Measurement of the CP-even Fraction of the $D^0 \to 2\pi^+2\pi^-$ Decay using Quantum Correlated $D\bar{D}$ Pairs at CLEO-c, and Real-time Alignment of the LHCb RICH Optical Systems

Claire Prouvé

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Ph.D. in the Faculty of Science.

School of Physics
September 2017

under supervision of
Dr. Jonas Rademacker

at the
University of Bristol
Senate House, Tyndall Avenue, Bristol, BS8 1TH, UK

Word count: 64 192

Abstract

This thesis covers three subjects, namely the measurement of the *CP*-even fraction of the $D^0 \to 2\pi^+2\pi^-$ decay, the real-time alignment of the LHCb RICH mirror systems and the estimation of the sensitivity to the CKM angle γ that can be obtained using $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ events at LHCb.

The *CP*-even fraction $F_{4\pi}^+$ of the $D^0 \to 2\pi^+2\pi^-$ decay is measured using a dataset corresponding to $818\,\mathrm{pb}^{-1}$ of quantum correlated $D\overline{D}$ decays produced in electron-positron collisions at the $\psi(3770)$ resonance collected by the CLEO-c experiment at Cornell University. In the analysis, one of the correlated D mesons is reconstructed as $D \to 2\pi^+2\pi^-$ while the other D meson is reconstructed as $D \to K^0_{\mathrm{S,L}}\pi^+\pi^-$. Sensitivity to the *CP*-even fraction of $D^0 \to 2\pi^+2\pi^-$ is obtained by determining the variation of yields over the $K^0_{\mathrm{S,L}}\pi^+\pi^-$ phase space, specifically the variation of yields between bins of the $K^0_{\mathrm{S,L}}\pi^+\pi^-$ phase space. The *CP*-even fraction is measured to be $F_{4\pi}^+ = 0.755 \pm 0.050\,(\mathrm{stat}) \pm 0.029\,(\mathrm{syst})$.

The LHCb RICH mirror alignment procedure using proton-proton collision data is implemented in the LHCb online reconstruction framework for Run II of the LHC. In Run II, all LHCb subdetectors are aligned and calibrated in real-time, that is between the two high level trigger stages HLT1 and HLT2. This enables the reconstruction in the high level trigger to be identical to the offline reconstruction and the direct use of the trigger output for physics measurements. The RICH mirror alignment was implemented and commissioned throughout 2015 and 2016 and the procedure was improved at different points. This reduced the time needed to perform an alignment from several days in Run I to about 20 min in Run II. The information gathered by the frequent running of the alignment in 2016 is used to further the understanding of the alignment procedure as well as the understanding of the LHCb detector itself.

The sensitivity to the CKM angle γ that can be achieved at the LHCb experiment with $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ decays, is studied. The study is performed using the $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ event yield in $8\,\mathrm{fb}^{-1}$ of the proton-proton collision data expected to be recorded by LHCb by the end of Run II of the LHC. The $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ event yield of the data collected in 2016 is determined and extrapolated to the expected full luminosity of Run II. A series of pseudo-experiments is used to simulate the distribution of $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ events over the bins of the $D \to 2\pi^+2\pi^-$ phase space and a fit is performed to extract the resulting uncertainty on the CKM angle γ. The statistical uncertainty is determined to be 20° for the 2016 dataset, 10° for the full expected Run II dataset and 9° for the combined Run I and Run II dataset. The current best measurements of γ from single analyses have a statistical uncertainty of $\approx 15^\circ$ while the combined uncertainty on γ is $\approx 7^\circ$ [1]. The measurement of the angle γ using binned $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ decays can thus significantly contribute to the overall constraint on γ.

Acknowledgements

First, I would like to thank Dr. Jonas Rademacker for giving me the opