure of the VELO software alignment algorithms.*

constants is particularly high, as the *VELO* is being moved for every fill of the *LHC*. The alignment constants most likely to change are those describing the relative position of the two halves, i.e. precise correction in addition to the measurements from the resolvers.

Second, as most particles will only produce hits in sensors of one half, a separation of the algorithms to align either half internally (see Sec. 3.2.3) and to obtain a relative alignment of the two halves (see Sec. 3.2.4) is a natural choice. This separation (see figure 3.2) also allows to optimise the data set for the specific alignment task as described below.

The `MILLEPEDE` algorithm (see Sec. 3.2.2) provides a framework for a fast and robust computation of alignment constants. However, its requirement of linearity of the relation of residuals, track and alignment parameters prevents the direct use of the *VELO* sensor information. The strips of the R and Φ sensors have a non-trivial connection to the misalignments which are translations along and rotations around axes of a Cartesian system. Hits of both sensors on a module have to be combined to a space point to reach a linear system. As both sensors are glued onto a common module, they are not expected to move significantly with respect to each other. In contrast to this, the possible movement of one module with respect to the other is much larger.

Initially, it was assumed that, given the structure of the *VELO*, an alignment algorithm to determine the relative sensor alignment was not necessary. After it had been shown that such an algorithm was easy to be implemented when based on the shape of residual distribution (see Sec. 3.2.1), the initial decision was revised. Subsequently, this algorithm has proved to be a very useful tool, not only in refining the alignment constants, but also in helping to understand the detector at the level of greatest detail (see e.g. section 3.3.3.1). Figure 3.2 shows the full structure of the *VELO* software alignment with the separation into the sensor alignment and the two `MILLEPEDE` based algorithms to align the modules within each half and the two halves with respect to each other.

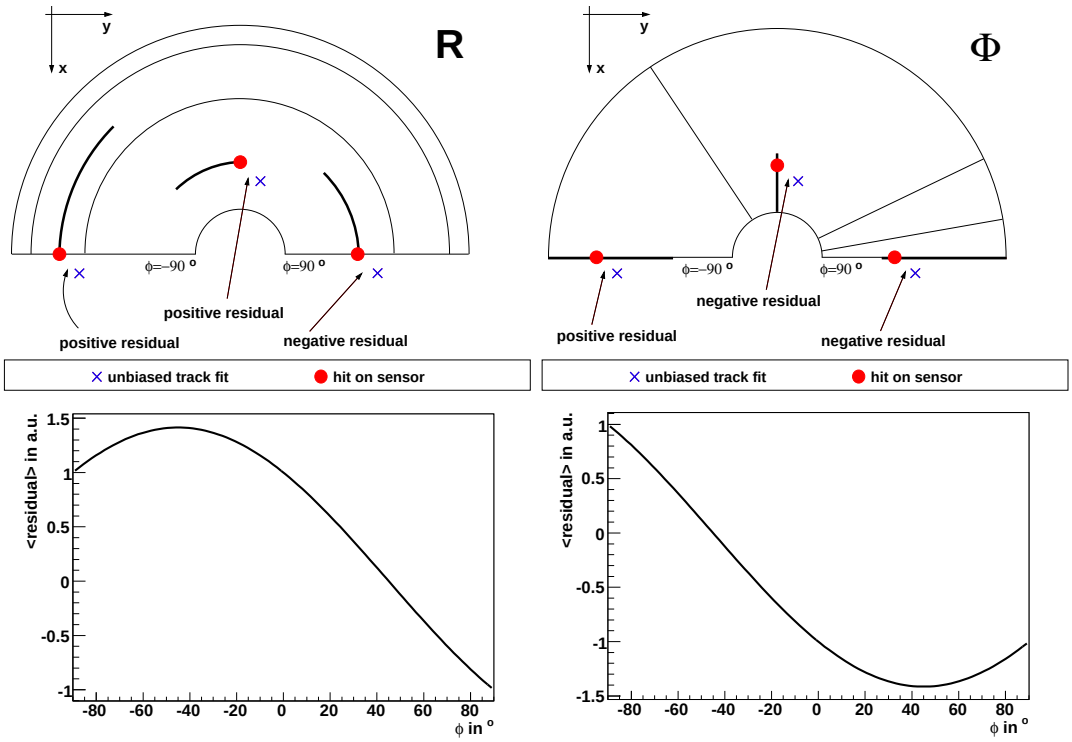


Figure 3.3: *Influence of misalignments on residuals of R and Φ sensors (top). The misalignment shown is a translation both along the negative x and y direction. Corresponding shape of the residual distribution as a function of ϕ (bottom).*

3.2.1 Relative Alignment of the VELO Sensors

For the determination of the relative misalignment of R and Φ sensors the problem is no longer linearisable, which is essential for exploiting a global matrix inversion technique like `MILLEPEDE`. Hence, an iterative approach is used that extracts the misalignment constants from the distribution of residuals plotted against position.

The characteristic shape of these distributions can easily be related to the misalignment of the sensors. In a non-misaligned geometry, the plane of the sensor surface is, to first order approximation, the x-y coordinate plane in the *LHCb* coordinate system. The y-axis runs along the straight edge of the sensor, while the x-axis lies on the symmetry axis of the sensor and defines $\phi = 0$ (see fig. 3.3). The sensors are slightly tilted inwards towards their straight edge by a rotation around their y axis of about 2 mrad to further reduce the risk of contact with the RF-foil.

Clearly, measurements will only be affected by misalignment translations that are non-parallel to the corresponding strip on the sensor. Thus, R sensors are most sensitive to x-translations around $\phi = 0$, whereas they are most sensitive to y-translations near

$\phi = \pm\pi/2$. The opposite is true for Φ sensors (see fig. 3.3).

Ideally, the method described below should be applied for each sensor in its local coordinate system, as it is sensitive to translations of the sensor in its own plane. However, to simplify the fit code, all fits are done in a common coordinate system. Therefore, the respective *VELO* half coordinate system has been chosen as it also allows the algorithm to work when the *VELO* halves are retracted. The simplification of using a common coordinate system is justified as explained in the following.

Defining the residual as the difference between the hit position and the extrapolated position of an unbiased track fit[2] one can write the relation between misalignments (Δ_i) and residuals ($\epsilon_{R/\Phi}$) as follows. Note that the track position is only extrapolated to the z-position of the sensor, i.e. neglecting the sensor tilts around the x and y axes. However, this effect is only of the order of the square of the tilts, hence in the sub-micron range.

$$\begin{aligned} \epsilon_R &= -\Delta_x \cos\phi_{track} + \Delta_y \sin\phi_{track} && (R \text{ sensor}), \\ \epsilon_\Phi &= +\Delta_x \sin\phi_{cluster} + \Delta_y \cos\phi_{cluster} + \Delta_\gamma r_{track} && (\Phi \text{ sensor}), \end{aligned} \tag{3.1}$$

where Δ_γ describes a misalignment in the form of a rotation around the z axis, which translates into a shift in ϕ by multiplication with the radial coordinate of the extrapolated track in the sensor plane. It is sufficient to leave it as a free parameter in the form of a constant when fitting the shape of the residual distribution as a function of ϕ, as the Δ_γ term does not contain any ϕ dependence. This formalism is slightly simplified as it ignores the stereo angle orientation of the Φ sensor strips whose treatment is explained below.

The sensor tilts around the x and y axes are neglected again as the residuals are determined in the sensor plane but plotted against r and ϕ in the respective *VELO* half coordinate system. Once more, this is justified as their effect on x- and y-translations is only of second order.

The value for Δ_γ can be directly extracted by fitting the residual distribution on the Φ sensor versus r (rather than versus ϕ as just discussed). Fitting a linear function to the residual distribution versus r gives Δ_γ as the slope. This value is used in the iterations of this sensor alignment procedure to improve the convergence of the algorithm. The final value for the z-rotation alignment constant will not be determined by this method, but is determined by the module alignment algorithm as described in section 3.2.3.

[2] Here, the residual is calculated (as provided by the `DeVelo[R/Phi]Type` classes) as the perpendicular distance to the strip hit in the sensor plane including inter-strip fractions (as provided by the `VeloClusterPositionTool` [67]).

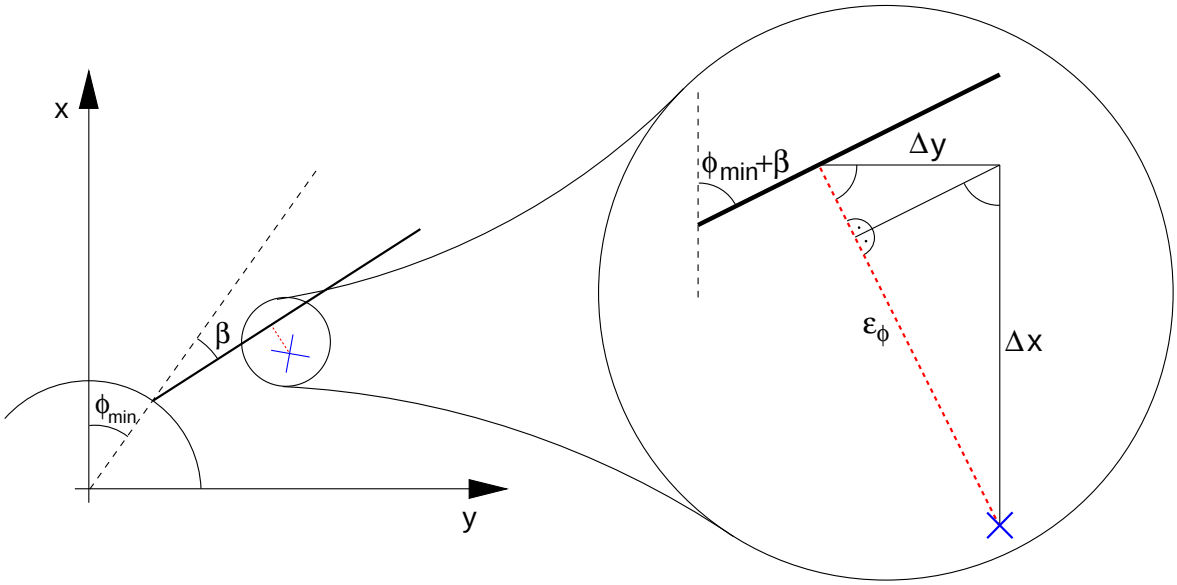


Figure 3.4: *Influence of x and y translation misalignments on residuals of Φ sensors taking into account the stereo angle strip geometry.*

In order to perform this fit the residuals of both R and Φ sensors have to be plotted as a function of ϕ (and for the Φ sensor only also as a function of r). For the R sensor, the ϕ coordinate of the residual is taken from the fitted track. Similarly, for the Φ sensor, the r coordinate is taken from the fitted track.

As an example of the resulting distributions, the misalignment introduced in the upper half of figure 3.3 would give rise to the shape of the residual distributions shown in the lower half of figure 3.3. The fit is performed in the respective *VELO* half frame to keep its code simple and general. One consequence is that the range in ϕ is $[-90°, 90°]$ for sensors in the *VELO* A-side (as shown in figure 3.3), and $[-180°, -90°]$ and $[90°, 180°]$ for sensors in the *VELO* C-side.

3.2.1.1 Treatment of Stereo Angle Strips

There is a further complication for the *VELO* Φ sensors arising from their stereo angle, β (see Fig. 3.4), that has been neglected in the discussion above. For misalignments in the form of x and y translations it is obvious that equation 3.1 stays valid with the replacement $\phi \rightarrow \phi' = \phi_{min} + \beta$ leading to:

$$\epsilon_\Phi = \Delta_x \sin(\phi_{min} + \beta) + \Delta_y \cos(\phi_{min} + \beta) \qquad (\Phi \text{ sensor}). \tag{3.2}$$

The case of z rotation misalignments (Δ_γ) requires a closer look. As the Φ sensor strips are tangents to circles with a radius, d (see figure 3.5), of 2.8 mm (inner region)

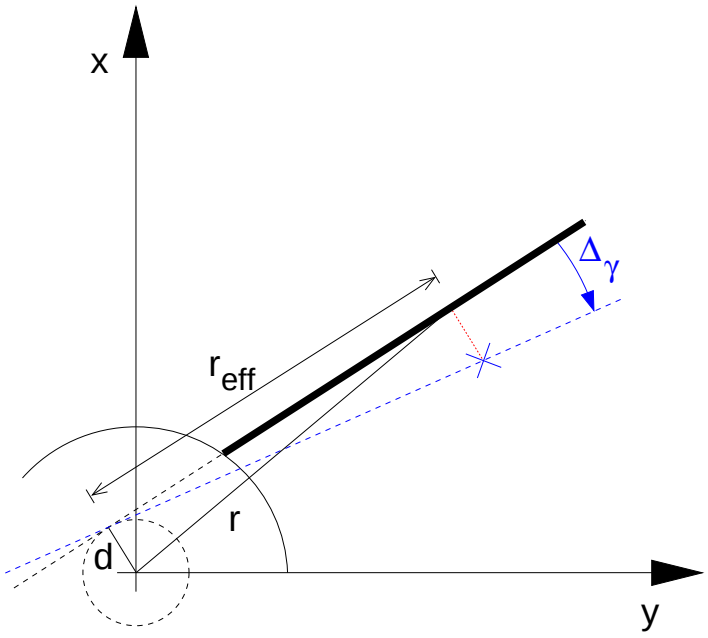


Figure 3.5: *Influence of z rotation misalignments on residuals of Φ sensors. The effective radius is defined as the distance between the hit and the point where the extrapolated strip meets the circle that defines its stereo angle.*

or 3.1 mm (outer region), the dependence of the residual is not exactly linear to the radial hit position. For small values of Δ_γ the linear relation between the residual and Δ_γ holds for the effective radius, which is given by $r_{eff} = \sqrt{r^2 - d^2}$. Hence, the Φ sensor residuals are given by

$$\epsilon_\Phi = \Delta_x \sin(\phi_{min} + \beta) + \Delta_y \cos(\phi_{min} + \beta) + \Delta_\gamma \sqrt{r^2 - d^2} \qquad (\Phi \text{ sensor}). \tag{3.3}$$

For reasons of simplicity, the radial dependence of the residual is fitted with a linear function which is a good approximation and leads to a relative bias of Δ_γ of less than 1%. As the value for Δ_γ is only used internally to the sensor alignment and not propagated any further, as explained below, this simplification has no consequence on the performance of the alignment algorithm.

3.2.1.2 Weakly Constrained Degrees of Freedom

Thus far, the alignment of three degrees of freedom has been described. The remaining ones, z translations and rotations around the x and y axes, are not fitted due to a lack of sensitivity.

Rotations around the x and y axes, called α and β respectively, only cause a second order effect on residuals. For R sensors one obtains

$$\begin{aligned}\epsilon_R &= (1-\cos\Delta_\alpha) r_{track}\sin^2\phi_{track} + (1-\cos\Delta_\beta) r_{track}\cos^2\phi_{track} \qquad (R \text{ sensor})\\ &\approx \tfrac{1}{2}\Delta_\alpha^2 r_{track}\sin^2\phi_{track} + \tfrac{1}{2}\Delta_\beta^2 r_{track}\cos^2\phi_{track},\end{aligned} \tag{3.4}$$

where Δ_α and Δ_β are the rotation misalignments around x and y, respectively, which are of the order of 1 mrad according to survey measurements. For Φ sensors the sensitivity is even further reduced as the rotation axis for y rotations is defined as the straight edge of the sensors where the sensitivity would be largest.

The case is somewhat different for z translations. The effective z translation is given by $\Delta_{z,eff} = \Delta_z + y_{track}\tan\Delta_\alpha + x_{track}\tan\Delta_\beta$, as rotations around x and y cause local z translations. The measured residual then depends on the angle of the track with respect to the z axis, θ, and the angle of the track projection in the sensor plane with a tangent to the strip at the hit position, ρ:

$$\epsilon = \tan\theta_{track}\sin\rho_{track}(\Delta_z + y_{track}\tan\Delta_\alpha + x_{track}\tan\Delta_\beta), \tag{3.5}$$

which is valid for both R and Φ sensors.

For large statistics with a track sample that covers the full sensor at a range of impact angles θ it should be possible to constrain also these degrees of freedom. This process has not been considered yet, as this requires collision data. For data collected so far, only tracks with small θ angles have been used leading to a $\tan\theta$ suppression of effects from the misalignments discussed here.

3.2.1.3 Implementation

The sensor alignment fit is based on residual distributions as it has been explained above and illustrated in figure 3.3. The residuals are calculated in a linear track fit (see section 3.2.1.4) and plotted as function of r and ϕ for each sensor.

A cleaning procedure is applied to the residual distributions to improve the fit stability. First, a minimum number of entries (200 by default) is required for the whole distribution. In order to suppress outliers in the distribution of the residual means, a minimum number of entries is also required for each bin in ϕ. This threshold is 1/10 of the threshold for the whole distribution.

The shape of the residual distributions is determined by calculating the profile of the histograms (see e.g. figure 3.12 on page 97). These profile histograms are then used to fit the alignment parameters using the residual equations derived above.

After Δ_x,Δ_y, and Δ_γ have been determined in each iteration, the new alignment constants are applied to the Φ sensor. For x and y translations, the difference in

the misalignment between the Φ and R sensor is applied as the Φ sensor's alignment constants, while the common misalignment will be attributed to the module by the module alignment algorithm. For rotations around the z-axis, the alignment constant is taken directly from the Φ sensor as the R sensor is insensitive to this motion.

A two-dimensional unbinned likelihood fit has been implemented to test the validity of the one-dimensional binned fits. No difference has been observed and, hence, the faster and less complicated one-dimensional fits are kept.

In the context of the *LHCb* alignment software, the *VELO* sensor alignment algorithm is implemented as part of the `VeloAlignment` package inside the `Alignment` project. Its iterations are controlled by a python script while all the actual code is a `C++` implementation of a `GaudiTupleAlg`.

The sensor alignment will be run as the first step of the *VELO* alignment. It produces alignment constants that reflect the relative x and y position of the Φ sensor with respect to the R sensor. These are then used as input values for the module alignment algorithm. By definition, the R sensor is kept perfectly aligned with the module and hence its alignment constants are all set to 0. The output of the module alignment algorithm is then used to update the module alignment constants before aligning the two *VELO* halves. In practise, the three algorithms will be run repeatedly and their order can be changed as needed.

As explained below, the most time consuming part of this alignment algorithm is the track fit. Depending on the complexity of the events, one iteration of the track fit and relative sensor alignment algorithm using 20000 tracks takes about one minute on a single CPU[3].

3.2.1.4 Track Fits

For each iteration, unbiased residuals have to be determined from track fits excluding hits on the sensors of the module under study. This means that the set of hits used for fitting one track will vary when the residuals for sensors of different modules are calculated. It turns out that the resulting large number of track fits accounts for the bulk of the time consumption of the algorithm.

Two different track fits have been studied. The bi-directional Kalman track fit, as it is used by the main reconstruction software, and a straight line track fit, which fits a straight line to a set of at least four space-points made of an (r/ϕ) pair. Both fits show no significant difference in their results for high momentum tracks[4]. However, the

[3] 1 CPU = 1000 SpecInt2000 units

[4] The track sample studied had a flat momentum distribution between 1 and 100 GeV.

Kalman track fit appears to be roughly 100 times slower than the straight-line track fit.

A tool is available to extract unbiased residuals from a single Kalman fit using all hits. As this tool will, in the best case, leave the Kalman fit a factor 5 slower than the straight line fit, the latter was used to carry out the larger scale studies presented in section 3.2.5. In addition, ideally the hits from both sensors should be excluded from the track fit to avoid any bias when determining the relative sensor alignment. This is not possible with a single Kalman fit.

3.2.2 The Millepede Algorithm

As mentioned above, the relative alignment of the *VELO* modules is performed using an approach based on the `MILLEPEDE` algorithm [65]. This algorithm is designed to solve large sets of linear equations.

The individual measurements of R and Φ sensors have to be combined to space points in order to reach a linear problem for the *VELO*. With the two sensors of each module being separated by 2 mm, the Φ sensor measurement is projected onto the R sensor. In this projection, the change in the ϕ coordinate is estimated by a track fit. This is a valid procedure as the influence of misalignments on the track slope is small and hence their influence on the ϕ correction is negligible. Using the projected ϕ coordinate, ϕ_{corr}, a space point is calculated as

$$\begin{aligned} x &= r \cdot \cos(\phi_{corr}) \\ y &= r \cdot \sin(\phi_{corr}) \\ z &= z(R\ sensor), \end{aligned} \tag{3.6}$$

where $z(R\ sensor)$ is the z coordinate of the R sensor. Using these space points, a general track equation is given by

$$\mathbf{Y} = f(\mathbf{X}) + \epsilon, \tag{3.7}$$

where $\mathbf{Y}$ denotes the vector of measurements, $f(\mathbf{X})$ is a function of track parameters and derivatives, and ϵ is the vector of residuals. In the case of a linear track model, one has

$$f(\mathbf{X}) = X \cdot \alpha, \tag{3.8}$$

where α is the vector of track parameters, i.e. slopes and offsets, and X denotes the matrix of derivatives, in this case z for the slopes and 1 for the offsets. The track parameters are also called local parameters as they are different for each track.

For an alignment problem, the aim is to find a relation between track residuals and misalignments in order to fit the latter simultaneously with the track parameters. In order to use `MILLEPEDE`, this relation has to be linear, resulting in

$$\epsilon = C \cdot \mathbf{\Delta}, \tag{3.9}$$

where $\mathbf{\Delta}$ is the vector containing the alignment parameters, and C the matrix containing their derivatives. As the parameters are independent of the individual tracks they are called global parameters and global derivatives, respectively.

Defining the covariance matrix V^i for the i^{th} measurement $\mathbf{Y}^i$ on a track, the χ^2 can be written as

$$\chi^2 = \sum_i \left(\mathbf{Y}^i - X^i\alpha - C^i\mathbf{\Delta}\right)^T (V^i)^{-1} \left(\mathbf{Y}^i - X^i\alpha - C^i\mathbf{\Delta}\right). \tag{3.10}$$

When summing over the whole track sample with index t, this leads to the full alignment χ^2:

$$\chi^2 = \sum_{t,i} \left(\mathbf{Y}^{t,i} - X^{t,i}\alpha^t - C^{t,i}\mathbf{\Delta}\right)^T (V^{t,i})^{-1} \left(\mathbf{Y}^{t,i} - X^{t,i}\alpha^t - C^{t,i}\mathbf{\Delta}\right). \tag{3.11}$$

Minimisation with respect to both local ($\mathbf{\Delta}$) and global ($\mathbf{\Delta}$) parameters leads to the following matrix equation:

$$\begin{pmatrix} \mathbf{\Delta} \\ \cdots\cdots \\ \alpha \end{pmatrix} = \begin{pmatrix} A_{11} & \vdots & A_{12} \\ \cdots\cdots\cdots & & \\ A_{21} & \vdots & A_{22} \end{pmatrix}^{-1} \cdot \begin{pmatrix} \sum_{t,i}(C^{t,i})^T(V^{t,i})^{-1}\mathbf{Y}^{t,i} \\ \cdots\cdots\cdots\cdots\cdots \\ \sum_{t,i}(X^{t,i})^T(V^{t,i})^{-1}\mathbf{Y}^{t,i} \end{pmatrix}, \tag{3.12}$$

with

$$\begin{aligned} A_{11} &= \textstyle\sum_{t,i}(C^{t,i})^T(V^{t,i})^{-1}C^{t,i} \\ (A_{12})_t &= \textstyle\sum_{i}(C^{t,i})^T(V^{t,i})^{-1}X^{t,i} \\ A_{21} &= A_{12}^T \\ (A_{22})_{tt} &= \textstyle\sum_{i}(X^{t,i})^T(V^{t,i})^{-1}X^{t,i}, \end{aligned} \tag{3.13}$$

where the sub-matrix A_{12} has to be seen as a vector of n_{track} blocks of size $n_{global} \times n_{local}$, and the sub-matrix A_{22} is 0 apart from blocks of size $n_{local} \times n_{local}$ along its diagonal.

As the aim is predominantly the solution for the global parameters, i.e. for the alignment constants, one obtains

$$\begin{aligned} \mathbf{\Delta} &= \left(A_{11} - A_{12}A_{22}^{-1}A_{12}^T\right)^{-1} \\ &\quad \cdot \left[\left(\sum_{t,i}(C^{t,i})^T(V^{t,i})^{-1}\mathbf{Y}^{t,i}\right) - A_{12}A_{22}^{-1}\left(\sum_{t,i}(X^{t,i})^T(V^{t,i})^{-1}\mathbf{Y}^{t,i}\right)\right]. \end{aligned} \tag{3.14}$$

The idea behind the MILLEPEDE algorithm is to solve the required components of this system of equations exploiting the special structure of the A matrix. This involves the inversion of A_{22} which consists of small sub-matrices along its diagonal and, hence, is trivial to invert. The only additional inversion needed is the one of the first bracket in equation 3.14 which is of the size $n_{global} \times n_{global}$. This significantly reduces the resources needed compared to the inversion of the full matrix A which would be needed to solve the complete system, as n_{global} is independent of the number of tracks used in the alignment. In the *VELO*, n_{global} is typically 126, whereas usually several 10000 tracks are used, leading to a size of the matrix A of $\mathcal{O}(10^5)$.

3.2.3 Relative Alignment of the VELO Modules

The module alignment is based on measurements of space points as defined in equation 3.6. For each module the algorithm has to determine six alignment constants, three translational degrees of freedom (Δ_x, Δ_y, Δ_z) and three rotational degrees of freedom (Δ_α, Δ_β, Δ_γ). For the latter, α, β, and γ denote the rotations around the x, y, and z axes, respectively. As the module alignment computes the relative positions of the modules in each half separately, there are in total 126 parameters per half.

For small rotational misalignments the equivalent of equation 3.7 is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} z & 1 & 0 & 0 \\ 0 & 0 & z & 1 \end{pmatrix} \cdot \begin{pmatrix} m_x \\ b_x \\ m_y \\ b_y \end{pmatrix} + \begin{pmatrix} -1 & 0 & m_x & m_x y & m_x x & y \\ 0 & -1 & m_y & m_y y & m_y x & -x \end{pmatrix} \cdot \begin{pmatrix} \Delta_x \\ \Delta_y \\ \Delta_z \\ \Delta_\alpha \\ \Delta_\beta \\ \Delta_\gamma \end{pmatrix}, \tag{3.15}$$

where the first term on the right hand side shows a standard straight line track model while the second shows the influence of misalignments. The straight line model is justified as there is nearly no magnetic field inside the *VELO* volume[5]. Small deviations that are caused for low-momentum particles near the downstream end of the *VELO*

[5]Inside the *VELO* volume the main magnetic field component is $B_y < 0.05$ T.

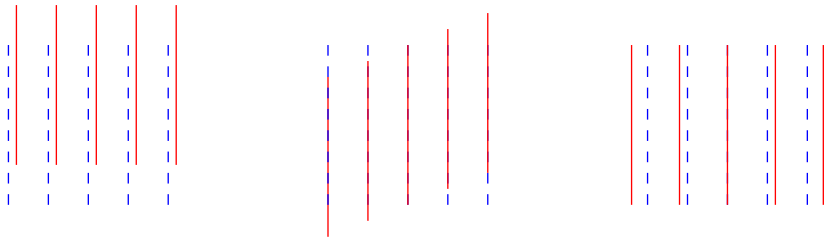


Figure 3.6: *Examples for weak modes in alignment algorithms: average translation (left), shearing (centre), and scaling (right). The original position of the detector elements is shown as blue (dashed) lines and the displaced as red (solid) lines.*

are suppressed by having high momentum tracks as well as tracks of both charges in the sample.

Global derivatives that depend on track parameters, i.e. on local parameters that are free parameters in the overall fit, introduce non-linearities (these are the derivatives for Δ_z, Δ_α, and Δ_β). This impedes the straightforward use of the `MILLEPEDE` algorithm as the minimisation would use only the starting values of the derivatives to compute the best fit result. The track slopes which are in any case rather small for *VELO* tracks are set to 0 in the first pass of the fit when used as global derivatives to account for these non-linearities. A second iteration uses the fit results of the first pass to compute the derivatives.

A well chosen sample of tracks is essential for reaching an optimal alignment result (see section 3.2.6). Tracks parallel to the z axis connect the highest number of modules and hence produce the most powerful constraints for the major degrees of freedom, Δ_x, Δ_y, and Δ_γ. On the other hand, only tracks with non-negligible slopes are able to constrain the remaining three degrees of freedom. The optimal track sample is obtained by combining tracks parallel to the z axis (see section 3.2.6) with tracks from collisions in the primary interaction region which naturally have larger slopes.

As mentioned in section 3.1.3.5, changes to the alignment constants which leave the χ^2 unchanged (weak modes) are a particular danger to alignment algorithms as they can introduce systematic biases. Possible weak modes (see figure 3.6) include global translations and rotations, scaling of the system in one or more dimensions, shearing (translation in x or y of individual modules as a function of z), and twist (rotation around z as a function of z).

In order to constrain weak modes, the alignment system, i.e. the χ^2 function, has to be constrained. One approach would be to constrain all tracks by external tracking systems. However, as the *VELO* resolution is a factor of 6 more precise than any

z axis rotation	$\sum_i \Delta^i_\gamma = 0$
x, y, z translations	$\sum_i \Delta^i_{(x,y,z)} = 0$
xz and yz shearing	$\sum_i \Delta^i_{(x,y)} \cdot (z^i - \overline{z}) = 0$
z axis scaling	$\sum_i \Delta^i_z \cdot (z^i - \overline{z}) = 0$

Table 3.1: *List of constrained weak modes in the VELO module alignment with their constraint equations.*

other sub-detector, this approach is not considered. There are two equivalent options for internal constraints: fixing two modules, or fixing the average movement of all modules using Lagrange multipliers.

For the implementation in the `MILLEPEDE` algorithm, a constraint equation like $\sum_i \Delta^i_x = 0$, which constrains the overall movement in x to 0, has to be written in the form

$$\mathbf{X}^T \cdot \mathbf{f} = f_0, \tag{3.16}$$

where $\mathbf{X}^T = (\mathbf{\Delta}, \alpha)$ is the parameter vector. Thus, the χ^2 as defined in equation 3.11 acquires additional terms with the Lagrange multiplier λ^i for the i^{th} constraint, leading to

$$\chi^2 \rightarrow \chi^2 + \sum_i \lambda^i \left(\mathbf{X}^T \cdot \mathbf{f}^i - f^i_0\right). \tag{3.17}$$

For equation 3.12 this leads the vectors $\mathbf{f}^i$ being appended as new columns to A_{12}, the Lagrange multipliers extend the parameter vector, and the constraint values f^i_0 extend the vector on the right hand side. Table 3.1 shows the weak modes which are constrained using this approach.

The strategy outlined above explained how weak modes can be constrained. However, despite them leaving the χ^2 of the alignment problem unchanged, they can have an impact on physics observables. One possibility to measure weak modes is by moving to alignment systems in which these modes do have an impact on the χ^2 of the alignment problem. This can be done by combining several sub-detectors. In this case, for example a shearing of the planes of a single sub-detector is replaced by a shearing of the combined system. The increased size of the system to be aligned leads to a smaller size of potentially unresolved weak modes.

Another possibility to measure weak modes is by using their impact on physics observables in calibration measurements. Measuring the mass scale of known particles to high precision in combination with their momentum dependence allows the distinction of effects of the magnetic field as well as alignment weak modes.

3.2.4 Relative Alignment of the Two VELO Halves

Due to the *VELO* being moved in and out by 3 cm for each fill of the *LHC*, a precise control of the position of the individual halves is of utmost importance. The two halves are centred around the beam during the closing procedure and their individual position with respect to a reference is measured by resolvers to a precision of 10 μm (see page 53). However, the relative position of the two halves with respect to each other is intrinsically only known from metrology at a precision of about 100 μm. Hence, a precise measurement of the relative position of the two halves has to come from a track based software alignment, both for an initial precision determination and for long term checks in conjunction with the frequent movements. Two complementary approaches to a software alignment solution of aligning the *VELO* halves have been developed. They are described in the following sections. One is based on tracks that produce hits in sensors of both halves. The other uses the reconstruction of the interaction point of the primary collision.

3.2.4.1 Alignment with Overlap Tracks

The first approach is to use tracks that have space points in modules of both *VELO* halves. Due to the shape of the *RF* foil the straight edge of the sensors of both halves overlap in the x-y projection (see page 53). In x, this overlap region has a size of up to 1.7 mm when the *VELO* is fully closed.

Tracks that pass this region with no slope in x can produce several space points in either of the two halves. They give access only to the relative position of the two halves, however, due to their rather even distribution in z, they provide good sensitivity also to relative rotations around x and y. These tracks can come from the PV with an appropriate slope in y, or from beam gas interactions further away from *LHCb* and hence moving parallel to the beam. While PV interactions are collected by the standard data acquisition strategy, tracks parallel to the beam are not. An approach for collecting a sufficient sample is described in section 3.2.6.

Under the assumption of a perfect relative module alignment, the equivalent of equation 3.7 is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} z & 1 & 0 & 0 \\ 0 & 0 & z & 1 \end{pmatrix} \cdot \begin{pmatrix} m_x \\ b_x \\ m_y \\ b_y \end{pmatrix} + \begin{pmatrix} -1 & 0 & m_x & m_x y & m_x x + z & y \\ 0 & -1 & m_y & m_y y & m_y x + z & -x \end{pmatrix} \cdot \begin{pmatrix} \Delta_x \\ \Delta_y \\ \Delta_z \\ \Delta_\alpha \\ \Delta_\beta \\ \Delta_\gamma \end{pmatrix}, \quad (3.18)$$

where the Δ_i are the relative misalignments of the two halves. The only difference to equation 3.15 is the z dependence of $\Delta_{(\alpha,\beta)}$ which is responsible for the high sensitivity to these parameters. Using this equation, the `MILLEPEDE` algorithm is used in exactly the same way as described for the module alignment in the previous section.

3.2.4.2 Alignment with Primary Vertices

The second approach uses tracks originating from the main proton-proton collision, the so-called primary vertex (PV). Fitting common vertices with tracks from both halves leads to sensitivity to their relative position. As the primary vertices have a small range in z ($\sigma(z_{PV}) = 5.3$ cm) the lever arm, and with it the sensitivity, for resolving rotations around x or y is small. A major advantage of this method is that it also works with a partially or fully retracted *VELO*, which is not the case for the approach using overlap tracks.

It is important to note that in this approach the measurements are complete tracks, (m_x, b_x, m_y, b_y), rather than individual space points. The local parameters are the vertex coordinates (v_x, v_y, v_z). Individual space points no longer enter the equation.

Apart from this change, the application of the `MILLEPEDE` method remains as described before for the module alignment. The equivalent of equation 3.7 is then

$$\begin{pmatrix} b_x \\ b_y \end{pmatrix} = \begin{pmatrix} 1 & 0 & -m_x \\ 0 & 1 & -m_y \end{pmatrix} \cdot \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}$$
$$+ \begin{pmatrix} -1 & 0 & m_x & m_x v_y & m_x v_x + v_z & -v_y \\ 0 & -1 & m_y & m_y v_y & m_y v_x + v_z & v_x \end{pmatrix} \cdot \begin{pmatrix} \Delta_x \\ \Delta_y \\ \Delta_z \\ \Delta_\alpha \\ \Delta_\beta \\ \Delta_\gamma \end{pmatrix}, \quad (3.19)$$

where the Δ_i are the relative misalignments of the two halves.

As both methods for measuring the relative position of the two *VELO* halves share the set of global parameters it is straightforward to combine them within the `MILLEPEDE` algorithm.

3.2.5 Simulation Results

The aim of the alignment algorithms is to ensure that the remaining misalignments are significantly smaller than the single hit resolution. As shown in section 5.4 misalignments that are smaller than one third of the single hit resolution have negligible impact on physics results. This section describes the results from MC simulation studies used to evaluate the performance of the *VELO* software alignment.

3.2.5.1 Relative Sensor Alignment Results

The sensor alignment method has been tested with 10 samples of randomly generated misalignments. All sensors have been misaligned individually, thus generating a scenario equivalent to simultaneous module to module and sensor to sensor misalignments. Each of the 10 samples consists of 20000 tracks with small slopes, thus passing through all sensors of one *VELO* half and evenly distributed across the sensor surface. Typically three iterations of the alignment procedure are required to obtain the best resolution.

Figure 3.7 shows the generated and the remaining misalignments after all iterations. The resolution on the relative x and y translation of the sensors of one module is 1.3 μm, i.e. a significant improvement over the survey precision of approx 3 μm.

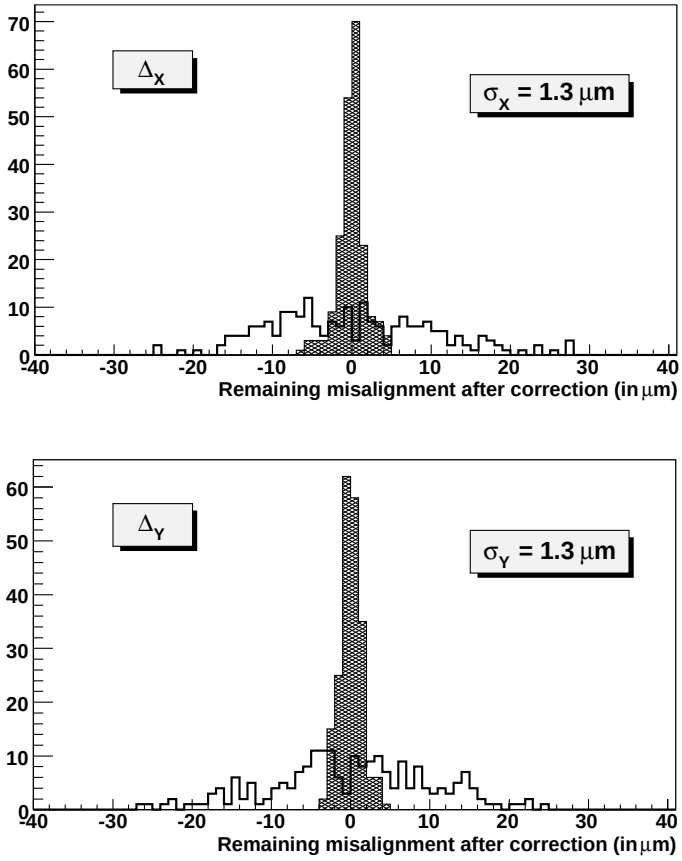


Figure 3.7: *Misalignment values before (□), and after (■) relative sensor alignment.*

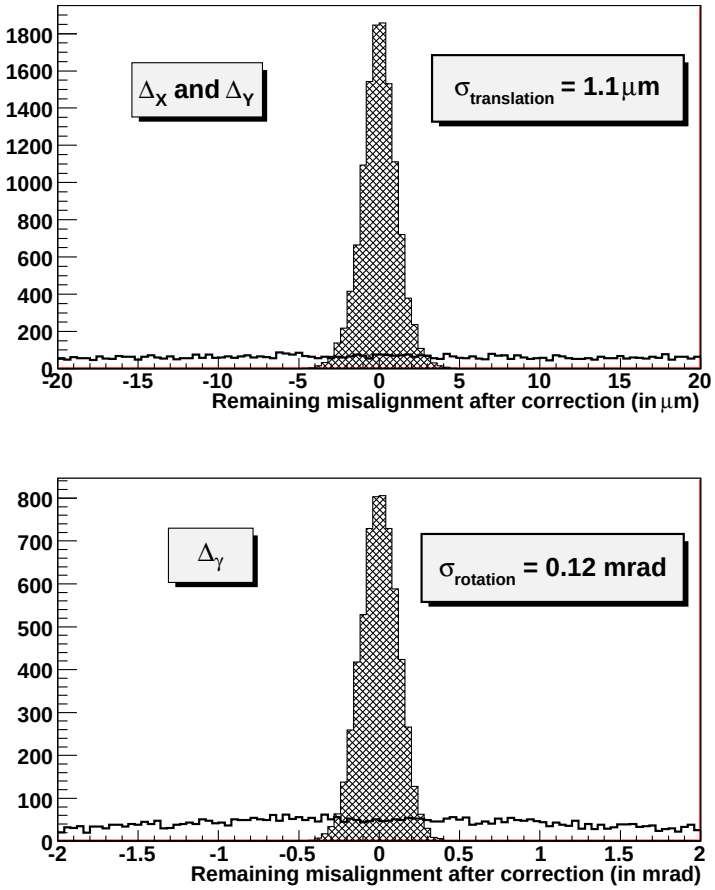


Figure 3.8: *Misalignment values before (□), and after (■) relative module alignment.*

3.2.5.2 Relative Module Alignment Results

A simulation of 200 samples of 25000 events each has been produced to test the relative module alignment. Each sample, which comprises a mixture of 5000 minimum bias events ($\approx$ 100000 tracks from primary vertex interactions) plus 20000 beam gas interaction like events, was produced with a different set of alignment constants. All the module and detector-half degrees of freedom have been misaligned.

The internal alignment of the modules in each detector-half is primarily sensitive to translations of the modules in x and y and rotations around the z axis. In figure 3.8 the alignment constants for 200 event samples are shown before and after correction. Resolutions on the x and y translation alignment parameters of 1.1 μm and on rotations around the z axis of 0.12 mrad are obtained.

Concerning the non-linear degrees of freedom, the observed sensitivity is as expected worse than for the other parameters. However, some results were obtained for the

modules which are close to the interaction region, i.e. where track slopes are larger. Restricting the study to these stations (1 to 14), one obtains a reasonable sensitivity to Δ_z (28 μm) and a fair sensitivity to Δ_α and Δ_β (0.8 mrad and 1.1 mrad respectively). This sensitivity is worse than the survey precision, but will provide a cross-check of this survey information.

The performance of this algorithm has been evaluated separately, as it is, in general, run independently of the relative sensor alignment algorithms In the presence of relative misalignments of the sensors on a given module the module's position will be aligned to the average position of the two sensors[6].

3.2.5.3 Relative VELO Half Alignment Results

Although the three alignment steps can be performed independently, in practice it is expected that steps two and three will be run consecutively. Hence, the results presented in this section are for the realistic case of performing both of these alignment steps on misaligned samples. The tracks are refitted after the module alignment procedure in order to update the track parameters. The results presented here have been obtained with about 300 tracks in the overlap region.

The results of the study are shown in figure 3.9. The resolution on the x and y translation alignment parameters is 12 μm for x and y translations, and the resolution on the x and y tilts is 36 μrad.

Some of the degrees of freedom are more difficult to constrain, similar to the case of the module alignment. In the *VELO* half alignment these weakly constrained misalignments are the z translation and the rotation around the z axis. The relative rotation around z between the two halves is constrained using the overlap tracks. Translations along z are estimated through the vertex fitting technique leading to a resolution of about 40 μm.

3.2.6 The Data Acquisition Strategy for the VELO Software Alignment

It is necessary to have tracks that traverse the *VELO* parallel to the beam axis in addition to those coming from interaction vertices inside the *VELO* volume, as explained in the previous sections. Such tracks can originate from beam gas interactions away from the *LHCb* detector or from beam halo. These tracks should arrive in time with

[6]This requires a track sample with a sufficiently flat distribution in ϕ which is given for the samples used for *VELO* alignment.

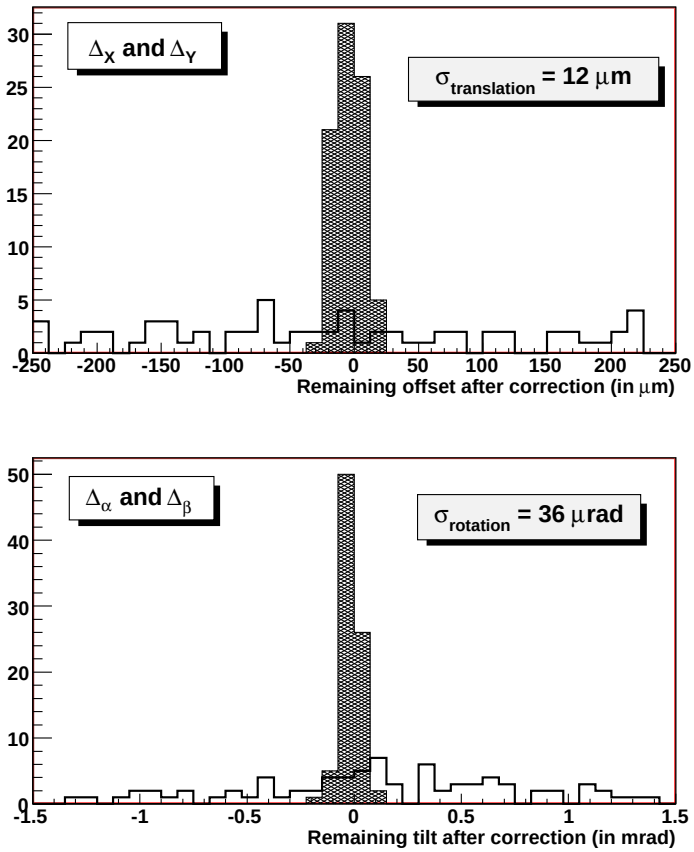


Figure 3.9: *Misalignment values before (□), and after (■) VELO half alignment.*

the main collisions, as both sources are beam induced. Hence, they should be present in a random sub-sample of the acquired events.

It is foreseen to run all *LHCb* alignment on data collected with a special calibration stream. This stream has an output rate of 5 Hz. The amount of events containing tracks parallel to the beam axis will be tuned to be about 10% of the calibration stream (0.5 Hz). Thus, it should be possible to acquire a sample comparable to those used in the simulation studies within each run.

Tracks running parallel to the beam are unlikely to fire the first trigger level (*L0*, see section 2.3.1). Events with this kind of tracks that have already passed the *L0* trigger have to be selected in the high level trigger to reach the required output rate. A special algorithm has been designed to find isolated tracks that are parallel to the beam. It is based on first selecting tracks in the x-y projection of all sensors and thereafter confirming with the distribution of space points along z. Thus, avoiding any track fit, this algorithm is particularly fast and, hence, suited to be run at high rate in the trigger.

3.3 Test Beam Results

The first opportunity to test the *VELO* software alignment was with data from a test beam that took place in November 2006. The following sections will describe the apparatus and various studies that demonstrate the impact and quality of the alignment.

3.3.1 Test Beam Apparatus

A partially equipped *VELO* detector half was tested in November 2006 in a 180 GeV/p hadron and muon beam at the *CERN SPS*. The mechanical suspension, cooling system and vacuum operation were designed to provide a good representation of the conditions expected from the final experiment. Ten of the 21 modules in one half of the detector were installed in their final position. The readout electronics chain, as well as all software, were the same as used in the final installation.

Six out of the ten installed modules were read out simultaneously. Data was taken with several different cabling configurations for the module readout. Particles were observed directly from the beam or from interactions of the beam with a series of targets. The 1 mm radius 300 μm thick circular lead targets were installed to represent the primary vertex location that will be obtained in the final experiment. Figure 3.10

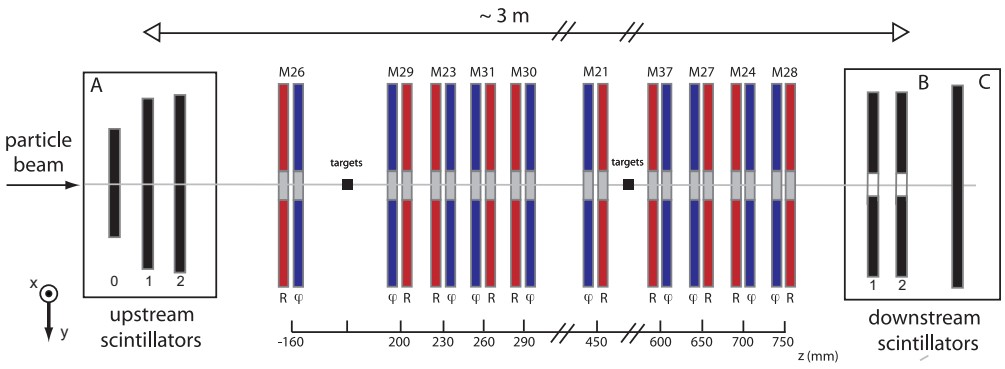


Figure 3.10: *Schematic top view of the test beam apparatus. A total of ten modules were mounted in the detector half. The module numbers are indicated and the location of the R and Φ sensors in the modules. The location of the targets is also shown.*

shows a schematic overview of the mounted modules. The coordinate system used, as indicated, is equivalent to that used in the final *LHCb* detector.

The electronics readout system and prototype data processing algorithms of the final experiment were applied. The detector half was operated under vacuum (10^{-3} mbar) with modules cooled down ($< 0°$C).

3.3.2 Alignment Quality

The *LHCb VELO* alignment algorithms are presented in section 3.2. The only algorithms applied were the relative sensor and relative module alignment, as the test beam configuration contained only modules of one *VELO* half.

The results presented here used the data from two readout cabling configurations and primarily used data in which the beam passed through the targets, as this contained a complimentary set of tracks both perpendicular and at small angles to the sensors.

The relative positions of the R and Φ sensors inside the individual modules and the relative position of the modules were initially assumed to be at their nominal design positions. Corresponding alignment constants were applied as the starting point for the alignment procedure. The software algorithms to determine the relative alignment of the R and Φ sensors and the relative alignment of the modules were then applied and the results are presented in the following sections.

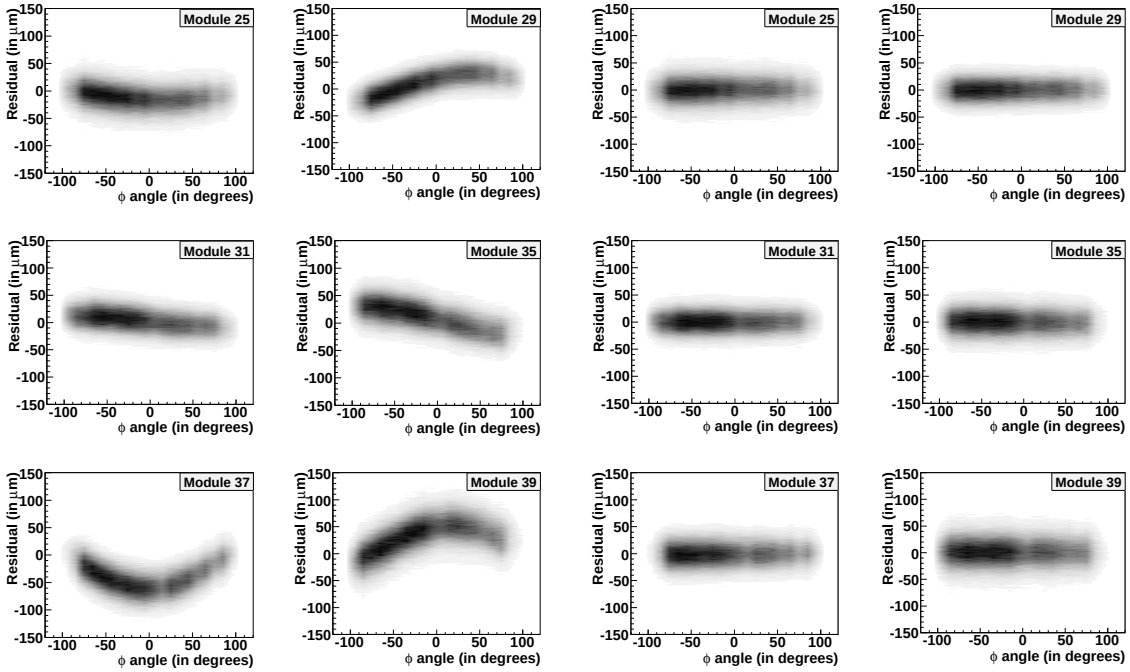


Figure 3.11: *Unbiased* Φ *sensor residuals as a function of the* ϕ *coordinate without any alignment information (left) and including survey and alignment results (right).*

3.3.2.1 Residual Distributions

The distribution of the residuals across the sensor surface is sensitive to misalignments. For example, as described in section 3.2.1, plotting the unbiased Φ and R sensor residuals as a function of the ϕ coordinate gives direct information on the relative x-y translations of the sensors. In the case of a perfect alignment these distributions should be flat when plotted against any coordinate variable.

Figure 3.11 shows the distribution of residuals on the Φ sensor plotted against the ϕ coordinate before the alignment procedure has been performed, assuming the alignment constants are as in the nominal detector design, as well as after applying the alignment procedure. Equivalent results were obtained for R sensors. As expected, applying the alignment information results in reducing the deformations in the distributions (which result primarily from the x and y displacements of the sensors) and moving the mean of the residuals towards 0.

In figure 3.12 the mean of the residual distributions for the Φ and R sensors in one typical module are shown plotted against both r and ϕ coordinates after the alignment procedure has been applied. The distribution of residuals on the Φ sensor plotted against r is seen to have a small change at the transition radius ($r = 17.25$ mm) between

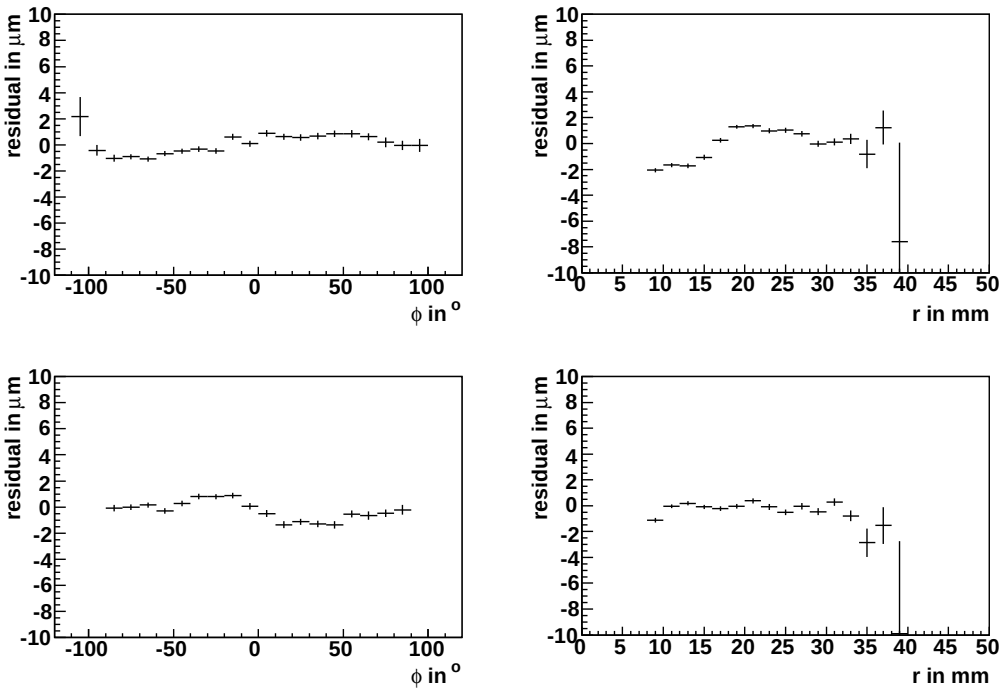


Figure 3.12: *The distribution of* Φ *(top) and* R *sensor (bottom) residuals in module 25 as a function of* ϕ *(left) and* r *coordinate (right).*

the inner and outer sectors of the sensor. This effect is understood to originate from a z translation of the sensor (see section 3.2.1.2), however, it has not been corrected in the alignment that was applied to test beam data.

The mean of the residual distributions can be projected on the residual axis to assess the quality of the alignment achieved, i.e. projections of the plots shown in figure 3.12. The spread of this distribution then provides information on the remaining misalignments. These distributions obtained from all six sensors read out in a particular cabling configuration are shown in figure 3.13.

The left hand plot in figure 3.13 shows the projections of the residual means as a function of ϕ for all twelve sensors (six R and six Φ sensors) under study. The r.m.s. of this distribution confirms that the x and y translations of all sensors are known to a precision of 2.1 μm, in agreement with the combined precision of 1.1 μm for the module alignment and 1.3 μm for the sensor alignment as obtained from the simulation studies reported in section 3.2.5.

The right hand plot in figure 3.13 shows the projections of the residual means as a function of r. This plot is primarily sensitive to rotations around the z axis. The projection is made only for the data from the six Φ sensors since the R sensors are insensitive to z rotations. The measured r.m.s. of 1.1 μm relates to the quality of

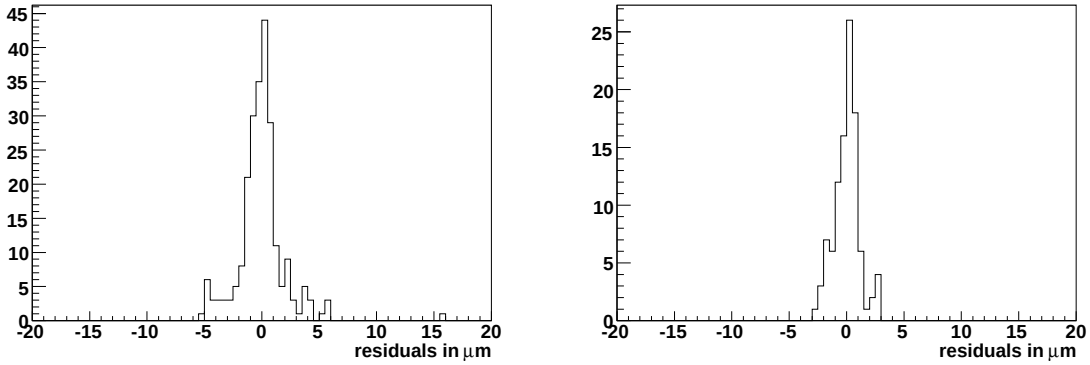


Figure 3.13: *Alignment precision of x and y translations (left) and of rotations around the z axis (right).*

constraining the rotations around the z axis at an effective radius. This effective radius was determined by simulation to be 11 mm and hence these rotations are known to a precision of 0.1 mrad, matching that reported in section 3.2.5.2.

3.3.2.2 Comparison with Metrology

During the module production, an optical metrology survey of the relative positions of the R and Φ sensors inside the individual modules was performed. Another quality measure of the alignment can be obtained from the comparison of the alignment constants as determined by the software alignment with those measured by the metrology of the individual modules. The comparison between the two sets of measured constants for the relative sensor translations on each of the modules is shown in figure 3.14. Agreement between both methods at a level of about 5 μm is obtained, which is equivalent to the combination of the expected precision of the metrology measurements (3 μm) and the software alignment (1.3 μm).

3.3.3 Effect of Alignment on Detector Performance

This section reports on critical elements of the detector performance that are strongly affected by the alignment precision. The use of the alignment to diagnose a cluster centre reconstruction bias is discussed first. Then, a qualitative demonstration of the impact of the alignment on vertexing is shown. Finally, the sensor hit resolution after alignment is reported.

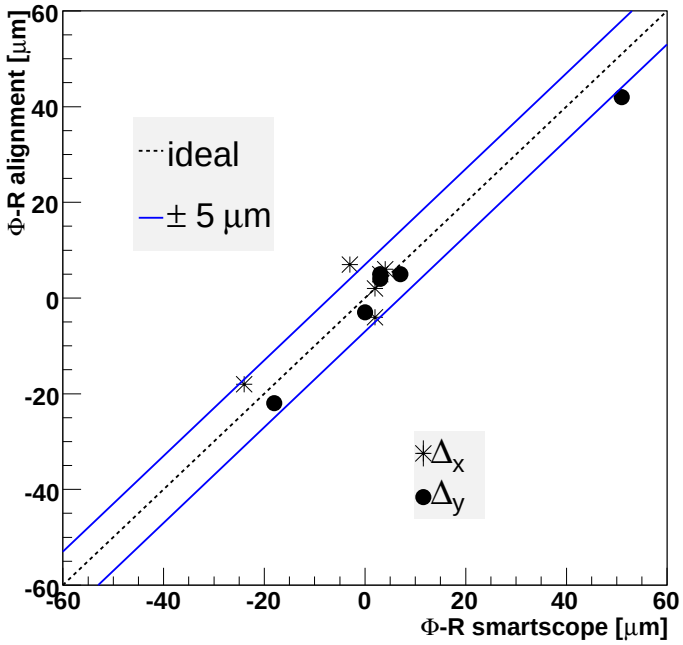


Figure 3.14: *Comparison of optical metrology results and software alignment results for relative sensor translations.*

3.3.3.1 FIR Filter

It has become possible to check for small scale effects in the detector geometry or biases in the cluster reconstruction position, as a result of the high precision obtained from the detector alignment. A cluster centre reconstruction bias was observed and removed through the application of a Finite Impulse Response (FIR) filter [68].

The *VELO R* sensors contain four approximately 45° sectors each with 512 strips. The strips are connected to bond pads at the outer edge of the sensor through the use of a double metal routing layer. In the first sector, the strips 127 to 0 are read out first, followed by those from 128 to 511 (see figure 3.15). This pattern is reversed in adjacent sectors.

Forward cross-talk between the analogue chip output signals in the cables to the readout board gave rise to a bias in the reconstructed signals and, hence, cluster positions. However, as a result of the readout pattern, the direction of this bias reverses for sensor strips 0-127 and 128-511. The residual bias is clearly visible in figure 3.16 before the correction is applied. The alignment quality was improved, once a FIR filter was applied to remove this effect. This correction will be included for the final experiment. However, as this requires a time consuming data reprocessing, this correction has not been included for the results given in the previous section and only been applied to a fraction of the data sample. Hence, figure 3.16 after the correction is shown with

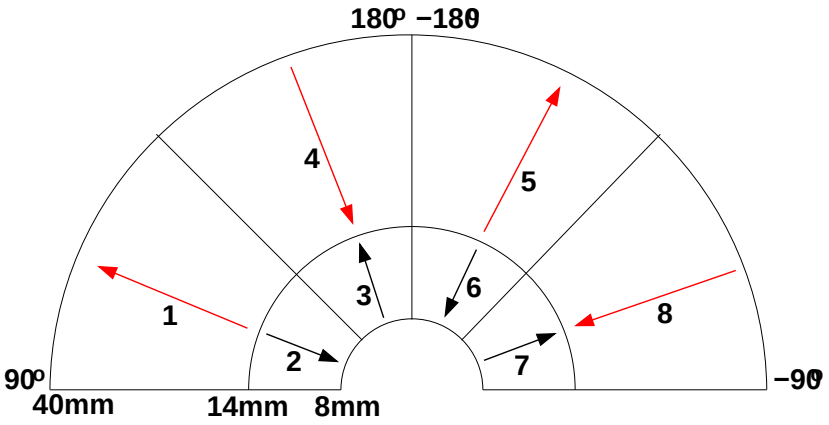


Figure 3.15: *R sensor readout order. The inner sectors (black) are read out by one Beetle chip each, while the outer sectors (red) are read out by three Beetle chips each.*

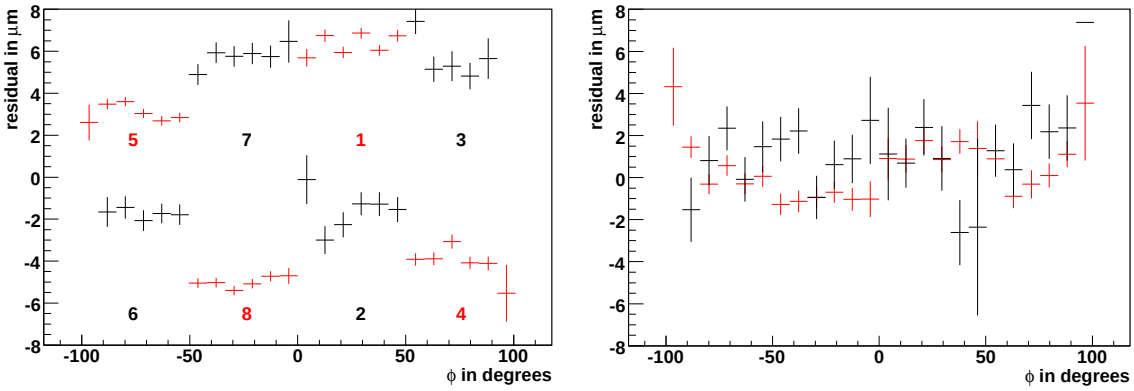


Figure 3.16: *R sensor residuals plotted against ϕ position across sensor, before and after a FIR filter is applied. The residuals are split in groups of equal readout direction with the same numbering scheme used in figure 3.15. In both cases the alignment procedure has been applied. The right hand figure has reduced statistics.*

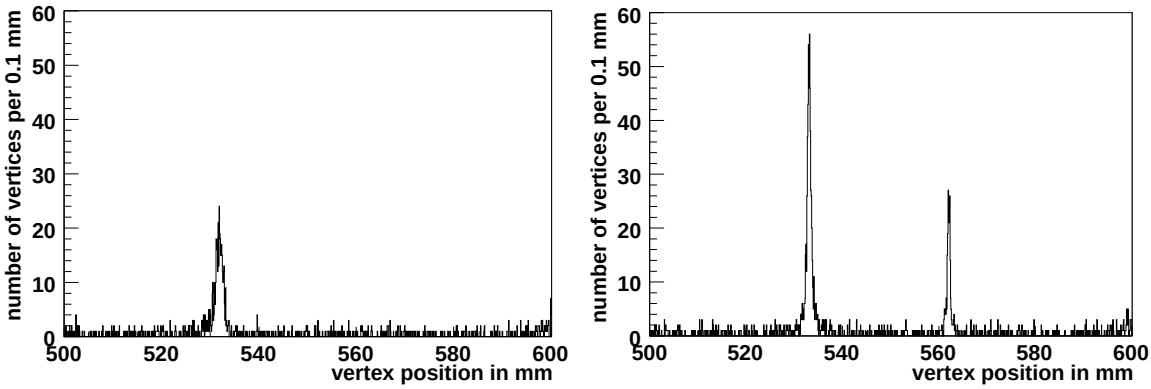


Figure 3.17: *Vertices reconstructed in targets before (left) and after (right) applying the alignment procedure.*

reduced statistics.

3.3.3.2 Vertexing Performance

The *LHCb* trigger is based on a precise separation of B hadron decay vertices from the primary interaction vertices. Hence, a very good vertexing accuracy is crucial for the experiment.

The alignment quality of the *VELO* has a critical effect on the vertexing performance. This is demonstrated in figure 3.17. These plots were obtained on the same set of target events, using the same standard tuning for the pattern recognition. In the left hand plot the nominal design positions of the modules are applied, and it is apparent that only vertices from one of the two expected targets are observed. The right hand plot shows the situation after the alignment procedure has been applied: the second target appears, and for the first target the precision of the vertex reconstruction improves significantly which also leads to a larger number of reconstructed vertices.

3.3.3.3 Sensor Resolution

The *VELO* sensor resolution has been determined using tracks of perpendicular incidence. The resolution has been determined from the sigma of a Gaussian fit to the distribution of the unbiased residuals. Both the R and Φ sensors on the module under study were excluded from the track fit. The resolution is extracted as a function of the local strip pitch at the track intercept point.

The widths of these residual distributions reflect the actual resolution of the sensor and a contribution due to the track extrapolation error. This contribution was computed (see reference [69]) under the assumption that all sensors have equal per-

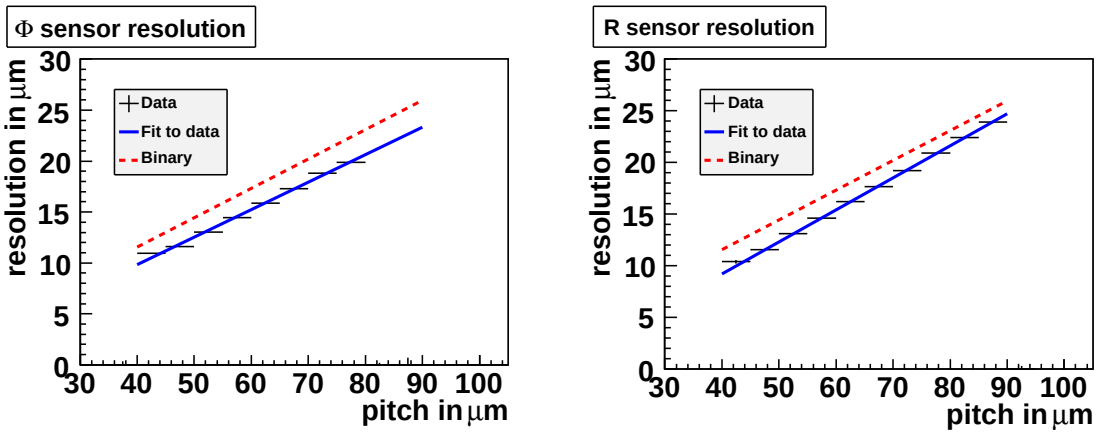


Figure 3.18: *Φ (left) and R (right) sensor resolutions, averaged over 6 sensors.*

formance and that the local pitch of all the R or Φ sensors on a track is the same. This assumption is valid as the track angles are less than 2 mrad. A relative correction was calculated and applied to the width of the distributions for each sensor such that the corrected width reflects the sensor resolution. The resolutions for all sensors were found to be in excellent agreement.

Figure 3.18 shows the *VELO* sensor resolution, averaged over 6 R and Φ sensors. A single hit precision of roughly $9.5 + 0.3 \times (\text{pitch} - 40)$ μm is obtained for both R and Φ sensors, for normal incidence tracks. A significantly better resolution is expected for tracks at angles around 8° for which the charge sharing between adjacent strips is optimal. The resolution improves significantly with track angle as the charge sharing is increased. The resolution has been extracted using a simple weighted pulse height algorithm for reconstruction of the cluster position. Additional development of the clustering algorithm is expected to further improve the precision.

3.3.4 Alignment Stability

The *VELO* modules are placed as close as possible to the *LHC* beam, to optimise the vertex reconstruction capability. As a result, the sensors will be operated in vacuum separated from the primary vacuum of the *LHC* by an RF foil (see section 2.2.1). Furthermore, due to the *LHC* beam stability during injection, the modules must be retracted by 30 mm and reinserted for each fill. Hence, the stability of the alignment to pressure variations and mechanical movements is of great importance.

Carbon fibre structures, such as the *VELO* module bases, may deform when pumped down in a vacuum chamber due to the release of humidity absorbed in the mesh. The left hand plot in figure 3.19 shows the alignment constants determined with data

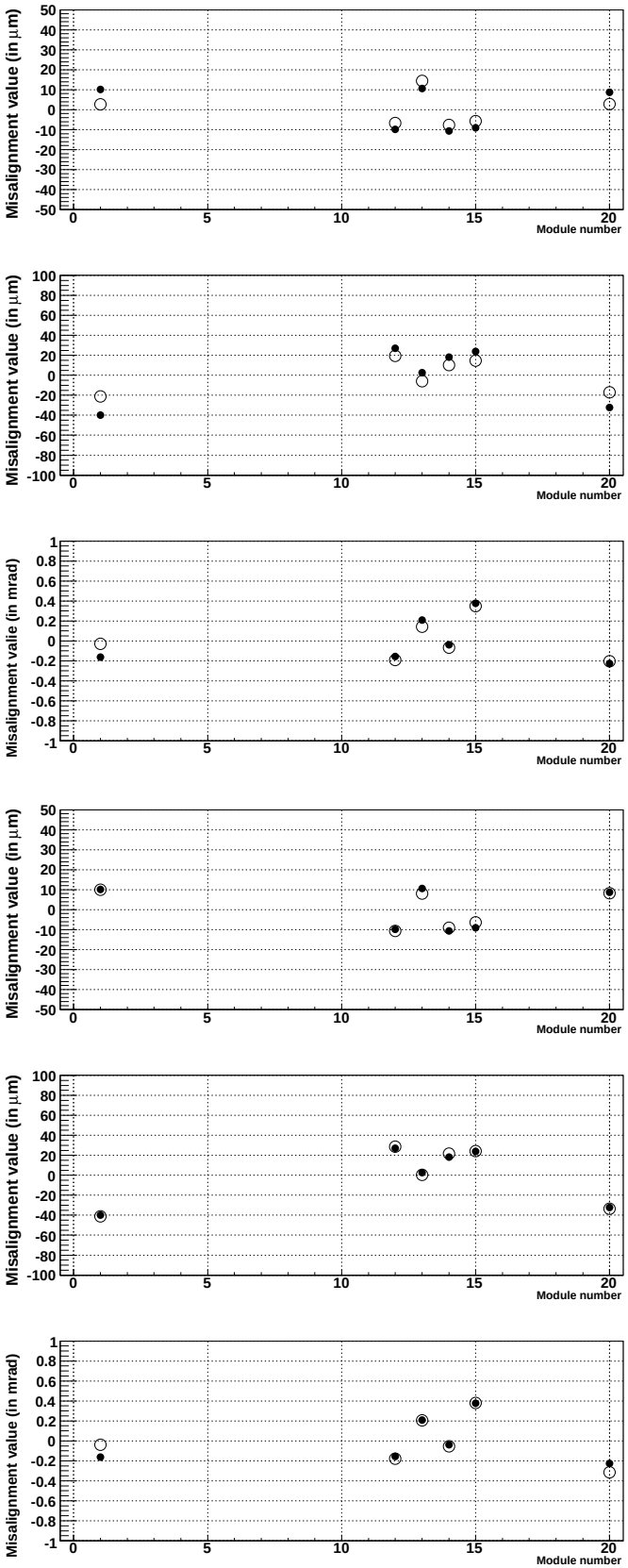


Figure 3.19: *Misalignment values in air (•), and in vacuum (○) (top), and before (○), and after (•) detector halves movement (bottom).*

taken at atmospheric pressure and re-determined using data collected after the air was pumped out of the *VELO* vacuum vessel to a pressure of 10^{-3} mbar. The three major degrees of freedom are shown: translations along the x and y-axis (top and middle plots), and rotations around the z axis (bottom plot). These plots show that module movements as a result of the pumping operation are small (< 10 μm). This is an important result, particularly for the x translations, where the mechanical constraints are very tight for the distance between the sensors and their surrounding RF foil. The effects studied here, however, must not be confused with the expected movements during operation under vacuum when the pressure will remain within 10^{-3} mbar with respect to the beam vacuum (see section 2.2.1.1).

In order to move the beam from passing straight through the modules to hitting the targets, a shift of the experimental apparatus along the x axis was made. This movement is equivalent to the *VELO* half retraction, albeit with a different mechanical construction. The alignment constants before and after the movement are shown in the right hand plot of figure 3.19. Again, a very good stability of the module positions within their respective *VELO* half is observed. This result justifies the baseline assumption that the alignment procedure will not need to be performed on-line for each fill (for use in the trigger system). Instead, it is expected that the previously determined alignment constants for the modules can be used, while the *VELO* half alignment constants are updated to sufficient precision ($5-10$ μm) by knowledge from mechanical sensors of the *VELO* movement system. The alignment constants can then be refined for use in the off-line physics analysis.

3.4 Results from LHC Synchronisation Tests

A proton beam from the *SPS* was guided towards the *LHC* through the transfer line *TI8* (see figure 2.1) and stopped by a beam dump, as part of the *LHC* commissioning. This beam dump, known as the *TED* (Transfer line External beam Dump), is located close to the end of the transfer line and, hence, approximately 340 m away from the *LHCb* experiment.

Every 48 s a bunch of about 5×10^9 protons was shot onto the dump, producing a shower of particles. The muon part of this shower and its reaction products could still be registered by the *LHCb* detector.

In the first of these tests the *VELO* was operated with five active modules in each half. This allowed for the first time the reconstruction of tracks induced by the *LHC* beam on 22^{nd} August 2008. The first event recorded by the *VELO* is shown in figure 3.20.

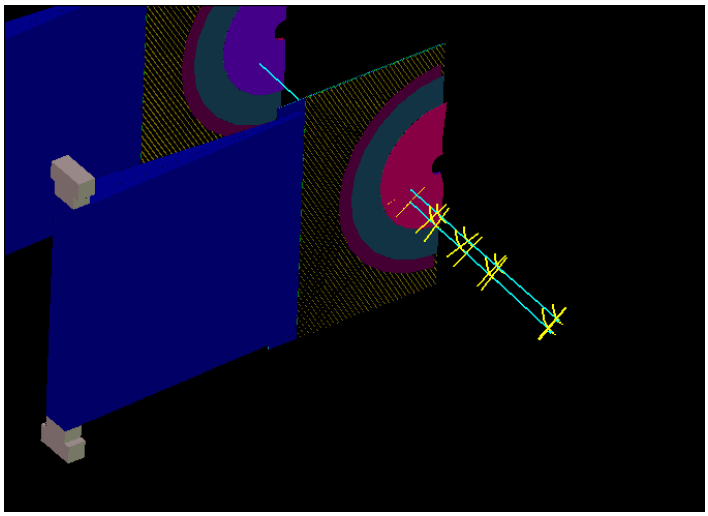


Figure 3.20: *Event display of the first event reconstructed by the VELO on* 22^{nd} *August 2008.*

After this successful initial test, all available (76) *VELO* sensors were read out[7] for extended periods of shots on the *TED*. In this configuration, data was taken on 24^{th} August and on the 5^{th} and 6^{th} September 2008. In the following, the terms August and September data samples refer to these periods with 76 sensors being read out. The next section describes the data taking conditions in detail, followed by demonstrations of the alignment and detector performances.

3.4.1 Data Taking Conditions

The data samples in August and September were taken under rather different conditions. As the time alignment was only very roughly known initially, data was taken with 15 consecutive triggers. This meant that, in addition to the main time sample for which the trigger decision was given by the calorimeter, the seven previous and seven subsequent time samples were read out additionally. Furthermore, to improve the robustness against time misalignments, the front-end chip control parameters were adjusted. This broadened the pulse shape of the analogue signals by about a factor two in time, with the aim that all *VELO* sensors would give a signal in the same time sample.

This proved to be a very successful approach and the signal was found to be predominantly in the default time sample. Therefore, it was decided to revert back to the standard signal shape for the September data taking and to reduce the number

[7]Eight sensors were not read out due to broken *TELL1* readout boards or problems with the power supplies. Meanwhile, all sensors are fully operational and can be read out.

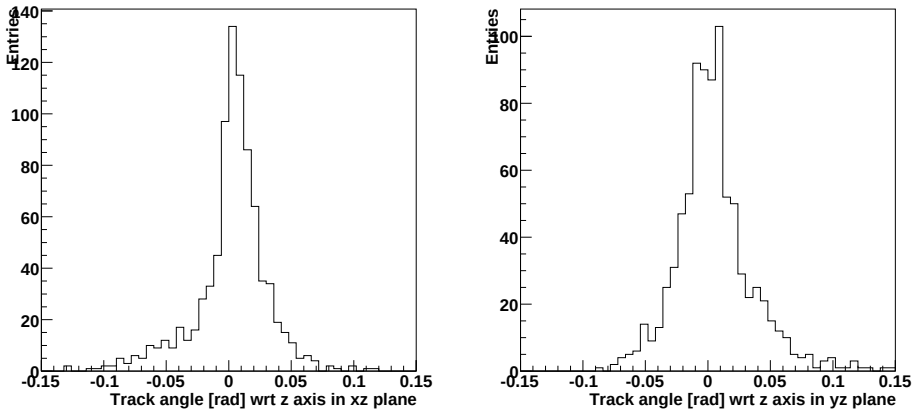


Figure 3.21: *Track angle distribution in the x-z plane (left) and in the y-z plane (right).*

of consecutive triggers to 11. In addition, the clustering thresholds were increased to rather harsh values in order to better suppress noise that was observed to be at a high level in August. This, however, led to the fact that most clusters were single-strip clusters as the smaller signal on the second strip would not pass the threshold.

As the individual sensors were not yet fully time aligned, the default length of the pulse shape meant that some sensors (about one third) had their main signal in the first time sample before the default one while other sensors had their main signal in the second time sample before the default one. For all track reconstruction algorithms the optimal time sample was chosen on a sensor by sensor level. In the August data sample 790 tracks with at least five space points were reconstructed, while 1370 were found for the September sample. Due to the significantly larger statistics only data from the September sample will be presented below unless otherwise stated.

Figure 3.21 shows the track angle distribution in the x-z and y-z planes, respectively. It should be noted that particularly the x slope shows a positive mean, consistent with the direction of the *TED* beam dump being roughly 12 mrad.

Figure 3.22 shows the number of space points per track for A side and C side, respectively. The highest entries of each distribution are equivalent to the number of active modules, i.e. modules with both sensors being read out. Normally, the distributions would be expected to peak at this value for tracks that are nearly parallel to the nominal beam axis. However, due to the time alignment conditions described above in combination with the harsh clustering thresholds, it was more likely to miss out one or more sensors and, hence, to reconstruct shorter tracks.

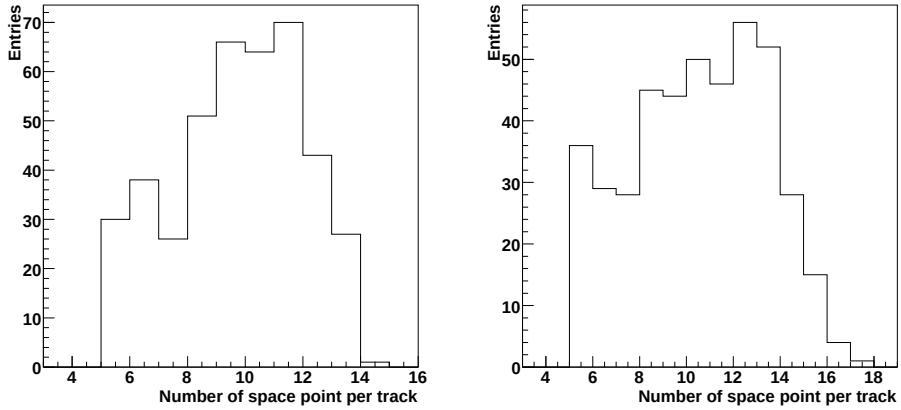


Figure 3.22: *Number of space points per track for September, shown for A side (left) and C side (right).*

3.4.2 Alignment

The data taken in August and September 2008 provided the first opportunity to exercise the *VELO* software alignment on the full system after the assembly at its final position in the *LHCb* detector. However, as the *VELO* was operated with both halves fully retracted, i.e. about 6 cm apart, and as all tracks had rather small angles with respect to the z axis, no alignment of the two halves with respect to each other could be performed.

The relative alignment of the sensors on each module requires large statistics in order to reach a precision that can improve the alignment constants obtained from optical survey measurements. Hence, no results from this method are reported either.

The relative alignment of the modules within each half was performed on all modules with a sufficient number of simultaneous hits on both their R and Φ sensors. Figure 3.23 shows the difference of the alignment constants obtained from software alignment with respect to the initial constants from metrology as a function of the module z position. The results are shown for both August and September data samples.

In order to assess the quality of the software alignment, the difference in the constants obtained with the August and September data sets is shown in figure 3.24. The quality of the constants for x and y translations is about 3 μm, while 140 μrad are obtained for rotations around the z axis. Both numbers are in agreement with the expectations for this number of tracks.

Another possible check for the reliability of the alignment constants obtained is to compare two methods of computing the constants. Therefore, the results of the `MILLEPEDE` algorithm were compared to those obtained by the one based on the Kalman

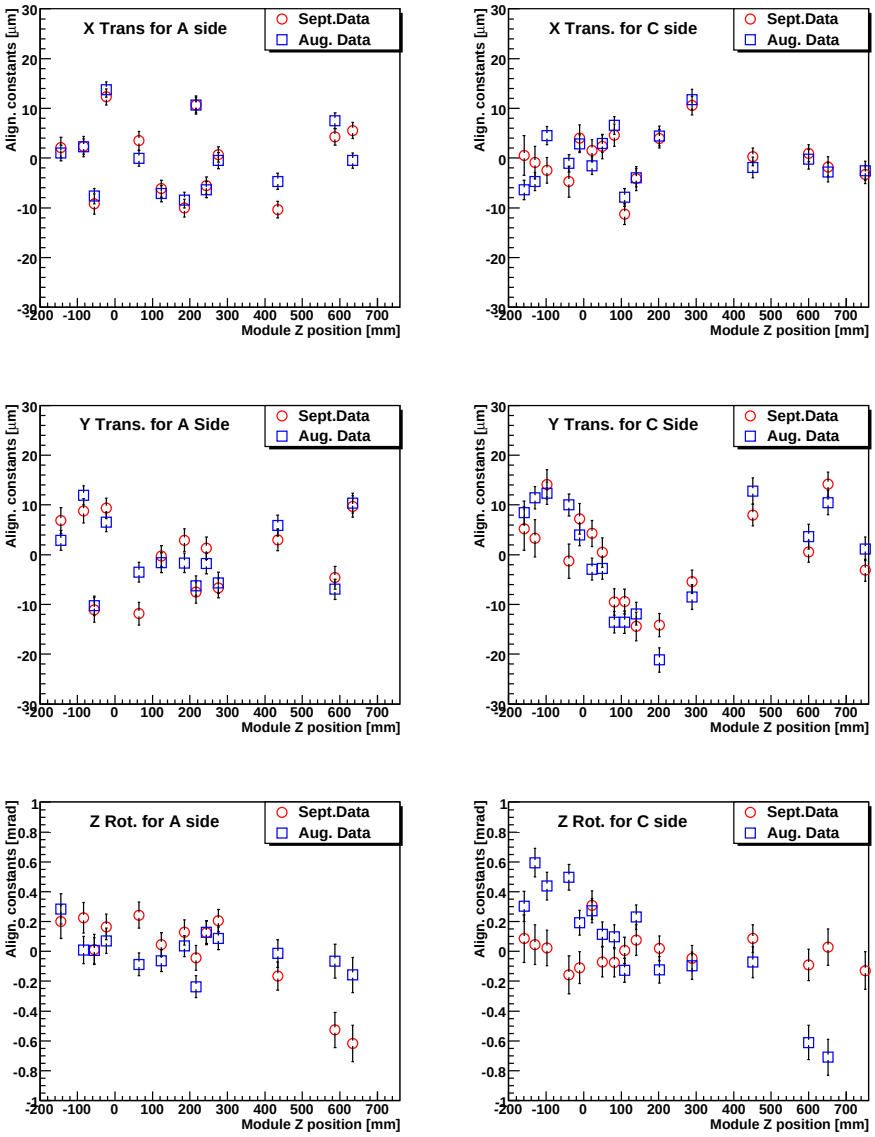


Figure 3.23: *Comparison of the alignment constants with respect to metrology for August (blue squares) and September (red circles) data samples. The alignment constants from top to bottom are x translations, y translations, and rotations around the z axis, shown as a function of the module z position for A side (left) and C side (right).*

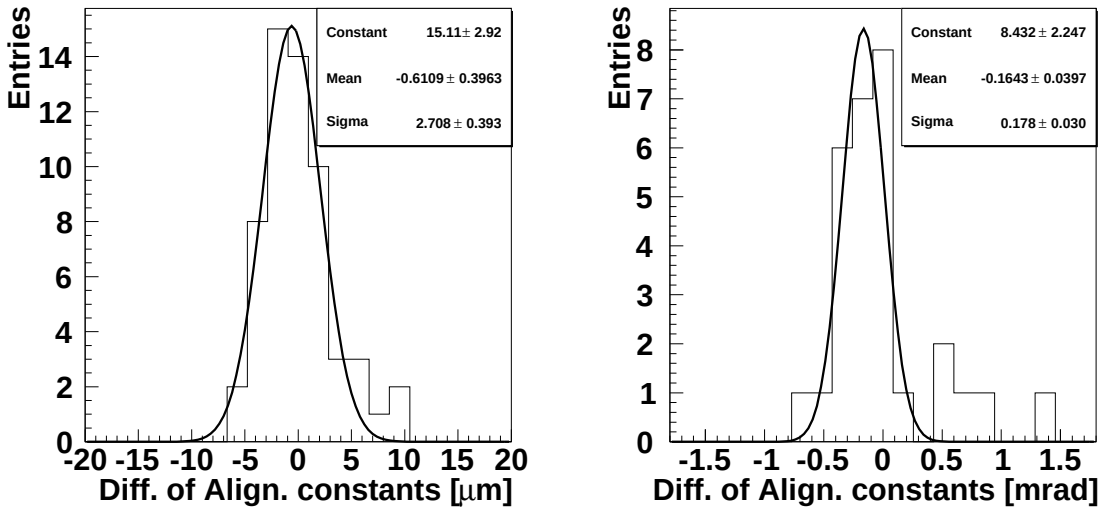


Figure 3.24: *Difference in alignment constants from August and September samples for all x and y translations (left) and all rotations around the z axis (right).*

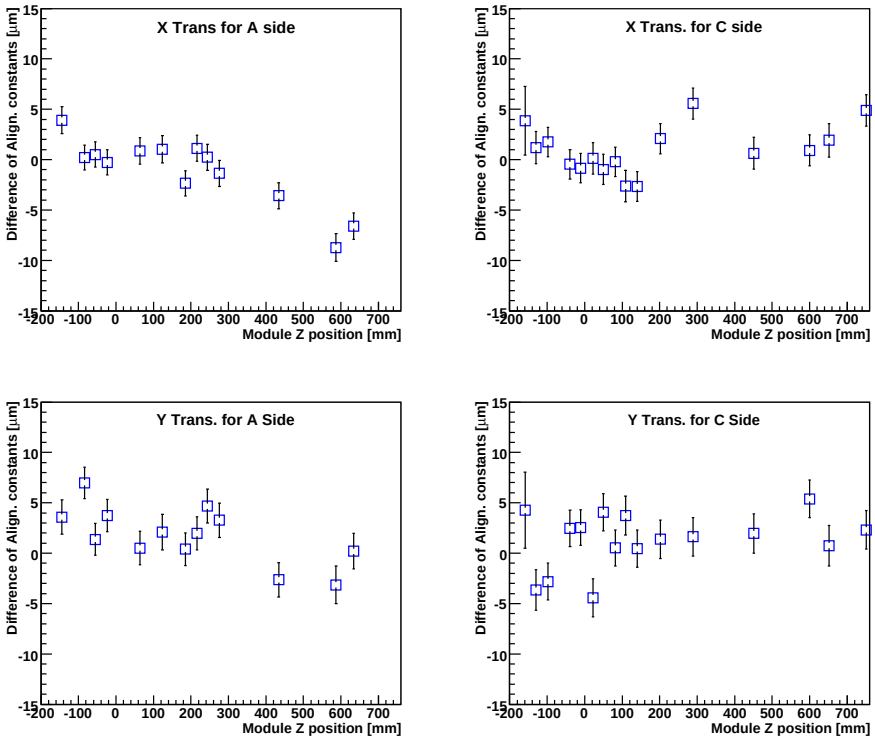


Figure 3.25: *Difference in alignment constants between those determined by the* `MILLEPEDE` *algorithm and those determined by an algorithm based on the Kalman filter track fit as a function of module z position. The alignment constants are x translations (top) and y translations (bottom), shown for A side (left) and C side (right).*

filter track fit (see page 73). The differences in the results are shown in figure 3.25 for x and y translations only as rotations around the z axis were not fitted by the Kalman filter based approach. All results are in very good agreement, thus, confirming the quality measured by comparing the August and September data sets.

3.4.3 Detector Performance

The basic quantity to assess the detector performance of a tracking detector is its single hit resolution. Its measurement with test beam data has been presented in section 3.3.3.3. A resolution measurement with the small sample of *TED* data is challenging, however, a minimal track selection is possible.

It was decided to select a narrow window around the direction towards the *TED* since tracks with larger angles had scattered and were, hence, more likely to have a lower momentum. This selection was necessary as the momentum distribution of the tracks observed in the *VELO* was a priori unknown. Tracks with low momenta also tend to be more prone to multiple scattering leading to significant changes in the track direction. All tracks were required to have at least 16 hits to avoid such effects, i.e. they must have passed through at least 8 *VELO* modules.

The residuals were plotted in five bins of increasing strip pitch. A correction has been applied to remove the contribution of the track extrapolation error, equivalent to the procedure discussed in section 3.3.3.3. As the statistics were not sufficient for a measurement on a per sensor level, the residuals from all sensors were combined. In this combination, any average offset in the individual distributions was removed to achieve results comparable to those obtained from the test beam where the resolution was measured for individual sensors. The resolution was then determined as the width of a single Gaussian fit to the residual distribution in each bin of strip pitch.

The result is shown in figure 3.26. For both R and Φ sensors the resolutions agree with the resolution expected for a binary readout. Given that the error on the fit is roughly 1 μm and that the data set contains $85\% - 90\%$ single strip clusters, this measurement is in agreement with the expectation from the previous test beam.

With appropriately lowered clustering thresholds and an improved time alignment it should be possible to reach the resolution measured in the test beam. Additional *TED* collision runs before the *LHC* re-start allow to further tune the *VELO* in order to have a high precision vertex detector at the start of physics data taking.

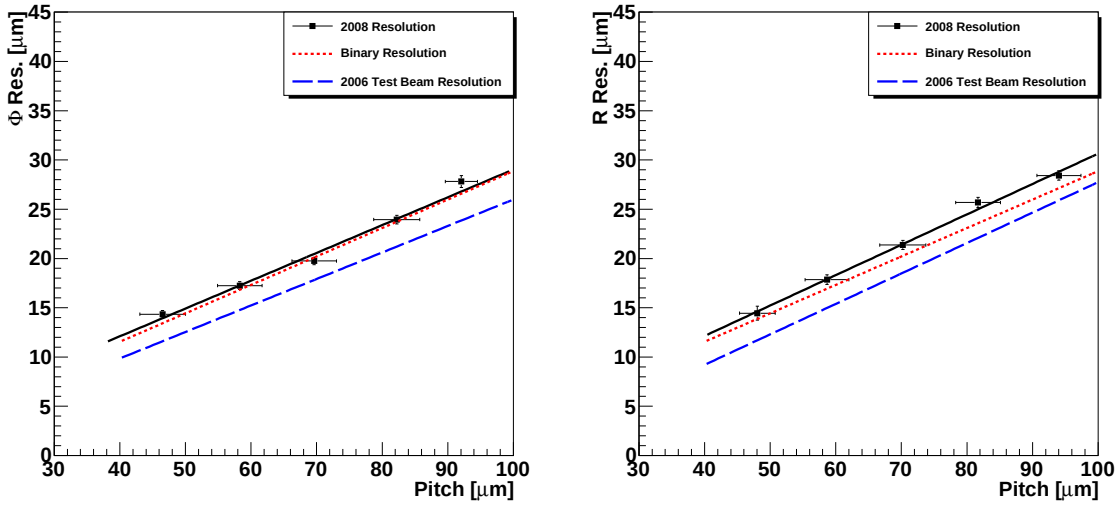


Figure 3.26: *Resolution as measured with data from the September data sample as a function of pitch for Φ sensors (left) and R sensors (right).*

4 Lifetime Fitting

> While the individual man is an insoluble puzzle, in the aggregate he becomes a mathematical certainty. You can, for example, never foretell what any one man will be up to, but you can say with precision what an average number will be up to. Individuals vary, but percentages remain constant. So says the statistician.
>
> *Arthur Conan Doyle*

Measuring the lifetime of a decaying particle can be a challenging task. However, decaying particles follow much simpler rules than the human organism. The main difference between the two is that the probability of dying is independent of time for a particle. For a human, by contrast, this probability tends to increase steadily with age (ignoring the decrease after a slightly higher mortality rate for newborns).

Figure 4.1 shows these probabilities. The constant decay probability for particles leads to an exponential lifetime distribution, whereas the human lifetime only begins to fall off rather sharply after the age of 60. The exponential is the basic function that is used to determine the lifetime of a particle.

4.1 Lifetime Fitting with Hadronic Channels

Lifetime fitting techniques in high energy physics are dominated by methods to correct for effects that distort the measured lifetime from its usual exponential shape. Such effects can, for example, be caused by changes in the geometrical acceptance as a function of lifetime. This is mostly the case for fixed target experiments which study long lived particles, such as kaons or hyperons, with typical decay regions of several tens of metres in length.

Other effects that distort the natural lifetime distribution can occur during the event selection, both at trigger level and during the off-line reconstruction. A very common and effective method to distinguish heavy flavour decays from other events is to apply a

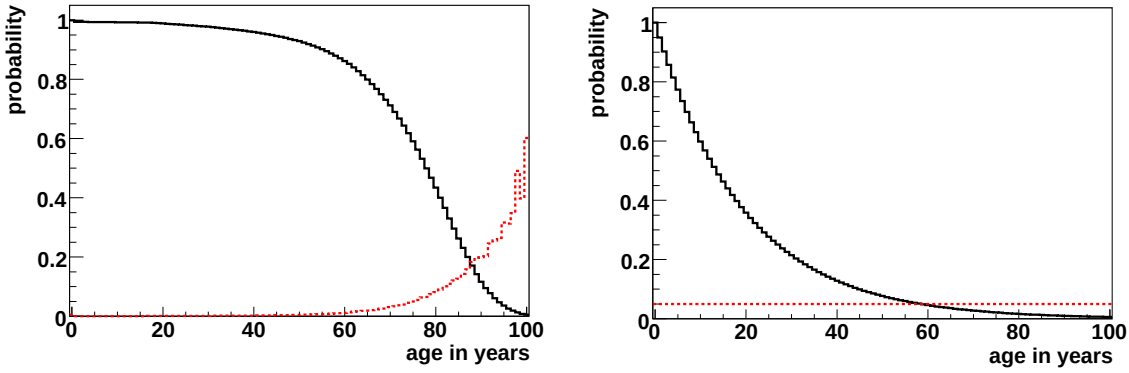


Figure 4.1: *Left: Relative number of survivors (black, solid line) and mortality rate (red, dashed line) of a Scottish male as a function of age. Input data from [70]. Right: The same plot for a decaying particle with an average lifetime of* 20 *years.*

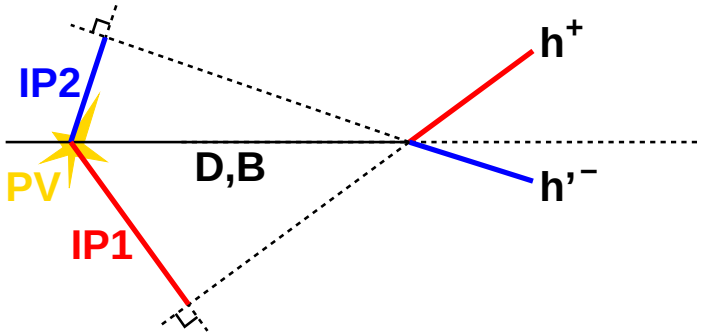


Figure 4.2: *Definition of impact parameter (IP) as the shortest distance between an extrapolated track and the primary vertex (PV) shown for decays of D or B mesons in two hadrons.*

cut on the minimal impact parameter (IP) of final state tracks. The impact parameter (see figure 4.2) is the shortest distance between an extrapolated track and the primary vertex (PV).

For D and B mesons, which have lifetimes of the order of 1 ps, the average IP of the tracks produced by their decay products is significantly non-zero. The ability to resolve these small quantities is guaranteed by the performance of the $VELO$, as discussed in section 2.2.4. Requiring a minimum value for the IP (significance) strongly suppresses promptly decaying background, however, it also rejects true signal events that decay at short lifetimes. This introduces a distortion in the reconstructed lifetime distribution from the true exponential shape.

The correction of this bias is the main difficulty in lifetime measurements of heavy flavour mesons. Before discussing methods to remove this lifetime bias, it should be noted that there are also ways of avoiding it. Leptonic or semi-leptonic decays are often able to reach sufficient purity only by particle identification, i.e. without using lifetime biasing selection criteria. For fully hadronic decays in $LHCb$ the only source of particle identification are the two $RICH$ detectors, which are currently unavailable at trigger level as their reconstruction is too time consuming. Hence, lifetime measurements using hadronic decays generally suffer from a biased lifetime distribution.

A bias due to the off-line event selection can be corrected in a reliable way as all quantities that are cut on are known. In addition, the event sample prior to application of the cuts is available for comparison. Contrary to that, all selections at trigger level are based on a preliminary, simplified reconstruction for reasons of the computing time available. Also, events not accepted by the trigger are not recorded and, hence, a comparison of the samples before and after selection similar to the off-line case is not possible. The quantities used at trigger level, called on-line quantities, cannot be directly linked to those from the off-line reconstruction. Hence, it is not trivial to deduce the impact that lifetime biasing cuts on on-line quantities have on off-line quantities.

In general, there are two possibilities to correct for the lifetime bias in the reconstructed lifetime distribution. One is to determine an average lifetime acceptance function from MC simulation (see figure 4.3). This is a function that determines the relative quantity of accepted events for a given lifetime. Such an approach relies on the correct simulation of the relation between off-line and on-line reconstruction, as well as on the correct description of the data by the MC simulation.

The alternative to using an average acceptance function is to determine an event-by-event acceptance function without the use of MC simulation. The following section

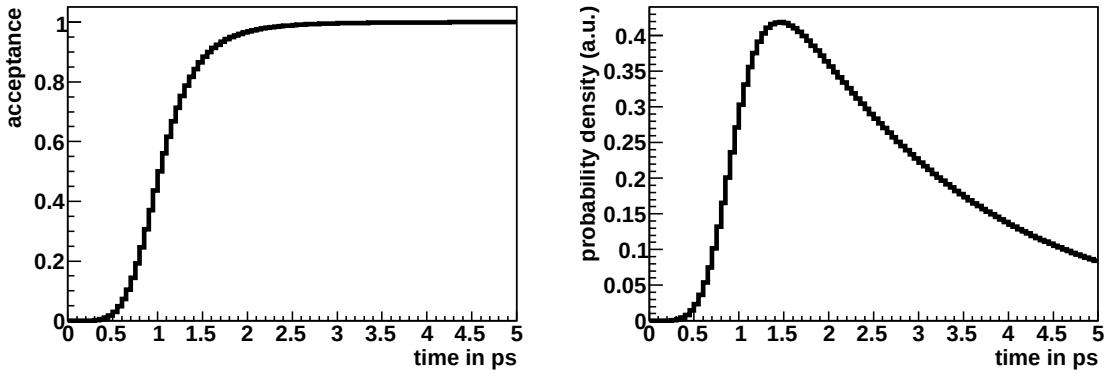


Figure 4.3: *Example of an average acceptance function due to a cut suppressing events at short lifetimes (left) and the corresponding measured lifetime distribution (right).*

describes how this function is obtained within the *LHCb* software. As will be seen, this method can also be used to determine the average acceptance function in a MC independent manner. The last section (4.3) in this chapter completes the method of measuring lifetimes by presenting a fitter based on the use of event-by-event acceptance functions.

4.2 Obtaining an Event-By-Event Acceptance Function

The idea of measuring event-by-event acceptance functions without the use of simulation is based on an approach introduced by Jonas Rademacker to the *CDF* experiment [71]. It has since been adapted to the possibilities offered by the *LHCb* software [72, 73].

The method follows the principle that the probability density for observing a decay with a set of kinematic variables kin at time t, $f(t, kin)$, can be factorised as

$$f(t, kin) = f(t|kin) \cdot f(kin), \tag{4.1}$$

where the key element of the method is to realise that the probability density of the event kinematics, $f(kin)$, is independent of the measured lifetime. The event kinematics depend on the phase space of the decay and potential form factors describing non-uniformities in the decay distributions. The lifetime at which a decay occurs only depends on the physics processes that lead to the decay, as discussed in section 1.3. The connection between the measured lifetime and the event kinematics is made by

selection cuts, e.g. at the trigger level. Due to these cuts, the event kinematics enter as a condition in $f(t|kin)$. The implementation of the *LHCb* trigger that leads to the lifetime bias is described in the following section, while section 4.2.2 explains the method of obtaining an event-by-event acceptance function and its use in a lifetime fit.

4.2.1 Lifetime Bias in the LHCb Trigger Selections

The *LHCb* trigger system has been introduced in section 2.3.1. Only the path of hadronic final states is discussed here as only decays into such states are studied in this thesis. Their main way of being triggered is through so-called hadronic triggers. Other trigger decisions are rare enough that they can be ignored in the analysis.

Hadrons are triggered in *L0* if they exceed a minimum transverse energy measured by the *HCAL*. By design, this decision has to be confirmed by the hadron alley in *HLT1*. This confirmation is reached if several requirements are fulfilled:

- Tracks are reconstructed in the *VELO* as r-z ($2D$) tracks and at least one has to be matched to an *HCAL* cluster that triggered *L0*.
- For any such track a $3D$ track fit is performed in the *VELO* and an *IP* of at least 0.1 mm with respect to any primary vertex is required.
- Successful tracks are extended to the other tracking stations which allows a momentum measurement. The transverse momentum of these tracks is required to be larger than 2.5 GeV/c.

After this stage there are two possibilities of completing the *HLT1* hadron alley:

- **Single Hadron Trigger**: This trigger requirement is fulfilled if the track in question has a transverse momentum above 5 GeV/c.
- **Di-Hadron Trigger**: This trigger requires a second track with an impact parameter of more than 0.1 mm and a transverse momentum above 1 GeV/c. The two tracks have to form a good vertex and their combination is required to point towards a primary vertex.

The requirements introducing a lifetime bias are the impact parameter cuts on the tracks. From figure 4.2 it is obvious that the distance of flight of a particle has to be larger than (or equal to) the impact parameters of its daughter tracks. Hence, requiring a minimum impact parameter implies a lower cut on the particles lifetime. It should be noted that this cut depends on the angular distribution of the tracks and, hence,

there is no straightforward relation between an impact parameter cut and the resulting average lifetime acceptance function.

The selection during the *HLT2* stage is not discussed here as its actual implementation is changing rapidly. The basic concept of lifetime biasing cuts and the way to compensate them, as described below, will, however, be very similar to the situation for *HLT1*.

4.2.2 Measuring an Event-By-Event Acceptance Function with the LHCb Trigger

The single event probability density of measuring a decay at time t, ignoring measurement errors, is given by

$$f(t|A) = \frac{\frac{1}{\tau}e^{-t/\tau}A(t)}{\int_{-\infty}^{\infty}\frac{1}{\tau}e^{-t'/\tau}A(t')dt'}, \tag{4.2}$$

where τ is the average lifetime of the decay and $A(t)$ is the lifetime acceptance function. The factor $\frac{1}{\tau}$ is kept as it normalises numerator and denominator separately in the absence of an acceptance function. The lifetime acceptance function depends on the event kinematics, hence $f(t|A) = f(t|kin)$ from equation 4.1. As described above, an average acceptance function like the one shown in figure 4.3 is not straightforwardly obtained from data. However, a lifetime acceptance function can be determined on an event-by-event basis.

Ignoring an overall efficiency for detecting and reconstructing an event, which is independent of the lifetime, this acceptance acquires values of either zero (rejected) or one (accepted). For example, for an event with given kinematics, i.e. fixed track slopes and momenta, there is a direct relation between the lifetime and the impact parameters of the tracks. Hence, cuts on impact parameters directly translate into a discrete decision about acceptance or rejection of an event as a function of its lifetime. This is illustrated in figure 4.4, where a lower impact parameter cut is assumed. None of the selections (trigger and off-line) considered in this thesis uses upper impact parameter cuts. Hence, the event-by-event acceptance function takes the shape of a step function.

Using such an event-by-event acceptance function in equation 4.2 leads to

$$f(t|A) = \frac{\frac{1}{\tau}e^{-t/\tau}\Theta(t-t_{min})}{\int_{t_{min}}^{\infty}\frac{1}{\tau}e^{-t'/\tau}dt'}, \tag{4.3}$$

where t_{min} is the turn-on point of the acceptance, and $\Theta(t-t_{min})$ is the Heaviside function. The denominator of this expression could be further simplified by evaluating the

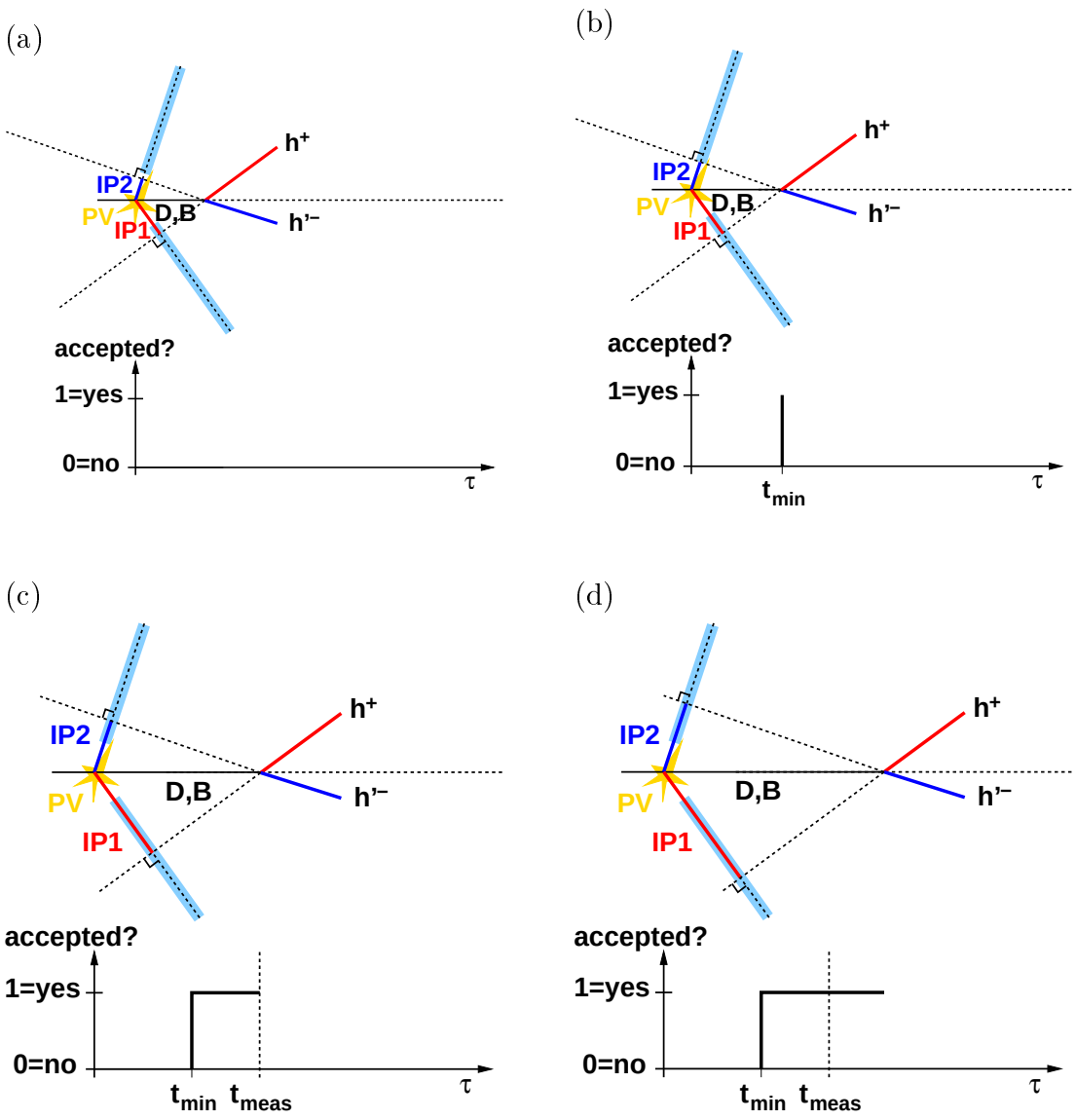


Figure 4.4: *Lifetime acceptance function for an event of a two-body hadronic decay. The shaded, light blue regions show the bands for accepting a track IP. After IP2 is too low in (a) it reaches the accepted range in (b). The actual measured lifetime lies in the accepted region (c), which continues to larger lifetimes (d).*

integral, however, this becomes non-trivial when taking into account the measurement resolution as discussed in section 4.3.1.1.

To actually determine the event-by-event acceptance function the trigger decision has to be evaluated for all lifetimes. Since, as described above, this function is a step function this task boils down to the determination of the position of the step. This requires the ability to manipulate the lifetime of an event and to re-evaluate the trigger decision thereafter.

In *LHCb* this can be done using an interface to the trigger software that has been implemented for this purpose. Thus, the implementation of this method is reasonably decoupled from the actual implementation of the trigger selection. It works as follows:

- The tracks used to reconstruct the decay off-line have to be associated to their counterparts in the on-line environment. As the on-line track information is retained in the off-line data structure this can be done with associator tools.

- The position of the primary vertex as reconstructed in the trigger is changed along the direction of flight of the particle under study[1].

- The trigger decision is evaluated for each new PV position using only the tracks from the signal decay.

- Once a change in the trigger decision, i.e. a step in the acceptance, is found the procedure is repeated around this position with a refined step size to increase precision.

The fact that the primary vertices are moved rather than the tracks themselves is an approximation that greatly simplifies the implementation. It is an approximation since the tracks would produce hits at different parts of the *VELO* sensors when originating from a particle decay at a different lifetime. This leads to changes in the extrapolation distances and measurement errors involved. As these effects are themselves small, they have negligible effects on the determination of the position of the turning points of the acceptance.

One caveat is the case of events with multiple primary vertices. In these events the method should still be applicable as described above. However, as discrepancies might occur from ambiguities in choosing the PV with respect to which the impact parameters have to be evaluated, cases with more than one primary vertex have not

[1] The direction of flight is defined as the line between the primary vertex and the decay vertex of the particle.

yet been studied in detail. This does not reflect a significant loss in events as around 80% of all recorded events in *LHCb* have exactly one primary vertex.

The lifetime bias caused by the off-line selection (this will be described in detail in section 5.1.4) can be treated in an equivalent way. However, one can take advantage of the exact knowledge of the signal tracks to analytically determine the turning point position.

After having obtained the event-by-event acceptance function for a data sample the average lifetime acceptance function can be extracted as well. Integrating out the kinematic (acceptance) part in the total probability density yields

$$\begin{aligned} f(t) &= \int f(t,A)dA \\ &= \int f(t|A)f(A)dA. \end{aligned} \tag{4.4}$$

Assuming a single turning point in the event-by-event acceptance leads to

$$\begin{aligned} f(t) &= \int \frac{\frac{1}{\tau}e^{-t/\tau}\Theta(t-t_{min})}{\int_{t_{min}}^{\infty}\frac{1}{\tau}e^{-t'/\tau}dt'} f(t_{min})dt_{min} \\ &= \frac{1}{N}\sum_i \frac{\Theta(t-t_{min,i})}{\int_{t_{min,i}}^{\infty}\frac{1}{\tau}e^{-t'/\tau}dt'}\frac{1}{\tau}e^{-t/\tau} \\ &= \frac{1}{N}\sum_i \frac{\Theta(t-t_{min,i})}{e^{-t_{min,i}/\tau}}\frac{1}{\tau}e^{-t/\tau}, \end{aligned} \tag{4.5}$$

where the sum goes over all N events. The average acceptance function follows as

$$A(t) = \frac{\sum_i \Theta(t-t_{min,i})e^{t_{min,i}/\tau}}{\sum_i e^{t_{min,i}/\tau}}. \tag{4.6}$$

Figure 4.5 shows the average acceptance function obtained using fully reconstructed $B_d^0 \to \pi^+\pi^-$ events. To evaluate the quality of this acceptance function it is overlaid with the measured lifetime distribution divided by an exponential with the true average lifetime. By definition, this results in the acceptance function. The slight discrepancy at small lifetimes results from an incomplete treatment of the determination of the acceptance at small lifetimes which will be corrected in an ongoing re-implementation of the method. This shows that the method described here can also be used to determine the average acceptance function purely from data for use in other analyses.

4.3 A Lifetime Fitter for Multi-Signal Environments

In the analysis of hadronic decays, channels often appear in groups with similar final states. One such example are the $B_{(s)}^0 \to h^+h'^-$ decays which are decays of B_d^0 and B_s^0

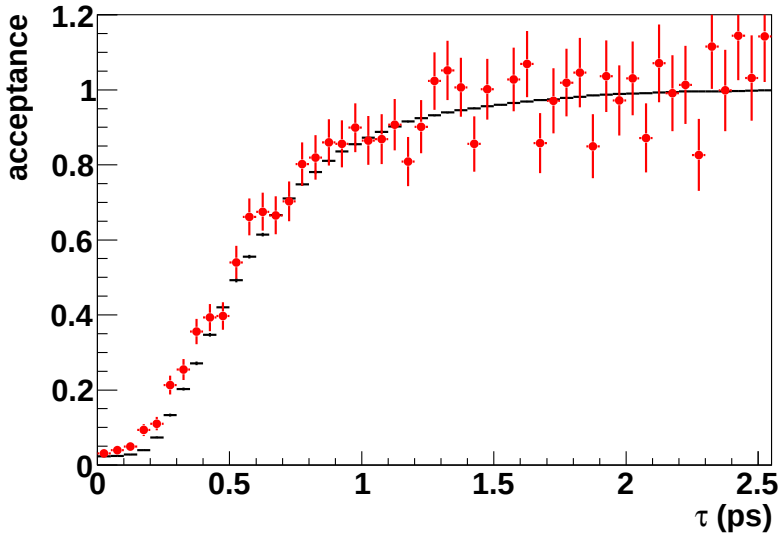


Figure 4.5: *Average lifetime acceptance function as obtained from per-event acceptance functions (black histogram) for $B_d^0 \to \pi^+\pi^-$ events compared to the measured lifetime distribution divided by an exponential (red dots).*

mesons into final states of two hadrons, mostly kaons and pions. Four of these channels, $B_d^0 \to \pi^+\pi^-$, $B_d^0 \to K^+\pi^-$, $B_s^0 \to \pi^+K^-$, and $B_s^0 \to K^+K^-$occur in similar quantities. Hence, their invariant masses when reconstructed under a common hypothesis partially overlap. A detailed study of these channels is presented in section 5.1.

For lifetime studies with any of these channels it is important to have a fitter that is able to account for several channels of specific background. Additionally, combinatorial background, i.e. background from partially or wrongly reconstructed decays, has to be taken into account. This section presents a lifetime fitter that has been designed for exactly these requirements.

The observables used in the fitter are the reconstructed lifetime t, the turning point from the event-by-event acceptance function t_{min}, and variables to identify the various signal channels $\mathbf{X}$. The most prominent variable to distinguish signal classes is the reconstructed invariant mass m. Other quantities that could be used for this purpose are single particle identification (*PID*) observables or angular distributions. The detailed use of these variables is described in section 4.3.3.

The total probability density for measuring a set of observables can be factorised as

$$f(t, t_{min}, \mathbf{X}) = f(t, t_{min}|\mathbf{X}) \cdot f(\mathbf{X}). \tag{4.7}$$

This equation has one term depending on the lifetime observables and a second depending on the variables to distinguish signal classes. As the fitter follows the concept of a

maximum likelihood fit, i.e. maximising the negative logarithm of the total probability, these two terms result in two independent likelihood curves.

To speed up the fitting process the fit is split in two parts:

- a signal fraction fit that determines the relative fractions of all signal classes involved,
- a lifetime fit which uses the previously obtained signal fractions as fixed input parameters and which performs the actual lifetime measurement.

The mass fit probability density $f(\mathbf{X})$ can be written as the sum over the values of probability density functions (PDF) for all signal classes multiplied by the corresponding relative fractions of that class:

$$f(\mathbf{X}) = \sum_{class} f(\mathbf{X}|class) \cdot P(class). \tag{4.8}$$

The term signal classes refers to all possible classes, i.e. the main signal, specific background, and combinatorial background. Using Bayes theorem the probability of a single event to belong to a particular signal class can be written as

$$P(class|\mathbf{X}) = \frac{f(\mathbf{X}|class) \cdot P(class)}{f(\mathbf{X})}. \tag{4.9}$$

This expression is used in the lifetime part of the fit that is described in the following section. Details of the mass fit are given in section 4.3.3.

4.3.1 Fitting the Lifetime of Exclusive Channels

It has been shown above how the total probability density for one event splits up into a lifetime part and a part that is used to fit the relative signal fractions (see equation 4.7). The lifetime part can be expressed as a sum over the contributions of the different signal classes.

$$f(t, t_{min}|\mathbf{X}) = \sum_{class} f(t, t_{min}|class) \cdot P(class|\mathbf{X}), \tag{4.10}$$

where $P(class|\mathbf{X})$ is the probability of an event belonging to a certain signal class as defined in equation 4.9. The remaining term for the time probability density for a given signal class is given by

$$f(t, t_{min}|class) = f(t|t_{min}, class) \cdot f(t_{min}|class), \tag{4.11}$$

with the first term being the probability density of measuring time t given the turning point of the acceptance function at t_{min} and a particular signal class, and the second term being the probability density of having this turning point for the given signal class.

4.3.1.1 The Signal Time Probability

The probability density of observing a lifetime t for a given signal class with turning point t_{min} is given by the convolution of an exponential with a resolution function $R(t,\sigma)$, and normalisation $N(t,t_{min})$

$$f(t|t_{min},signal) = N(t,t_{min})\frac{1}{\tau}e^{-t/\tau}\Theta(t) \otimes R(t,\sigma). \tag{4.12}$$

The resolution function is assumed to be a single Gaussian with width σ_t

$$R(t,\sigma) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(t'-t)^2}{2\sigma^2}}, \tag{4.13}$$

leading to the convolution equation

$$f(t|t_{min},signal) = N(t,t_{min})\int_{-\infty}^{\infty}\frac{1}{\tau}e^{-t'/\tau}\Theta(t')\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(t'-t)^2}{2\sigma^2}}dt', \tag{4.14}$$

which, using the definition of the frequency function

$$F(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-y^2/2}dy, \tag{4.15}$$

can be written as

$$f(t|t_{min},signal) = N(t,t_{min})\frac{1}{\tau}e^{-t/\tau}e^{\frac{1}{2}\sigma^2/\tau^2}F\left(\frac{t}{\sigma}-\frac{\sigma}{\tau}\right). \tag{4.16}$$

The normalisation is given by

$$\begin{aligned} N(t,t_{min})^{-1} &= \int_{t_{min}}^{\infty}\frac{1}{\tau}e^{-t'/\tau}e^{\frac{1}{2}\sigma^2/\tau^2}F\left(\frac{t'}{\sigma}-\frac{\sigma}{\tau}\right)dt' \\ &= \left[-e^{-t'/\tau}e^{\frac{1}{2}\sigma^2/\tau^2}F\left(\frac{t'}{\sigma}-\frac{\sigma}{\tau}\right)+F\left(\frac{t'}{\sigma}\right)\right]_{t_{min}}^{\infty} \\ &= 1+e^{-t_{min}/\tau}e^{\frac{1}{2}\sigma^2/\tau^2}F\left(\frac{t_{min}}{\sigma}-\frac{\sigma}{\tau}\right)-F\left(\frac{t_{min}}{\sigma}\right). \end{aligned} \tag{4.17}$$

This leads to the final equation for the lifetime probability density

$$f(t|t_{min},signal) = \frac{\frac{1}{\tau}e^{-t/\tau}e^{\frac{1}{2}\sigma^2/\tau^2}F\left(\frac{t}{\sigma}-\frac{\sigma}{\tau}\right)}{1+e^{-t_{min}/\tau}e^{\frac{1}{2}\sigma^2/\tau^2}F\left(\frac{t_{min}}{\sigma}-\frac{\sigma}{\tau}\right)-F\left(\frac{t_{min}}{\sigma}\right)}. \tag{4.18}$$

This can be simplified since $\sigma \ll \tau$ and in general also $t_{min} \gg \sigma$, as the resolution is roughly 0.04 ps, whereas the turning point position and average lifetimes are of the order of 1 ps. For $t, t_{min} \gg \sigma$ this leads to the simplified formula

$$\begin{aligned} f(t|t_{min},signal) &\approx \frac{\frac{1}{\tau}e^{-t/\tau}}{e^{-t_{min}/\tau}} \\ &= \frac{1}{\tau}e^{-(t-t_{min})/\tau}. \end{aligned} \tag{4.19}$$

Introducing $\tilde{t} \equiv t - t_{min}$, this can be written as

$$f(\tilde{t}|signal) = \frac{1}{\tau}e^{-\tilde{t}/\tau}. \tag{4.20}$$

Nevertheless, the full equation 4.18 has been implemented for $f(t,t_{min}|class)$ in the fit.

4.3.2 Inclusion of Combinatorial Background

It is very difficult to estimate a proper parametrisation for the lifetime probability density of background from combinatorial events. Therefore, an approach has been developed here that does not make any a priori assumptions about the shape of the background lifetime probability density function (PDF).

The idea is to obtain the background lifetime PDF from the observed total lifetime distribution. This is based on an implementation of a non-parametric PDF for neural net outputs used in Higgs analyses at *LEP* [74]. A similar approach has been first introduced to an *LHCb* analysis to model the forward-backward asymmetry in the rare decay $B_d^0 \to K^* \mu^+ \mu^-$ [75].

4.3.2.1 Obtaining a Non-Parametric Distribution

The easiest non-parametric PDF is obtained by using a histogram. However, this does not yield a smooth distribution which is a requirement for its use in an unbinned likelihood fit. In the method exploited for this thesis, for each event a so-called kernel function is used to represent the measured lifetime. The sum of all kernel functions, weighted by the event's probability of being background, then describes the lifetime probability density for background events. Hence, for a pure background sample the observed time distribution would be computed as

$$\tilde{f}(t|bg) = \frac{\sum_i Gauss(t, t_i, \sigma) \times P(bg|X_i)}{\underbrace{\sum_i P(bg|X_i)}_{N_{bg}}}, \tag{4.21}$$

where $\tilde{}$ denotes the measured distribution. Here, a Gaussian is used as a kernel function. In principle any normalisable function can be used as a kernel. A Gaussian has the advantage of guaranteeing a continuous derivative of the obtained PDF.

Figure 4.6 shows an example for the result of such a PDF for 10 events such that the individual kernel functions are visible. Due to the very low statistics, the description of the true PDF is far from perfect, however it is already much better than e.g. when using a histogram with only 10 entries. With increasing statistics the contributions from individual kernels will be less pronounced and the description of the actual PDF will improve. The quality of this approach with high statistics is shown in section 4.3.4.

The width of the kernel functions plays a key role in the method. It must not be confused with the measurement error since it is a parameter used to achieve a smooth distribution for the PDF and, hence, is more comparable with the bin size in

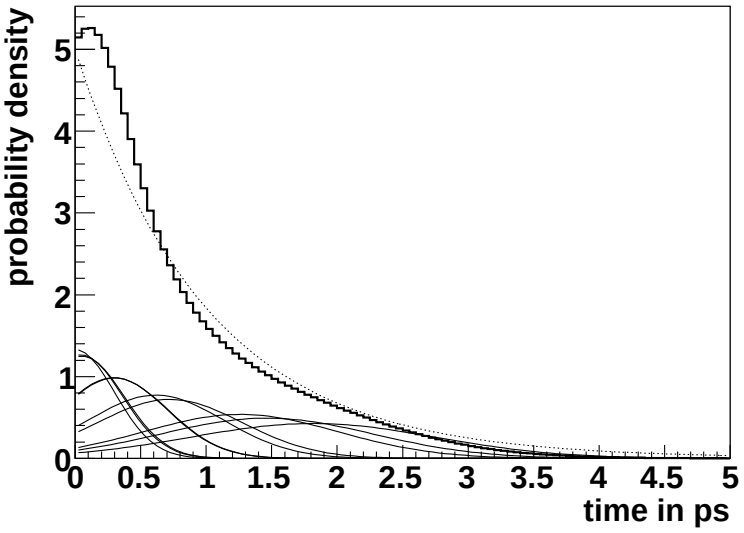


Figure 4.6: *Individual kernel contributions from* 10 *events (thin Gaussians) with the resulting PDF (thick histogram, not to scale) and the true distribution (dotted exponential).*

a histogram. There is no straightforward way of estimating the optimal width of a kernel. Cranmer [74] suggests to use

$$h(x_i) = \left(\frac{4}{3}\right)^{1/5} \sqrt{\frac{\sigma}{f_0(x_i)}} n^{-1/5}, \tag{4.22}$$

as a width estimate for the kernel for the i^{th} event, where σ is the measurement error, n the total number of events, and $f_0(x_i)$ an a priori estimate of the PDF at the measured value x_i [74]. This formula is derived from

$$h = \left(\frac{4}{3}\right)^{1/5} \sigma n^{-1/5}, \tag{4.23}$$

which minimises the mean integrated squared error for a normal distribution of width σ and $n \to \infty$.

Equation 4.22 introduces a recursive behaviour into the method. To avoid iterations of the rather CPU time intensive kernel method the term $f_0(x_i)$ of equation 4.22 is estimated by a histogram. This allows a generic implementation of the method, as no parameters need to be tuned with varying statistics.

4.3.2.2 Application to the Lifetime Fit

In the usual case of a data sample that comprises both signal and background events the sum in equation 4.21 would run over the full sample and hence the observed background

time distribution would get a contribution from signal events. To avoid this the signal contribution is subtracted from the total distribution using the known shape of the signal PDF and an estimate of the signal acceptance to yield a pure background time distribution. The signal acceptance ($A(t|class)$) has to be applied as the signal PDF describes the theoretical exponential distribution while the subtraction works with measured distributions. The resulting function is

$$\tilde{f}(t|bg) = \frac{\sum_i \left(Gauss(t, t_i, \sigma) - \sum_{class \neq bg} f(t|class) \cdot A(t|class) \cdot P(class|X_i) \right)}{\sum_i P(bg|X_i)}. \tag{4.24}$$

The actual background lifetime PDF, which is used in the likelihood formula, is obtained by dividing the observed background lifetime distribution by the background acceptance $A(t|bg)$

$$f(t|bg) = \tilde{f}(t|bg)/A(t|bg). \tag{4.25}$$

The average acceptance functions for the various signal classes are obtained by integrating the turning point distributions for each signal class using the weights as defined in equation 4.6. The individual classes are distinguished using $P(class|\mathbf{X})$ as defined in equation 4.9. The turning point distributions are measured using Gaussian kernel functions. This results in the function

$$A(t|class) = \frac{\sum_i F((t - t_{min,i})/\sigma) e^{t_{min,i}/\tau} P^2(class|\mathbf{X})}{\sum_i e^{t_{min,i}/\tau} P^2(class|\mathbf{X})}, \tag{4.26}$$

where F is the frequency function defined in equation 4.15 and σ the smoothing parameter for the turning point distributions. The respective class probability ($P(class|\mathbf{X})$) is used as an additional weight in order to increase the contribution of regions in phase space which have less background.

The signal models enter the calculation of the average acceptance function, thus introducing a dependence of the average acceptance functions on the fit parameters, i.e. the average lifetimes. Calculating the average acceptance functions for each set of fit parameters would introduce a quadratic dependence on the number of events. To avoid this, the average lifetimes used in the calculation of the average acceptance functions are fixed. A second iteration of the fit can be run with updated values of the average lifetimes to prove the stability of the results. The impact of this simplification is studied in detail in section 5.2.3.10.

The average acceptance function of the combinatorial background is described by a set of exponentials. Also this description can be updated (to a non-parametric description if necessary) after a first fit to real data has revealed the actual structure of background events.

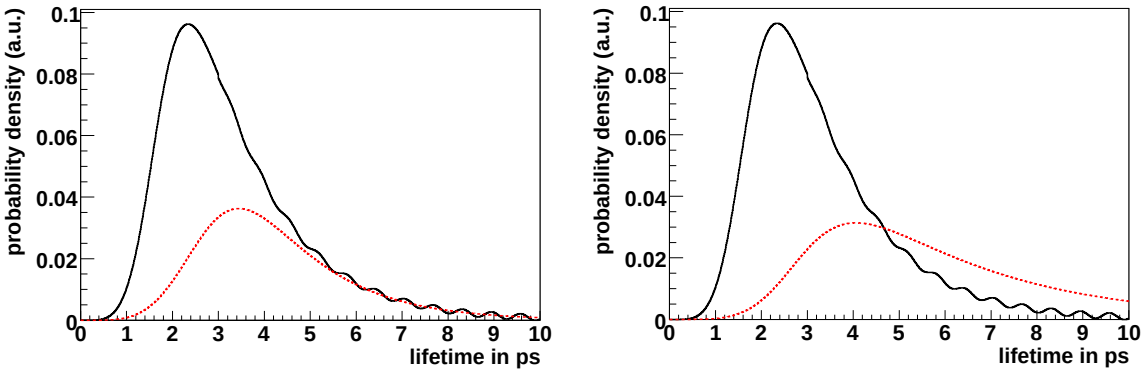


Figure 4.7: *Example for situations resulting in negative values for the background PDF when subtracting the signal contribution (red, dotted) from the total distribution (black) due to statistical fluctuations (left) or wrong fit values of the signal lifetime (right).*

4.3.2.3 Avoiding Negative Probability Densities

This method gives rise to an additional complication in its implementation. For vastly wrong signal lifetimes this can lead to negative values in certain regions of lifetime. The trial of incorrect signal lifetimes is a necessary part of the fitting process. Even when using the correct signal lifetime negative values are not excluded (see figure 4.7). They can occur due to statistical fluctuations that may not be fully smoothed out by the kernel approach in regions of lifetime where the signal fraction is high, and hence the subtraction can naturally produce small negative values.

As no lifetime region is physically excluded for background events, setting the background PDF to zero if the subtraction yields negative values is not an option. Instead, if the resulting background PDF has a fraction of the total PDF of less than a small value ϵ, all possible fractions between ϵ and $-\infty$ are continuously mapped to be inside the interval $[\epsilon, 0)$. The mapping uses the diverging behaviour of the tangent function to uniquely assign a positive value to every value between ϵ and $-\infty$. Using this method a stable behaviour of the fitter is obtained.

The value of ϵ has to be tuned. Too small values result in too little correction. Hence, the method is strongly affected by individual events producing probabilities close to zero, which vary with the fitting parameters. This yields a likelihood curve that is no longer parabolic around its minimum but which can have steps or fluctuations. This leads to a failure of the minimisation or wrong error estimates. Too large values of ϵ, on the other hand, introduce too large a correction leading to a bias in the fit result, as the background PDF is artificially set to significantly non-zero values where it should

be close to zero. A value of ϵ of 0.01 has proven to lead to a very stable behaviour of the fit, independently of the size of the data sample. It should be noted that even this value can lead to a small but non-negligible bias as discussed in detail in section 5.2.1.

4.3.2.4 Improved Implementation

The probability for each event of being background is used as a weight to enhance the contribution of regions with a cleaner background sample in the computation of the background time PDF. The resulting formula is

$$f(t|bg) = \frac{\sum_i \left(\left[Gauss(t, t_i, \sigma) - \sum_{class \neq bg} f(t|class) \cdot A(t|class) \cdot P(class|X_i) \right] \cdot P(bg|X_i) \right)}{(\sum_i P(bg|X_i)^2) \cdot A(t|bg)}. \tag{4.27}$$

This weighting also further reduces the cases where negative values occur for the background lifetime PDF. A disadvantage of this method is that it relies more strongly on the non-correlation of the background lifetime and the variables used to distinguish the signal classes. However, a method obtaining the background distribution from a side band in mass would depend even more on this assumption of non-correlation. It has been shown with data from *CDF* that this correlation is negligible [76].

Since the value of the kernel functions of all events has to be evaluated for each event this results in a quadratic CPU time dependence of the fit on the number of events. Therefore, a lookup table has been introduced which evaluates the kernel functions at discrete points before the start of the fit. A variable binning is used to allow an efficient coverage of a time range up to about 500 times the lifetime of a B meson. This reduces the CPU time dependence of the fitter to be linear with respect to the number of events.

A third order spline interpolation is used to obtain a smooth PDF from the discrete sampling points. A linear interpolation would be significantly faster, however its non-continuous derivative leads to fluctuations in the likelihood curve that prevent the fitter from converging properly. Quadratic spline interpolations are prone to large fluctuations in the interpolated curve and are hence useless. In the extremely rare case that the third order spline interpolation yields negative values the fitter reverts back to the linear interpolation.

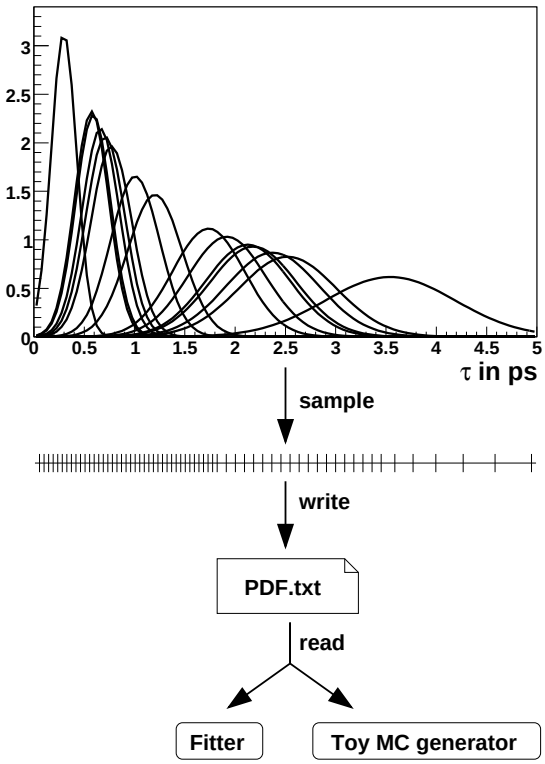


Figure 4.8: *Schema of applying the kernel method: after evaluation of the kernel for all events the distribution is sampled at varying intervals. These sampling values are stored in text files and used in the fit or for toy MC generation by using spline interpolation.*

4.3.3 Distinguishing Various Signal Classes

The lifetime fit as described in the previous section requires a statistical distinction of various signal classes by means of equation 4.9. Therefore the signal fractions, $P(class)$, are fitted prior to the lifetime fit. They are determined by a fit to the distributions of a set of variables that are particularly useful for distinguishing the various signal classes. As mentioned above, the most typical such variable is the invariant mass of the reconstructed daughter particles.

Another excellent candidate, particularly for hadronic channels, is the *PID* of the individual daughter particles. The channels considered in this thesis all have either pions or kaons in their final state. Hence, the *PID* variable used is the difference in the logarithm of the likelihood of the particles being either pion or kaon, $\delta \log \mathcal{L}_K$. This variable is dominated by information from the *RICH* detectors which underlines their importance to the experiment.

Following equation 4.8 the per-event probability density is given by

$$f(\mathbf{X}) = \sum_{class} f(\mathbf{X}|class) \cdot P(class), \tag{4.28}$$

where $\mathbf{X}$ is a vector of the variables used to distinguish the various signal classes, i.e. $\mathbf{X} = (m, \delta \log \mathcal{L}_K^+, \delta \log \mathcal{L}_K^-)$, with the *PID* variable $\delta \log \mathcal{L}_K$ for positive and negative particles, respectively. As the three variables used here are independent it follows that

$$f(\mathbf{X}|class) = \prod_i f(X_i|class). \tag{4.29}$$

This means that the ingredients needed for the signal fraction fitter are the PDF of the three variables for the individual signal classes. These can be obtained with various methods.

For the mass distribution an analytic expression for the line shapes of the individual components can be used [29]. Alternatively the shape of the mass distributions can be described by a template PDF. The latter approach is used in this thesis. To obtain the templates the mass distributions for the various signal channels are scanned using the full MC simulation. The use of MC simulation is not essential here, as the analytic descriptions could be used just as well.

The actual determination of the templates exploits the kernel method. Each event is assigned a kernel function completely analogously to the determination of the total time distribution in the lifetime fit (see section 4.3.2). The obtained distribution is sampled at variable intervals and stored in text files. Thus, the only CPU time consuming step in the fit is the spline interpolation during the initialisation. Hence, the evaluation of

the PDF values during the minimisation is faster than the evaluation of an analytic description of the PDF, as it only requires the evaluation of the spline values.

The distribution of the *PID* parameters cannot be described in an analytic form. There, the same approach with templates from the kernel PDF is used to describe the shapes of $\delta \log \mathcal{L}_K$. Four different templates have been obtained from full MC simulation: those for positive and negative pions from the decay $B^0_d \to \pi^+\pi^-$ and those for positive and negative kaons from $B^0_s \to K^+K^-$ events.

With data this approach will be replaced by a purely data driven technique. It relies on measuring the *PID* distribution in clean control channels and re-weighting these distributions according to the various distributions in momentum and transverse momentum. This method has been developed for and tested with the $\mathcal{CP}$ violation analyses in $B^0_{(s)} \to h^+h'^-$ decays [29].

4.3.4 Quality of the Kernel Method

The various applications of the kernel sampling technique are evaluated by comparing with the original distributions to establish the quality of the method. A χ^2 variable is defined as the square of the difference between the kernel result and the original distribution divided by the statistical error of the original distribution at the point of evaluation. Figure 4.9 shows three distributions, one where the kernel method has been applied to the lifetime distribution of an event sample, a second for a mass distribution and a third for a kaon *PID* distribution. In all cases the individual kernel widths have been determined using the method based on equation 4.22. The χ^2 variable divided by the number of non-empty bins for the three distributions is 115/104 (lifetime), 107/97 (mass), 59/95 (*PID*).

4.4 Conclusion

This chapter presented a method for performing lifetime measurements of decays into hadronic final states. This class of final states suffers from a lifetime biasing effect caused by the software trigger. A method has been described for determining this bias on an event-by-event level using an interface to the trigger software. This method does not require any input from MC simulation.

A lifetime fitter has been presented which is based on a non-parametric treatment of combinatorial background. This is achieved through the use of kernel functions to describe the total measured lifetime distribution and by subtracting the various signal

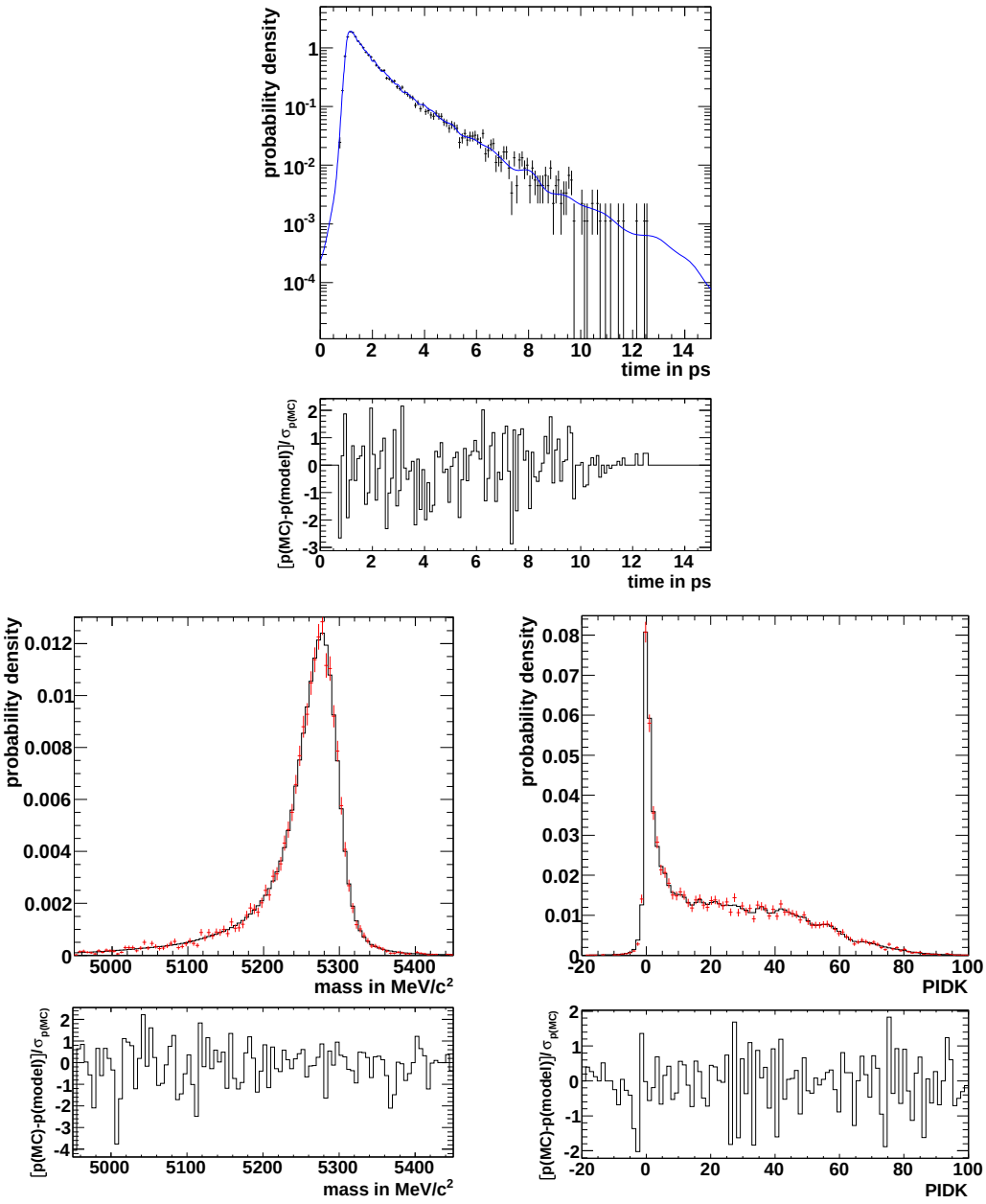


Figure 4.9: *Evaluation of the kernel method for lifetime distribution (top centre), mass (bottom left), and PID (bottom right). The upper plots show data as crosses and the kernel scan as a histogram. The lower plots show the difference between data and kernel scan divided by the statistical error of the data distribution.*

contributions in order to retrieve the background PDF. Detailed studies to evaluate the quality of the fitter and the sensitivity of lifetime measurements using this method are presented in the following chapter.

5 Lifetime Measurements in Two-Body B and D Decays

> A judicious man uses statistics, not to get knowledge, but to save himself from having ignorance foisted upon him.
>
> *Thomas Carlyle*

The main goal of this thesis is to provide a method and study its sensitivity for measuring $\Delta\Gamma_s$ with $B_s^0 \to K^+K^-$ events. The first section outlines the method for extracting $\Delta\Gamma_s$. Section 5.2.3 covers studies of the stability of the lifetime fitter used as well as the evaluation of potential systematic uncertainties. Section 5.3 outlines the possibility of applying the method to lifetime measurements in D mesons which provide excellent opportunities for measurements with early data. Finally, section 5.4 is dedicated to a study of the impact of misalignments on measurements with two-body decays. This misalignment study obviously has an impact on the $\Delta\Gamma_s$ extraction, however, it is also applicable to other measurements such as $\mathcal{CP}$ violation measurements.

5.1 Measuring $\Delta\Gamma_s$ with $B_s^0 \to K^+K^-$ Events

The extraction of $\Delta\Gamma_s$ from $B_s^0 \to K^+K^-$ events is based on a lifetime measurement which yields a result according to equation 1.80. This equation contains five unknown parameters: the decay width Γ_s, the width difference $\Delta\Gamma_s^{SM}$, the New Physics mixing phase ϕ^{NP}, the Standard Model $\mathcal{CP}$ violating phase σ^{SM}, and the phase responsible for $\mathcal{CP}$ violating NP effects σ^{NP}. Obviously, these cannot be constrained by a single measurement.

5.1.1 Measurements in a Standard Model Scenario

In section 1.4.2 it was already stated that Γ_s can be obtained from measurements in flavour specific channels such as $B_s^0 \to D_s^{\pm}\pi^{\mp}$. Using such a measurement the quantity

$$\tilde{y} \equiv \frac{\Gamma(B_s^0 \to K^+K^-)}{\Gamma(B_s^0 \to D_s^{\pm}\pi^{\mp})} - 1 = \frac{\tau(B_s^0 \to D_s^{\pm}\pi^{\mp})}{\tau(B_s^0 \to K^+K^-)} - 1 \tag{5.1}$$

can be defined. In the *SM* case, both ϕ^{NP} and σ^{NP} vanish and $\tilde{y}$ only depends on $y = \frac{\Delta\Gamma_s}{2\Gamma_s}$ and the *SM* $\mathcal{CP}$ violating phase σ^{SM}. Hence, equation 1.80 leads to the relation

$$\tilde{y} = -y\cos(2\sigma^{SM}) - y^2\sin^2(2\sigma^{SM}) = -y\cos(2\sigma^{SM}) - y^2(1-\cos^2(2\sigma^{SM})). \tag{5.2}$$

A value for y can be extracted from $\tilde{y}$ using the knowledge of $\cos(2\sigma^{SM})$ (see equation 1.79). The error on $\tilde{y}$ is given by

$$\sigma(\tilde{y}) = \sqrt{\delta(\tau(B_s^0 \to K^+K^-))^2 + \delta(\tau(B_s^0 \to D_s^{\pm}\pi^{\mp}))^2} \cdot \frac{\tau(B_s^0 \to D_s^{\pm}\pi^{\mp})}{\tau(B_s^0 \to K^+K^-)}. \tag{5.3}$$

The dominant contribution to the statistical error of $\tilde{y}$ will be the relative error of the lifetime measurement in $B_s^0 \to K^+K^-$. However, the measurement in $B_s^0 \to D_s^{\pm}\pi^{\mp}$ will have a significant contribution as its annual yield is less than a factor of three higher than the one of $B_s^0 \to K^+K^-$ [77].

Given the constraints on $\cos(2\sigma^{SM})$ the relation between the observable $\tilde{y}$ and y is shown in figure 5.1. The value of y is extracted from a measured value of $\tilde{y}$ by solving equation 5.2 for y.

Theory favours a value for y around 0.05, i.e. $\Delta\Gamma_s/\Gamma_s = 0.10$. Therefore, a measurement of y with an absolute error at the per cent level would be desirable to achieve a 5σ significance. Within the allowed range for $\cos(2\sigma^{SM})$ a measurement of $\tilde{y} = 0.050 \pm 0.010$ would translate into $y = 0.052 \pm 0.010$. Assuming $\Gamma_s = 0.68$ ps^{-1} (from $\tau(B_s^0) = 1.47$ ps [12]) this would translate into $\Delta\Gamma_s = (0.071 \pm 0.014)$ ps^{-1}. This means that a measurement of $\Delta\Gamma_s$ at the 5σ level requires a 0.8% measurement of the $B_s^0 \to K^+K^-$ lifetime, when taking into account the error on the Γ_s measurement from $B_s^0 \to D_s^{\pm}\pi^{\mp}$.

5.1.2 Measurements in a New Physics Scenario

In the presence of *NP* the interpretation of $\tilde{y}$ is less straightforward. With two additional phases entering the equation additional constraints are needed. Figure 5.2 shows the impact of a large *NP* mixing phase on the extraction of y from a measurement of

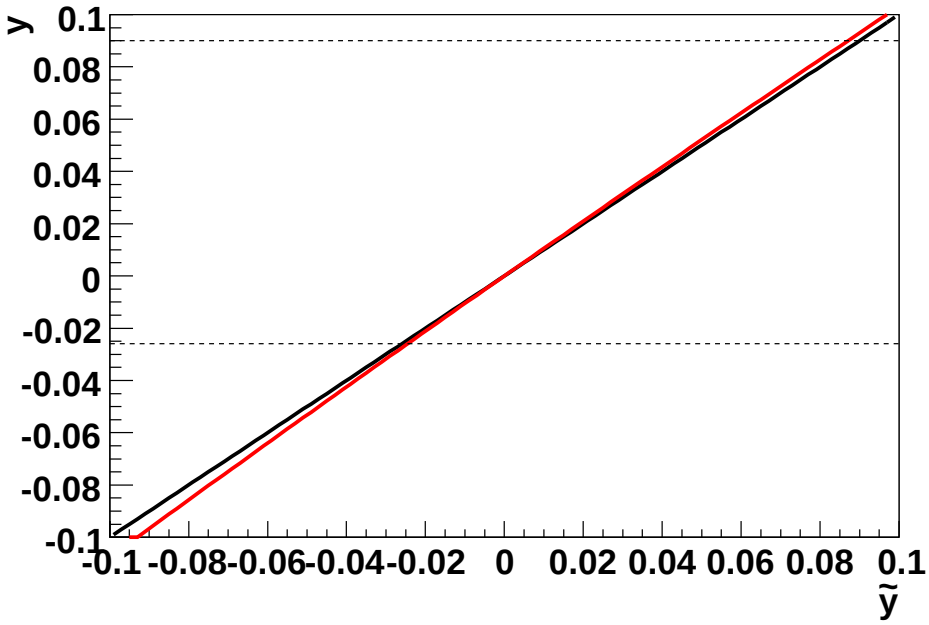


Figure 5.1: *Relation of y and $\tilde{y}$ (for definition see text) for the SM cases of $\cos(2\sigma^{SM}) = -1$ (black) and $\cos(2\sigma^{SM}) = -0.95$ (red, steeper). Also shown is the current 95% C.L. for y.*

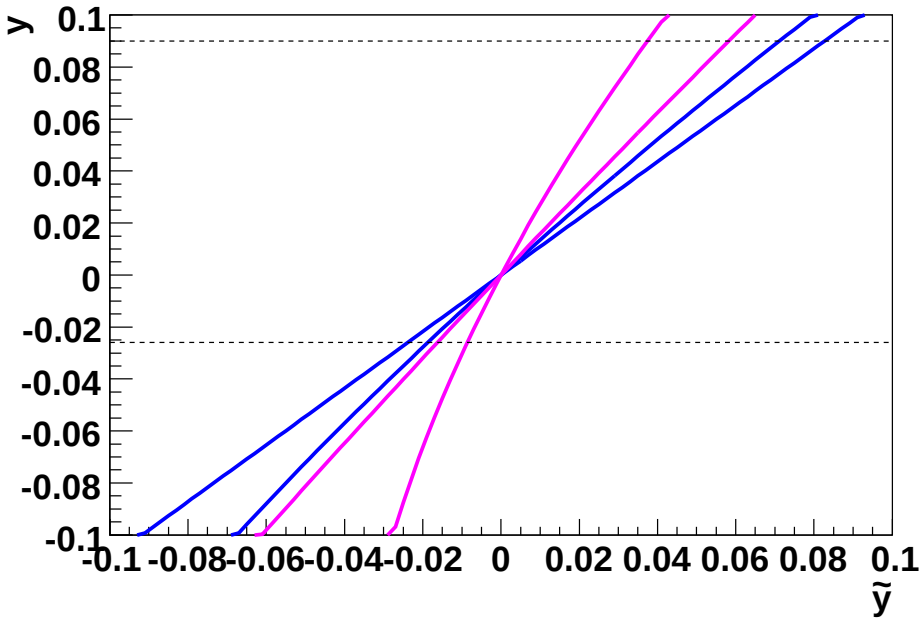


Figure 5.2: *Relation of y and $\tilde{y}$ (for definition see text) for NP phases of $\phi^{NP} = 0.4$ (blue) and $\phi^{NP} = 0.8$ (purple). The two lines of each colour mark the allowed region due to the uncertainty on σ^{SM}. Also shown is the current 95% C.L. for y.*

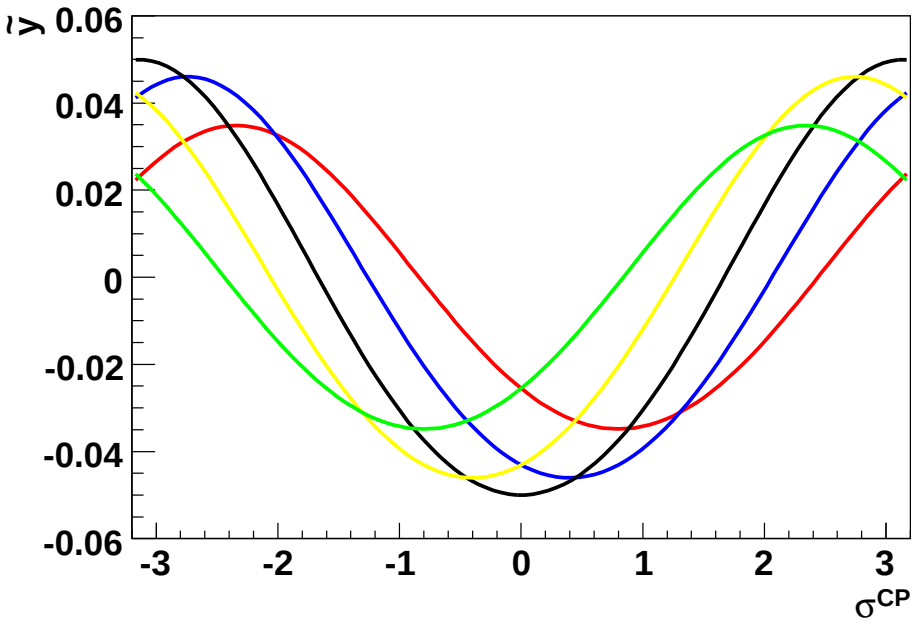


Figure 5.3: *Relation of* $\tilde{y}$ *and* $\sigma^{CP} \equiv 2\sigma^{SM} + 2\sigma^{NP}$ *for NP phases of* $\phi^{NP} = -0.8, -0.4, 0.0, 0.4, 0.8$ *(minima from left to right).*

$\tilde{y}$. As opposed to the *SM* $\mathcal{CP}$ violating phase, a large *NP* phase significantly degrades the sensitivity on y and hence on $\Delta\Gamma_s$.

As described in section 1.4.2, a measurement of the lowest angular momentum component of the decay $B_s^0 \to J/\psi\phi$ is a good candidate for a comparison with a $B_s^0 \to K^+K^-$ measurement. The only difference between the observable $\tilde{y}$ for the two channels is the $\mathcal{CP}$ violating phase contributing to $B_s^0 \to K^+K^-$. However, a full angular analysis of the decay $B_s^0 \to J/\psi\phi$ can lead to measurements of y and ϕ^{NP} on its own. It is therefore of interest to study the sensitivity to the $\mathcal{CP}$ violating *NP* phase entering the decay $B_s^0 \to K^+K^-$.

Figure 5.3 shows the dependence of the observable $\tilde{y}$ on the combined *SM* and *NP* $\mathcal{CP}$ violating phases. The relation is given for five different values of ϕ^{NP} showing that a reasonably good sensitivity can be obtained if ϕ^{NP} is known to about 0.2 rad or better. It should be noted that the actual extraction of the *NP* $\mathcal{CP}$ violating phase σ^{NP} requires the subtraction of the *SM* component. The *SM* is constrained by equation 1.79 to $|2\sigma^{SM}| < 0.3$ rad. Assuming $\phi^{NP} = 0$, a measurement of $\tilde{y} = 0.00 \pm 0.01$ would lead to two constraints on σ^{CP} of $|\sigma^{CP}| = 1.67 \pm 0.20$. Combined with the constraint on $|2\sigma^{SM}|$ this yields $|2\sigma^{NP}| = 1.67 \pm 0.36$.

5.1.3 The LHCb Sensitivity to $\Delta\Gamma_s$

LHCb can measure $\Delta\Gamma_s$ in various channels. The sensitivity to $\Delta\Gamma_s$ in the *SM* scenario for a measurement based on $B_s^0 \to K^+K^-$ events is 0.015 ps^{-1} for one nominal year of data taking (following from section 5.1.1) as will be discussed in more detail below (see section 5.2.2.2).

A measurement with the channel $B_s^0 \to \phi\phi$ is sensitive to $\Delta\Gamma_s$ as well (see section 1.4.2.2). The branching ratio for $B_s^0 \to \phi\phi$ is lower than that of $B_s^0 \to K^+K^-$ by about a factor two. In addition, this channel has a significantly lower trigger efficiency due to the lower momentum of the four final state particles. A study of this channel [78] estimates the yield in a nominal year to 3100 events. Although the study does not quote a sensitivity to $\Delta\Gamma_s$ it is clear that it cannot compete with that from $B_s^0 \to K^+K^-$ discussed above due to the number of events being lower by about a factor of 20.

With a branching ratio of two orders of magnitude above that of $B_s^0 \to K^+K^-$, $B_s^0 \to J/\psi\phi$ is the prime candidate for measuring $\Delta\Gamma_s$ (see section 1.4.2.1). As this, however, requires an angular analysis, similar to $B_s^0 \to \phi\phi$, the analysis is significantly more complex than a lifetime measurement in the $B_{(s)}^0 \to h^+h'^-$ environment. After a nominal year of data taking a sensitivity to $\Delta\Gamma_s$ of 0.010 ps^{-1} is expected. Hence, despite the significantly smaller data set the $B_s^0 \to K^+K^-$ measurement will still make an important contribution to the determination of $\Delta\Gamma_s$.

ATLAS and *CMS* also have a B physics programme. This is based on channels involving muons due to the requirements of their triggers. Therefore, they will also perform measurements of $B_s^0 \to J/\psi\phi$ decays as the J/ψ decays into two muons with a large branching fraction of 6% [12]. *CMS* quote a sensitivity for $\Delta\Gamma_s$ of 0.009 ps^{-1}[1] for one nominal year of data taking [79]. As opposed to *LHCb*, both *ATLAS* and *CMS* assume a luminosity of 10^{33} cm^{-2} s^{-1}, i.e. a factor of 5 higher than *LHCb*, during their first three years of data taking. This is still a factor of 10 below their maximum design luminosity and is therefore the preferred period for B physics measurements.

CDF and *D0* have published measurements of $\Delta\Gamma_s$ using $B_s^0 \to J/\psi\phi$ decays. Their results are $\Delta\Gamma_s = 0.076^{+0.059}_{-0.063}(stat.) \pm 0.006(syst.)$ ps^{-1} [80] and $\Delta\Gamma_s = 0.13 \pm 0.09$ ps^{-1} [81], respectively. Extrapolating from the data sample used for these measurements, both experiments may reach a sensitivity of 0.03 ps^{-1} with their full data set assuming a collected integrated luminosity of 9 fb^{-1}. This underlines that measurements at the *LHC* will be able to quickly improve on the current precision.

[1] Statistical error only, extrapolated from 1.3 fb^{-1}.

5.1.4 The Selection of $B^0_{(s)} \to h^+h'^-$ Events

The selection of signal events follows the one developed for the analysis to extract the *CKM* angle γ with $B^0_{(s)} \to h^+h'^-$ decays [29]. Its basic principle is to provide a common selection for all $B^0_{(s)} \to h^+h'^-$ channels. Therefore, all events are reconstructed assuming the pion mass for the two selected charged particles. This leads to all $B^0_{(s)} \to h^+h'^-$ overlapping in their reconstructed invariant mass.

The selection has been optimised to have a high efficiency on signal events while suppressing the dominant sources of background. There are three classes of background to $B^0_{(s)} \to h^+h'^-$ events:

- Tracks from partially reconstructed decays of B mesons into three-body final states. These are decays such as $B^0 \to \rho^\pm\pi^\mp$, with $\rho^\pm \to \pi^\pm\pi^0$, where only the pair of charged pions is reconstructed. As all decays of this class have in common that one of their particles is not reconstructed, the invariant mass of the two tracks is bound to be below the nominal B mass.

- Decays of Λ^0_b baryons into two-body final states. The two final states in question are $\Lambda^0_b \to p^\pm K^\mp$ and $\Lambda^0_b \to p^\pm\pi^\mp$. The underestimation of the invariant mass due to the assignment of the pion mass to both daughters is compensated by the higher mass of the Λ^0_b compared to B^0_d and B^0_s. Both channels have their main distribution at a slightly higher mass than the $B^0_{(s)} \to h^+h'^-$ modes.

- The most important background class is combinatorial background. It results from combining particles that do not originate from the same mother particle. In order to fulfil the selection criteria for $B^0_{(s)} \to h^+h'^-$ as outlined below, these tracks must either be produced away from the primary vertex (decay of long-lived particles or conversion in material) or be mis-reconstructed such that they fake tracks with large impact parameters. Events in this category originate from purely random processes which lead to a continuous, falling spectrum in the reconstructed invariant mass.

The selection developed for the $B^0_{(s)} \to h^+h'^-$ channels is a cut based selection (see table 5.1). It does not require any particle identification of the daughter tracks in order to select all $B^0_{(s)} \to h^+h'^-$ modes. For the daughter tracks, minimal requirements for the impact parameter significance ($IP/\sigma(IP)$) as well as the transverse momentum (p_T) have to be met by at least one or both tracks. The B candidate is defined by the daughter Lorentz vectors. It has to satisfy requirements on the transverse momentum, the flight distance significance ($L^B/\sigma(L^B)$), the impact parameter significance and the

Cut type	Cut value
$min[(IP/\sigma(IP))^h, (IP/\sigma(IP))^{h'}]$	> 6
$max[(IP/\sigma(IP))^h, (IP/\sigma(IP))^{h'}]$	> 12
$min[p_T^h, p_T^{h'}]$ in GeV/c	> 1
$max[p_T^h, p_T^{h'}]$ in GeV/c	> 3
p_T^B in GeV/c	> 1
$L^B/\sigma(L^B)$	> 18
$(IP/\sigma(IP))^B$	< 2.5
$\chi^2(vertex)$	< 5
$m_{\pi\pi}$ in GeV/c^2	$[5.0, 5.8]$

Table 5.1: *Cuts for the selection of $B^0_{(s)} \to h^+h'^-$ events using the full reconstruction on events that have passed the trigger.*

χ^2 of the fit of the decay vertex. Finally, a mass window is defined between 5.0 GeV/c^2 and 5.8 GeV/c^2.

This selection has been evaluated using MC simulation. For the $B^0_{(s)} \to h^+h'^-$ channels the selection efficiency of selected events with respect to all generated events (without geometrical constraints at generator level) is between 3.7% and 3.9%. For the two decay modes of Λ_b^0 mentioned above the efficiency is about 3.3%.

A similar selection has been developed to select events at trigger level. As the exact implementation of the software trigger has yet to be finalised the trigger selection is not discussed in detail. However, the effect of the trigger implementation on the overall efficiency of selecting events should be small. Hence, the total selection efficiency that has been obtained with the preliminary implementation of the trigger will be used for estimating yields. For $B^0_{(s)} \to h^+h'^-$ decays the overall selection efficiency (ϵ_{tot}) is 1.4% and it is 1.2% for $\Lambda_b^0 \to p^\pm h^\mp$ decays.

The annual yield for each channel is calculated according to

$$Y = L \cdot \sigma_{b\bar{b}} \cdot f_{hadr} \cdot 2 \cdot \mathcal{BR} \cdot \epsilon_{tot}, \tag{5.4}$$

where L is the annual integrated luminosity of 2 fb^{-1}, $\sigma_{b\bar{b}}$ is the $b\bar{b}$ production cross section which is assumed to be 0.5 mb, and the hadronisation fraction f_{hadr} and the branching ratios $\mathcal{BR}$ are given in table 5.2. Most relevant for this thesis is the estimate of roughly $72k$ $B_s^0 \to K^+K^-$ events that will be selected for analysis per nominal year.

The amount of background has been estimated by the analysis of a sample of MC simulation that contains all known processes in their relative fractions (minimum bias

Channel	f_{hadr}	$\mathcal{BR} \times 10^6$	Annual yield
$B_d^0 \to \pi^+\pi^-$	0.403	5.16	$58.8k$
$B_d^0 \to K^+\pi^-$	0.403	19.4	$216.6k$
$B_s^0 \to K^+K^-$	0.101	25.8	$71.9k$
$B_s^0 \to \pi^+K^-$	0.101	5.27	$15.1k$
$\Lambda_b^0 \to p^\pm K^\mp$	0.092	5.0	$10.9k$
$\Lambda_b^0 \to p^\pm \pi^\mp$	0.092	3.1	$7.0k$

Table 5.2: *Annual yield for $B^0_{(s)} \to h^+h'^-$ channels and $\Lambda_b^0 \to p^\pm h^\mp$ channels assuming an annual integrated luminosity of* 2 fb^{-1}.

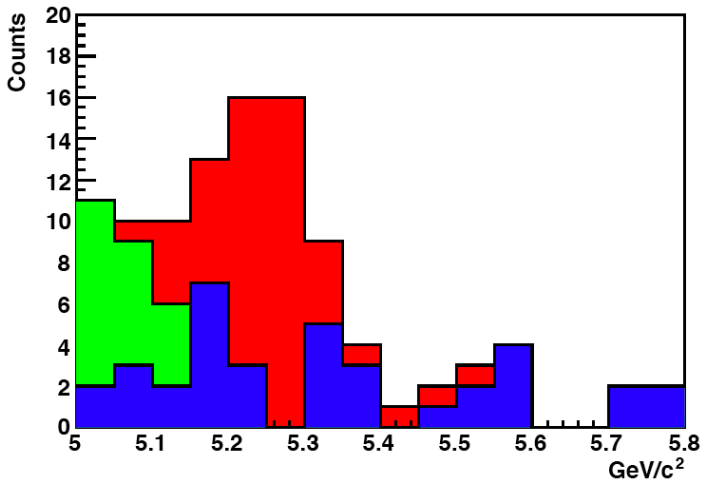


Figure 5.4: *Distribution of events selected by the $B^0_{(s)} \to h^+h'^-$ selection applied to minimum bias MC. Signal events (B and Λ_b^0) are shown in red, partially reconstructed three-body decays are shown in green, and combinatorial background is shown in blue.*

simulation). Due to the low branching ratios of the processes studied here the number of selected events is very low and the uncertainties on the relative fractions are high. Figure 5.4 shows the distribution of selected events.

For the tests of the lifetime fitter it has been assumed that the relative fraction of events with partially reconstructed three-body decays is 14% and that one third of all events originate from combinatorial background. These numbers are taken for the full mass window from 5.0 GeV/c^2 to 5.8 GeV^2 and are in agreement with the intervals deduced from minimum bias MC simulation.

5.1.5 Lifetime Fitting with $B^0_{(s)} \to h^+h'^-$ Events

The lifetime fitting procedure is split in three parts:

- The data sample to be used in the fitting process is obtained by the $B^0_{(s)} \to h^+h'^-$ selection described above.
- The event-by-event lifetime acceptance functions are determined for all selected $B^0_{(s)} \to h^+h'^-$ candidates by the method described in the previous chapter.
- The lifetime fit is performed using the event-by-event acceptance function as well as other observables as input.

The complete set of input variables to the lifetime fit is given by the measured lifetime and the event-by-event acceptance functions for the actual lifetime fit, and by the invariant mass using a di-pion hypothesis together with the *PID* values for the two daughter tracks for the fit of the signal fractions.

The lifetime is not a genuine observable of *LHCb*. It is defined as $t = d \cdot m/p$ where d is the measured distance between the primary and the decay vertex, m is the mass of the decaying particle and p is its momentum. In the selection described above the events are reconstructed assuming they were $B^0_d \to \pi^+\pi^-$ decays. This implies that the B^0_d mass is used to calculate the measured lifetime. For decays of other particles than B^0_d this leads to a biased determination of the lifetime. The observable $\xi \equiv t/m = d/p$ can be introduced to avoid this dependence on the particle mass. Replacing t by ξ the fit can be performed as described in the previous chapter with the fit parameters turning into the ratios of lifetime and mass. As this approach only reflects a scaling of the observables by the B^0_d mass it has not been considered for the toy studies presented here. In fact, it has been assumed that all individual lifetimes can be measured directly which means that the scaling $m(H^0_b)/m(B^0_d)$ for decays of H^0_b ($H^0_b = B^0_d, B^0_s, \Lambda^0_b$) has

been applied implicitly. As all masses involved in this study are very similar, the scaling is very small, and its effect is neglected in the sensitivity studies presented here.

The lifetime fitting technique described in the previous chapter has been extensively tested with simulated events. To the extent that it was possible the method for obtaining the event-by-event lifetime acceptance functions has been tested with the full *LHCb* MC data. This resulted in the average lifetime acceptance function shown in figure 4.5, which confirmed good agreement was obtained between the measured function and the true distribution. A signal only fit has been performed using these data yielding an unbiased result.

The actual fitter has been tested far more extensively with toy MC simulations as already described in section 5.2. Equation 5.3 states that, to a good approximation, the relative error on the lifetime measured with $B_s^0 \to K^+K^-$ determines the absolute error on the observable $\tilde{y}$. Therefore, studying the stability of the fits of the $B_s^0 \to K^+K^-$ lifetime is of highest interest. A detailed account of these checks is given in the following section.

5.2 Test Results from Toy Monte Carlo Studies

Toy simulation studies are indispensable in the development of a fitter. They provide an easy way of testing a fitter for particular features. Toy simulations are designed to generate effective distributions of the observables used in the fit. Hence, they do not need to go through the time consuming cycle of simulating and reconstructing events in a particular detector.

A major advantage of the fast generation of toy simulations is that one can rule out statistical fluctuations in the fit results. Therefore, usually at least several hundred equivalent data sets are generated and fitted in one test. Most tests used in the development of this fitter used 1000 toy experiments with of the order of 10^4 events each.

In the toy generator developed for this fitter the lifetimes are simulated as exponentials with a Gaussian resolution function. The acceptance function is applied by simulating a Gaussian distribution for the turn-on points of the per-event acceptance functions. Mass and *PID* are simulated according to the templates described in the previous section.

5.2.1 Toy Monte Carlo Tests for $B^0_{(s)} \to h^+h'^-$ Events

The fit to $B^0_{(s)} \to h^+h'^-$ events is the most challenging as it involves a significant number of specific channels with similar branching ratios. For the evaluation of the quality of the fitter, seven exclusive signal channels have been simulated together with combinatorial background. Four of the seven channels are the four $B^0_{(s)} \to h^+h'^-$ channels, i.e. $B^0_d \to K^+\pi^-$, $B^0_d \to \pi^+\pi^-$, $B^0_s \to K^+K^-$, and $B^0_s \to \pi^+K^-$. In addition, the decays $\Lambda^0_b \to p^\pm K^\mp$ and $\Lambda^0_b \to p^\pm\pi^\mp$ are reconstructed in the upper mass side band. The seventh channel is $B^0_d \to \pi^+\pi^-\pi^0$, which has been chosen to represent B^0_d and B^0_s decays with three particles in the final state where one has not been reconstructed. All individual channels, the three-body background, and the combinatorial background are simulated at their expected levels as discussed in detail in section 5.1.4. The combinatorial background is simulated to be flat in the invariant mass. This is in reasonable agreement with the available simulation of combinatorial background in the B mass region.

The fit is split into two parts as explained in the previous chapter. First, the signal fractions are determined within limits of 0 and 1 from a fit to the vector $\mathbf{X}$ of distinguishing variables, i.e. mass and $\delta \log \mathcal{L}_K$. Thereafter, these fractions are fed into the lifetime fitter. It has been tested that these limits have no impact on a standard fit, which is confirmed by the width of the pull distributions. The minimisation is done by the `Minuit` minimiser `Migrad`.

Figure 5.5 shows the mass distribution with the fitted components of a typical toy experiment. The lower plot shows the significance of the difference between the fitted templates and the distribution of generated data. A χ^2 defined as the sum of the squares of the entries of the lower plot yields a value of 103 for 100 non-zero entries. The agreement confirms the quality of the kernel method presented above.

The results of the signal fraction fit are shown in figures 5.6, 5.7, and 5.8 and summarised in table 5.3. All signal fractions are unbiased. The pull widths are all in agreement with unity. Hence, the results of this fit provide the necessary input for the lifetime fitter.

The fitted lifetime distribution is shown in figure 5.9. The two sources of non-zero differences between the input data and the fit models are fluctuations in data that are smoothed out by the kernel method and the mechanism to prevent the background distribution from acquiring values close to or below zero (characterised by the fitter parameter ϵ, see section 4.3.2.3). A χ^2 defined as above for the mass plot yields a value of 145 for 143 non-zero entries.

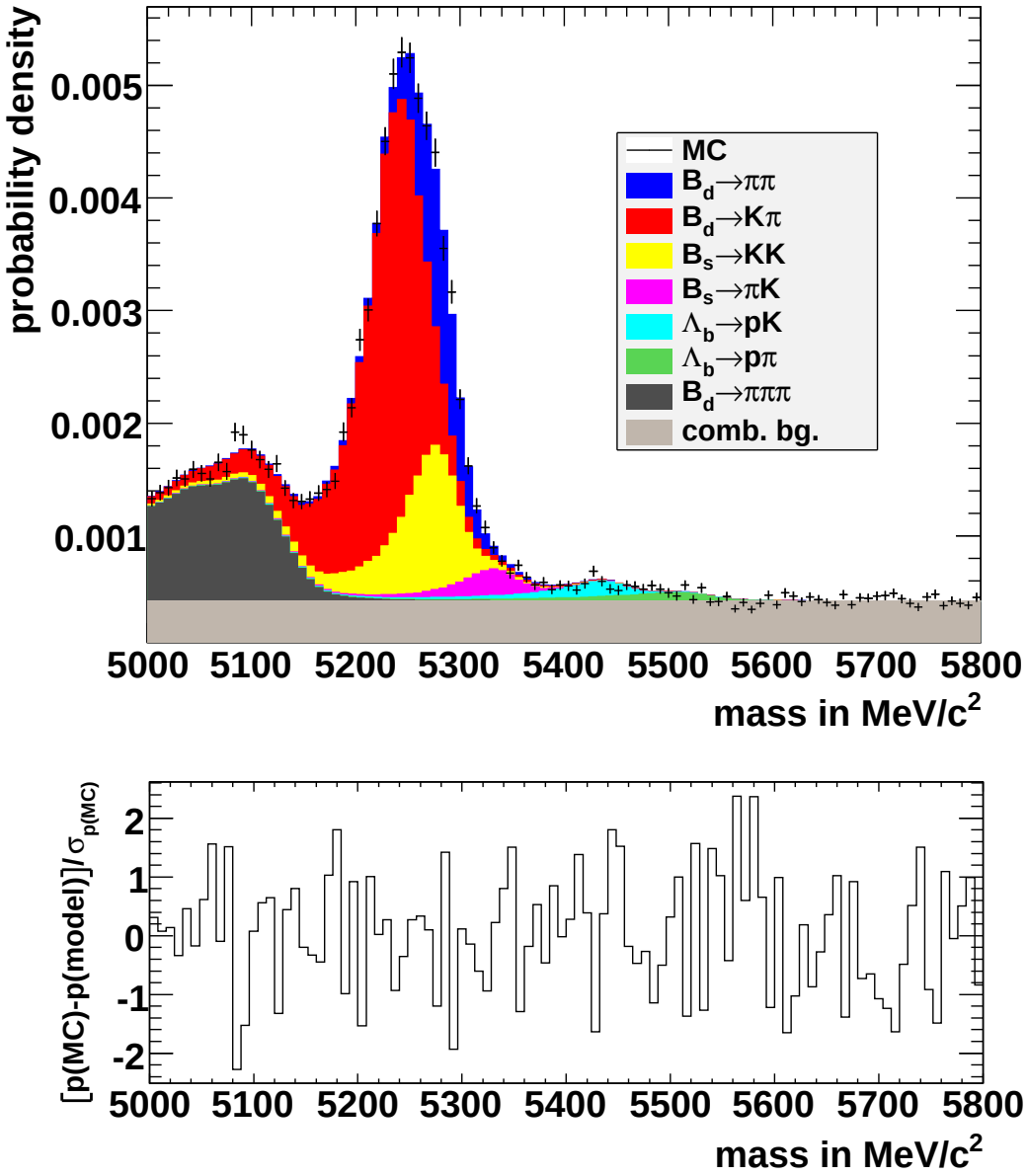


Figure 5.5: *Result of a $B^0_{(s)} \to h^+h'^-$ signal fraction fit. The upper plot shows the generated distribution as crosses (MC) and the cumulative distributions of the model using the fitted signal fractions as shaded areas. The lower plot shows the difference per bin divided by the statistical error of the data sample.*

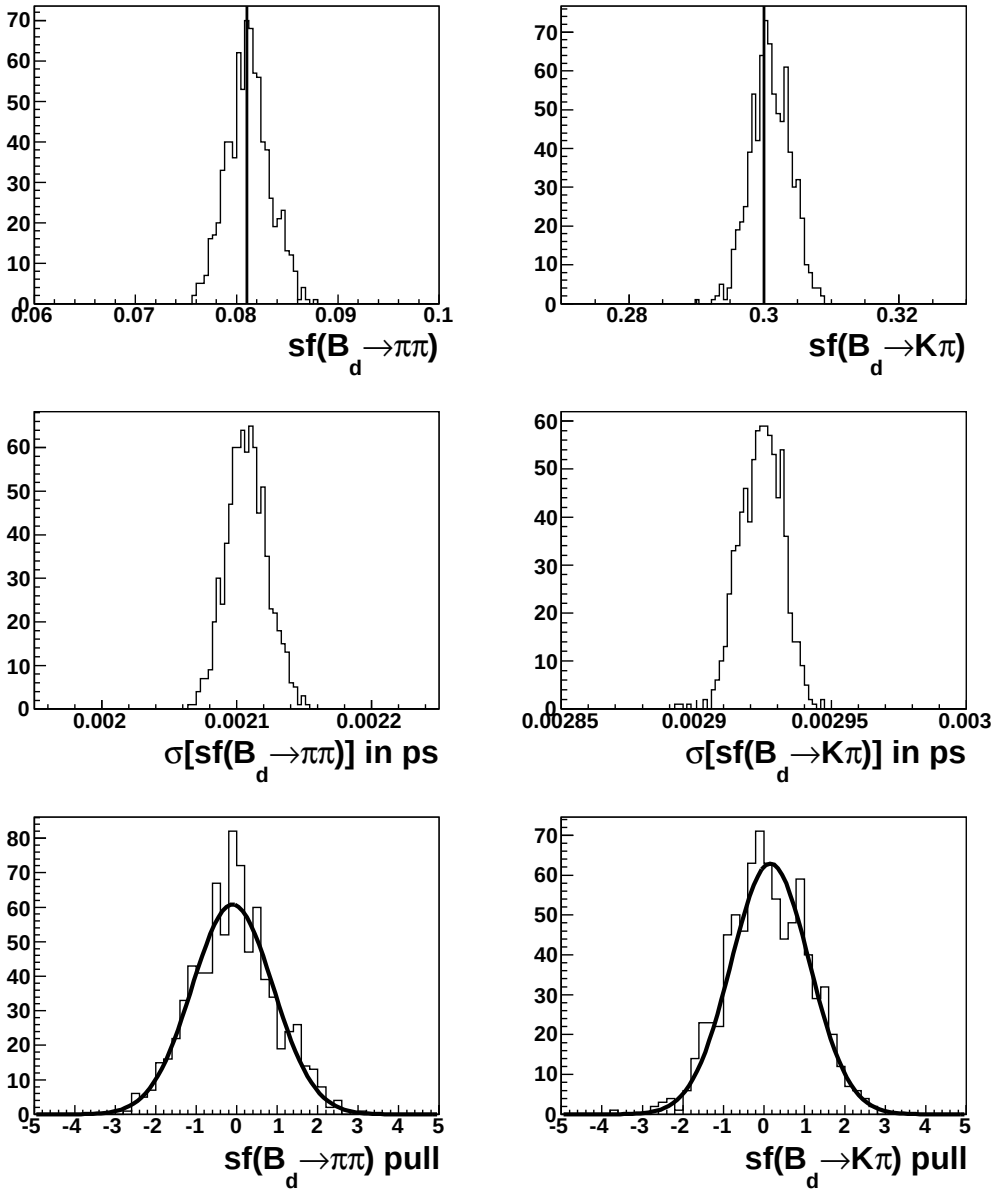


Figure 5.6: *Toy study results for $sf(B_d^0 \to \pi^+\pi^-)$ (left) and $sf(B_d^0 \to K^+\pi^-)$ (right). The plots show from top to bottom the fit results where the line marks the average true value, the error distributions, and the pull distributions. The results are summarised in table 5.3.*

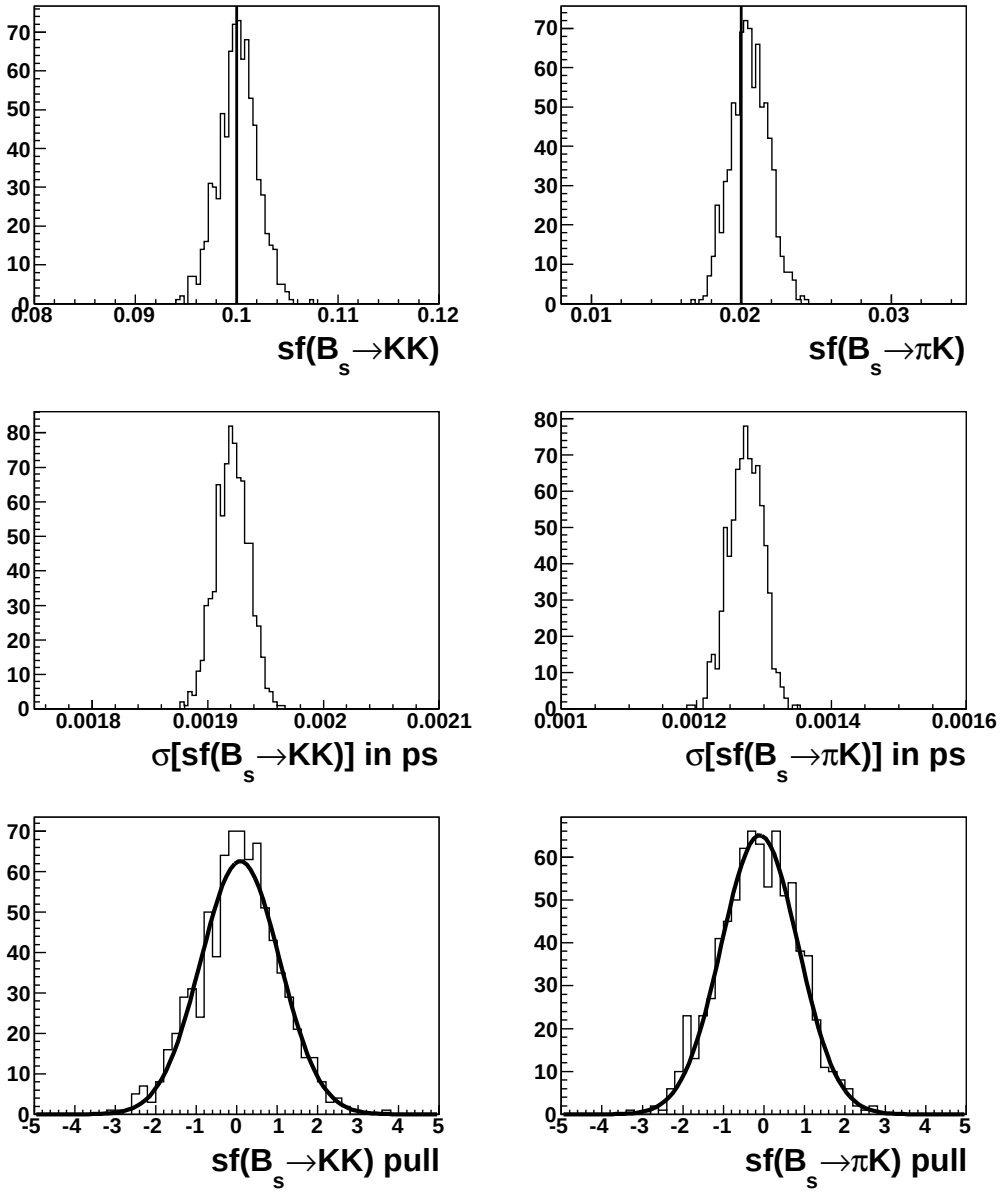


Figure 5.7: *Toy study results for $sf(B_s^0 \to K^+K^-)$ (left) and $sf(B_s^0 \to \pi^+K^-)$ (right). The plots show from top to bottom the fit results where the line marks the average true value, the error distributions, and the pull distributions. The results are summarised in table 5.3.*

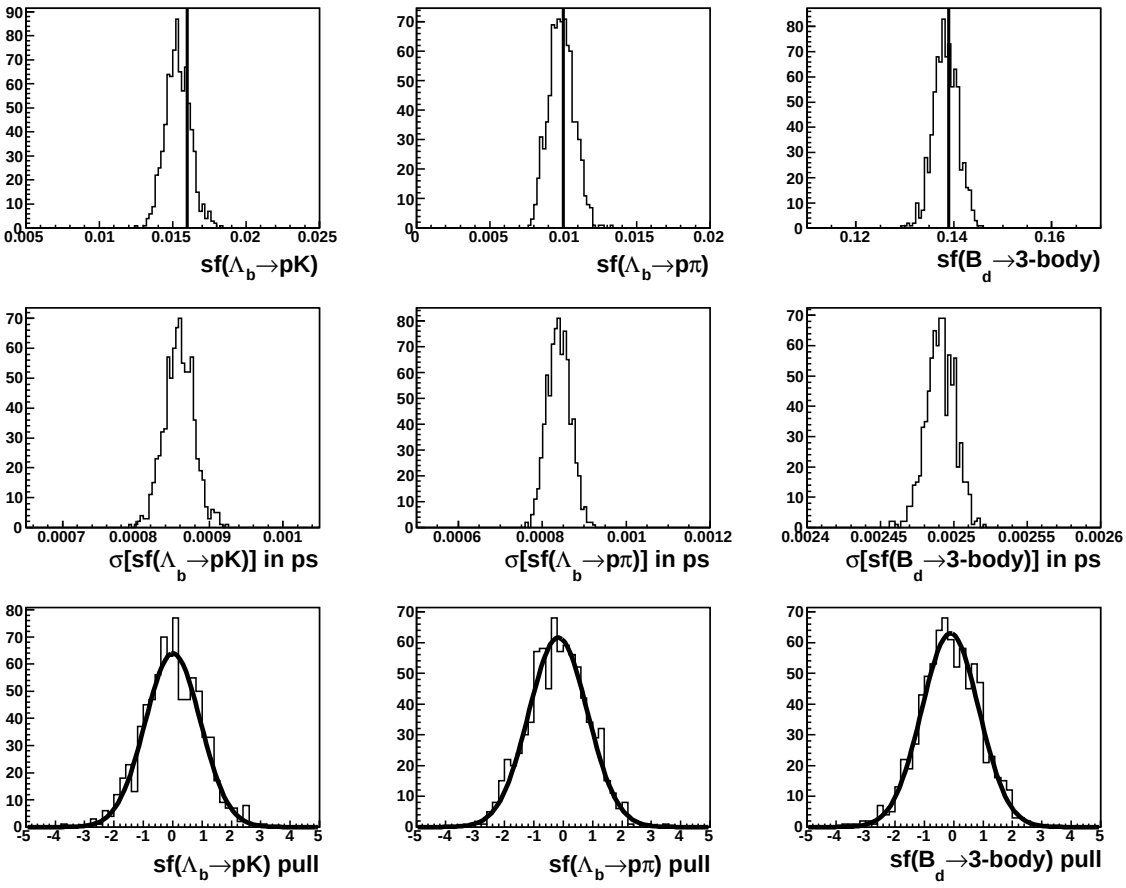


Figure 5.8: *Toy study results for $sf(\Lambda_b^0 \to p^{\pm}K^{\mp})$ (left), $sf(\Lambda_b^0 \to p^{\pm}\pi^{\mp})$ (centre), and $sf(B \to 3-body)$ (right). The plots show from top to bottom the fit results where the line marks the average true value, the error distributions, and the pull distributions. The results are summarised in table 5.3.*

quantity	true value	avg. result	avg. error	pull mean	pull width
$sf(B_d^0 \to \pi^+\pi^-)$	0.081	0.081	0.002	-0.07 ± 0.03	1.01 ± 0.03
$sf(B_d^0 \to K^+\pi^-)$	0.300	0.301	0.003	0.10 ± 0.03	0.99 ± 0.02
$sf(B_s^0 \to K^+K^-)$	0.100	0.100	0.002	0.07 ± 0.03	0.97 ± 0.02
$sf(B_s^0 \to \pi^+K^-)$	0.020	0.021	0.001	-0.10 ± 0.03	0.93 ± 0.02
$sf(\Lambda_b^0 \to p^{\pm}K^{\mp})$	0.016	0.015	0.001	-0.00 ± 0.03	0.97 ± 0.02
$sf(\Lambda_b^0 \to p^{\pm}\pi^{\mp})$	0.010	0.010	0.001	-0.13 ± 0.03	1.01 ± 0.03
$sf(B \to 3-body)$	0.139	0.138	0.002	-0.11 ± 0.03	0.98 ± 0.02

Table 5.3: *Toy study results for the fit of the signal fractions of the seven exclusive channels used in the $B_{(s)}^0 \to h^+h'^-$ fit.*

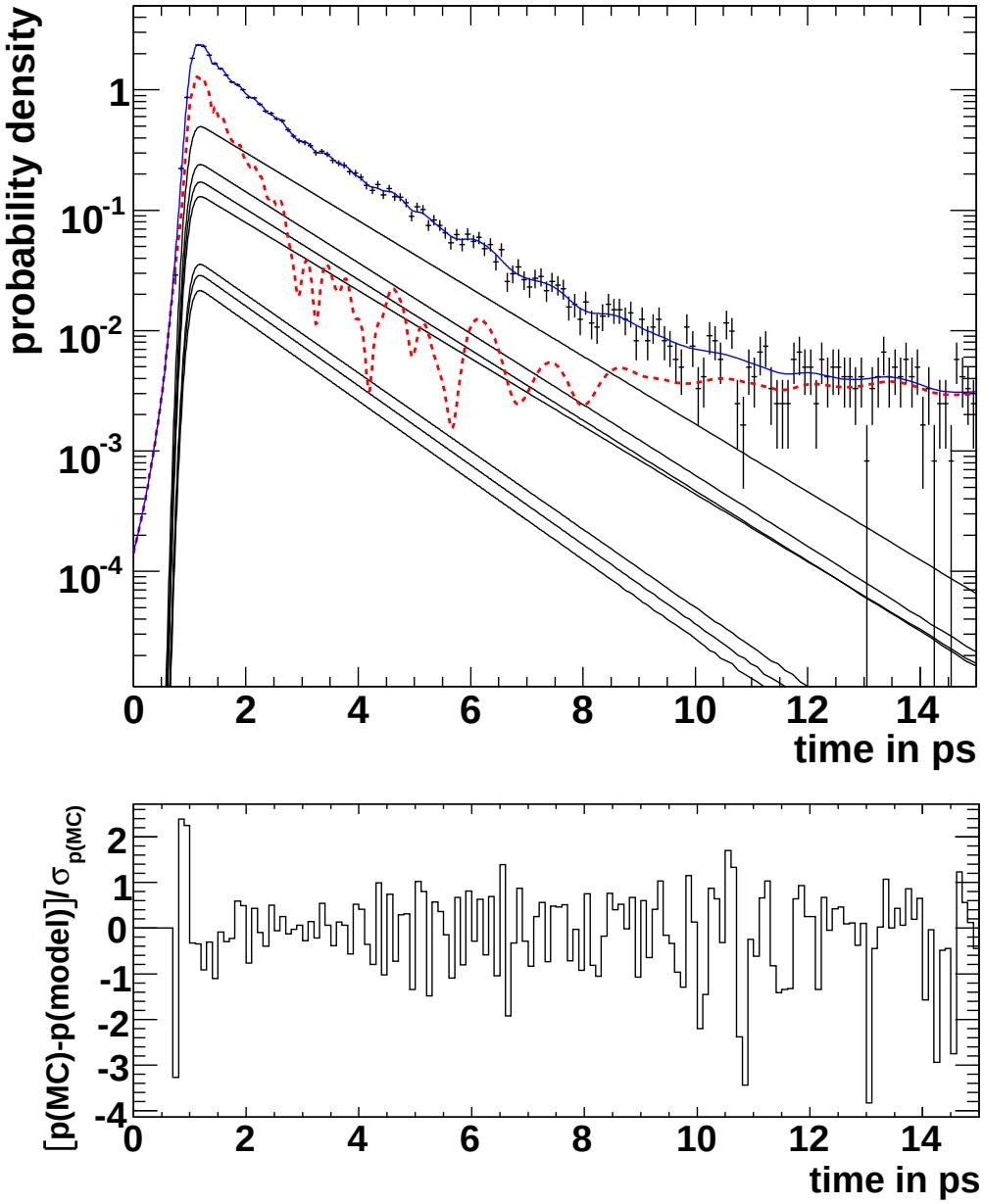


Figure 5.9: *Result of a $B^0_{(s)} \to h^+h'^-$ lifetime fit. The upper plot shows the generated distribution as crosses (MC), the distributions of the signal models using the fitted lifetimes as black lines, the background distribution as a red (dashed) line, and the total distribution according to the fit results as a blue line. The lower plot shows the difference per bin between generated data and the total distribution of the model divided by the statistical error of the data sample.*

To evaluate the quality of the lifetime fit a set of toy experiments has been generated that represents the data expected for about 0.10 fb^{-1}. This amount is 1/20 of a nominal year of data taking, and is of the order of the amount of data expected up to the end of 2010. 1000 toy experiments have been generated with 36000 events each. 12000 events were simulated as combinatorial background and 24000 were shared between the four $B^0_{(s)} \to h^+h'^-$ channels, the two Λ^0_b decays, and the three-body decays in their nominal proportions.

The sample size equivalent to the luminosity of 0.10 fb^{-1} is valid for all studies described in the following unless otherwise stated. This sample size has been chosen as it reflects roughly the available data set expected by the end of 2010 and as these event numbers present a size for which large numbers of example fits can be performed on a reasonable time scale. The CPU time consumption of the fitter has been confirmed to be linear with the number of events in the input sample, as expected by the layout of the algorithm. Data sets of different sizes ranging from those equivalent to the luminosities of 0.01 fb^{-1} to 2.00 fb^{-1} have been tested as well. Their results are documented in section 5.2.2.2.

Five fit parameters are used in the lifetime fit (see table 5.4). Both channels $B^0_d \to \pi^+\pi^-$ and $B^0_d \to K^+\pi^-$ are used to measure the B^0_d lifetime. The B^0_s lifetime is measured using $B^0_s \to \pi^+K^-$ decays only, where the theoretically more interesting lifetime ratio $\tau(B^0_s)/\tau(B^0_d)$ is left as a free parameter (see section 1.3.1). The third parameter is the lifetime measured from $B^0_s \to K^+K^-$ which is expected to deviate from the $\mathcal{CP}$ averaged B^0_s lifetime as outlined in section 1.4.1 and discussed further in section 5.1. The Λ^0_b lifetime is measured using the channels $\Lambda^0_b \to p^\pm K^\mp$ and $\Lambda^0_b \to p^\pm\pi^\mp$ where again the lifetime ratio with respect to the B^0_d lifetime $(\tau(\Lambda^0_b)/\tau(B^0_d))$ is left as a free parameter. Similar to the lifetime ratio $\tau(B^0_s)/\tau(B^0_d)$, more precise theoretical predictions exist for the lifetime ratio $\tau(\Lambda^0_b)/\tau(B^0_d)$ than for the Λ^0_b lifetime itself. Finally, the three-body modes are fitted using a single lifetime. This is an approximation as their contribution consists of decays of both B^0_d and B^0_s mesons, however, this approach is sufficient for dealing with this well separated background channel.

The results of the lifetime fit are shown in figures 5.10 and 5.11 and summarised in table 5.4. The lifetime fits of the variables of physical interest are unbiased. The errors are well estimated.

The lifetime for three-body decays is the only one for which the fit returns biased results. The mass distribution of this channel is the most similar to the combinatorial background and hence any effect related to the treatment of the background affects $\tau(B \to 3\text{–}body)$ more than other parameters. The lifetime measured in the three-body

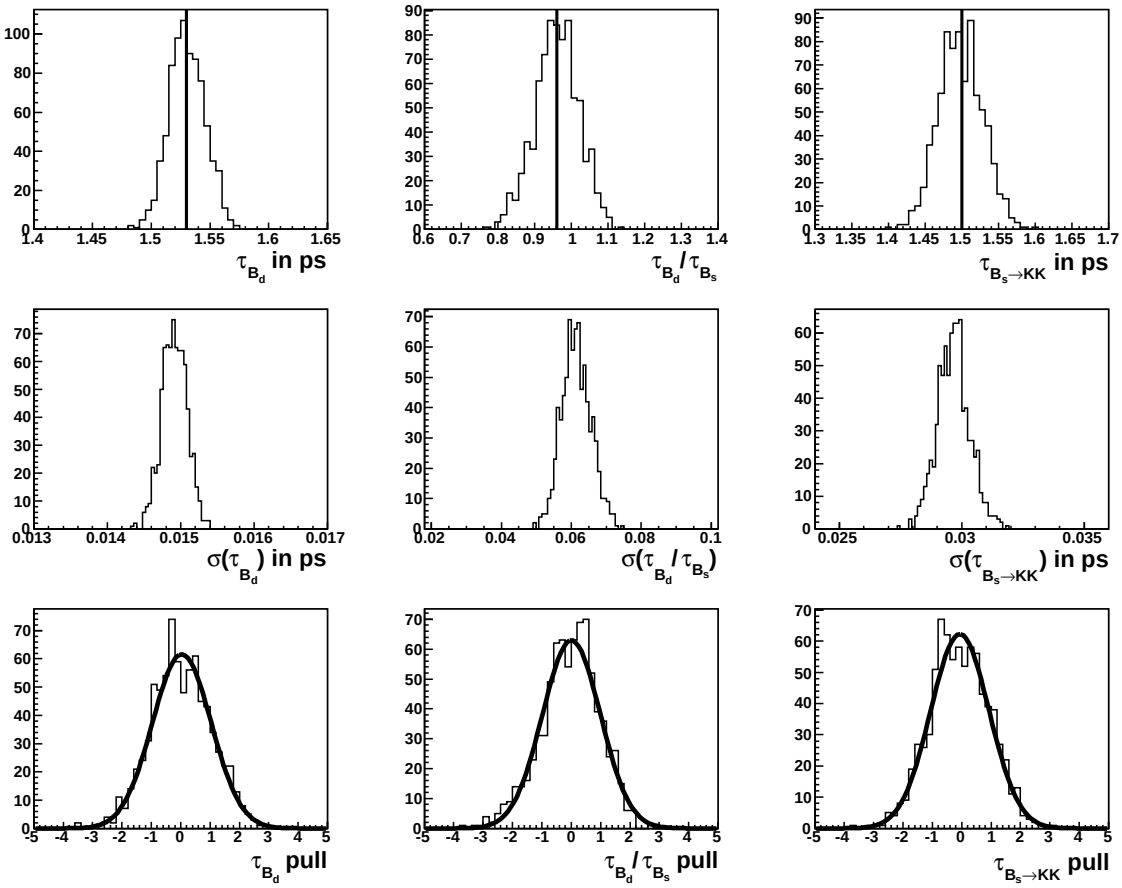


Figure 5.10: *Toy study results for $\tau(B_d^0)$ (left), $\tau(B_s^0)/\tau(B_d^0)$ (centre), and $\tau(B_s^0 \to K^+K^-)$ (right). The plots show from top to bottom the fit results where the line marks the average true value, the error distributions, and the pull distributions. The results are summarised in table 5.4.*

quantity	true value	avg. result	avg. error	pull mean	pull width
$\tau(B_d^0)$	1.530 ps	1.530 ps	0.015 ps	0.04 ± 0.04	1.00 ± 0.03
$\tau(B_s^0)/\tau(B_d^0)$	0.961	0.961	0.061	0.01 ± 0.04	0.99 ± 0.03
$\tau(B_s^0 \to K^+K^-)$	1.500 ps	1.498 ps	0.030 ps	-0.06 ± 0.04	0.99 ± 0.03
$\tau(\Lambda_b^0)/\tau(B_d^0)$	0.902	0.901	0.041	-0.07 ± 0.03	0.96 ± 0.03
$\tau(B \to 3-body)$	1.500 ps	1.507 ps	0.026 ps	0.26 ± 0.04	1.05 ± 0.03

Table 5.4: *Toy study results for the fit of the lifetimes of the seven exclusive channels used in the $B_{(s)}^0 \to h^+h'^-$ fit.*

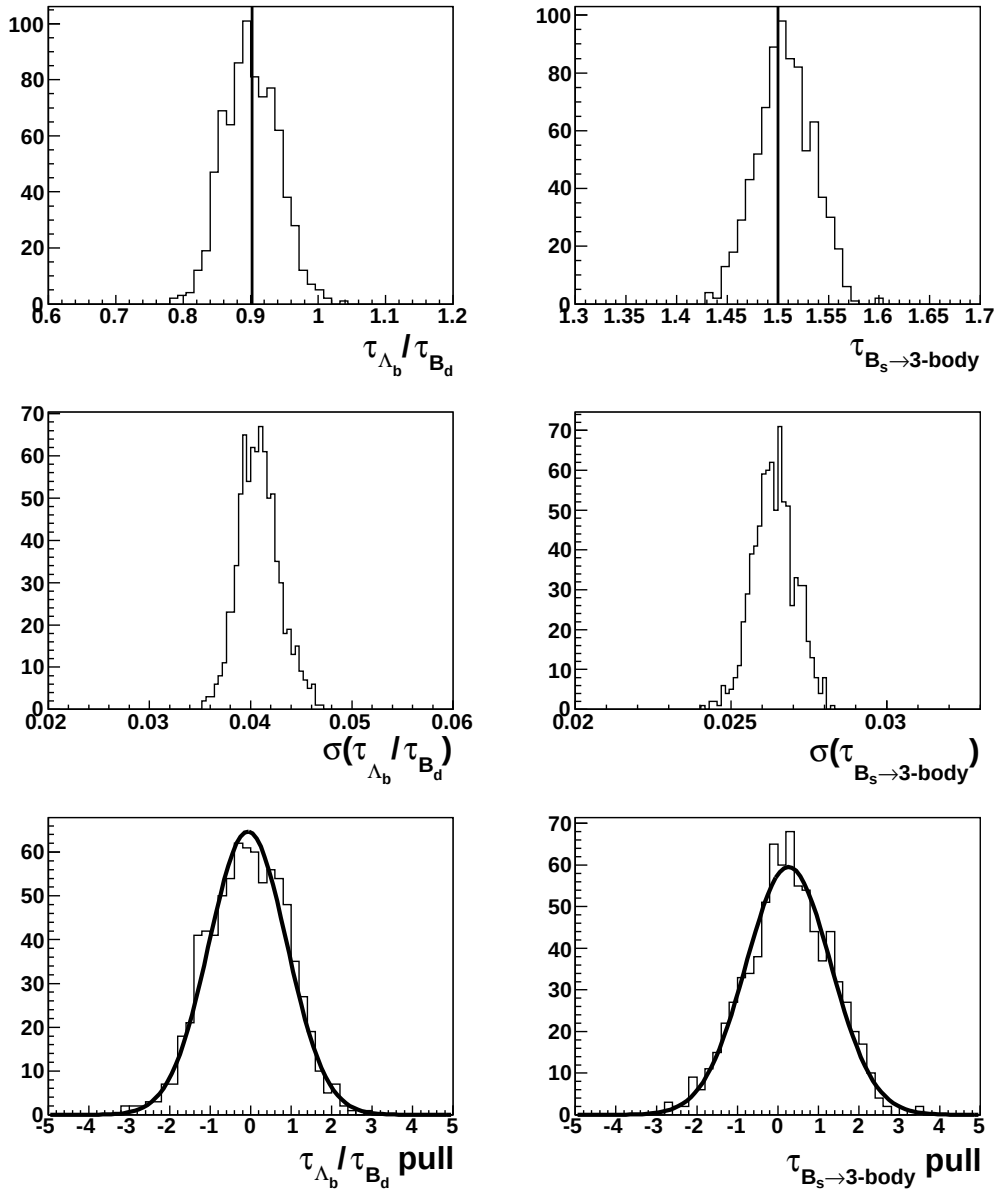


Figure 5.11: *Toy study results for $\tau(\Lambda_b^0)/\tau(B_d^0)$ (left) and $\tau(B \to 3-body)$ (right). The plots show from top to bottom the fit results where the line marks the average true value, the error distributions, and the pull distributions. The results are summarised in table 5.4.*

channels is not a physics parameter to be extracted from the fitter and hence only of secondary interest. Furthermore, it is possible to tighten the mass window on the lower side in order not to be affected by the contribution of three-body decays. As already described in section 4.3.2.3 this bias is connected to the correction limiting the background time PDF to positive values.

5.2.2 Stability of Lifetime Measurements with $B^0_{(s)} \to h^+h'^-$ Events

This section summarises tests that have been performed using toy MC data sets to evaluate the stability of the fitter. The procedure was simplified to allow more efficient testing. Where the input of the signal fraction fit is unchanged with respect to the default scenario, this stage has been omitted and the true average signal fractions have been used as input to the lifetime fit. Using the average true values implies that the values used on average deviate from the true fraction of the individual toy experiments by their usual statistical fluctuation. Similarly, the lifetimes of the Λ^0_b and three-body decays are fixed to their true average values.

To assess the impact of these simplifications a set of toy fits has been performed on the default data set using the simplified fitter. The average fit results are $\tau(B^0_d) = 1.530$ ps, $\tau(B^0_s)/\tau(B^0_d) = 0.961$, and $\tau(B^0_s \to K^+K^-) = 1.499$ ps. These are in excellent agreement with the results of the full fit given in table 5.4, thus confirming the validity of the simplifications made.

The studies presented in this section are made to evaluate the stability of the fitter with respect to controllable changes. The first test is the evaluation of the impact of various mass windows in the selection, followed by an evaluation of the scaling behaviour of the fitter with changes in the size of the data sample. Finally, the impact of a wrongly fixed signal fraction is tested. No systematic uncertainties are assigned for any of these checks. Systematic uncertainties, including those due to changes in the input templates that can lead to wrong signal fractions, are studied in section 5.2.3.

5.2.2.1 Variation of the Mass Window

The variation of the mass window is an important cross-check for the fitter. Varying the mass window allows to control the level of combinatorial background and how many events of various exclusive channels are included in the fit. If the fitter is stable with respect to a varied mass window this can be used to evaluate the size of potential correlations between the mass and lifetime distributions of combinatorial background

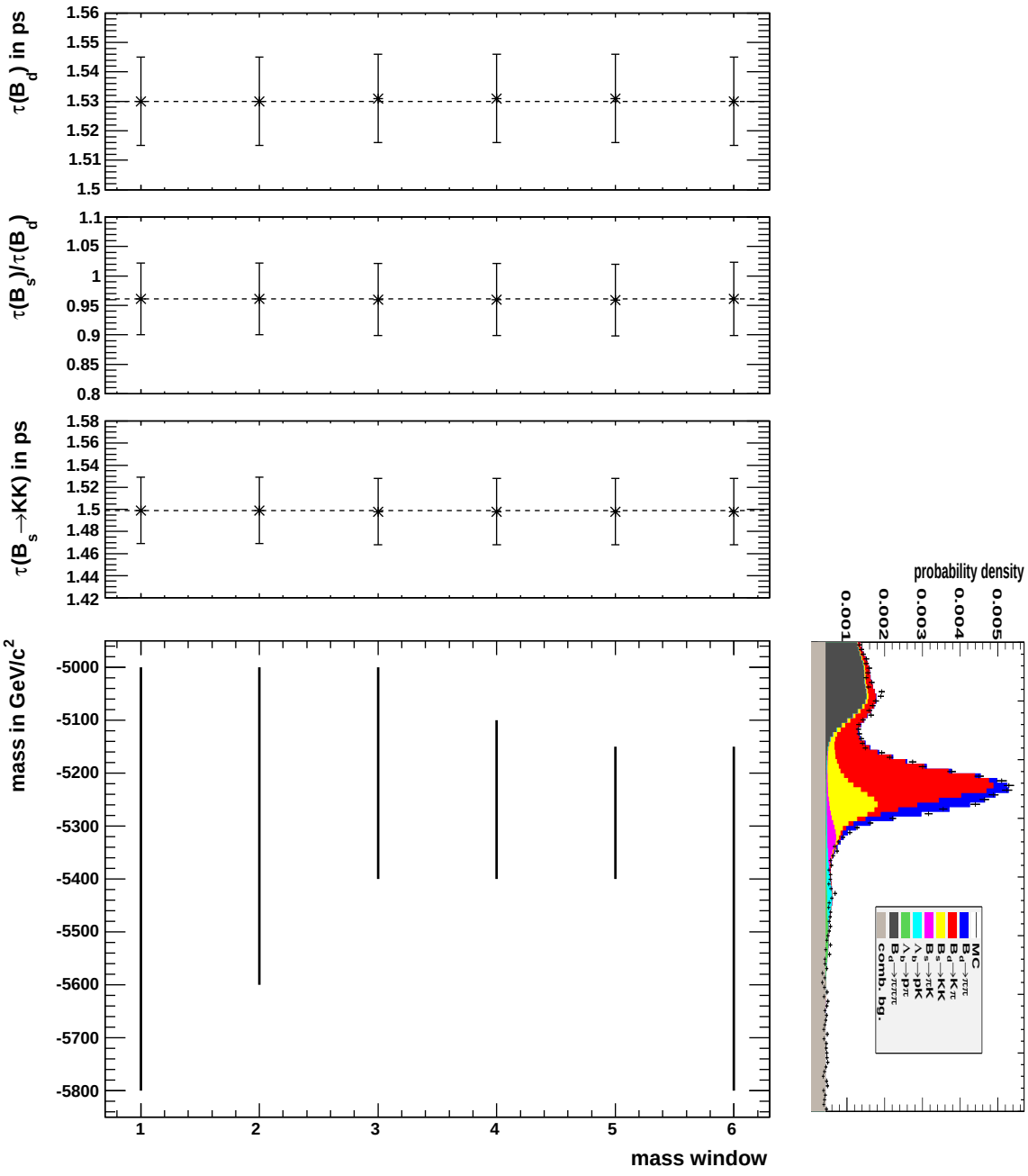


Figure 5.12: *Toy study results for the variation of the mass window (top three plots). Default is* $(5000-5800)$ GeV$/c^2$*. The bars in the bottom left plot indicate the mass windows and the bottom right shows the mass distribution of the various signal channels for reference. The dashed line indicates the average result of the default fit.*

(see section 4.3.2.4). No systematic error is assigned as this is a study of the capabilities of the fitter rather than an evaluation of any systematic effect.

The default mass window is between 5000 MeV/c^2 and 5800 MeV/c^2 in the invariant di-pion mass. Five additional mass windows have been evaluated with tighter cuts at the lower and/or upper end. Tightening the cuts at the lower end suppresses the contribution of background from partially reconstructed three-body decays. At the upper end, tighter cuts lead to a removal of combinatorial background as well as of the Λ_b^0 decays. The signal fraction fit has been performed for these fits as the correct signal fractions vary with the mass window.

Figure 5.12 shows the fit results for the various mass windows. The lower left plot indicates the allowed regions of the mass windows as bars and the mass distribution on the lower right shows how these cuts affect the individual signal channels. The fits show only very minor fluctuations which is confirmed by the numbers in tables 5.5 to 5.7. The fits were performed on the same data set, hence, the errors are fully correlated.

5.2.2.2 Variation of the Size of the Data Sample

The next test is the behaviour of the fit with the variation of the size of the data sample. This is of particular interest as it shows the stability of the fitter with very small data samples, i.e. smaller signal significances. At the same time it allows the study of any bias of the fitter with high statistics samples. The sample size has been varied between those equivalent to integrated luminosities of 0.01 fb^{-1} to 2 fb^{-1}.

Figure 5.13 and tables 5.8 to 5.10 summarize the results. Only the B_d^0 lifetime measurement maintains a significant bias. However, even this bias is smaller than the statistical uncertainty for 2 fb^{-1}, i.e. about 0.2%. In conclusion, there is no problem for a high precision lifetime measurement using this fitter.

Note that the lifetime ratio measurement remains totally unbiased for all sample sizes. With a statistical uncertainty of about 1.5%, even the $B_s^0 \to \pi^+K^-$ decay offers promising opportunities for precision tests of *HQE* and the *SM*.

The stability of the fitter with very small data samples allows its commissioning with the first physics data recorded by *LHCb*. The samples equivalent to a luminosity of 0.01 fb^{-1} comprise only 3600 events each.

5.2.2.3 Impact of Wrong Signal Fractions

The signal fraction fit has been fixed for most of the systematic checks discussed in this chapter. It is therefore important to study the impact that wrong fit results for the signal fractions have on the lifetime fit. The signal fraction for $B_s^0 \to K^+K^-$ events has

mass window [GeV/c^2]	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
$5000 - 5800$	1.530	1.530	0.015	-0.01	1.01
$5000 - 5600$	1.530	1.530	0.015	0.01	1.02
$5000 - 5400$	1.530	1.531	0.015	0.07	0.99
$5100 - 5400$	1.530	1.531	0.015	0.04	1.01
$5150 - 5400$	1.530	1.531	0.015	0.08	0.98
$5150 - 5800$	1.530	1.530	0.015	0.05	0.96

Table 5.5: *Toy study results for $\tau(B_d^0)$ for the variation of the mass window. Default is* $(5000 - 5800)$ GeV/c^2.

mass window [GeV/c^2]	true value	avg. result	avg. error	pull mean	pull width
$5000 - 5800$	0.961	0.961	0.061	0.00	0.98
$5000 - 5600$	0.961	0.961	0.061	0.00	0.98
$5000 - 5400$	0.961	0.960	0.061	-0.03	0.94
$5100 - 5400$	0.961	0.960	0.061	-0.01	0.97
$5150 - 5400$	0.961	0.959	0.061	-0.05	0.94
$5150 - 5800$	0.961	0.961	0.062	0.00	0.98

Table 5.6: *Toy study results for $\tau(B_s^0)/\tau(B_d^0)$ for the variation of the mass window. Default is* $(5000 - 5800)$ GeV/c^2.

mass window [GeV/c^2]	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
$5000 - 5800$	1.500	1.499	0.030	-0.05	0.99
$5000 - 5600$	1.500	1.499	0.030	-0.04	0.99
$5000 - 5400$	1.500	1.498	0.030	-0.06	1.00
$5100 - 5400$	1.500	1.498	0.030	-0.08	0.99
$5150 - 5400$	1.500	1.498	0.030	-0.08	1.01
$5150 - 5800$	1.500	1.498	0.030	-0.08	1.00

Table 5.7: *Toy study results for $\tau(B_s^0 \to K^+K^-)$ for the variation of the mass window. Default is* $(5000 - 5800)$ GeV/c^2.

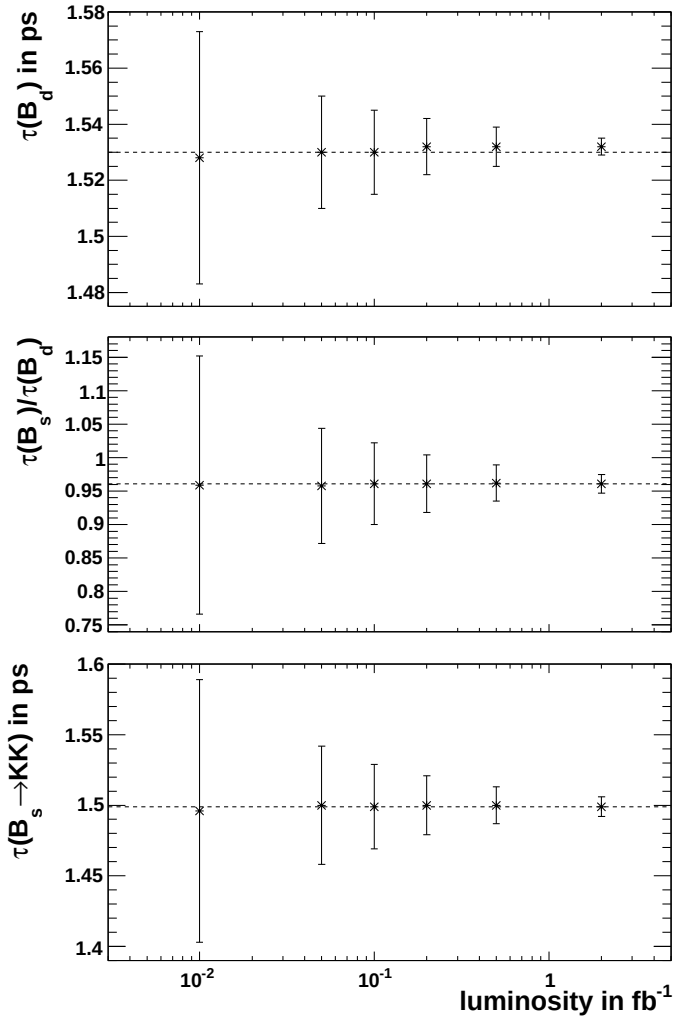


Figure 5.13: *Toy study results for fits with differently sized data samples. Default is* 0.10 fb^{-1}. *The dashed line indicates the average result of the default fit.*

luminosity [fb^{-1}]	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
0.01	1.530	1.528	0.045	−0.06	1.05
0.05	1.530	1.530	0.020	−0.03	1.00
0.10	1.530	1.530	0.015	0.01	1.01
0.20	1.530	1.532	0.010	0.19	0.98
0.50	1.530	1.532	0.007	0.29	0.99
2.00	1.530	1.532	0.003	0.62	0.99

Table 5.8: *Toy study results for* $\tau(B_d^0)$ *for fits with differently sized data samples. Default is* 0.10 fb^{-1}.

luminosity [fb^{-1}]	true value	avg. result	avg. error	pull mean	pull width
0.01	0.961	0.959	0.193	−0.11	0.93
0.05	0.961	0.958	0.086	−0.08	0.96
0.10	0.961	0.961	0.061	0.00	0.98
0.20	0.961	0.961	0.043	−0.02	0.98
0.50	0.961	0.962	0.027	0.02	0.98
2.00	0.961	0.961	0.014	0.03	0.96

Table 5.9: *Toy study results for* $\tau(B_s^0)/\tau(B_d^0)$ *for fits with differently sized data samples. Default is* 0.10 fb^{-1}.

luminosity [fb^{-1}]	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
0.01	1.500	1.496	0.093	−0.08	0.99
0.05	1.500	1.500	0.042	0.03	0.94
0.10	1.500	1.499	0.030	−0.06	1.00
0.20	1.500	1.500	0.021	−0.01	1.07
0.50	1.500	1.500	0.013	−0.04	1.03
2.00	1.500	1.499	0.007	−0.10	0.97

Table 5.10: *Toy study results for* $\tau(B_s^0 \to K^+K^-)$ *for fits with differently sized data samples. Default is* 0.10 fb^{-1}.

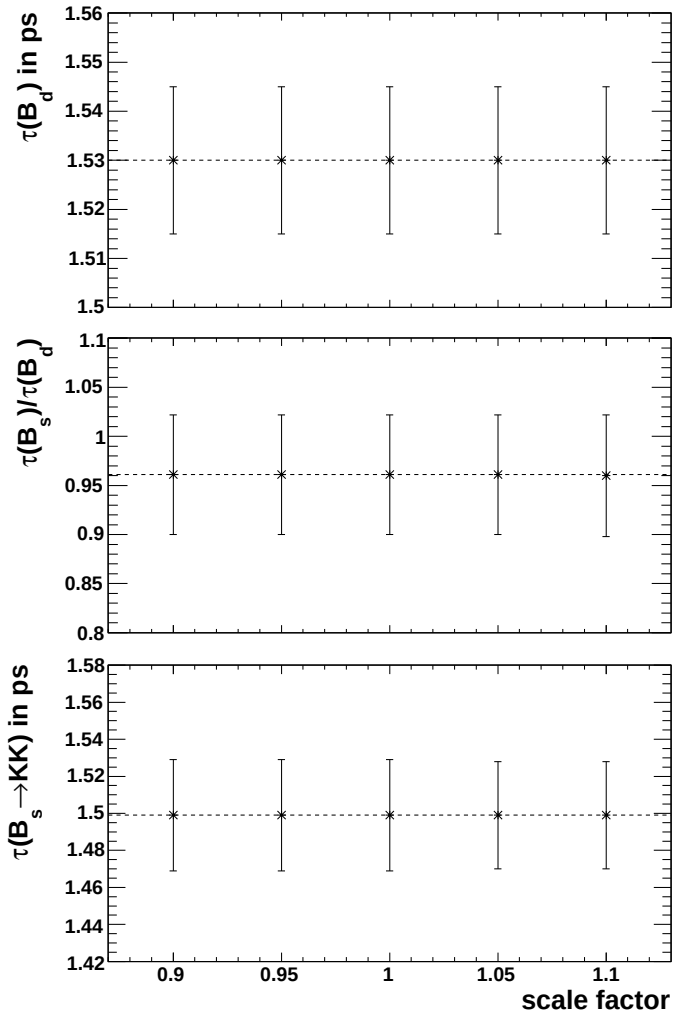


Figure 5.14: *Toy study results for fits with a wrong signal fraction for* $B_s^0 \to K^+K^-$. *Default is* 1.00. *The dashed line indicates the average result of the default fit.*

been varied by a scale factor between 0.9 and 1.1. This has been compensated by all other channels while keeping their relative fractions. No systematic error is assigned as this is merely a cross-check of the simplified fitting approach. Systematic effects leading to wrong signal fractions can originate from wrong templates for the variables used in the signal fraction fits. These will be studied in sections 5.2.3.1 and 5.2.3.2.

The results are summarised in figure 5.14 and tables 5.11 to 5.13. The fits were performed on the same data set, hence, the errors are fully correlated. All fit parameters show excellent stability which confirm that statistical fluctuations in the signal fractions have no influence on the lifetime fit. This also underlines the validity of the approach of fixing the signal fractions to their true average values in all toy fits that share the true signal fractions with the standard sample.

$sf(B_s^0 \to K^+K^-)$	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
0.90	1.530	1.530	0.015	0.00	1.01
0.95	1.530	1.530	0.015	0.00	1.01
1.00	1.530	1.530	0.015	0.01	1.01
1.05	1.530	1.530	0.015	0.01	1.01
1.10	1.530	1.530	0.015	0.00	1.00

Table 5.11: *Toy study results for $\tau(B_d^0)$ for fits with a wrong signal fraction for $B_s^0 \to K^+K^-$. Default is* 1.00.

$sf(B_s^0 \to K^+K^-)$	true value	avg. result	avg. error	pull mean	pull width
0.90	0.961	0.961	0.061	0.00	0.98
0.95	0.961	0.961	0.061	0.00	0.98
1.00	0.961	0.961	0.061	0.00	0.98
1.05	0.961	0.961	0.061	0.00	0.98
1.10	0.961	0.960	0.062	−0.01	0.99

Table 5.12: *Toy study results for $\tau(B_s^0)/\tau(B_d^0)$ for fits with a wrong signal fraction for $B_s^0 \to K^+K^-$. Default is* 1.00.

$sf(B_s^0 \to K^+K^-)$	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
0.90	1.500	1.499	0.030	−0.04	1.00
0.95	1.500	1.499	0.030	−0.05	1.00
1.00	1.500	1.499	0.030	−0.06	1.00
1.05	1.500	1.499	0.029	−0.06	1.00
1.10	1.500	1.499	0.029	−0.06	1.00

Table 5.13: *Toy study results for $\tau(B_s^0 \to K^+K^-)$ for fits with a wrong signal fraction for $B_s^0 \to K^+K^-$. Default is* 1.00.

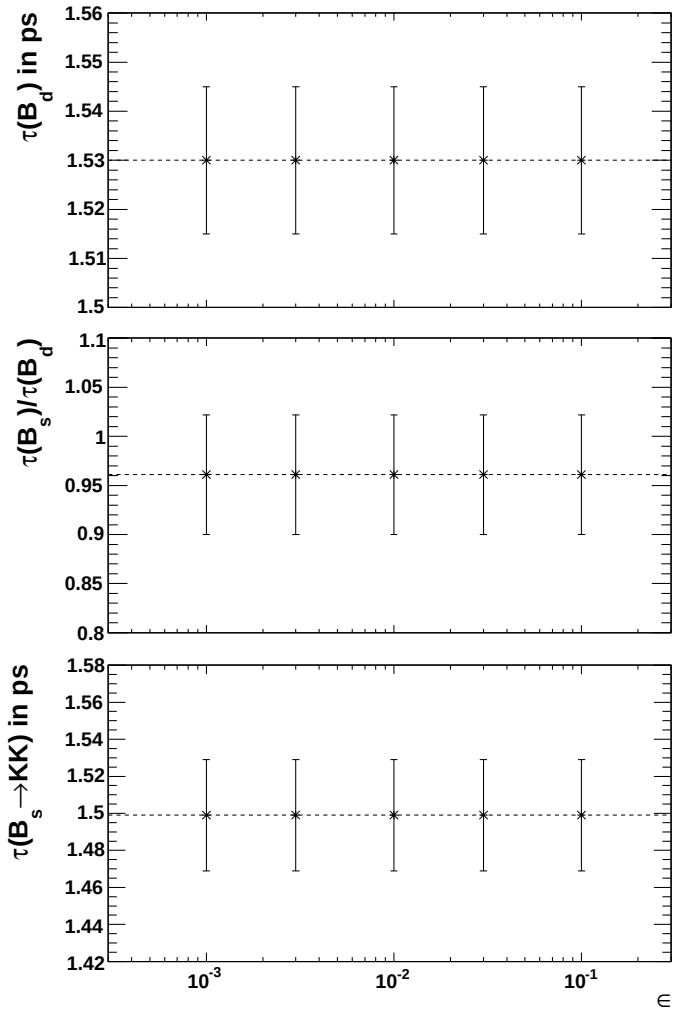


Figure 5.15: *Toy study results for the variation of the fitter parameter ϵ. Default is $\epsilon = 1 \times 10^{-2}$. The dashed line indicates the average result of the default fit.*

5.2.2.4 Variation of the ϵ Parameter

The last stability check is the variation of the parameter ϵ, which is used in the method to avoid negative probability densities (see section 4.3.2.3). The default value used in the fitter is $\epsilon = 0.01$. This value has been varied by one order of magnitude in either direction.

Figure 5.15 shows the toy results for the three fit parameters as a function of ϵ. The error bars shown are the full statistical error for each fit. For the variation of ϵ all fit parameters show excellent stability.

Tables 5.14 to 5.16 summarise the numerical results of the fits with different values of ϵ. In addition to the values shown in figure 5.15 they also give the respective values for pull mean and pull sigma as obtained by a Gaussian fit.

ϵ	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
1×10^{-1}	1.530	1.530	0.015	0.01	1.00
3×10^{-2}	1.530	1.530	0.015	0.01	1.01
1×10^{-2}	1.530	1.530	0.015	0.01	1.01
3×10^{-3}	1.530	1.530	0.015	0.01	1.01
1×10^{-3}	1.530	1.530	0.015	0.01	1.01

Table 5.14: *Toy study results for $\tau(B_d^0)$ for the variation of the fitter parameter ϵ. Default is $\epsilon = 1 \times 10^{-2}$.*

ϵ	true value	avg. result	avg. error	pull mean	pull width
1×10^{-1}	0.961	0.961	0.061	0.00	0.98
3×10^{-2}	0.961	0.961	0.061	0.00	0.98
1×10^{-2}	0.961	0.961	0.061	0.00	0.98
3×10^{-3}	0.961	0.961	0.061	0.00	0.98
1×10^{-3}	0.961	0.961	0.061	0.00	0.98

Table 5.15: *Toy study results for $\tau(B_s^0)/\tau(B_d^0)$ for the variation of the fitter parameter ϵ. Default is $\epsilon = 1 \times 10^{-2}$.*

ϵ	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
1×10^{-1}	1.500	1.499	0.030	-0.06	1.00
3×10^{-2}	1.500	1.499	0.030	-0.06	1.00
1×10^{-2}	1.500	1.499	0.030	-0.06	1.00
3×10^{-3}	1.500	1.499	0.030	-0.06	1.00
1×10^{-3}	1.500	1.499	0.030	-0.06	1.00

Table 5.16: *Toy study results for $\tau(B_s^0 \to K^+K^-)$ for the variation of the fitter parameter ϵ. Default is $\epsilon = 1 \times 10^{-2}$.*

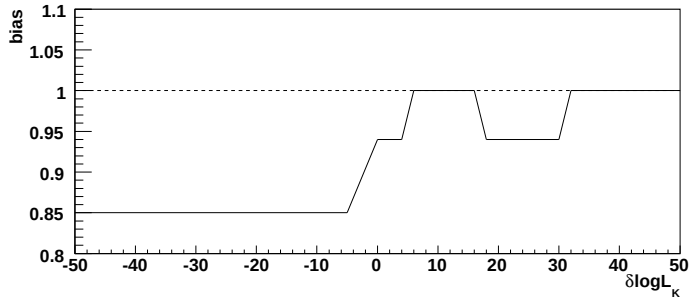


Figure 5.16: *Bias applied to $\delta \log \mathcal{L}_K$ template to model the difference between the re-weighted and the true distribution (solid line). The dashed line indicates the unbiased case.*

5.2.3 Systematic Uncertainties of Lifetime Measurements with $B^0_{(s)} \to h^+h'^-$ Events

This section summarises tests that have been performed using toy MC data sets to evaluate the size of potential systematic uncertainties of the lifetime fitting method. The fitting process has been simplified, as discussed in the previous section, by fixing the input signal fractions as well as the Λ_b^0 lifetime and that of three-body decays to their respective true average values. Since all fits for one given test are performed on the same data sample, the errors shown are fully correlated and any fluctuation has to be seen as a systematic effect.

5.2.3.1 Impact of Wrong $\delta \log \mathcal{L}_K$ Templates

Related to the variation of the signal fractions is the study of the impact of wrong templates for the variables used to distinguish the various signal classes. The first to be checked is the template for $\delta \log \mathcal{L}_K$. The *PID* templates are obtained by re-weighting the distributions measured in $D^0 \to \pi^+K^-$ decays according to the different distributions in momentum and transverse momentum. It has been shown that this method is able to determine the correct distributions for $B^0_{(s)} \to h^+h'^-$ decays up to a small uncertainty [29]. This uncertainty has been modelled and applied to the templates used in the fitter. Three scenarios have been studied in addition to the default one: one with the full bias (as shown in figure 5.16), one with 2/3 of the bias, and one with 1/3 of the bias. The $\delta \log \mathcal{L}_K$ templates are among the ingredients for the signal fraction fit, hence, for this study, both signal fraction and lifetime fits have been performed.

The results of the tests are summarised in figure 5.17 and tables 5.17 to 5.19. The fit results show some sensitivity to mis-calibrations of the $\delta \log \mathcal{L}_K$ distribution. The

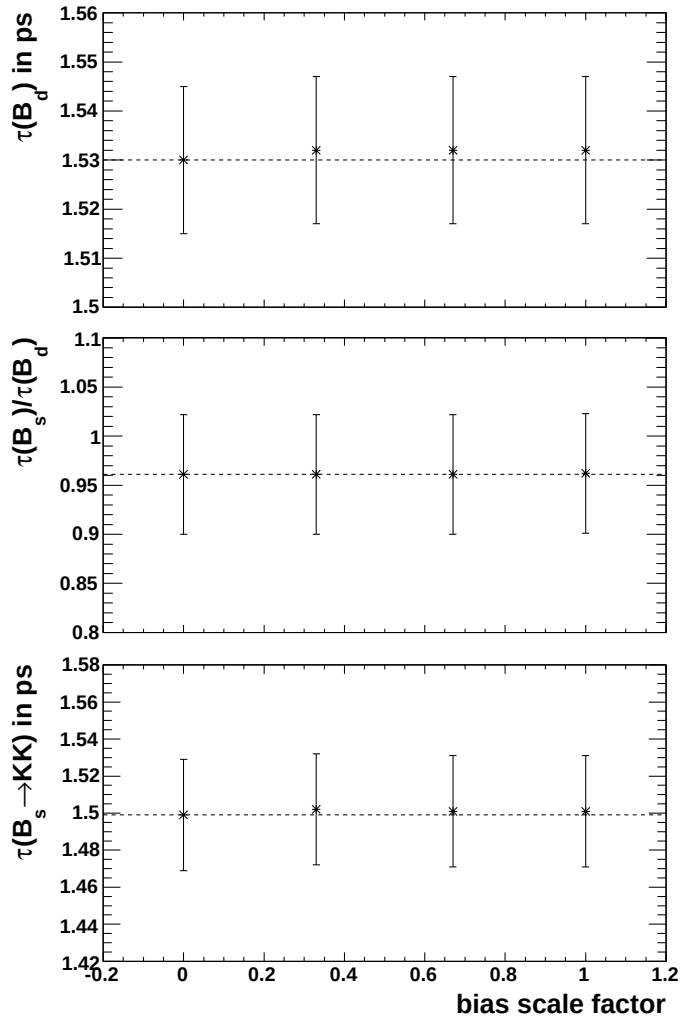


Figure 5.17: *Toy study results for fits with a biased $\delta \log \mathcal{L}_K$ template. Default is* 0.00. *The dashed line indicates the average result of the default fit.*

bias scale	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
0.00	1.530	1.530	0.015	0.01	1.01
0.33	1.530	1.532	0.015	0.12	0.97
0.67	1.530	1.532	0.015	0.15	0.96
1.00	1.530	1.532	0.015	0.16	0.96

Table 5.17: *Toy study results for $\tau(B_d^0)$ for fits with a biased $\delta \log \mathcal{L}_K$ template. Default is* 0.00.

bias scale	true value	avg. result	avg. error	pull mean	pull width
0.00	0.961	0.961	0.061	0.00	0.98
0.33	0.961	0.961	0.061	0.00	0.97
0.67	0.961	0.961	0.061	0.00	0.98
1.00	0.961	0.962	0.061	0.00	0.97

Table 5.18: *Toy study results for $\tau(B_s^0)/\tau(B_d^0)$ for fits with a biased $\delta \log \mathcal{L}_K$ template. Default is* 0.00.

bias scale	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
0.00	1.500	1.499	0.030	−0.06	1.00
0.33	1.500	1.502	0.030	0.04	1.03
0.67	1.500	1.501	0.030	0.04	1.03
1.00	1.500	1.501	0.030	0.04	1.04

Table 5.19: *Toy study results for $\tau(B_s^0 \to K^+K^-)$ for fits with a biased $\delta \log \mathcal{L}_K$ template. Default is* 0.00.

maximal deviations from the respective default values are taken as systematic uncertainties.

It may be expected that the method to obtain the *PID* distributions from data can be improved, e.g. by the usage of equally populated bins in the p-p_T plane. Therefore, these systematic uncertainties have the potential to be reduced in the future. Eventually, a study with the measured differences between the re-weighted and the true distributions should be used instead of the parametrisation used here.

5.2.3.2 Impact of Wrong $m(B_s^0 \to K^+K^-)$ Templates

The invariant mass templates are another ingredient to the signal fraction fit. A scale factor between 0.9 and 1.1 has been applied to the mass axis of the template for the mass distribution of $B_s^0 \to K^+K^-$ events, such that

$$f(m_{KK}) \to f(m_0 + scale \times (m_{KK} - m_0)), \tag{5.5}$$

where the centre of the scaling, m_0, has been set to the peak of the distribution such that its position remains unchanged. A change of the mass scale can be expected from mis-calibration of the magnetic field or from misalignments such as x translations of the stations of *IT* and *OT* which are a linear function of their respective z position.

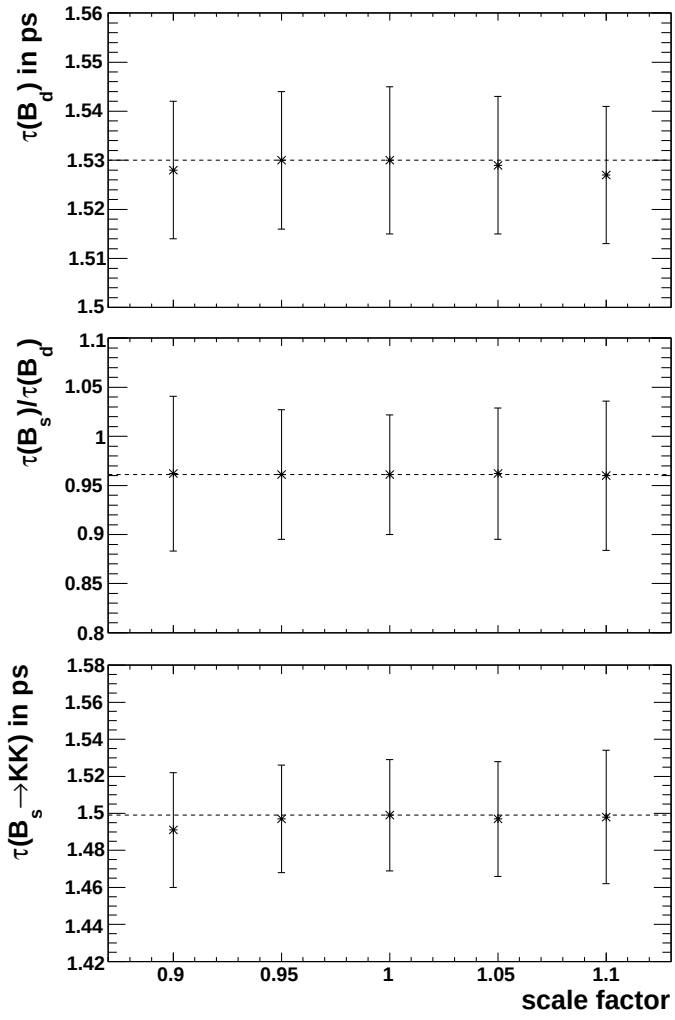


Figure 5.18: *Toy study results for fits with a wrongly fixed $m(B_s^0 \to K^+K^-)$ scale. Default is* 1.00*. The dashed line indicates the average result of the default fit.*

These effects would affect all channels alike. Differences between the channels may arise from effects linked to the daughter particles such as final state radiation. Thus, the study of the scaling of the $B_s^0 \to K^+K^-$ template alone is clearly a worst case scenario.

Figure 5.18 and tables 5.20 to 5.22 summarise the results. Some dependency on the scale factor can be observed, particularly for the largest deviations from nominal. The mass line shape is expected to be well under control. Mis-calibrations beyond the few per-cent level should be visible in the mass projection of the signal fraction fit. Hence, the maximal deviations for a scaling of $\pm 5\%$ are taken as systematic uncertainties to account for a possible lack of understanding of the mass line shape.

$m(B_s^0 \to K^+K^-)$ scale	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
0.90	1.530	1.528	0.014	−0.20	0.99
0.95	1.530	1.530	0.014	−0.05	0.99
1.00	1.530	1.530	0.015	0.01	1.01
1.05	1.530	1.529	0.014	−0.08	1.01
1.10	1.530	1.527	0.014	−0.20	1.01

Table 5.20: *Toy study results for $\tau(B_d^0)$ for fits with a wrongly fixed $m(B_s^0 \to K^+K^-)$ scale. Default is* 1.00.

$m(B_s^0 \to K^+K^-)$ scale	true value	avg. result	avg. error	pull mean	pull width
0.90	0.961	0.962	0.079	−0.01	0.97
0.95	0.961	0.961	0.066	0.00	0.97
1.00	0.961	0.961	0.061	0.00	0.98
1.05	0.961	0.962	0.067	0.00	0.97
1.10	0.961	0.960	0.076	−0.05	0.97

Table 5.21: *Toy study results for $\tau(B_s^0)/\tau(B_d^0)$ for fits with a wrongly fixed $m(B_s^0 \to K^+K^-)$ scale. Default is* 1.00.

$m(B_s^0 \to K^+K^-)$ scale	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
0.90	1.500	1.491	0.031	−0.32	1.02
0.95	1.500	1.497	0.029	−0.13	1.00
1.00	1.500	1.499	0.030	−0.06	1.00
1.05	1.500	1.497	0.031	−0.12	0.99
1.10	1.500	1.498	0.036	−0.07	1.01

Table 5.22: *Toy study results for $\tau(B_s^0 \to K^+K^-)$ for fits with a wrongly fixed $m(B_s^0 \to K^+K^-)$ scale. Default is* 1.00.

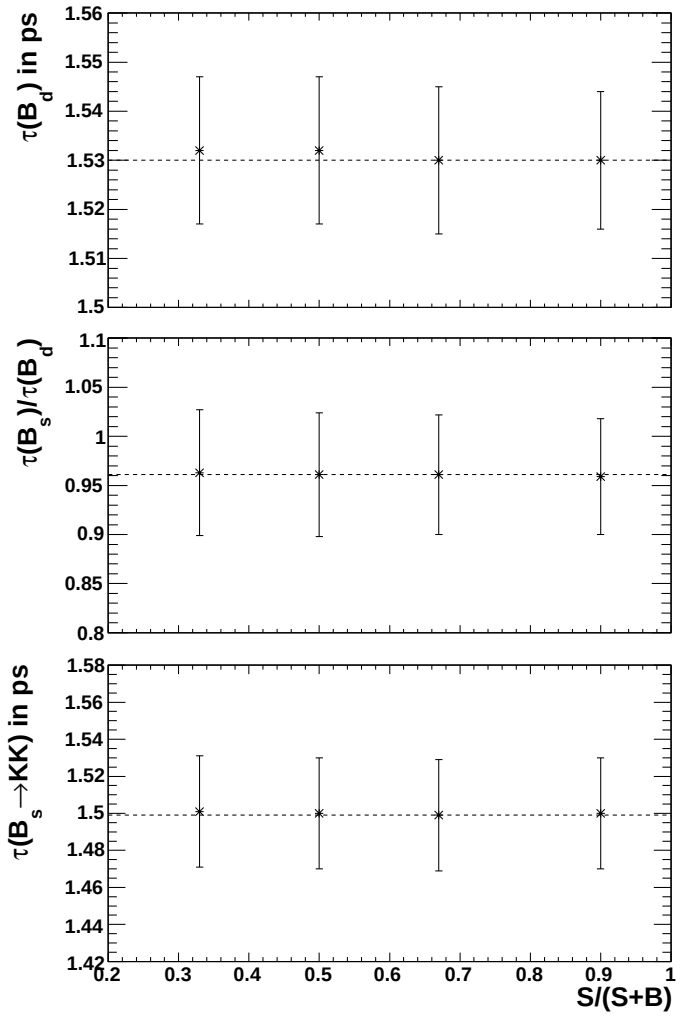


Figure 5.19: *Toy study results for fits with varying signal fractions. Default is* 0.67*. The dashed line indicates the average result of the default fit.*

5.2.3.3 Variation of the Level of Combinatorial Background

The treatment of combinatorial background is a central part of the lifetime fitter (see section 4.3.2). Therefore, it has been studied how different levels of combinatorial background affect the lifetime fitting process. In the full mass window the level of combinatorial background is estimated to amount to 1/3 of all events, which corresponds to a signal fraction (the sum of all exclusive channels) of $S/(S+B) = 2/3$. This signal fraction has been varied between 1/3 and 9/10 by varying the amount of combinatorial background, i.e. by maintaining the number of signal events.

The results are summarised in figure 5.19 and tables 5.23 to 5.25. The dependency on the level of combinatorial background can be understood as the intrinsic systematic uncertainty of the fitting method, since a pure signal fit is unbiased. Given that the level of combinatorial background can be controlled to some extent by varying the

$S/(S+B)$	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
0.33	1.530	1.532	0.015	0.11	0.97
0.50	1.530	1.532	0.015	0.14	0.98
0.67	1.530	1.530	0.015	0.01	1.01
0.90	1.530	1.530	0.014	−0.04	1.01

Table 5.23: *Toy study results for $\tau(B_d^0)$ for fits with varying signal fractions. Default is* 0.67*.*

$S/(S+B)$	true value	avg. result	avg. error	pull mean	pull width
0.33	0.961	0.963	0.064	0.01	0.97
0.50	0.961	0.961	0.063	−0.04	0.96
0.67	0.961	0.961	0.061	0.00	0.98
0.90	0.961	0.959	0.059	−0.06	1.00

Table 5.24: *Toy study results for $\tau(B_s^0)/\tau(B_d^0)$ for fits with varying signal fractions. Default is* 0.67*.*

$S/(S+B)$	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
0.33	1.500	1.501	0.030	0.01	1.00
0.50	1.500	1.500	0.030	−0.04	0.95
0.67	1.500	1.499	0.030	−0.06	1.00
0.90	1.500	1.500	0.030	0.00	1.00

Table 5.25: *Toy study results for $\tau(B_s^0 \to K^+K^-)$ for fits with varying signal fractions. Default is* 0.67*.*

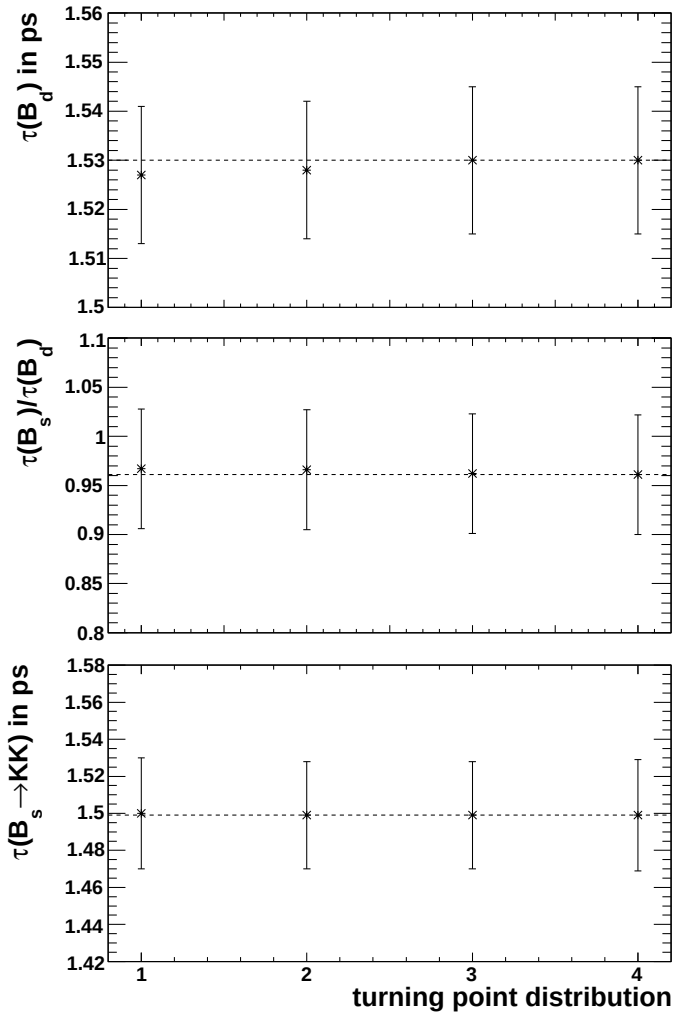


Figure 5.20: *Toy study results for fits with varying turning point distributions. See text for details. Default is all* 1.00*. The dashed line indicates the average result of the default fit.*

mass window, a systematic uncertainty of ± 0.001 is assigned to all three observables. The fact that the errors are almost constant for the different levels of combinatorial background underlines the high statistical power of the fitting method.

5.2.3.4 Different Lifetime Acceptance Functions

It has been discussed in section 4.3.2.2 how the average acceptance functions are obtained from data. In the standard toy MC generator the turning points defining the event-by-event acceptance functions have been generated using a Gaussian distribution with mean 1.0 ps and width 0.1 ps for all channels. However, as the turning points are, to first order, given by the lifetime biasing cuts (common to all decays) and the angular distributions of the tracks involved, which may differ for different decay modes, it is

turning point mean [ps]	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
bg. 0.90	1.530	1.527	0.014	−0.17	0.98
bg. 0.90, $B_s^0 \to K^+K^-$ 0.95	1.530	1.528	0.014	−0.18	1.02
$B_s^0 \to K^+K^-$ 0.90	1.530	1.530	0.015	−0.01	1.03
all 1.00	1.530	1.530	0.015	0.01	1.01

Table 5.26: *Toy study results for $\tau(B_d^0)$ for fits with varying turning point distributions. See text for details. Default is all* 1.00.

turning point mean [ps]	true value	avg. result	avg. error	pull mean	pull width
bg. 0.90	0.961	0.967	0.061	0.08	0.95
bg. 0.90, $B_s^0 \to K^+K^-$ 0.95	0.961	0.966	0.061	0.05	0.96
$B_s^0 \to K^+K^-$ 0.90	0.961	0.962	0.061	−0.04	0.95
all 1.00	0.961	0.961	0.061	0.00	0.98

Table 5.27: *Toy study results for $\tau(B_s^0)/\tau(B_d^0)$ for fits with varying turning point distributions. See text for details. Default is all* 1.00.

not guaranteed that all turning point distributions are the same. If one distribution is distinct from all others this acts as an additional distinguishing variable leading to a quasi background-free fit for that particular channel. Therefore, the only thing to be studied is different but overlapping turning point distributions.

Three scenarios have been studied in addition to the default one. In all cases the mean of some of the turning point distributions has been varied while the width was kept constant. The first scenario was produced with events where the turning point distribution for combinatorial background had a mean of 0.90 ps. In the second scenario the $B_s^0 \to K^+K^-$ distribution was generated with a mean of 0.95 ps, in addition to the changes of the first scenario. Finally, for the third scenario, only the $B_s^0 \to K^+K^-$ distribution was altered from its nominal shape to a mean of 0.90 ps.

The results are summarised in figure 5.20 and tables 5.26 to 5.28. It is apparent that the most critical turning point distribution is that of the combinatorial background. In general, it is not surprising that the B_d^0 lifetime measurement is more affected by combinatorial background than the $B_s^0 \to K^+K^-$ lifetime measurement. This is because the combinatorial background is simulated to consist purely of pions. While this is true to a large extent, the effect of combinatorial background on the $B_s^0 \to K^+K^-$

turning point mean [ps]	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
bg. 0.90	1.500	1.500	0.030	−0.03	0.96
bg. 0.90, $B_s^0 \to K^+K^-$ 0.95	1.500	1.499	0.029	−0.04	1.00
$B_s^0 \to K^+K^-$ 0.90	1.500	1.499	0.029	−0.06	0.98
all 1.00	1.500	1.499	0.030	−0.06	1.00

Table 5.28: *Toy study results for $\tau(B_s^0 \to K^+K^-)$ for fits with varying turning point distributions. See text for details. Default is all* 1.00.

measurement might be underestimated. Therefore, a systematic uncertainty due to the treatment of the average acceptance functions in the fitter of ±0.002 ps is assigned to the $B_s^0 \to K^+K^-$ measurement. The value of 0.002 ps is taken as the mean between the observed (and potentially underestimated) uncertainty for $B_s^0 \to K^+K^-$ of 0.001 ps and that for the B_d^0 lifetime of 0.003 ps. The other uncertainties are assigned as the maximal deviations observed.

5.2.3.5 Different Background Lifetime Distributions

The lifetime distribution for combinatorial background has been simulated with a short lived component of $\tau = 0.5$ ps (relative fraction 90%) and a long lived component of $\tau = 10$ ps (relative fraction 10%). This roughly matches the distribution of the full MC simulation. The lack of a long lived component can potentially cause more problems related to avoiding negative probability densities (see section 4.3.2.3).

In addition to the default scenario, three other distributions for combinatorial background have been studied. The first scenario contains a larger amount of the long lived component (1/3 instead of 10%). However, this is expected not to be a realistic scenario. The second scenario contains a component with a signal-like lifetime of $\tau = 1.5$ ps in addition to the short and long components. All have a relative fraction of 1/3. The third scenario only contains the short lived component.

The results for the various scenarios are summarised in figure 5.21 and tables 5.29 to 5.31. The fit results appear to be reasonably stable with respect to varying lifetime distributions of combinatorial background. The differences between the 'default' and the 'short only' scenario are taken as systematic uncertainties. Again, the $B_s^0 \to K^+K^-$ uncertainty is increased to 0.002 ps to account for the potentially larger influence of combinatorial background.

Different mass shapes for the combinatorial background have not been studied. The

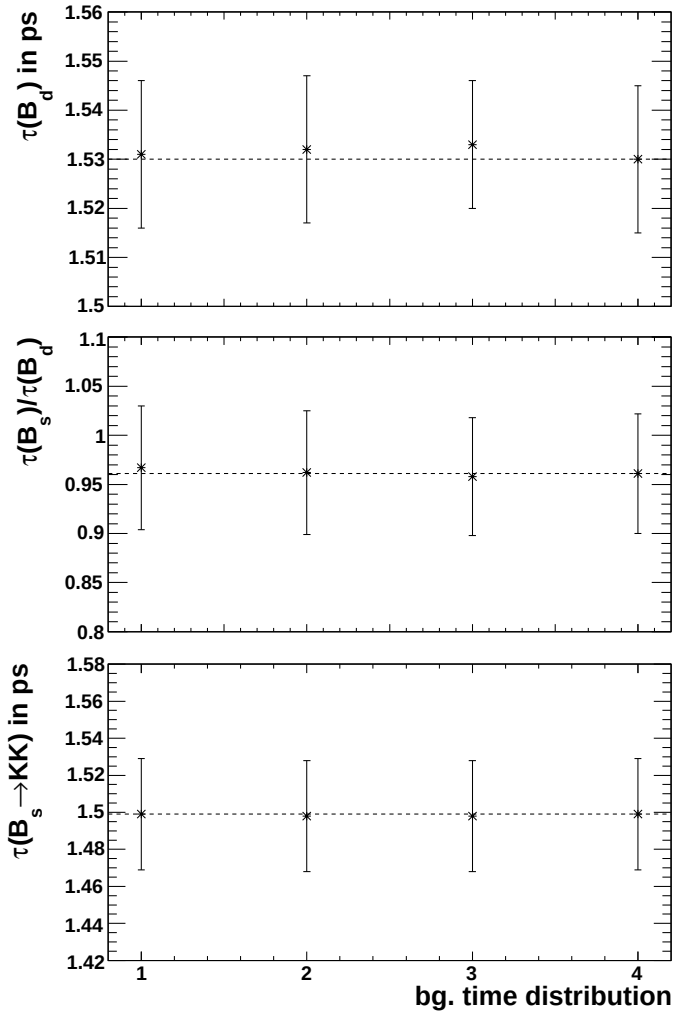


Figure 5.21: *Toy study results for fits with different background lifetime distributions. See text for details. The dashed line indicates the average result of the default fit.*

bg. distribution	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
short, long	1.530	1.531	0.015	0.07	0.99
short, medium, long	1.530	1.532	0.015	0.11	1.03
short only	1.530	1.533	0.013	0.26	1.15
default	1.530	1.530	0.015	0.01	1.01

Table 5.29: *Toy study results for $\tau(B_d^0)$ for fits with different background lifetime distributions. See text for details.*

bg. distribution	true value	avg. result	avg. error	pull mean	pull width
short, long	0.961	0.967	0.063	−0.03	0.99
short, medium, long	0.961	0.962	0.063	−0.02	0.95
short only	0.961	0.958	0.060	−0.09	0.97
default	0.961	0.961	0.061	0.00	0.98

Table 5.30: *Toy study results for $\tau(B_s^0)/\tau(B_d^0)$ for fits with different background lifetime distributions. See text for details.*

flat distribution is expected to be reasonably close to the one that will be observed in data. A change in this distribution can effectively be seen as a change in the signal fraction which has been studied in section 5.2.3.3.

5.2.3.6 Impact of Wrong Proper Time Resolutions

The nominal proper time resolution is about 40 fs. This can change due to miscalibrations of the detector such as misalignments (see section 5.4). This section studies the effect of an increase of the proper time resolution up to 80 fs. The values studied here roughly correspond to those observed in the misalignment studies as discussed

bg. distribution	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
short, long	1.500	1.499	0.030	−0.04	0.97
short, medium, long	1.500	1.498	0.030	−0.09	0.96
short only	1.500	1.498	0.030	−0.06	1.03
default	1.500	1.499	0.030	−0.06	1.00

Table 5.31: *Toy study results for $\tau(B_s^0 \to K^+K^-)$ for fits with different background lifetime distributions. See text for details.*

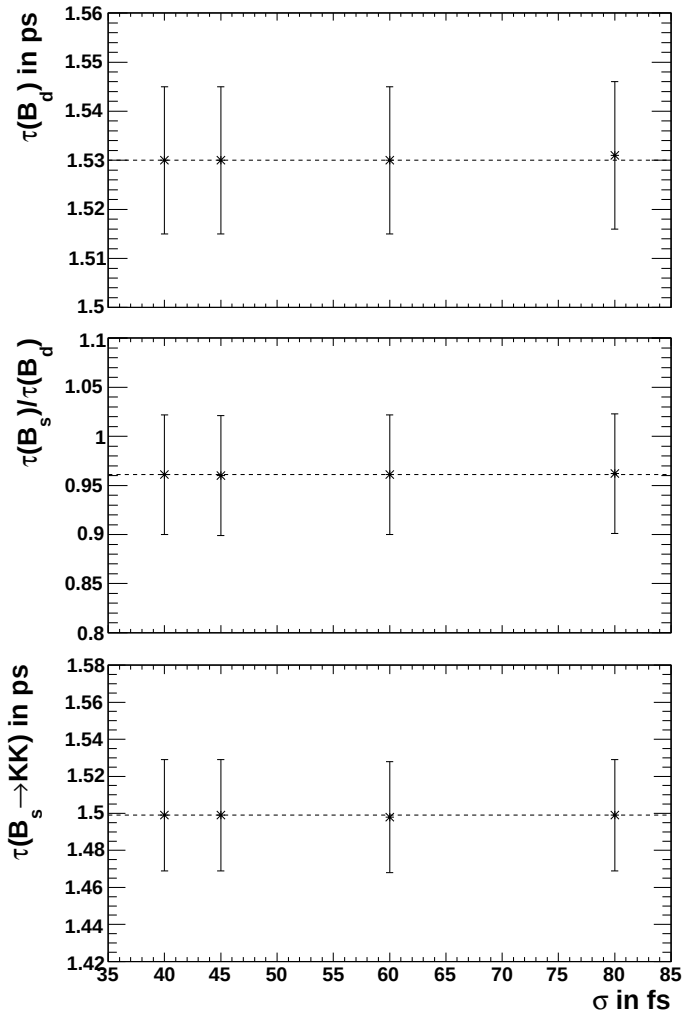


Figure 5.22: *Toy study results for fits with wrong proper time resolutions. The values are only varied in the generator while the default of* 40 fs *is kept in the fitter. The dashed line indicates the average result of the default fit.*

in section 5.4.2.3. While the input data are generated with the varied proper time resolutions, the fitter assumes the nominal resolution of 40 fs.

Figure 5.22 and tables 5.32 to 5.34 summarise the results. Since the worsening of the proper time resolution due to misalignments is expected to be at most 10%, any deviations of the fits with a resolution of 45 fs are taken as systematic uncertainties. Despite the fact that excellent proper time resolution is less important for lifetime measurements, it is crucial for measurements aiming at resolving oscillations in flavour-tagged B_s^0 decays.

proper time resolution [fs]	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
40	1.530	1.530	0.015	0.01	1.01
45	1.530	1.530	0.015	−0.01	1.00
60	1.530	1.530	0.015	0.02	0.99
80	1.530	1.531	0.015	0.05	1.00

Table 5.32: *Toy study results for $\tau(B_d^0)$ for fits with wrong proper time resolutions. The values are only varied in the generator while the default of* 40 fs *is kept in the fitter.*

proper time resolution [fs]	true value	avg. result	avg. error	pull mean	pull width
40	0.961	0.961	0.061	0.00	0.98
45	0.961	0.960	0.061	−0.02	0.95
60	0.961	0.961	0.061	−0.03	0.97
80	0.961	0.962	0.061	−0.01	0.98

Table 5.33: *Toy study results for $\tau(B_s^0)/\tau(B_d^0)$ for fits with wrong proper time resolutions. The values are only varied in the generator while the default of* 40 fs *is kept in the fitter.*

proper time resolution [fs]	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
40	1.500	1.499	0.030	−0.06	1.00
45	1.500	1.499	0.030	−0.06	0.98
60	1.500	1.498	0.030	−0.06	0.98
80	1.500	1.499	0.030	−0.04	0.97

Table 5.34: *Toy study results for $\tau(B_s^0 \to K^+K^-)$ for fits with wrong proper time resolutions. The values are only varied in the generator while the default of* 40 fs *is kept in the fitter.*

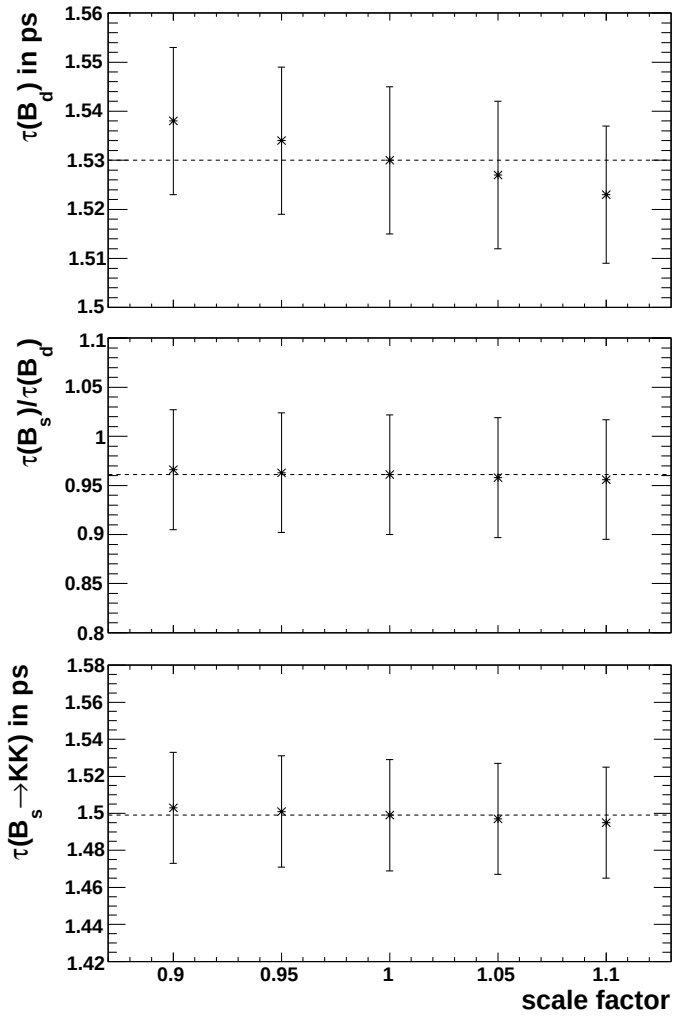


Figure 5.23: *Toy study results for fits with a wrongly fixed Λ_b^0 lifetime. Default is* 1.00. *The dashed line indicates the average result of the default fit.*

5.2.3.7 Impact of Wrongly Fixed Λ_b^0 Lifetime

The Λ_b^0 lifetime has been fixed in the fits used for the systematic studies. Therefore, the impact of a wrong value of the Λ_b^0 lifetime has been studied. Its value has been scaled by a factor between 0.9 and 1.1 compared to the true average value.

The results are summarised in figure 5.23 and tables 5.35 to 5.37. A strong dependence on the Λ_b^0 lifetime value is observed for all three fit parameters, strongest for the B_d^0 lifetime. The Λ_b^0 lifetime is known to a precision of about 3% [12]. Hence, fixing this lifetime in the real fit would lead to a systematic uncertainty of about 0.002 for all observables. However, as the full fit did not show any bias compared to the reduced fit, no systematic uncertainty is assigned.

As an alternative to fixing the Λ_b^0 lifetime the mass window can be tightened to an upper limit of 5400 MeV/c^2. This would exclude most of the events from Λ_b^0 decays.

$\tau(\Lambda_b^0)$ scale	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
0.90	1.530	1.538	0.015	0.53	1.01
0.95	1.530	1.534	0.015	0.24	1.01
1.00	1.530	1.530	0.015	0.01	1.01
1.05	1.530	1.527	0.015	−0.24	1.02
1.10	1.530	1.523	0.014	−0.46	1.03

Table 5.35: *Toy study results for $\tau(B_d^0)$ for fits with a wrongly fixed Λ_b^0 lifetime. Default is* 1.00.

$\tau(\Lambda_b^0)$ scale	true value	avg. result	avg. error	pull mean	pull width
0.90	0.961	0.966	0.061	0.06	0.95
0.95	0.961	0.963	0.061	0.00	0.93
1.00	0.961	0.961	0.061	0.00	0.98
1.05	0.961	0.958	0.061	−0.05	0.99
1.10	0.961	0.956	0.061	−0.09	0.97

Table 5.36: *Toy study results for $\tau(B_s^0)/\tau(B_d^0)$ for fits with a wrongly fixed Λ_b^0 lifetime. Default is* 1.00.

$\tau(\Lambda_b^0)$ scale	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
0.90	1.500	1.503	0.030	0.06	1.00
0.95	1.500	1.501	0.030	0.01	1.00
1.00	1.500	1.499	0.030	−0.06	1.00
1.05	1.500	1.497	0.030	−0.12	1.00
1.10	1.500	1.495	0.030	−0.19	1.00

Table 5.37: *Toy study results for $\tau(B_s^0 \to K^+K^-)$ for fits with a wrongly fixed Λ_b^0 lifetime. Default is* 1.00.

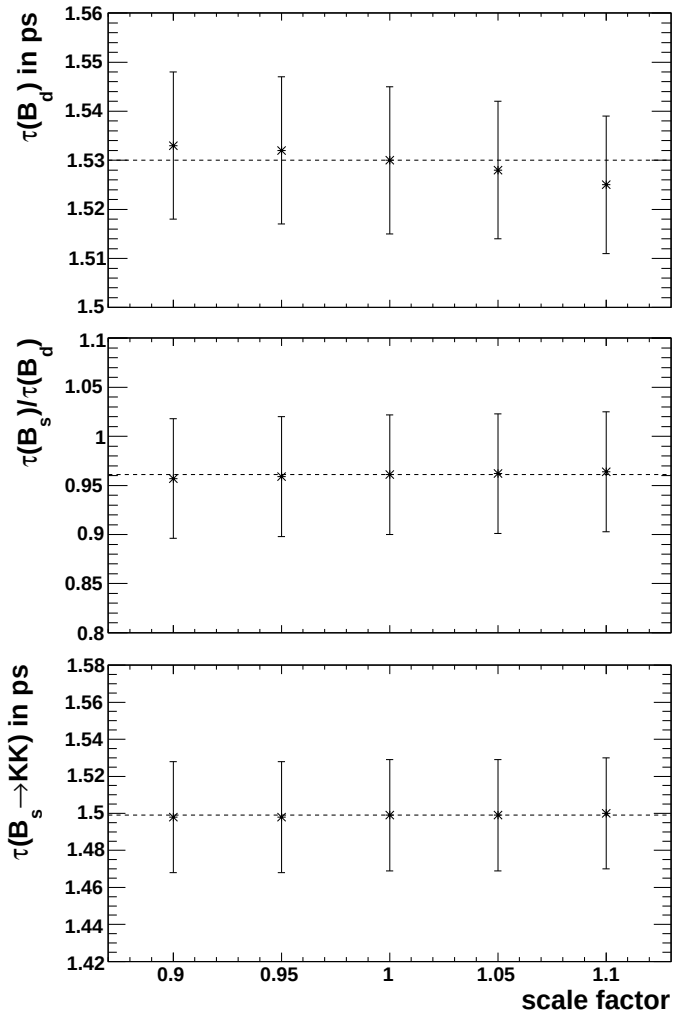


Figure 5.24: *Toy study results for fits with a wrongly fixed* 3*−body lifetime. Default is* 1.00*. The dashed line indicates the average result of the default fit.*

In this case, a systematic uncertainty of about 0.001 would have to be assigned to all observables according to section 5.2.2.1.

5.2.3.8 Impact of Wrongly Fixed Three-Body Lifetime

Similarly to the previous section, the lifetime for three-body decays has been fixed. It has been varied by a scale factor between 0.9 and 1.1.

Figure 5.24 and tables 5.38 to 5.40 summarise the results. A similar, however weaker, dependency compared to the Λ_b^0 lifetime is observed. Again, no systematic uncertainty is assigned, as the full fit is unbiased. A tightening of the mass window to exclude most of the three-body background would lead to the assignment of a systematic error of 0.001 to all fit parameters, as discussed in the previous section.

$\tau(3-body)$ scale	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
0.90	1.530	1.533	0.015	0.22	0.99
0.95	1.530	1.532	0.015	0.12	1.02
1.00	1.530	1.530	0.015	0.01	1.01
1.05	1.530	1.528	0.014	−0.15	1.02
1.10	1.530	1.525	0.014	−0.36	1.04

Table 5.38: *Toy study results for $\tau(B_d^0)$ for fits with a wrongly fixed 3–body lifetime. Default is* 1.00.

$\tau(3-body)$ scale	true value	avg. result	avg. error	pull mean	pull width
0.90	0.961	0.957	0.061	−0.08	0.97
0.95	0.961	0.959	0.061	−0.03	0.98
1.00	0.961	0.961	0.061	0.00	0.98
1.05	0.961	0.962	0.061	0.00	0.94
1.10	0.961	0.964	0.061	0.03	0.96

Table 5.39: *Toy study results for $\tau(B_s^0)/\tau(B_d^0)$ for fits with a wrongly fixed $3-body$ lifetime. Default is* 1.00.

$\tau(3-body)$ scale	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
0.90	1.500	1.498	0.030	−0.07	1.00
0.95	1.500	1.498	0.030	−0.06	1.00
1.00	1.500	1.499	0.030	−0.06	1.00
1.05	1.500	1.499	0.030	−0.03	1.00
1.10	1.500	1.500	0.030	−0.02	1.00

Table 5.40: *Toy study results for $\tau(B_s^0 \to K^+K^-)$ for fits with a wrongly fixed $3-$ body lifetime. Default is* 1.00.

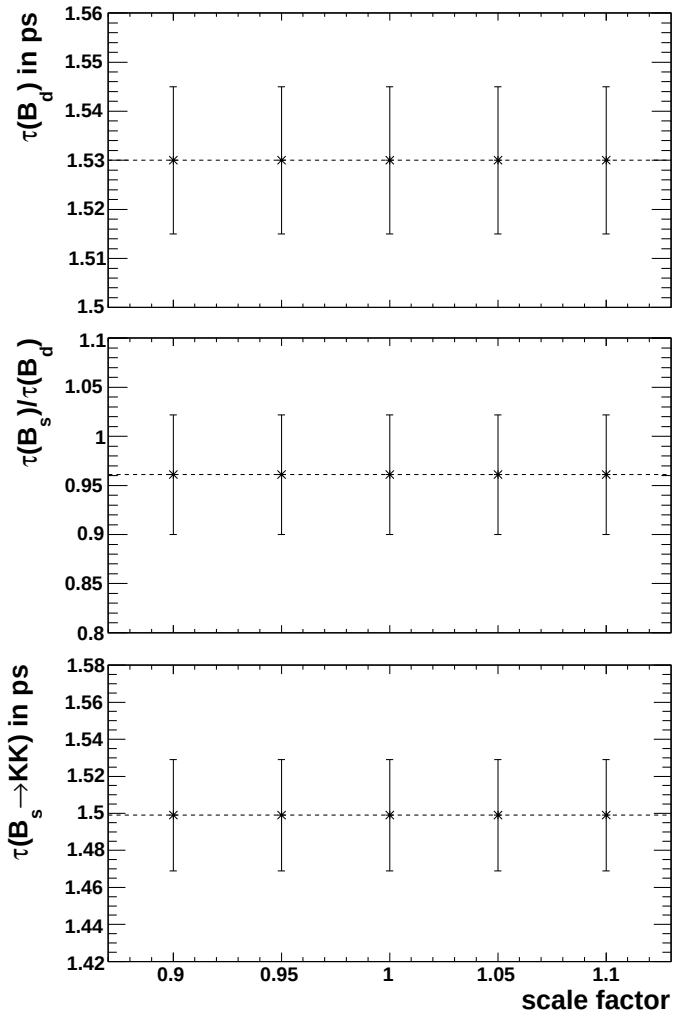


Figure 5.25: *Toy study results for fits with a wrongly fixed $\tau(B_s^0 \to K^+K^-)$ scale. Default is* 1.00*. The dashed line indicates the average result of the default fit.*

5.2.3.9 Impact of Wrongly Fixed Average $B_s^0 \to K^+K^-$ Lifetime

Fixed values for the average signal lifetimes are used in the calculation of the average acceptance functions. It has been discussed in section 4.3.2.2 that these can be updated in a second iteration of the fit if their values prove to be wrong in the first iteration. In order for this to work the dependence on wrong input values has to be sufficiently small. In this section, the average $B_s^0 \to K^+K^-$ lifetime has been scaled by a factor between 0.9 and 1.1.

The results are summarised in figure 5.25 and tables 5.41 to 5.43. Perfect stability is observed such that no systematic uncertainty has to be assigned.

$\tau(B_s^0 \to K^+K^-)$ scale	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
0.90	1.530	1.530	0.015	0.01	1.01
0.95	1.530	1.530	0.015	0.01	1.01
1.00	1.530	1.530	0.015	0.01	1.01
1.05	1.530	1.530	0.015	0.01	1.01
1.10	1.530	1.530	0.015	0.01	1.00

Table 5.41: *Toy study results for $\tau(B_d^0)$ for fits with a wrongly fixed $\tau(B_s^0 \to K^+K^-)$ scale. Default is* 1.00.

$\tau(B_s^0 \to K^+K^-)$ scale	true value	avg. result	avg. error	pull mean	pull width
0.90	0.961	0.961	0.061	0.00	0.98
0.95	0.961	0.961	0.061	0.00	0.98
1.00	0.961	0.961	0.061	0.00	0.98
1.05	0.961	0.961	0.061	0.00	0.98
1.10	0.961	0.961	0.061	0.00	0.98

Table 5.42: *Toy study results for $\tau(B_s^0)/\tau(B_d^0)$ for fits with a wrongly fixed $\tau(B_s^0 \to K^+K^-)$ scale. Default is* 1.00.

$\tau(B_s^0 \to K^+K^-)$ scale	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
0.90	1.500	1.499	0.030	−0.06	1.00
0.95	1.500	1.499	0.030	−0.06	1.00
1.00	1.500	1.499	0.030	−0.06	1.00
1.05	1.500	1.499	0.030	−0.06	1.00
1.10	1.500	1.499	0.030	−0.06	1.00

Table 5.43: *Toy study results for $\tau(B_s^0 \to K^+K^-)$ for fits with a wrongly fixed $\tau(B_s^0 \to K^+K^-)$ scale. Default is* 1.00.

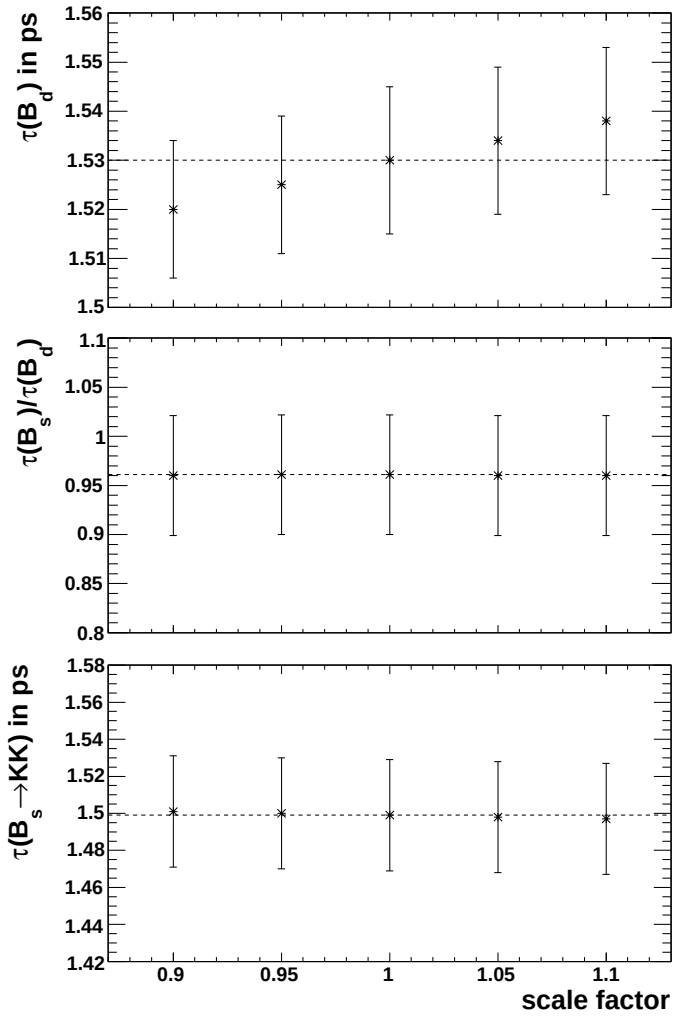


Figure 5.26: *Toy study results for fits with a wrongly fixed $\tau(bg)$ scale. Default is* 1.00. *The dashed line indicates the average result of the default fit.*

5.2.3.10 Impact of Wrongly Fixed Average Background Lifetime

Similar to the previous section, the values for the average lifetimes of combinatorial background components have been varied. The variation is again given by a scale factor between 0.9 and 1.1.

Figure 5.26 and tables 5.44 to 5.46 summarise the results. In this case, a rather strong dependency is observed, particularly for the B_d^0 lifetime measurement. However, with the size of the background sample used in the toy experiments of this study it should be feasible to measure the background lifetime distribution to about 1% statistical precision. Using a second iteration of the fit, this would lead to a reduced dependency. Hence, a systematic uncertainty of only 0.001 ps is assigned to the two direct lifetime measurements while no uncertainty is assigned to the lifetime ratio measurement.

$\tau(bg)$ scale	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
0.90	1.530	1.520	0.014	−0.72	1.03
0.95	1.530	1.525	0.014	−0.34	1.02
1.00	1.530	1.530	0.015	0.01	1.01
1.05	1.530	1.534	0.015	0.30	1.02
1.10	1.530	1.538	0.015	0.55	1.00

Table 5.44: *Toy study results for $\tau(B_d^0)$ for fits with a wrongly fixed $\tau(bg)$ scale. Default is* 1.00.

$\tau(bg)$ scale	true value	avg. result	avg. error	pull mean	pull width
0.90	0.961	0.960	0.061	−0.05	0.96
0.95	0.961	0.961	0.061	0.00	1.00
1.00	0.961	0.961	0.061	0.00	0.98
1.05	0.961	0.960	0.061	−0.01	0.98
1.10	0.961	0.960	0.061	−0.02	0.97

Table 5.45: *Toy study results for $\tau(B_s^0)/\tau(B_d^0)$ for fits with a wrongly fixed $\tau(bg)$ scale. Default is* 1.00.

$\tau(bg)$ scale	true value [ps]	avg. result [ps]	avg. error [ps]	pull mean	pull width
0.90	1.500	1.501	0.030	0.03	0.99
0.95	1.500	1.500	0.030	−0.01	0.99
1.00	1.500	1.499	0.030	−0.06	1.00
1.05	1.500	1.498	0.030	−0.08	1.00
1.10	1.500	1.497	0.030	−0.11	1.01

Table 5.46: *Toy study results for $\tau(B_s^0 \to K^+K^-)$ for fits with a wrongly fixed $\tau(bg)$ scale. Default is* 1.00.

	$\tau(B_d^0)$ [ps]	$\tau(B_s^0)/\tau(B_d^0)$	$\tau(B_s^0 \to K^+K^-)$ [ps]
statistical uncertainty	0.015	0.061	0.030
source of systematic uncertainty			
$\delta \log \mathcal{L}_K$ templates	0.002	0.001	0.003
$m(B_s^0 \to K^+K^-)$ template	0.001	0.001	0.002
combinatorial background level	0.001	0.001	0.001
avg. lifetime acceptance functions	0.003	0.006	0.002
background lifetime shape	0.003	0.003	0.002
misalignments/ proper time error	0.000	0.001	0.000
Λ_b^0 lifetime	0.000	0.000	0.000
three-body lifetime	0.000	0.000	0.000
average $B_s^0 \to K^+K^-$ lifetime	0.000	0.000	0.000
average background lifetime	0.001	0.000	0.001
total	0.005	0.007	0.005

Table 5.47: *Summary of systematic uncertainties discussed in this section.*

5.2.3.11 Summary of Systematic Uncertainties

The systematic uncertainties studied and discussed in the previous sections are summarised in table 5.47. The dominant uncertainty for the $B_s^0 \to K^+K^-$ lifetime measurement is the knowledge of the $\delta \log \mathcal{L}_K$ template. Variations in the average lifetime acceptance functions are another major source of systematic uncertainties (dominant for the B_d^0 lifetime and the lifetime ratio measurement). Future improvements in the method of extracting the average lifetime acceptance functions and eventually the knowledge of the actual distributions observed in data will allow a more reliable assessment of this uncertainty.

The last significant source of systematic uncertainties is the lifetime shape of the combinatorial background. This essentially describes the quality of the method of treating combinatorial background developed for this fitter (see section 4.3.2). Also here, the knowledge of the real distribution from data will allow a more reliable estimation of the systematic uncertainty.

The only sources of systematic uncertainties which can be expected to have a significant correlation are the level of combinatorial background and the shape of the lifetime distribution of combinatorial background. As the uncertainty assigned to the first is small this correlation is ignored and all uncertainties are added in quadrature. The

total systematic uncertainties are given in table 5.47.

Using the fit results quoted in table 5.4 and the systematic uncertainties given in table 5.47 a fit of a data sample equivalent to an integrated luminosity of 0.1 fb^{-1} would yield the following result:

$$\begin{aligned} \tau(B_d^0) &= (1.530 \pm 0.015_{stat.} \pm 0.005_{syst.}) \text{ ps} \\ \tau(B_s^0)/\tau(B_d^0) &= 0.961 \pm 0.061_{stat.} \pm 0.007_{syst.} \\ \tau(B_s^0 \to K^+K^-) &= (1.498 \pm 0.030_{stat.} \pm 0.005_{syst.}) \text{ ps}. \end{aligned} \tag{5.6}$$

The studies presented here estimate the effects of one level of lifetime biasing selection cuts and their correction within the lifetime fitting method. However, there are several levels of selections in the *LHCb* data taking, i.e. at trigger level and during the off-line selection. The selections are similar and so should be their systematic effects. But as the lifetime acceptance has to be evaluated for all selection levels, some systematic uncertainties may be slightly larger for the full fit. Since the acceptances should mostly overlap, it is expected that the change in the systematic uncertainties is small, which is therefore ignored at this stage.

The measurement of $\Delta\Gamma_s$ in the *SM* scenario, as discussed in section 5.1.1, would have roughly the same relative uncertainty as $\tau(B_s^0 \to K^+K^-)$. Hence, the size of the data set studied here would be sufficient to match the precision expected for the TeVatron experiments at 9 fb^{-1} (see section 5.1.3). To achieve a measurement of $\tau(B_s^0 \to K^+K^-)$ with a precision of 0.8%, sufficient for a 5σ measurement of $\Delta\Gamma_s$ assuming $\Delta\Gamma_s/\Gamma_s$=0.10, a data sample of 0.7 fb^{-1} is needed. According to the current *LHC* running scenarios such a data sample would be available in 2011. The next section discusses opportunities for interesting measurements with early data.

5.3 Early Physics Measurements

Lifetime measurements using $B^0_{(s)} \to h^+h'^-$ events require a reasonably long period of data taking in order to acquire a sample of sufficient size for measurements with sensitivity to New Physics. However, while B physics is clearly the main priority at *LHCb*, also decays of other particles can be studied. The cross-section for the production of $c\bar{c}$ pairs is about five times higher than that for producing $b\bar{b}$ pairs. Hence, D mesons are produced in abundance at the *LHC*. The decays of D^0 mesons into two charged hadrons have been studied due to their similarity to $B^0_{(s)} \to h^+h'^-$ decays.

Cut type	Cut value
$(IP/\sigma(IP))^h$	> 4
p_T^h *in* GeV/c	> 0.5
p^h *in* GeV/c	> 5
p_T^D *in* GeV/c	> 1.5
$L^D/\sigma(L^D)$	> 4
$(IP/\sigma(IP))^D$	< 3
$\chi^2(vertex)$	< 10
$\cos\theta$	> 0.99993

Table 5.48: *Cuts for the selection of $D^0 \to h^+h'^-$ events using the full reconstruction on events that have passed the trigger.*

$D^0 \to h^+h'^-$ decays have higher branching ratios than their B partners, however with larger differences among them. The most abundant is the Cabibbo allowed decay $D^0 \to \pi^+K^-$ with a branching ratio of about 4% [12]. The same mode with the charge conjugate final state ($D^0 \to K^+\pi^-$) is doubly Cabibbo suppressed and has a branching ratio of roughly 0.015%. Finally, the two modes with two hadrons of the same flavour are singly Cabibbo suppressed and have branching ratios of 0.4% ($D^0 \to K^+K^-$) and 0.14% ($D^0 \to \pi^+\pi^-$), respectively.

Several observables that are accessible via lifetime measurements in $D^0 \to h^+h'^-$ decays have been introduced in section 1.5. It has been pointed out that these measurements as well as others in the charm sector have a high potential for revealing New Physics [82].

5.3.1 Selecting Prompt $D^0 \to h^+h'^-$ Events

The first selection of prompt $D^0 \to h^+h'^-$ events in *LHCb* has been designed for the studies presented here. It follows the selection for $B^0_{(s)} \to h^+h'^-$ events. All cuts apart from the one on the transverse momentum of the mother particle have been loosened since the D has a shorter lifetime compared to the B and the daughter tracks carry less momentum. One cut has been added to suppress the contamination from true D decays originating from B decays: as the B decays into at least one other particle when decaying into a D, the direction of flight of the D usually differs from that of the B and hence does not point towards the PV. Therefore, a powerful cut is that on the angle θ between the D momentum and the line connecting the D decay vertex with the PV.

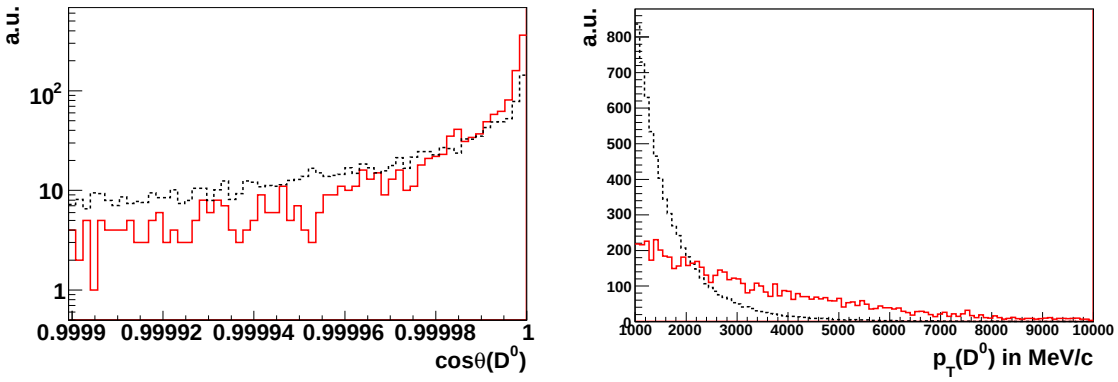


Figure 5.27: *Distributions for* $\cos\theta(D^0)$ *(left) and* $p_T(D^0)$ *(right) for* $D^0 \to \pi^+K^-$ *events (red, solid line) and combinatorial background (black, dashed line).*

The cut values are summarised in table 5.48. The cuts on the transverse momentum of the mother particle and on $\cos\theta$ are not very efficient for the signal channels, however they were found to be very powerful in suppressing background (see figure 5.27). Due to the abundance of D decays in *LHCb* such a tight selection still results in high signal rates with very low background rates.

Figure 5.28 shows the invariant mass distribution after the $D^0 \to h^+h'^-$ selection has been applied to a sample of about 1.8 million minimum bias events that have passed the *L0* trigger. The events have been reconstructed under the hypothesis that the positive track has the mass of a charged pion while the negative is a charged kaon. Excellent separation between signal and background is achieved despite the fact that no *PID* information has been exploited at this point. Within ± 25 MeV/c^2 around the nominal D^0 mass, 151 signal events are reconstructed and 39 events are identified to be background.

Unlike $B^0_{(s)} \to h^+h'^-$, the mass of charged kaons is assigned to both particles when calculating the invariant mass for all further studies. Compared to the π^+K^- hypothesis it has the advantage that it does not suppress the respective charge-conjugate decays, i.e. in $K^+\pi^-$. The distributions for $D^0 \to K^+K^-$ and $D^0 \to K^+\pi^-$ decays are close together to allow a tight mass window, while still being separated due to the good mass resolution of roughly 8 MeV/c^2. The disadvantage of using the K^+K^- hypothesis (which is also why it is not used for measurements involving $B^0_s \to K^+K^-$) is that three-body decays into final states involving particles lighter than kaons can be reconstructed with invariant masses inside the mass window. One possibility to further suppress background, which has not been exploited, is to apply a cut on the

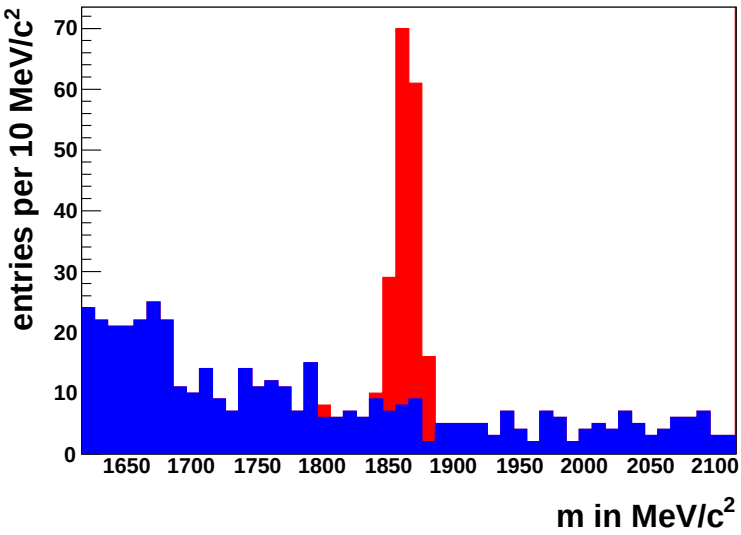


Figure 5.28: *Invariant mass distribution after application of the $D^0 \to h^+h'^-$ selection. The events have been reconstructed under the hypothesis that the positive track has the mass of a charged pion while the negative is a charged kaon. In the cumulative distribution the red area corresponds to true $D^0 \to \pi^+K^-$ decays while the blue depicts background events.*

larger of the two $\delta \log \mathcal{L}_K$ values of the daughter tracks, as this should be a kaon for both channels of interest.

Due to the lack of a full simulation for prompt three-body D decays, the contribution of these channels had to be evaluated by a generator level study. The distributions of several possible three-body D decays were studied by generating events according to a flat phase space, adding resolution effects and reconstructing the invariant mass of two charged tracks using the K^+K^- hypothesis. It was found that all distributions lie either outside the mass window or stay constrained to the lower mass region. Their fractions inside the mass window and thus their effective branching ratios inside the mass window are given in table 5.49.

The signal yields have been estimated by applying the off-line selection to a sample of so-called *L0* trigger-stripped minimum bias events, i.e. events that reflect the data sample at the input to the *HLT*. Assuming the nominal *HLT* input rate of 1 MHz and an efficiency of the *HLT* selection of roughly 1/3, the frequency of signal events written to tape can be directly calculated from the number of selected signal events. For example, three selected events in an input sample of 10^6 events would lead to a frequency of 1 Hz of recorded signal events.

Table 5.50 shows the estimated yields for the charm decays under study, when apply-

Channel	$\mathcal{BR}$ in %	fraction in mass window in %	$\mathcal{BR}_{eff}$ in %
$D^0 \to K^-\pi^+\pi^0$	14	2.16	0.30
$D^0 \to K^-e^+\nu$	3.6	6.92	0.25
$D^0 \to K^-\mu^+\nu$	3.3	6.76	0.22
$D^0 \to K_S\pi^+\pi^-$	3	0	0.00
$D^+ \to (K^-\pi^+)\pi^+$	9	1.9	0.17
$D^+ \to (K^-\pi^+)e^+\nu$	4	0.26	0.01
$D^+ \to (K^-\pi^+)\mu^+\nu$	4	0.05	0.00
$D^+ \to K^+(K^-\pi^+)$	1	0	0.00
total			0.95

Table 5.49: *Contributions from three-body background reconstructed under the K^+K^- mass hypothesis. $\mathcal{BR}_{eff}$ denotes the effective $\mathcal{BR}$ inside the mass window. The brackets for the D^+ decays denote the particles that have been used to calculate the invariant mass.*

Channel	$\mathcal{BR}$ in %	$\mathcal{BR}_{eff}$ in %	Frequency to tape in Hz
$D^0 \to \pi^+K^-$	4	2.4	20
$D^0 \to K^+K^-$	0.4	0.4	3.3
$D^0 \to \pi^+\pi^-$	0.14	0.0	0
three-body		1.2	10

Table 5.50: *Estimated yields for two-body and three-body D decays. $\mathcal{BR}_{eff}$ is defined as in table 5.49.*

ing the mass window as described in table 5.48. The three-body contribution has been slightly increased compared to the study presented above to be conservative. Other modes which have been neglected here for this initial study but which have to be included in a full simulation study are decays of the $D_s^{\pm}$ meson, e.g. into $K^+K^-\pi^{\pm}$ or $\pi^+\pi^-\pi^{\pm}$.

5.3.2 Lifetime Measurements with $D^0 \to h^+h'^-$ Events

Lifetime measurements with $D^0 \to h^+h'^-$ events offer several opportunities for interesting measurements. The earliest accessible quantity is $y_{\mathcal{CP}}$ which, as defined in equation 1.90, can be measured through a lifetime ratio measurement of $D^0 \to \pi^+K^-$ and $D^0 \to K^+K^-$ events.

The fitter described in this and the previous chapter for $B^0_{(s)} \to h^+h'^-$ events has been adapted for $D^0 \to h^+h'^-$ events. The fit parameters are the $D^0 \to \pi^+K^-$ lifetime, $y_{\mathcal{CP}}$, and a lifetime assigned to the background from three-body D decays. New templates have been obtained to describe the invariant mass distributions.

The templates for $\delta \log \mathcal{L}_K$ where taken from the $B^0_{(s)} \to h^+h'^-$ studies. Eventually, the *PID* templates will be obtained from D decays directly as they can be selected in a clean way. For the simulation studies presented here, the impact of the differences between the *PID* templates for D and B on the lifetime fits should be negligible.

5.3.2.1 Background from Secondary D Decays

An additional complication, compared to the fit to $B^0_{(s)} \to h^+h'^-$ events, arises from D decays for which the D does not originate from the PV. Since most B decays involve a D meson in their decay chain this background is a priori very high. The cuts which ensure the D direction of flight to be compatible with an origin at the PV suppress the background from secondary D decays. However, for small B lifetimes or for D decays in the direction of flight of the B mesons, secondary D decays are an indistinguishable background.

The fitter has to be adapted to account for the double decay chain of secondary D mesons. Therefore, the PDF for a simple particle decay,

$$f(t) = \frac{1}{\tau} e^{-t/\tau}, \tag{5.7}$$

has to be replaced by the convolution of two exponential decay functions

$$f(t) = \frac{1}{\tau_B - \tau_D} \left(e^{-t/\tau_B} - e^{-t/\tau_D} \right). \tag{5.8}$$

This model has been tested to be unbiased with a sample of pure secondary D decays. The assumption that allows this simple extension of the formalism is that both decays occur in the same direction such that the total measured distance of flight is the sum of the flight distances of the B and the D mesons.

One caveat arises from the fact that lifetimes cannot be measured directly. It has been discussed previously that the lifetime may be replaced by the ratio of lifetime and mass, which leaves the ratio of flight distance and momentum as a measurable observable (see section 5.1.5). However, in this case only the D momentum is known and furthermore it is not clear which part of the flight distance belongs to the B and which to the D. If the overall assumption of the sum of two exponential decays remains valid for the quantity $(L_B + L_D)/p_D$, this formula can still be used since there is no physics interest in either of the lifetimes. Alternatively, one can attempt to reconstruct the decay chain $B \to D^*(D\pi)X$, which would allow a splitting of the distance of flight by using the D^* vertex.

Using the 151 signal events selected from the MC sample mentioned in section 5.3.1 it has been found that the remaining background due to secondary D decays amounts to about 10% of the selected events. This fraction has been fixed to its correct value in the fitter.

5.3.2.2 Toy Monte Carlo Studies with $D^0 \to h^+h'^-$ Events

The sensitivity to $y_{\mathcal{CP}}$ has been evaluated using a toy MC study. 1000 data sets have been generated with $125k$ events each. The $125k$ events split up into $42k$ combinatorial background, and $83k$ signal events which are split according to table 5.50. This data set is equivalent to some 40 minutes of data taking under nominal experimental conditions or to an integrated luminosity of 0.5 pb^{-1}.

The fitter first fits the signal fractions which were then used as input to the lifetime fit. The three-body lifetime has been fixed to its true average input value such that the D lifetime, obtained from $D^0 \to \pi^+K^-$ decays, and $y_{\mathcal{CP}}$, obtained from $D^0 \to K^+K^-$ decays, were left as free parameters.

The result of a typical signal fraction fit is shown in figure 5.29. The χ^2 describing the agreement between fit model and input data is 103 for 100 contributing entries. The large fluctuations in the template for $D \to 3\text{–}body$ decays are due to low statistics at this extreme upper end of the mass distribution. No larger sample was generated as this study has to be repeated with full MC simulation for the three-body decays. The fit results of the signal fraction fitter are unbiased.

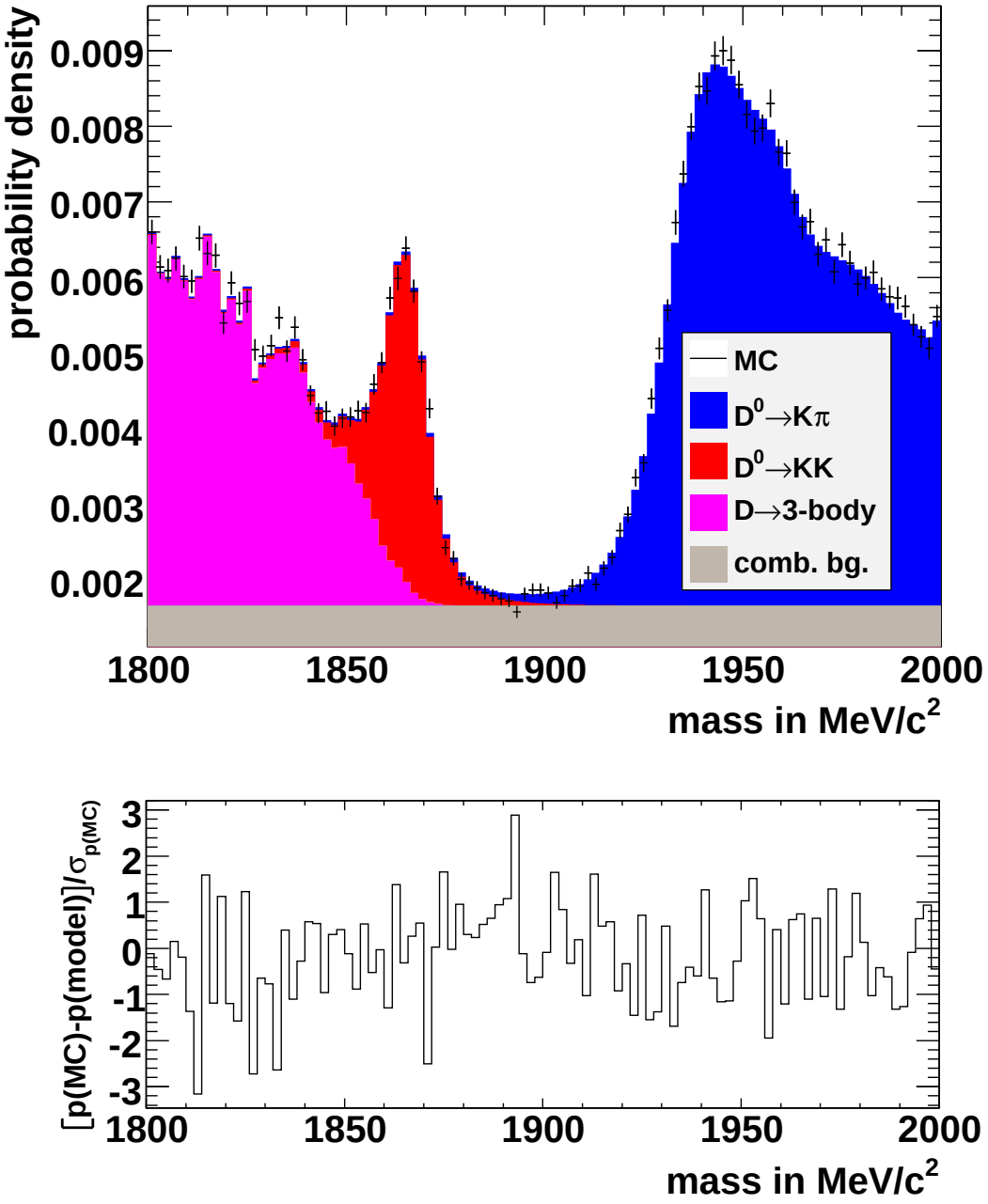


Figure 5.29: *Result of a $D^0 \to h^+h'^-$ signal fraction fit. The upper plot shows the generated distribution as crosses (MC) and the cumulative distributions of the model using the fitted signal fractions as shaded areas. The lower plot shows the difference per bin divided by the statistical error of the data sample.*

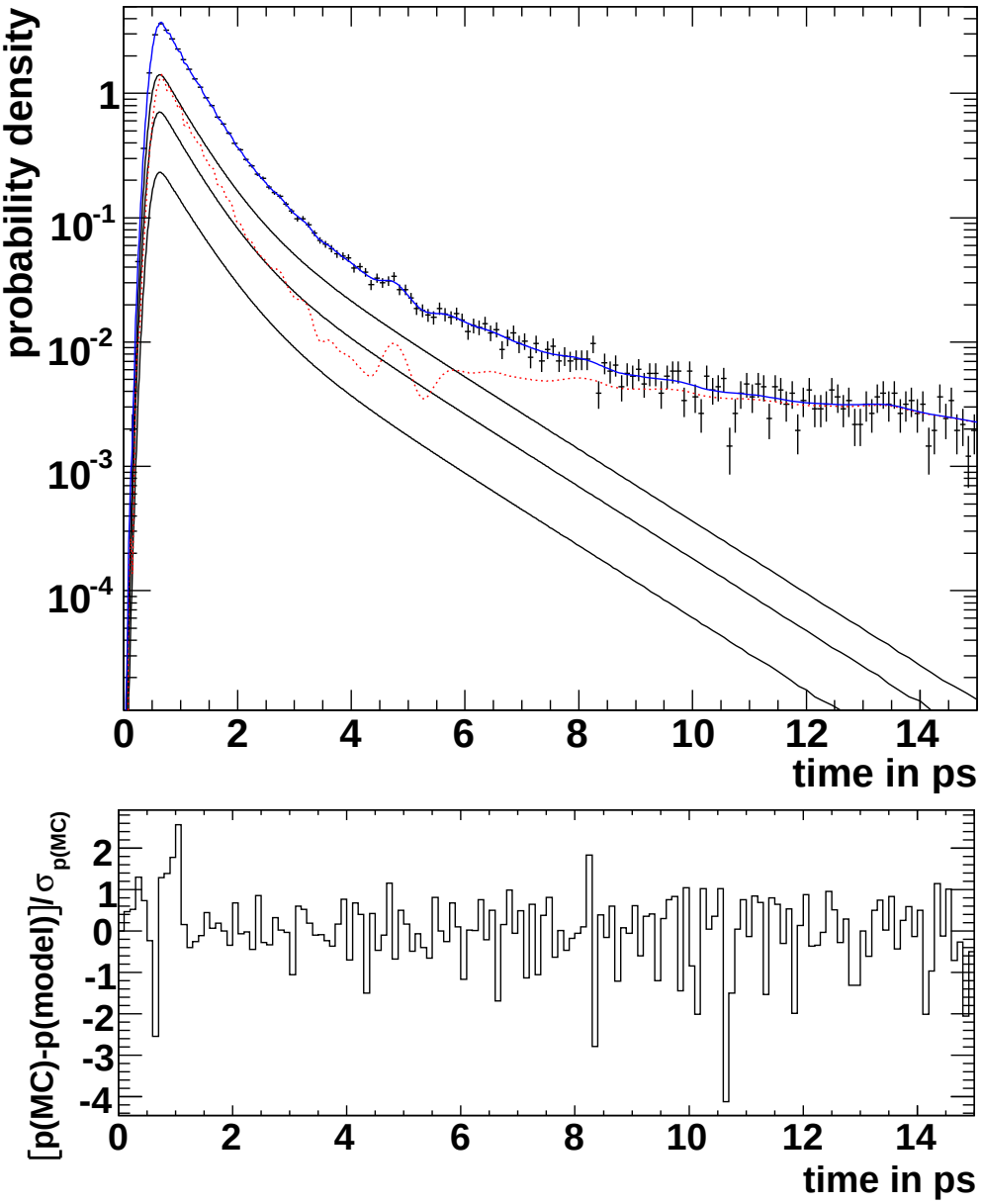


Figure 5.30: *Result of a $D^0 \to h^+h'^-$ lifetime fit. The upper plot shows the generated distribution as crosses (MC), the distributions of the signal models using the fitted lifetimes as black lines, the background distribution as a red (dashed) line, and the total distribution according to the fit results as a blue line. The lower plot shows the difference per bin between generated data and total distribution of the model divided by the statistical error of the data sample.*

An example for a lifetime fit result is shown in figure 5.30. The χ^2 describing the agreement between fit model and input data is 123 for 149 contributing entries.

In the full study, the D lifetime was measured on average as 0.593 ps with an average uncertainty of 0.004 ps. The input value was 0.500 ps. $y_{\mathcal{CP}}$ was determined as 0.057 with an error of 0.017 and an input value of 0.060.

The large bias is due to the contamination from secondary D decays. They are not distinguished from prompt D decays in the fitter such that every event is treated as having a 90% probability of being prompt and a 10% probability of being a secondary D decay. The biases originating from the prompt being treated as secondary and vice versa do not cancel. This results in the overall bias reported above.

5.3.2.3 Prospects for D Lifetime Measurements at LHCb

The presence of secondary D decays is a challenge for lifetime measurements with prompt D decays. It is beyond the scope of this thesis to study lifetime fitting with prompt D decays in more detail. However, a few remarks are given below on how to deal with this background.

Key to a successful fit with prompt D decays is a way to distinguish decays of D mesons from the PV from those originating in B decays. Any variable used for this purpose will have to exploit the fact that D mesons from B decays are not bound to point back to the PV. Hence, one possible variable for a distinction on a statistical basis is the impact parameter of the D meson. Such a technique has already been successfully exploited by the *CDF* collaboration [83].

Figure 5.31 shows the *IP* distribution of the D mesons for prompt and secondary $D^0 \to h^+h'^-$ decays after a reduced selection. The two distributions clearly differ which allows their exploitation for the statistical distinction of the two samples. The main aspect to be studied is the need for relaxation of selection cuts, particularly the cut on $\cos\theta$, in order to have access to this distribution. This can have implications on the assumption that B and D fly in the same direction when calculating the convoluted PDF (see equation 5.8).

A possible way of suppressing secondary D decays is by reconstructing the decay chain $D^* \to D\pi_s$, where π_s denotes a slow pion. Requiring that the D^* vertex is in agreement with the PV can suppress D^* decays from detached vertices. However, it remains to be studied whether the slow pion can be reconstructed with high efficiency among the numerous tracks coming from the PV. Furthermore, the vertex resolution of the D^* decay is rather poor due to the low momentum pion. Therefore, it needs to

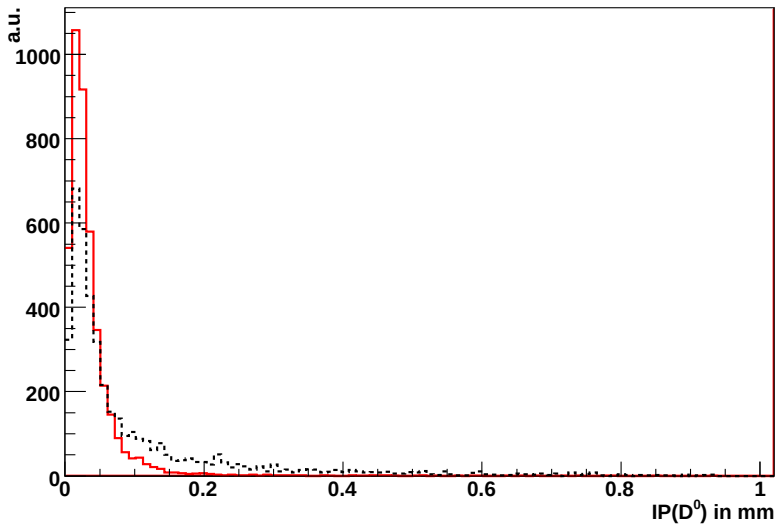


Figure 5.31: *D^0 IP distribution for prompt (red, solid line) and secondary (black, dashed line) $D^0 \to h^+h'^-$ decays.*

be evaluated whether the suppression of secondary background outweighs the loss in statistics due to the additional constraints.

A toy study similar to the one reported above has been performed to estimate the improvement necessary to achieve a precise measurement of $y_{\mathcal{CP}}$. The fraction of secondary D decays has been reduced from 10% to 1% while all other settings have been kept. With the same input values the average D lifetime result is 0.514 ps with an error of 0.002 ps. The result for $y_{\mathcal{CP}}$ is 0.056 with an average error of 0.014. This shows that the bias is significantly reduced, however, still present. Hence, a statistical treatment of prompt and secondary D decays is mandatory.

The latest HFAG average for $y_{\mathcal{CP}}$ has an uncertainty of 0.0026 [12]. To reach this level of sensitivity, the uncertainty quoted above of 0.017 has to be reduced by at least a factor of 7. Hence, the size of the data sample necessary for this measurement is equivalent to only about 30 pb^{-1}. This data sample is currently expected to be available in 2010.

5.4 Impact of Misalignments On Measurements with Two-Body Decays

An accurate and efficient tracking system is of crucial importance to the success of the *LHCb* experiment. The alignment of the tracking system is of great importance, as misalignments potentially cause losses in tracking efficiencies and, hence, physics

performance.

First studies of the deterioration of the *LHCb* tracking and (software) trigger performance due to residual misalignments in the Vertex Locator (*VELO*) were discussed in [84] [2]. A study of the consequences of a misaligned Outer Tracker on the signal and background separation of $B^0_{(s)} \to h^+h'^-$ decays can be found in [85].

Here, the effects of misalignments of both the Vertex Locator and the tracking T-stations on the analysis of the $B^0_{(s)} \to h^+h'^-$ decays are investigated. The effects on the pattern recognition performance, and also on the event selection efficiencies and reconstruction performance are described.

The next section details the implementation of misalignments and the data samples used for the study. Section 5.4.2 presents the impact of random misalignments of the tracking detectors on the analysis of $B^0_{(s)} \to h^+h'^-$ decays. Section 5.4.3 presents the impact of z-scaling effects in the *VELO*.

5.4.1 Implementation of Misalignments

5.4.1.1 Misalignment Scales

The effects of misalignments are assessed in this section as a function of their magnitude. No assumptions are made based on the quality of the metrology or the expected performance of the alignment algorithms.

The misalignment effects are looked at as a function of a "misalignment scale". The scales were chosen to be roughly 1/3 of the detector single-hit resolution – called "1σ". Misalignments were then applied to each *VELO* module and sensor, each IT box and OT layer following a Gaussian distribution with a sigma corresponding to the 1σ values (see table 5.51).

For each sub-detector 10 sets of such 1σ misalignments were generated, to avoid any potentially "friendly" or "catastrophic" set of misalignments. Likewise, this procedure was repeated with the creation of 10 similar sets for each *VELO* module and sensor and each IT box and OT layer with misalignment scales increased by factors of 3 (3σ) and 5 (5σ).

Each of these 10 misalignment sets were implemented and stored in dedicated (conditions) databases. In total 9 databases were produced, corresponding to the 1σ, 3σ and 5σ misalignments for the *VELO*, IT and OT detectors.

[2] Note that these studies relate to a rather old and obsolete version of the trigger.

Detector	Translations (μm)			Rotations (mrad)		
	Δ_x	Δ_y	Δ_z	R_x	R_y	R_z
VELO modules	3	3	10	1.00	1.00	0.20
VELO sensors	3	3	10	1.00	1.00	0.20
IT boxes	15	15	50	0.10	0.10	0.10
OT layers	50	0	100	0.05	0.05	0.05

Table 5.51: *Misalignment "1σ" scales for the VELO modules and sensors, the IT boxes and OT layers.*

5.4.1.2 Data Samples

The study was performed with a 20 k sample of $B^0_d \to \pi^+\pi^-$ events[3] for each scenario:

- perfect alignment (denoted 0σ in the rest of the note);
- 1σ, 3σ, and 5σ misalignments for the following cases:
  - *VELO* misalignments,
  - IT and OT misalignments,
  - and misalignments of *VELO*, IT and OT.

Each 20 k sample consists in reality of 10 sub-samples of 2 k events, each of which was processed with a different one of the 10 sets of a particular misalignment scenario. In addition, the effects of a systematic change in the *VELO* z-scale have also been studied.

5.4.1.3 Event Processing

All the events were generated and digitized with a perfect geometry (`Gauss` generation program version v25r8 and `Boole` digitization program version v12r10). Starting always from the same digitized data samples, the misalignments were only introduced at reconstruction level, where pattern recognition, track fitting, primary vertexing and particle identification are performed. The version `v32r2` of the `Brunel` reconstruction software was used for this task. The physics analysis was later performed with the `DaVinci` program version `v19r9`.

[3]For the sake of simplicity only one of the $B^0_{(s)} \to h^+h'^-$ family of decays was considered, as their different final states and B-mother are not relevant in the present study.

5.4.2 Impact of Random Misalignments

The effects of misalignments of the tracking detectors have been studied separately for *VELO* and T-stations as well as for their combination. It was found that the effects clearly split into those due to *VELO* misalignments and those due to T-station misalignments. Therefore, only the results of combined misalignments are presented below with remarks as to which sub-detector dominates the particular effect. A detailed account of the effects of the individual sub-detector misalignments is given in reference [86].

5.4.2.1 Effect on the Pattern Recognition

Once the misalignments are introduced at the reconstruction level, as explained in section 5.4.1, their effects need to be studied both on the pattern recognition (track finding efficiencies) and on the event selection (efficiency for finding the correct decay).

The pattern recognition algorithms[4] considered are the ones that find:

- tracks in the *VELO* detector in r-z and 3D-space. The algorithms are hereafter denoted by `VeloR` and `VeloSpace`, respectively;
- tracks that traverse the whole *LHCb* detector (called "long tracks"). The two existing long tracking algorithms are hereafter denoted `Forward` and `Matching`.

In table 5.52 the `VeloR`, `VeloSpace`, `Forward` and `Matching` pattern recognition efficiencies for all long tracks in the event with no momentum cut applied at selection level are shown for the 0σ, 1σ, 3σ and the 5σ scenarios. For the set of misalignments considered there is a relative loss of 8.6% for the `Forward` efficiency and of 12.9% for the `Matching` efficiency between the 0σ and the 5σ case. These numbers roughly correspond to the combined losses due to the misalignments applied independently in the *VELO* and in the T-stations, shown in the previous subsections. The loss of efficiency in the `Forward` pattern recognition is dominated by *VELO* misalignments as the method is based on extrapolating *VELO* tracks. As the `Matching` method combines track seeds from both *VELO* and T-stations with similar extrapolation it is no surprise that misalignments in both sub-detectors contribute in roughly equal amounts.

5.4.2.2 Effect on the Event Selection

In table 5.53 the number of selected events is shown for the different misalignment scenarios of both the *VELO* and the T-stations. If only the T-stations misalignments

[4]For more details about the definitions of the pattern recognition efficiencies see [87].

Misalignment scenario	VeloR efficiency (%)	VeloSpace efficiency (%)	Forward efficiency (%)	Matching efficiency (%)
0σ	98.0 ± 0.1	97.0 ± 0.1	85.9 ± 0.2	81.1 ± 0.2
1σ	98.0 ± 0.1	96.8 ± 0.1	85.6 ± 0.2	80.8 ± 0.2
3σ	98.0 ± 0.1	94.3 ± 0.4	83.3 ± 0.5	77.3 ± 0.7
5σ	97.8 ± 0.2	90.1 ± 1.7	78.5 ± 1.8	70.6 ± 1.9

Table 5.52: VeloR, VeloSpace, Forward *and* Matching *pattern recognition efficiencies for various misalignment scenarios of both the VELO and the T-stations.*

are considered, the loss in the number of selected events amounts to 4.2%, while in the *VELO* case, the loss in number of selected events amounts to 73.9%. It can be concluded that the 75.6% loss in number of selected events, here seen in the worst-case scenario, is mostly due to losses induced by misalignments in the *VELO*.

Studies with *VELO* misalignments only have shown that the variables that are affected most strongly are the B impact parameter significance followed by the B decay vertex χ^2 (see figure 5.32). The daughter impact parameter significance cuts showed visible effects while the other cut variables appeared to be robust against misalignments. It should be kept in mind that although the effects of misalignments on the performance of the particle identification have not been studied here, the latter is expected to be influenced mainly by T-stations misalignments.

Misalignment scenario	Number of selected events
0σ	4141 (100%)
1σ	3807 (91.9%)
3σ	2041 (49.3%)
5σ	1009 (24.4%)

Table 5.53: *Number of selected events after running the* $B^0_{(s)} \to h^+h'^-$ *selection for the different misalignment scenarios of both the VELO and the T-stations considered.*

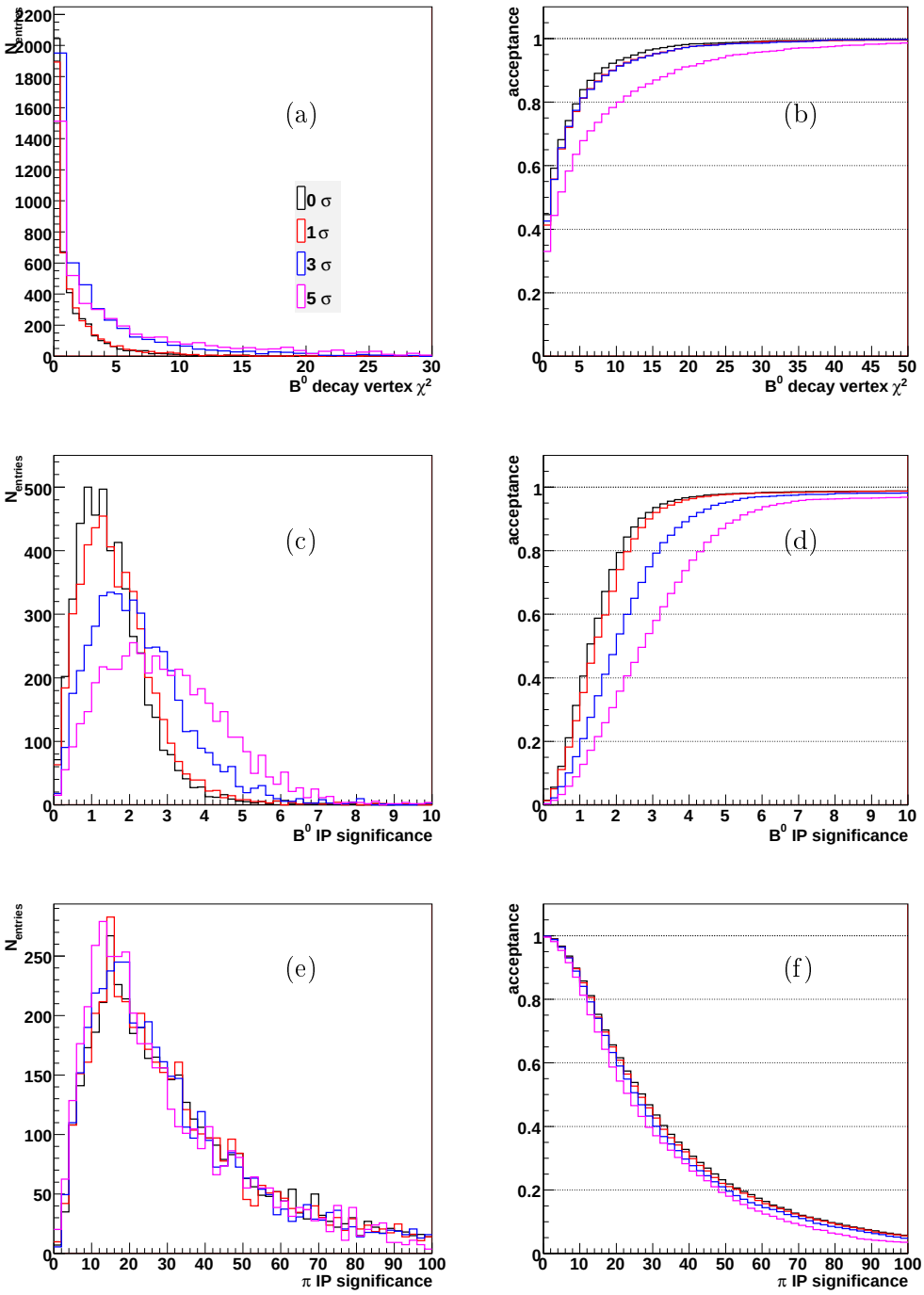


Figure 5.32: *Effect of VELO misalignments on the B^0 decay vertex χ^2 and impact parameter significances for the B^0 candidate and its daughter pions. The right-hand-side distributions correspond to the integrated left-hand-side distributions. The black line represents the 0σ misalignment scenario; the red line represents the 1σ scenario; the blue line represents the 3σ scenario and the magenta line represents the 5σ scenario.*

5.4.2.3 Effect on Resolutions

In figures 5.33 and 5.34 the B^0 daughters' momentum, B^0 mass and proper time, and primary vertex and B^0 vertex resolutions are shown for the 0σ, 1σ, 3σ and 5σ cases. The values of the resolutions (the sigmas of single-Gaussian fits) are summarised in tables 5.54 and 5.55.

Misalignment scenario	Momentum resolution (%)	Mass resolution (MeV)	Proper time resolution (fs)
0σ	0.49	22.5	37.7
1σ	0.50	22.3	40.9
3σ	0.56	25.1	58.0
5σ	0.63	25.5	78.6

Table 5.54: *Values of the resolutions on the daughters' momentum, the B^0 mass and the B^0 proper time for the different misalignment scenarios of both the VELO and the T-stations. The resolutions correspond to the sigmas of single-Gaussian fits. The errors on all resolutions are around 1-1.5 %.*

Comparing these results with the ones obtained for independent misalignments of the *VELO* and of the T-stations, it can be seen that while *VELO* misalignments strongly influence the primary and the B^0 vertex resolutions, and consequently the proper time resolution, T-stations misalignments have an effect on the daughters' momentum resolution and therefore on the B^0 mass resolution. Both misalignments have complementary effects.

Finally, the effect of misalignments on the B proper time has been studied (see table 5.56). figure 5.35 shows the distribution of the proper time error for the various misalignment scenarios as well as the respective pull distributions. A bias is observed in the estimation of the proper time, and the proper time errors are under-estimated. The degradation seen here is dominated by the misalignments in the *VELO*.

Misalignment scenario	Primary vertex (resolutions μm)			B^0 vertex resolutions (μm)		
	x	y	z	x	y	z
0σ	9	9	41	14	14	147
1σ	10	10	48	15	15	159
3σ	14	17	84	20	21	214
5σ	23	27	153	26	31	260

Table 5.55: *Values of the position resolutions on the primary and the B^0 decay vertices for the different misalignment scenarios of both the VELO and the T-stations. The errors on all resolutions are around 1-2 %.*

Misalignment scenario	Mean	Sigma
0σ	0.06 ± 0.02	1.14 ± 0.01
1σ	0.05 ± 0.02	1.22 ± 0.02
3σ	0.11 ± 0.04	1.63 ± 0.03
5σ	0.15 ± 0.07	2.10 ± 0.06

Table 5.56: *Values for the mean and sigma of the proper time pulls for the different misalignment scenarios of both the VELO and the T-stations.*

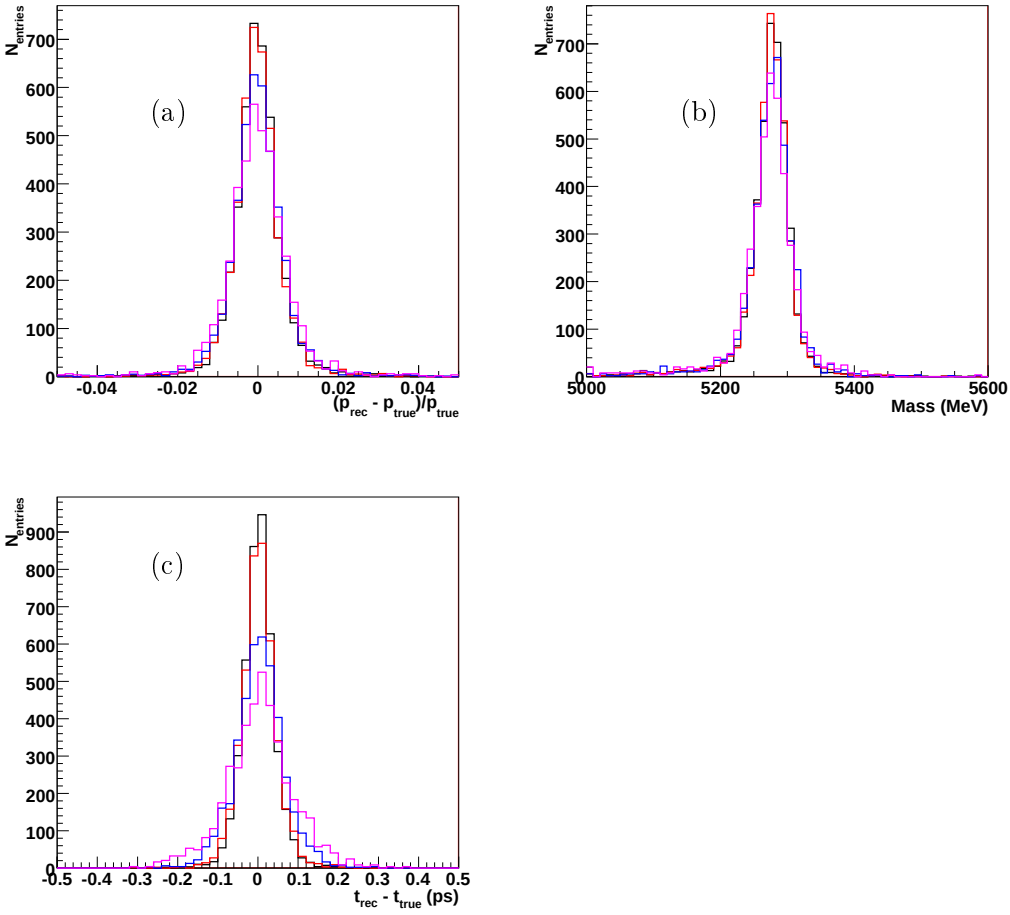


Figure 5.33: *Effect of VELO and T-stations misalignments on the resolutions in (a) momentum of the daughter pions, in (b) B^0 invariant mass and in (c) B^0 proper time. The various line styles are as explained in figure 5.32.*

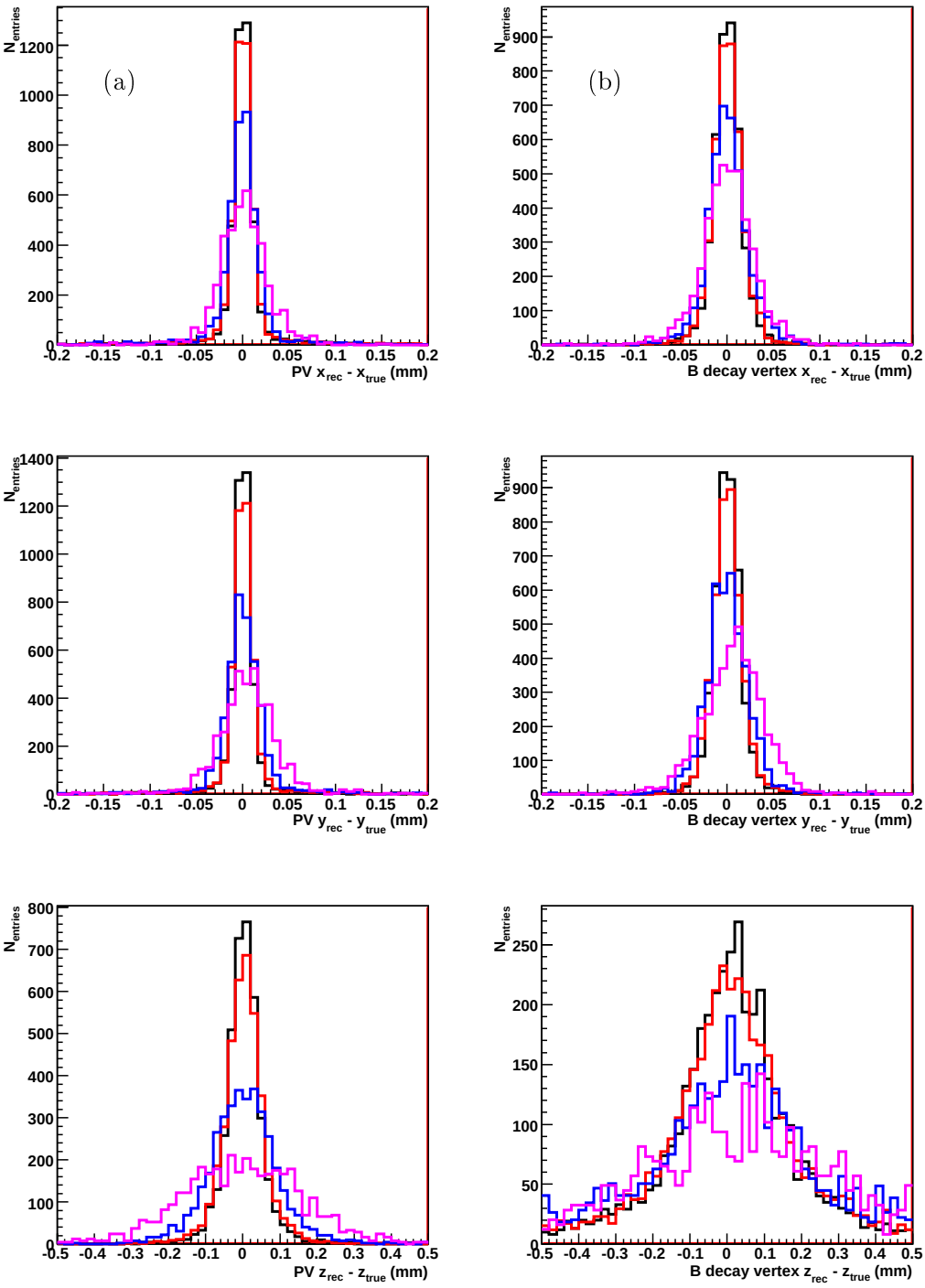


Figure 5.34: *Effect of VELO and T-stations misalignments on the resolutions of the (a) primary vertex and (b) the B^0 vertex. The plots show from top to bottom the x, y, and z component of the vertex resolutions. The various line styles are as explained in figure 5.32.*

5.4.3 Impact of a Systematic VELO z-Scaling

In addition to studying the effects of random misalignments, the change of the *VELO* z-scale has been examined. This is of particular interest to lifetime measurements as it potentially directly introduces a bias in the measured proper time.

A z-scaling effect could be expected from an expansion due to temperature variations of the *VELO* components, particularly the Aluminium base plate onto which the individual modules are screwed. However, the base plate is kept constant at 20°C by additional local heating. In addition, the scaling should be limited by the carbon-fibre constraint system that keeps the modules in place with a precision of 100 μm and which is less prone to temperature-induced expansion given its material[5].

To assess the influence of an incorrect knowledge of the *VELO* z-scale, four scenarios with different z-scales have been simulated and studied. For each scenario the z-position of each module has been changed according to the equation

$$z_{module} \rightarrow z_{module} \cdot (1 + scale), \tag{5.9}$$

where *scale* takes the four values $\frac{1}{3}10^{-4}$, 10^{-4}, $\frac{1}{3}10^{-3}$, and 10^{-3} for the four scenarios, respectively.

5.4.3.1 Effect on the Pattern Recognition

The first quantities to be studied with a changed *VELO* z-scale were the pattern recognition efficiencies. As shown in table 5.57 no deterioration has been observed up to a change in the z-scale of $1/3 \times 10^{-3}$. This is expected for the *VELO*-based pattern recognitions as a z-scaling effectively only changes the track slopes. For the largest z-scaling under study small losses in the *VELO*-based pattern recognition efficiencies are observed. These also propagate to the `Forward` and `Matching` efficiencies.

5.4.3.2 Effect on the Event Selection

When studying the influence of various z-scales on the event selection the situation observed for the pattern recognition performances repeats itself. The overview of the number of selected events is given in table 5.58. The first four scales under study show only a minor loss in the number of selected events, while a relative loss of about 20%

[5] A conservative estimate using a temperature change of 10 K yields a scaling in the z-direction of 2×10^{-5}. The 10 K is estimated as a maximal change in the temperature of the constraint system as it has a large area contact to the base plate at 20°C and only a small cross-section with the *VELO* modules at about −5°C.

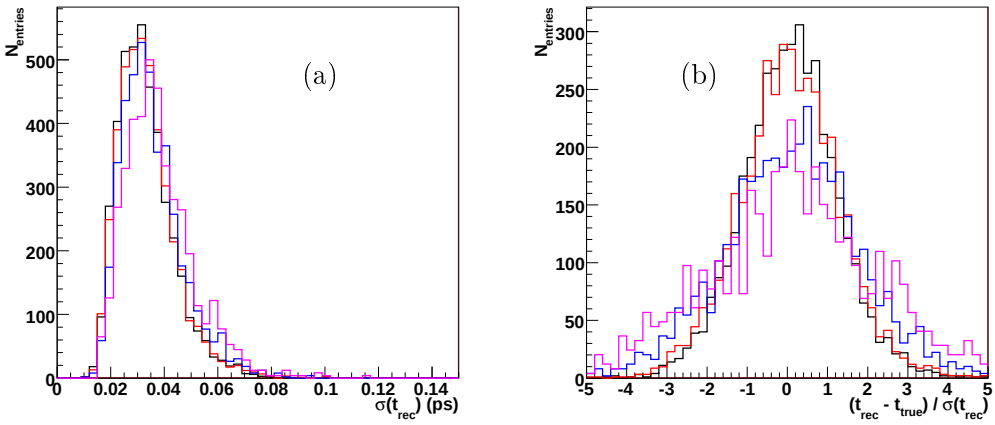


Figure 5.35: *Effect of VELO and T-stations misalignments on (a) the B^0 proper time error and (b) on the pull distribution of the B^0 proper time. The various line styles are as explained in figure 5.32.*

z-scale	`VeloR` efficiency (%)	`VeloSpace` efficiency (%)	`Forward` efficiency (%)	`Matching` efficiency (%)
1.00000	98.0	97.0	85.9	81.1
1.00003	98.0	97.0	85.9	81.2
1.00010	98.0	97.0	85.9	81.2
1.00033	98.0	96.8	85.7	81.0
1.00100	96.5	94.3	83.8	79.0

Table 5.57: `VeloR`, `VeloSpace`, `Forward` *and* `Matching` *pattern recognition efficiencies for the various VELO z-scaling misalignment scenarios.*

is observed for the largest z-scale. As for the studies in the previous chapters, this is due to a worsening in the resolution of the various cut parameters, where particularly the *VELO*-related quantities have shown great sensitivity.

z-scale	Number of selected events
1.00000	4141 (100.0%)
1.00003	4137 (99.9%)
1.00010	4142 (100.0%)
1.00033	4063 (98.1%)
1.00100	3273 (79.0%)

Table 5.58: *Number of selected events after running the $B^0_{(s)} \to h^+h'^-$ selection for the various VELO z-scaling misalignment scenarios.*

5.4.3.3 Effect on Resolutions

The effect of an incorrectly known *VELO* z-scale on the resolutions of various physics quantities is summarised in tables 5.59 and 5.60. The relevant resolution distributions are pictured in figures 5.36 and 5.37.

For the first three z-scaling scenarios the observed changes in the resolutions are minimal. Only for the two largest z-scaling cases one observes a sizeable deterioration in particular of the proper time and vertex resolutions.

Looking at the pull distributions for the reconstructed proper time shown in figure 5.38 and their summary in table 5.61, it appears that there is no significant change in the proper time bias due to a change in the z-scale. This is expected as, even for the largest z-scale under study, the estimated effect on the pull mean is of the order of its uncertainty.

5.4.4 Summary of Misalignment Effects

It has been shown above how misalignments of various sizes impact the physics observables used in the analyses presented here. The evaluation of the quality of the *VELO* alignment as presented in chapter 3 confirms that random misalignments will be constrained to the equivalent of the 1σ level of the studies presented here. Hence, no

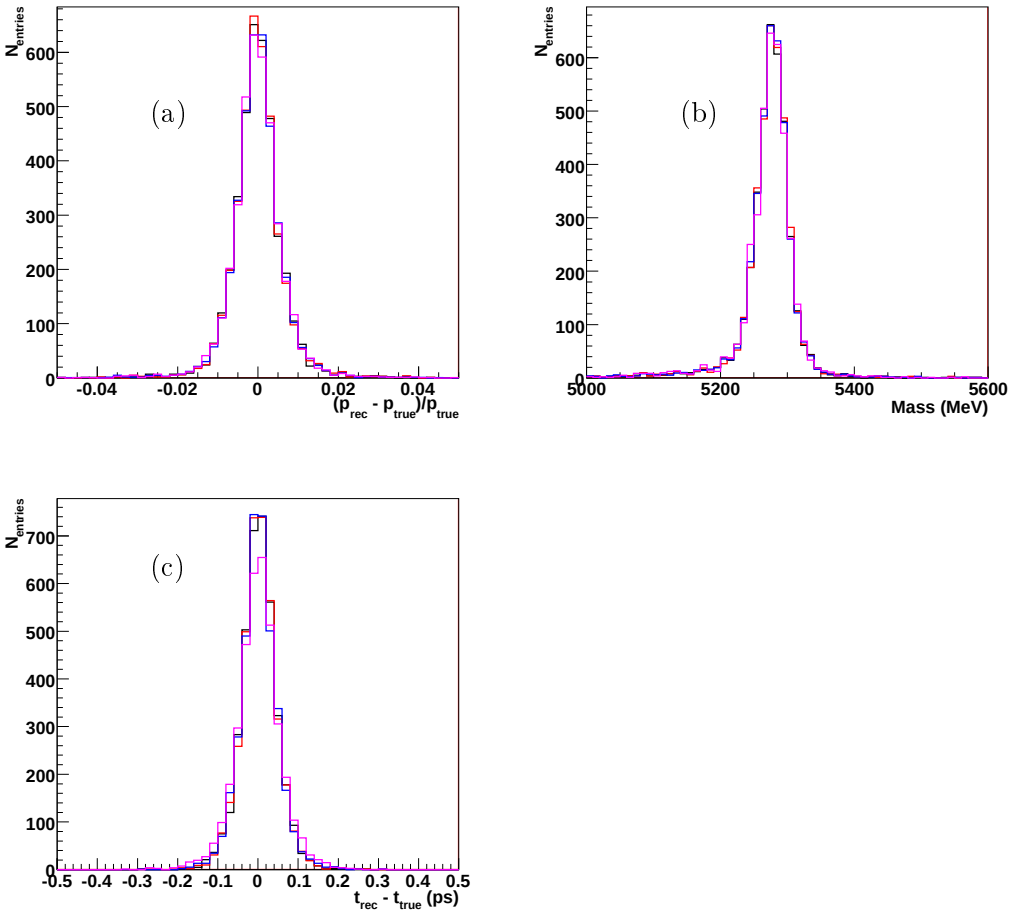


Figure 5.36: *Effect of VELO z-scaling misalignments on the resolutions in (a) momentum of the daughter pions, (b) B^0 invariant mass and in (c) B^0 proper time. The black line corresponds to a z-scale of* 1.00003*; the red line corresponds to a z-scale of* 1.00010*; the blue line corresponds to a z-scale of* 1.00033 *and the magenta line corresponds to a z-scale of* 1.00100*.*

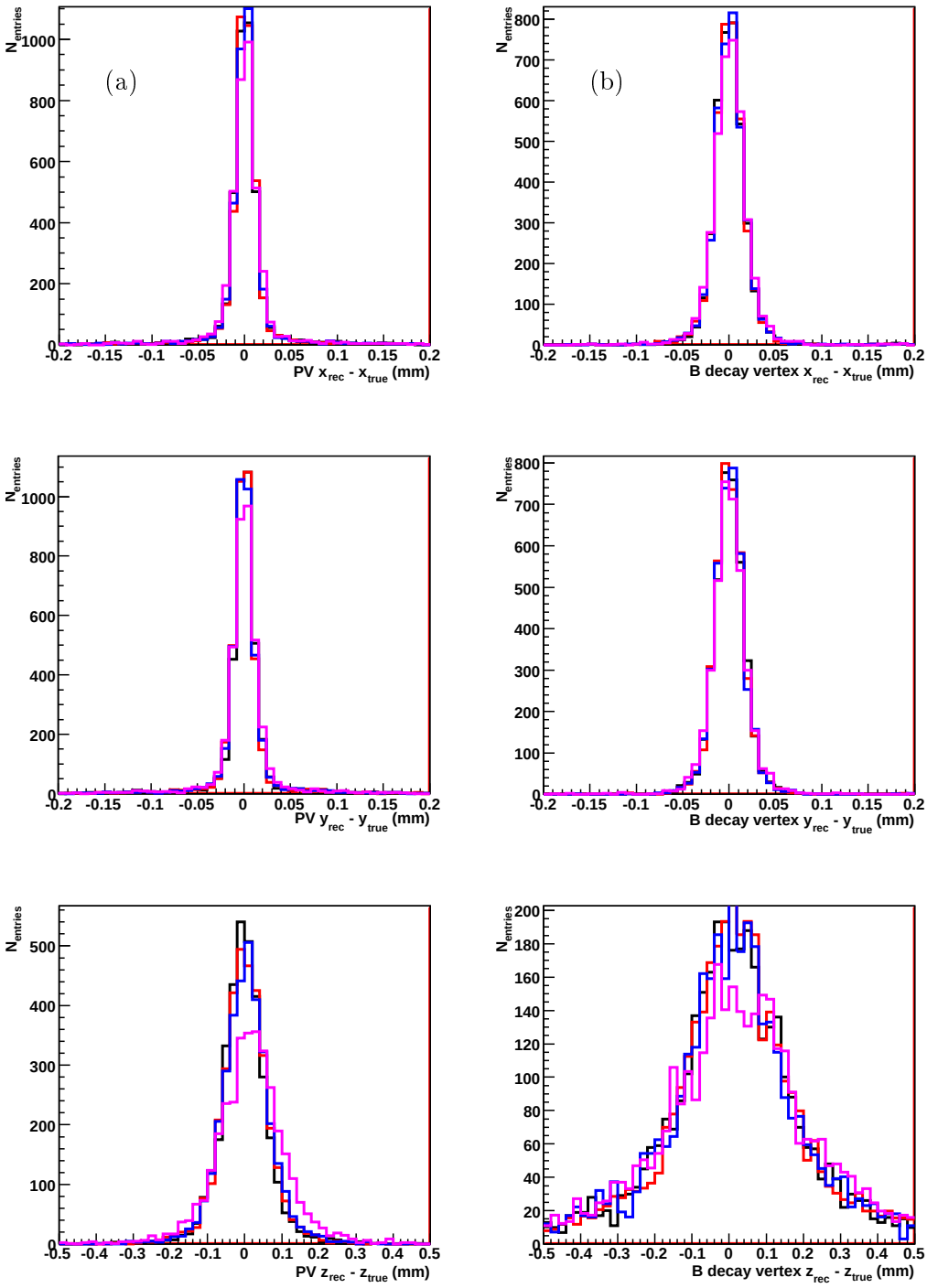


Figure 5.37: *Effect of VELO z-scaling misalignments on the resolutions of the (a) primary vertex and (b) the B^0 vertex. The plots show from top to bottom the x, y, and z component of the vertex resolutions. The various line styles are as explained in Figure 5.36.*

z-scale	Momentum resolution (%)	Mass resolution (MeV)	Proper time resolution (fs)
1.00000	0.49	22.5	37.7
1.00003	0.49	22.2	37.7
1.00010	0.49	22.1	37.7
1.00033	0.49	22.0	38.5
1.00100	0.50	22.0	46.8

Table 5.59: *Values of the resolutions on the daughters' momentum, the B^0 mass and the B^0 proper time for the various VELO z-scaling misalignment scenarios. The resolutions correspond to the sigmas of single-Gaussian fits. The errors on all resolutions are around 1-1.5 %.*

z-scale	Primary vertex resolutions (μm)			B^0 vertex resolutions (μm)		
	x	y	z	x	y	z
1.00000	9	9	41	14	14	147
1.00003	9	9	42	14	14	147
1.00010	9	9	42	14	14	145
1.00033	9	9	46	14	14	149
1.00100	11	11	72	16	15	184

Table 5.60: *Values of the resolutions of the primary and the B^0 decay vertices for the various VELO z-scaling scenarios. The errors on all resolutions are around 1-1.5 %.*

z-scale	Mean	Sigma
1.00000	0.06 ± 0.02	1.14 ± 0.01
1.00003	0.06 ± 0.02	1.15 ± 0.01
1.00010	0.07 ± 0.02	1.15 ± 0.02
1.00033	0.07 ± 0.02	1.15 ± 0.01
1.00100	0.05 ± 0.02	1.35 ± 0.02

Table 5.61: *Values for the mean and sigma of the proper time pulls for the various VELO z-scaling scenarios.*

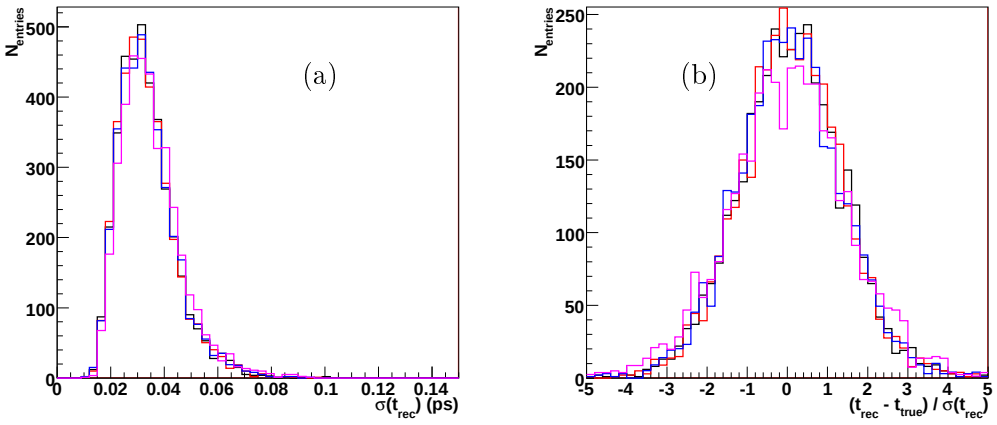


Figure 5.38: *Effect of VELO z-scaling misalignments on (a) the B^0 proper time error and (b) on the pull distribution of the B^0 proper time. The various line styles are as explained in Figure 5.36.*

significant systematic effect on lifetime measurements is expected due to any remaining random misalignments (see section 5.2.3.6). The only sizeable effect at the 1σ level is the reduction in the number of events, however the event selection will be tuned on the actual distributions measured in data, hence avoiding such dramatic losses.

Concerning systematic effects from wrong z-scales in the *VELO*, a scale of 1.0001 is a conservative limit as discussed above. No significant effects on resolutions or biases are expected at this scale. In summary, any residual misalignments in *LHCb* have negligible effects on lifetime measurements.

5.5 Conclusion

This chapter has described various opportunities for extracting physics observables based on lifetime measurements. It has been shown that $B_s^0 \to K^+K^-$ is an excellent channel to extract $\Delta\Gamma_s$ in a *SM* scenario. In the presence of *NP* the extracted value of $\Delta\Gamma_s$ will change significantly. Using external input, it will be possible to constrain a $\mathcal{CP}$ violating *NP* phase in $B_s^0 \to K^+K^-$ decays.

The lifetime fitter, described in the previous chapter, has been tested with simulation data and its potential for high precision lifetime measurements demonstrated. Using the fit results quoted in table 5.4 and the systematic uncertainties given in table 5.47 a fit of a data sample equivalent to an integrated luminosity of 0.1 fb^{-1} would yield the following result:

$$\begin{aligned} \tau(B_d^0) &= (1.530 \pm 0.015_{stat.} \pm 0.005_{syst.}) \text{ ps} \\ \tau(B_s^0)/\tau(B_d^0) &= 0.961 \pm 0.061_{stat.} \pm 0.007_{syst.} \\ \tau(B_s^0 \to K^+K^-) &= (1.498 \pm 0.030_{stat.} \pm 0.005_{syst.}) \text{ ps.} \end{aligned} \tag{5.10}$$

A competitive measurement of $\Delta\Gamma_s$ would require a data set equivalent to about 0.7 fb^{-1} of integrated luminosity.

With an integrated luminosity of only about 0.03 fb^{-1} it will be possible to make a competitive measurement of the D mixing parameter $y_{\mathcal{CP}}$ using prompt $D^0 \to h^+h'^-$ decays. A first event selection for prompt $D^0 \to h^+h'^-$ decays and an extension of the lifetime fitter for $B^0_{(s)} \to h^+h'^-$ decays to $D^0 \to h^+h'^-$ decays have been presented. The major hurdle for this measurement is suppressing the contribution from secondary D decays. Possible solutions have been discussed.

Finally, a study of the impact of misalignments of the tracking system has been presented. From its results it can be concluded that the remaining misalignments

after application of the alignment algorithms will not have any deteriorating effects on lifetime measurements. The results of this study are applicable to other measurements such as those of $\mathcal{CP}$ asymmetries.

6 Conclusion and Outlook

> Sobald jemand in einer Sache Meister geworden ist, sollte er in einer neuen Sache Schüler werden.
>
> *Gerhard Hauptmann*

6.1 Summary

Lifetime measurements form a sector of heavy flavour physics which offers a variety of interesting opportunities. Measurements of lifetime ratios allow precision tests of the *HQE* formalism. The determination of $\Delta\Gamma_s$ from a lifetime measurement of $B_s^0 \to K^+K^-$ gives access to *NP* effects. This was the main topic of the thesis presented here.

The *VELO* alignment was discussed in detail. A novel algorithm for the relative alignment of the *VELO* sensors has been presented. The high precision of the three alignment algorithms was confirmed with data from test beam and *LHC* injector commissioning runs. It has been shown that the remaining misalignments have no deteriorating effect on the measurements discussed in this thesis.

Chapter 1 gave an overview of the theory of the *SM* in general and lifetime measurements in particular. It discussed in detail how to interpret a lifetime measurement of $B_s^0 \to K^+K^-$ events. The final state of this channel is nearly $\mathcal{CP}$ even while, at the same time, it is loop dominated and thus sensitive to *NP* particles altering observables from their *SM* values. With these features, $B_s^0 \to K^+K^-$ is unique among the channels accessible in the first years of *LHCb* data taking. $\Delta\Gamma_s$ is extracted by the simple comparison of the lifetime measured in $B_s^0 \to K^+K^-$ decays to the $\mathcal{CP}$ averaged B_s^0 lifetime. The influence of *SM* and *NP* $\mathcal{CP}$ violating phases was discussed.

In addition, the differences between lifetime measurements of B and D mesons were demonstrated. The D system gives access to $\mathcal{CP}$ violation observables via simple lifetime ratio measurements. The formalism leading to the main observable, $y_{\mathcal{CP}}$, was presented.

The second chapter presented the accelerator facilities at *CERN*, from the proton source to the *LHC*, and the *LHCb* experiment. The *LHCb* detector is specialised on measuring particles from heavy flavour decays produced in high energy hadron collisions. In its forward geometry, it exploits the fact that $b\overline{b}$ pairs are produced co-linearly, close to the direction of either proton beam.

The sub-detectors of *LHCb* were introduced with a special focus on the *VELO*. Its layout with a set of r and ϕ measuring semi-circular silicon sensors follows the requirements for fast (r-z) tracking in the trigger within the constraints of the *LHCb* geometry. The high precision of the *VELO*, with a single hit resolution of better than 10 μm for the smallest strip pitches, is necessary to distinguish the secondary decay vertices from the primary collision point already at trigger level. Many analyses also benefit from the excellent proper time resolution of about 40 fs.

Central to achieving a good detector performance is the calibration. For most detectors, this includes spatial alignment. The alignment strategy for the *VELO* was discussed in chapter 3. It is split in three stages: the relative alignment of the sensors, the relative alignment of the modules, and the relative alignment of the two *VELO* halves. The sensor alignment is based on fitting residual distributions while the other two steps use the `MILLEPEDE` algorithm based on a linear track fit.

The algorithms were explained in detail and their performance evaluated with test beam data. They were found to have a precision of about 2 μm, which is significantly below any single hit precision and thus does not have any deteriorating effects on physics performance. Commissioning runs of the injector line to the *LHC* in 2008 allowed a test of the full *VELO* after its installation in the *LHCb* experiment. The tests exploited muons coming from a beam stopper around 340 m away from the experiment. They confirmed the functioning of the alignment procedure and yielded a first set of alignment constants with a precision of roughly 5 μm.

A repetition of the injector commissioning runs in June 2009 yielded a new and, compared to 2008, significantly larger data set of tracks recorded by the *VELO*. The increase in the number of tracks allowed the sensor alignment algorithm to be run (see section 3.2.1) for the first time on data from the fully assembled *VELO*. Figure 6.1 shows a comparison of the residual distribution as a function of ϕ for sensor 18 before and after application of the alignment algorithm. It shows the improvement achieved in the first application of the sensor alignment algorithm to this data set. The remaining misalignment is only of the level of a few μm and will be resolved with a larger data set available from collision data after the start of the *LHC*.

In addition, the *VELO* was operated for the first time in a nearly closed position.

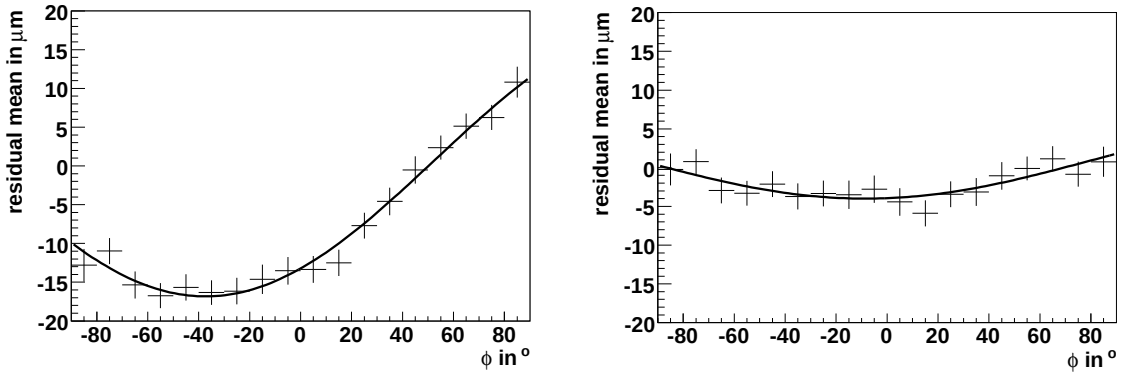


Figure 6.1: *Residual distribution as a function of ϕ for sensor* 18 *before (left) and after (right) application of the sensor alignment algorithm to data taken at the injector line commissioning in June 2009.*

This allowed the first measurement of the relative alignment of the two *VELO* which confirmed the metrology results.

Chapters 4 and 5 introduced a fitter for lifetime measurements and its application to *LHCb*, respectively. It was discussed how lifetime measurements in hadronic channels are affected by a bias caused in the *HLT*. The method discussed removes this bias by determining an event-by-event lifetime acceptance function and accounting for it in the fitting stage. This approach benefits from a direct interface to the *HLT* software and is fully independent of any MC simulation.

The fitter based on this method uses a two-stage un-binned maximum likelihood fit. In the first stage the signal fractions of the channels involved are fitted using distributions of the invariant mass and *PID*. The second stage is the actual lifetime fit. Its strength is that it does not rely on a parametrised model for the lifetime distribution of combinatorial background. This is determined by subtracting the measured signal distributions from the total distribution taking into account acceptance effects (see section 4.3.2). The shape of the various distributions is measured by applying a kernel method which guarantees a smooth distribution even on a small data set.

The main aim of this thesis was to study the measurement of $\Delta\Gamma_s$ with $B_s^0 \to K^+K^-$ decays. Section 5.1 described the extraction of $\Delta\Gamma_s$ from a lifetime measurement using $B_s^0 \to K^+K^-$ decays. The sensitivity to *NP* mixing and $\mathcal{CP}$ violating phases was discussed. The approach for a lifetime fit in the environment of $B_{(s)}^0 \to h^+h'^-$ decays was presented, including measurements of the B_d^0 lifetime and the B_s^0 to B_d^0 lifetime ratio in addition to the main measurement of the $B_s^0 \to K^+K^-$ lifetime.

The lifetime fitter was tested extensively with toy MC simulation data. It was found to perform unbiased fits of the B_d^0 lifetime, the $B_s^0 \to K^+K^-$ lifetime, and of the

B_s^0 to B_d^0 and Λ_b^0 to B_d^0 lifetime ratios. Tests of potential systematic effects revealed small systematic uncertainties. The dominant uncertainty for the $B_s^0 \to K^+K^-$ lifetime measurement is the knowledge of the $\delta \log \mathcal{L}_K$ template. Variations in the average lifetime acceptance functions are another major source of systematic uncertainties.

Using the fit results quoted in table 5.4 and the systematic uncertainties given in table 5.47 a fit of a data sample equivalent to an integrated luminosity of 0.1 fb^{-1} would yield the following result:

$$\begin{aligned} \tau(B_d^0) &= (1.530 \pm 0.015_{stat.} \pm 0.005_{syst.}) \text{ ps} \\ \tau(B_s^0)/\tau(B_d^0) &= 0.961 \pm 0.061_{stat.} \pm 0.007_{syst.} \\ \tau(B_s^0 \to K^+K^-) &= (1.498 \pm 0.030_{stat.} \pm 0.005_{syst.}) \text{ ps.} \end{aligned} \tag{6.1}$$

A competitive measurement of $\Delta\Gamma_s$ would require a data set equivalent to about 0.7 fb^{-1} of luminosity.

Due to the copious production of D mesons at the *LHC* and the high branching ratio of $D^0 \to h^+h'^-$ decays, these channels are prime candidates for early measurements at *LHCb*. The potential for lifetime measurements in this group of decays has been studied in section 5.3. With an integrated luminosity of only about 0.03 fb^{-1} it will be possible to make a competitive measurement of the D mixing parameter $y_{\mathcal{CP}}$ using prompt $D^0 \to h^+h'^-$ decays. A first event selection for prompt $D^0 \to h^+h'^-$ decays and an extension of the lifetime fitter for $B_{(s)}^0 \to h^+h'^-$ decays to $D^0 \to h^+h'^-$ decays have been presented. The major hurdle for this measurement is the suppression of the contribution from secondary D decays. Possible solutions have been discussed.

Finally, a study of the impact of misalignments of the tracking system has been presented. From these results it can be concluded that the remaining misalignments after application of the alignment algorithms will not have any deteriorating effects on lifetime measurements. Furthermore, the results of this study are applicable to other measurements such as those of $\mathcal{CP}$ asymmetries.

6.2 Outlook

The experiments at the *LHC* will have a great influence on flavour physics. Predominantly, this will be due to measurements of *LHCb* but also the *ATLAS* and *CMS* experiments will contribute with measurements of heavy flavour decays involving muons.

At the end of the first phase of the *LHC*, *LHCb* will have collected data with an integrated luminosity of 10 fb^{-1}. With this amount of data the *CKM* angle γ will be

constrained to $2^\circ - 3^\circ$ from measurements with tree level decays and to better than 5° with measurements from loop dominated decays. The weak mixing phase β_s will be measured to better than 0.01 using the full amount of $B_s^0 \to J/\psi\phi$ decays. The sensitivity on $\Delta\Gamma_s$ should be better than 0.005 ps^{-1}, i.e. about a factor of 30 better than its expected value.

Other areas of significant impact of *LHC* measurements will be rare decay searches which will reach sensitivities below the *SM* expectations, e.g. for $B_s^0 \to \mu^+\mu^-$. Measurements of radiative decays will increase existing data sets for exclusive $b \to s\gamma$ transitions by orders of magnitudes and will hence allow precise tests of *NP* models. Measurements of angular observables in the decay $B_d^0 \to K^*\mu^+\mu^-$ will yield precise information on their *SM* conformity or on the *NP* models involved.

Finally, the total data sample of D decays at the end of the first phase of *LHCb* will yield significant improvements in the $\mathcal{CP}$ violation and mixing measurements in this sector. This measurements will most likely be systematically limited at the per mill level. Searches for rare charm decays will be greatly advanced compared to their current limits.

6.2.1 LHCb Upgrade

Despite the significant increase with respect to existing data sets, particularly in the B_s^0 sector, the *LHC* will leave open questions. For example, *NP* influence on the *CKM* angle γ in loop decay measurements can only be established if it differs by significantly more than 20° from the *SM* measurement. For *NP* effects of only a few degrees a significantly larger data sample is needed.

The decay $B_s^0 \to \phi\phi$ is a loop mediated process that is highly sensitive to *NP* effects. During the nominal data taking period, *LHCb* expects to collect only $15.5k$ events of this mode. The measurement of a non-zero value of the sine-term of the time-dependent $\mathcal{CP}$ violation, $S(\phi\phi)$, would be an unambiguous sign of *NP*. A data set of several $100k$ of $B_s^0 \to \phi\phi$ events would allow a measurement of $S(\phi\phi)$ at the per-cent level [88].

An upgrade of the *LHCb* experiment is planned in order to reach the sensitivities needed to resolve the above issues. The measurements mentioned have in common that they involve fully hadronic final states. An upgrade to collect significantly more data inevitably involves a higher instantaneous luminosity. For *LHCb* this means that the number of visible interactions increases from around 1 to roughly 4 per bunch crossing. For these decays the trigger has to become more efficient to achieve the data reduction needed. This unavoidably involves the ability to select displaced vertices at the first trigger level.

The *LHCb* upgrade is currently in its design phase. It is planned to run at an average luminosity of $1-2\times10^{33}cm^{-2}s^{-1}$, i.e. $5-10$ times the current design luminosity. The trigger requirements will be met by reading out all sub-detectors at the full *LHC* bunch crossing rate of 40 MHz. Hence, one of the main challenges of the upgrade project is the design of new readout electronics that is able to cope with this high rate. Apart from this, some sub-detectors need to redesign their section close to the beam pipe to reduce the occupancy per channel.

It should be noted that the *LHCb* upgrade is rather independent of the planned upgrade of the *LHC*. Since the anticipated luminosity for the *LHCb* upgrade is below the design luminosity for the first phase of the *LHC*, the *LHCb* upgrade can be started prior to that of the *LHC*. However, as the installation of the new components require a significant amount of time the upgrade is ideally aligned with similar activities at the other experiments or longer technical shutdown periods of the accelerator.

6.2.2 Future Directions in Flavour Physics

Measurements of kaon decays have significantly contributed to experimentally constraining the *CKM* triangle for a long time. However, two channels that yield extremely high precision measurements sensitive to *NP* effects remain to be measured. They are the extremely rare decays of a neutral or charged kaon into a neutral or charged pion and a neutrino anti-neutrino pair [89]. The decays have *SM* branching ratios of 9×10^{-11} for $K^+\to\pi^+\nu\overline{\nu}$ and 3×10^{-11} for $K_L\to\pi^0\nu\overline{\nu}$. Thanks to the very good theoretical precision on these predictions of $<3\%$ and 1%, respectively, *NP* sensitivity can be obtained by branching ratio measurements alone. The *NA62* experiment is currently under construction at *CERN* and seeks to collect about 100 $K^+\to\pi^+\nu\overline{\nu}$ decays in a two-year data taking period (assuming the *SM* branching ratio). The even more challenging measurement of $K_L^0\to\pi^0\nu\overline{\nu}$ is planned at J-PARC [90].

A topic that will move to the centre of attention of future experiments is lepton flavour violation (*LFV*). It will already be studied at *LHCb* with decays like $\tau\to\mu\mu\mu$ or $\tau\to\mu\phi$. However, the sensitivity to these extremely rare processes will be very limited. The B factory experiments *BaBar* and *Belle* have already performed searches for these decays. The upgrade of the *Belle* experiment [91] and of the *KEK-B* accelerator and potentially the proposed *Super-B* factory in Italy [92] will intensify these searches and achieve significantly better sensitivity than *LHCb*.

Other decays that are only accessible to B factories are fully leptonic channels like $B^\pm\to\tau^\pm\nu$. These are of high theoretical interest as leptonic final states are free of

hadronic corrections and hence these processes can be calculated with high precision.

Apart from the *LFV* decays mentioned above, radiative lepton decays are highly sensitive to *NP* effects. The decay $\mu \to e\gamma$ is currently being searched for by the *MEG* experiment [93] which anticipates a sensitivity to the branching ratio of 10^{-14}. An experiment at J-PARC is foreseen to reach a sensitivity of up 10^{-16}. Despite the *SM* prediction being 10^{-55} these experiments have a realistic chance of observing this decay since several *NP* models involve an enhancement of the branching ratio to the level of the experimental sensitivity. The same effects are also studied in $\mu \to e$ conversions in muonic atoms. The $\mu 2e$ [94] and *COMET* [95] experiments at Fermilab and J-PARC, respectively, aim at a reduction of the limit to 10^{-16} while the proposed *PRISM* experiment [96] would go to 10^{-18}.

This short outline shows that the future direction of flavour physics experiments goes towards high precision measurements of extremely rare processes. Several aspects of future possibilities in flavour physics have not been discussed, e.g. mixing or even $\mathcal{CP}$ violation measurements in the neutrino sector at possible neutrino factories [97]. However, not only the experimental side of flavour physics will make significant progress over the next decades. The challenges on the theoretical side are in the explanation of the mass hierarchy of the hadronic and leptonic sector, revealing the source of the different mixing behaviours in the two sectors. The interpretation of flavour physics measurements, in particular with the increased precision of future experiments, also relies on advances in calculations based on lattice *QCD*. Finally, the origin of the matter dominance over anti-matter in the universe is the biggest riddle. Its answer may lie in baryogenesis via leptogenesis [98, 99]. It will certainly take decades to experimentally establish any model. Many questions related to flavour physics are yet to be answered and the measurements at *LHCb* will significantly advance the whole field.

Bibliography

[1] O. Philipsen. An introduction to quantum field theory. RAL-TR-2007-004.

[2] S.L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1961.

[3] A. Salam and J.C. Ward. Electromagnetic and weak interactions. *Phys. Lett.*, 13: 168–171, 1964.

[4] S. Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967.

[5] S. Davidson. The standard model. RAL-TR-2007-004.

[6] N. Cabibbo. Unitary Symmetry and Leptonic Decays. *Phys. Rev. Lett.*, 10:531–533, 1963.

[7] S.L. Glashow, J. Iliopoulos, and L. Maiani. Weak Interactions with Lepton-Hadron Symmetry. *Phys. Rev.*, D2:1285–1292, 1970.

[8] M. Kobayashi and T. Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.

[9] G.D. Rochester and C.C. Butler. Evidence for the Existence of New Unstable Elementary Particles. *Nature*, 160:855–857, 1947.

[10] K. Kleinknecht. CP Violation in the $K^0\,\overline{K}^0$ System. *Adv. Ser. Direct. High Energy Phys.*, 3:41–104, 1989.

[11] M.S. Sozzi. *Discrete symmetries and CP violation: From experiment to theory.* Oxford University Press, 2008.

[12] C. Amsler et al. Review of particle physics. *Phys. Lett.*, B667:1, 2008.

[13] G. Lüders. *Kgl. Danske Videnskab. Selskab, Matfys. Medd.*, 28(5):1, 1954.

[14] W. Pauli. *Niels Bohr and the Development of Physics.* Pergamon, Oxford, 1955. 30 pp.

[15] R. Jost. *Helv. Phys. Acta*, 30:209, 1957.

[16] J.H. Christenson, J.W. Cronin, V.L. Fitch, and R. Turlay. Evidence for the 2 pi Decay of the k(2)0 Meson. *Phys. Rev. Lett.*, 13:138–140, 1964.

[17] B. Aubert et al. Observation of CP violation in the B^0 meson system. *Phys. Rev. Lett.*, 87:091801, 2001.

[18] K. Abe et al. Observation of large CP violation in the neutral B meson system. *Phys. Rev. Lett.*, 87:091802, 2001.

[19] H. Burkhardt et al. First Evidence for Direct CP Violation. *Phys. Lett.*, B206: 169, 1988.

[20] V. Fanti et al. A new measurement of direct CP violation in two pion decays of the neutral kaon. *Phys. Lett.*, B465:335–348, 1999.

[21] A. Alavi-Harati et al. Observation of direct CP violation in K(S,L) –> pi pi decays. *Phys. Rev. Lett.*, 83:22–27, 1999.

[22] B. Aubert et al. Observation of direct CP violation in $B^0 \to K^+\pi^-$ decays. *Phys. Rev. Lett.*, 93:131801, 2004.

[23] Y. Chao et al. Evidence for direct CP violation in B0 –> K+ pi- decays. *Phys. Rev. Lett.*, 93:191802, 2004.

[24] J. Charles et al. CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories. *Eur. Phys. J.*, C41:1–131, 2005. http://ckmfitter.in2p3.fr.

[25] B. Aubert et al. Improved measurement of the CKM angle γ in $B^\mp \to D^{(*)}K^{(*\mp)}$ decays with a Dalitz plot analysis of D decays to $K^0_S\pi^+\pi^{-)}$ and $K^0_S K^+ K^{-)}$. *Phys. Rev.*, D78:034023, 2008.

[26] K. Abe et al. Updated Measurement of ϕ_3 with a Dalitz Plot Analysis of $B \to D^*K$ Decay. 2008.

[27] R. Fleischer. New strategies to extract beta and gamma from B/d –> pi+ pi- and B/s –> K+ K-. *Phys. Lett.*, B459:306–320, 1999.

[28] M. Adinolfi et al. The tree-level determination of γ at LHCb. Technical Report LHCb-ROADMAP5, CERN, Geneva. to be published.

[29] A. Bates et al. Road map for charmless charged two-body B decays at LHCb. Technical Report LHCb-ROADMAP6, CERN, Geneva. to be published.

[30] A. Lenz. Theoretical status of B_s -mixing and lifetimes of heavy hadrons. *Nucl. Phys. Proc. Suppl.*, 177-178:81–86, 2008.

[31] *CERN Workshop on Standard Model Physics (and more) at the LHC*, 2000. CERN-2000-004.

[32] W.S. Hou, M. Nagashima, and A. Soddu. Difference in B+ and B0 direct CP asymmetry as effect of a fourth generation. *Phys. Rev. Lett.*, 95:141601, 2005.

[33] C. Tarantino. FCNC Processes in the LHT Model: a 2009 Look. 2009.

[34] R. Fleischer and J. Matias. Exploring CP violation through correlations in $B \to \pi K$, $B_d \to \pi^+\pi^-$, $B_s \to K^+K^-$ observable space. *Phys. Rev.*, D66:054009, 2002.

[35] S. Bergmann et al. Lessons from CLEO and FOCUS Measurements of D0-anti-D0 Mixing Parameters. *Phys. Lett.*, B486:418–425, 2000.

[36] LHB collaboration. Letter of intent: measurement of CP violation in B meson decays with an extracted LHC beam. Technical Report CERN-LHCC-93-45. LHCC-I-5, CERN, Geneva, 1993.

[37] P. Schlein et al. COBEX: letter of intent for a collider beauty experiment at the Large Hadron Collider at CERN. Technical Report CERN-LHCC-93-50. LHCC-I-6, CERN, Geneva, 1993.

[38] GAJET collaboration. Letter of intent: study of CP violation in B-meson decays using an internal gas jet target at the LHC, (GAJET-Experiment). Technical Report CERN-LHCC-93-54. LHCC-I-7, CERN, Geneva, 1993.

[39] LHC-B collaboration. LHC-B: letter of intent. Technical Report CERN-LHCC-95-005. LHCC-I-8, CERN, Geneva, 1995.

[40] S. Amato et al. *LHCb : Technical Proposal.* Tech. Proposal. CERN, Geneva, 1998.

[41] R. Antunes-Nobrega et al. *LHCb reoptimized detector design and performance: Technical Design Report.* Technical Design Report LHCb. CERN, Geneva, 2003.

[42] C. Vanoli. The CERN accelerator complex. http://cdsweb.cern.ch/record/979035, Jun 2006.

[43] O.S. Bruning (Ed.) et al. LHC design report. Vol. I: The LHC main ring. CERN-2004-003-V-1.

[44] O.S. Buning (Ed.) et al. LHC Design Report. 2. The LHC infrastructure and general services. CERN-2004-003-V-2.

[45] M. Benedikt (Ed.), P. Collier (Ed.), V. Mertens (Ed.), J. Poole (Ed.), and K. Schindl (Ed.). LHC Design Report. 3. The LHC injector chain. CERN-2004-003-V-3.

[46] M. Brice. View of an open LHC interconnection. http://cdsweb.cern.ch/record/905940, Oct 2005.

[47] J.L. Caron. LHC Layout. http://cdsweb.cern.ch/record/841573, Sep 1997.

[48] AC Team. The four main LHC experiments. http://cdsweb.cern.ch/record/40525, Jun 1999.

[49] A. Augusto Alves et al. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008.

[50] Nikhef. VELO 3D Design Pictures. http://www.nikhef.nl/pub/departments/mt/projects/lhcb-vertex/design/DETECTORS/.

[51] M. Brice. Assembling the last module of the vertex locator for LHCb. http://cdsweb.cern.ch/record/1024838, Mar 2007.

[52] M. Brice. LHCb installs its precision silicon detector, the VELO. http://cdsweb.cern.ch/record/1068142, Oct 2007.

[53] P. Turner. Silicon Sensor Design and Geometry. https://edms.cern.ch/document/401568/4, Aug 2005.

[54] G. Haefeli et al. The LHCb DAQ interface board TELL1. *Nucl. Instrum. Meth.*, A560:494–502, 2006.

[55] T. Szumlak and C. Parkes. Description of the Vetra Project and its Application for the VELO Detector. Technical Report LHCb-2008-022. CERN-LHCb-2008-022, CERN, Geneva, May 2008.

[56] A. Bates et al. LHCb VELO module characterisation: System construction details and results from the Glasgow LHCb VELO module burn-in. *to be published in Nucl. Instrum. Meth.*, A, 2009.

[57] G. Barrand et al. GAUDI - A software architecture and framework for building HEP data processing applications. *Comput. Phys. Commun.*, 140:45–55, 2001.

[58] D.J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth.*, A462:152–155, 2001.

[59] S. Agostinelli et al. GEANT4: A simulation toolkit. *Nucl. Instrum. Meth.*, A506: 250–303, 2003.

[60] J. Allison et al. Geant4 developments and applications. *IEEE Trans. Nucl. Sci.*, 53:270, 2006.

[61] H. de Vries. VELO alignment report. https://edms.cern.ch/document/908162/1, Apr 2008.

[62] P. Sutcliffe. VELO Module Production - Final Module Metrology. Technical Report LHCb-2007-087. CERN-LHCb-2007-087, CERN, Geneva, Jan 2008.

[63] T. Huse. VELO Module Production - Sensor to Sensor Metrology. Technical Report LHCb-2007-085. CERN-LHCb-2007-085, CERN, Geneva, Nov 2007.

[64] M. Gersabeck. Initial LHCb VELO Alignment from Survey Measurements. Technical Report LHCb-2008-044. CERN-LHCb-2008-044, CERN, Geneva, Dec 2008.

[65] S. Blusk (Ed.) et al. Proceedings of the first LHC Detector Alignment Workshop, CERN, Geneva, Switzerland, 4-6 September 2006. CERN-2007-004.

[66] W. Hulsbergen. The global covariance matrix of tracks fitted with a Kalman filter and an application in detector alignment. 2008.

[67] C. Parkes, T. Ruf, and T. Szumlak. Reconstruction of Cluster Positions in the LHCb Velo. Technical Report LHCb-2007-151. CERN-LHCb-2007-151, CERN, Geneva, Dec 2007.

[68] K. Akiba, J. Borel, J. Buytaert, L. Eklund, and M. Gersabeck. The algorithm for FIR corrections of the VELO analogue links and its performance. Technical Report LHCb-2008-015. CERN-LHCb-2008-015, CERN, Geneva, 2009.

[69] P. Bartalini et al. VELO telescope resolution and efficiency measurements. Technical Report LHCb-2000-099, CERN, Geneva, Oct 2000.

[70] Government Actuary's Department. Life tables. http://www.gad.gov.uk/Demography_Data/Life_Tables, Apr 2009.

[71] F. Azfar et al. A Monte Carlo Independent Method for Lifetime Fits in Data biased by the Hadronic Trigger. Technical Report CDF/ANAL/BOTTOM/CDFR/6756, CDF, 2003.

[72] V. Gligorov and J. Rademacker. Monte Carlo Independent Lifetime Fitting at LHCb in Lifetime Biased Channels. Technical Report LHCb-2007-053. CERN-LHCb-2007-053, CERN, Geneva, Jun 2007.

[73] M. Gersabeck, V. Gligorov, J. Imong, and J. Rademacker. A Monte Carlo free method of measuring lifetimes using event-by-event acceptance functions at LHCb. Technical report, CERN, Geneva. to be published.

[74] K.S. Cranmer. Kernel estimation in high-energy physics. *Comput. Phys. Commun.*, 136:198–207, 2001.

[75] F. Marinho. A non-parametric method to estimate the Forward-Backward Asymmetry from the $B_d \to K^* \mu^+ \mu^-$. Technical Report LHCb-2009-004. CERN-LHCb-2009-004, CERN, Geneva, Jan 2009.

[76] N.L. Pounder. Measurement of the $\mathrm{B_s^0}$ lifetime in $\mathrm{B_s^0} \to \mathrm{K^+K^-}$ decays. FERMILAB-THESIS-2009-03.

[77] V. Gligorov. Reconstruction of the decay modes $B_d^0 \to D_d^{pm} \pi^{mp}$, $B_s^0 \to D_s^- \pi^+$, and $B_s^0 \to D_s^\pm K^\mp$ at LHCb. Technical Report LHCb-PUB-2009-003. CERN-LHCb-PUB-2009-003, CERN, Geneva, May 2009.

[78] S. Amato et al. LHCb's sensitivity to New CP-violating Phases in the Decay $B_s \to \phi\phi$. Technical Report LHCb-2007-047. CERN-LHCb-2007-047, CERN, Geneva, May 2007.

[79] A. de Roeck et al. *CMS physics: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 2006. revised version submitted on 2006-09-22 17:44:47.

[80] T. Aaltonen et al. Measurement of lifetime and decay-width difference in $B_s^0 \to J/\psi\phi$ decays. *Phys. Rev. Lett.*, 100:121803, 2008.

[81] V.M. Abazov et al. Measurement of B_s^0 mixing parameters from the flavor-tagged decay $B_s^0 \to J/\psi\phi$. *Phys. Rev. Lett.*, 101:241801, 2008.

[82] I.I. Bigi. Could Charm's 'Third Time' Be the Real Charm? – A Manifesto. 2009.

[83] T. Aaltonen et al. Evidence for $D^0 - \bar{D}^0$ mixing using the CDF II Detector. *Phys. Rev. Lett.*, 100:121802, 2008.

[84] D. Petrie, C. Parkes, and S. Viret. Study of the impact of VELO misalignments on the LHCb tracking and L1 trigger performance. Technical Report LHCb-2005-056. CERN-LHCb-2005-056, CERN, Geneva, Oct 2005.

[85] J. Nardulli. *Reconstruction of two-body B decays in LHCb*. PhD thesis, Vrije Univ. Amsterdam, Amsterdam, 2007. Presented on 04 Oct 2007.

[86] M. Gersabeck, J. Nardulli, and E. Rodrigues. Impact of misalignments on the analysis of B decays. Technical Report LHCb-2008-012. CERN-LHCb-2008-012, CERN, Geneva, Aug 2008.

[87] M. Needham. Performance of the LHCb Track Reconstruction Software. Technical Report LHCb-2007-144. CERN-LHCb-2007-144, CERN, Geneva, Jan 2008.

[88] Expression of Interest for an LHCb Upgrade. Technical Report LHCb-2008-019. CERN-LHCb-2008-019. CERN-LHCC-2008-007, CERN, Geneva, Apr 2008. revised version submitted on 2008-05-07 12:08:45.

[89] A. Ceccucci et al. Proposal to Measure the Rare Decay $K^+ \to \pi^+\nu\bar{\nu}$ at the CERN SPS. Technical Report CERN-SPSC-2005-013. SPSC-P-326, CERN, Geneva, Apr 2005.

[90] J. Comfort et al. Proposal for $K_L^0 \to \pi^0\nu\overline{\nu}$ Experiment at J-Parc. Proposal for Nuclear and Particle Physics Experiments at J-PARC, April 2006.

[91] A.G. Akeroyd et al. Physics at super B factory. 2004.

[92] M. Bona et al. SuperB: A High-Luminosity Asymmetric e^+e^- Super Flavor Factory. Conceptual Design Report. 2007.

[93] T. Mori et al. Search for $\mu^+ \to e^+\gamma$ down to 10^{-14} branching ratio. Research proposal to PSI, May 1999.

[94] R.M. Carey et al. Proposal to search for $\mu^- N \to e^- N$ with a single event sensitivity below 10^{-16}. FERMILAB-PROPOSAL-0973.

[95] K. Yuno et al. An Experimental Search for Lepton Flavor Violating $\mu^- \to e^-$ Conversion at Sensitivity of 10^{-16} with a Slow-Extracted Bunched Proton Beam. Experimental Proposal for Nuclear and Particle Physics Experiments at J-PARC, November 2007.

[96] S. Machida et al. An Experimental Search for the $\mu^- \to e^-$ Conversion Process at an Ultimate Sensitivity of the Order of 10^{-18} with PRISM. LOI for Nuclear and Particle Physics Experiments at the J-PARC, January 2003.

[97] A. Bandyopadhyay et al. Physics at a future Neutrino Factory and super-beam facility. 2007.

[98] M.A. Luty. Baryogenesis via leptogenesis. *Phys. Rev.*, D45:455–465, 1992.

[99] W. Buchmuller, P. Di Bari, and M. Plumacher. Leptogenesis for pedestrians. *Ann. Phys.*, 315:305–351, 2005.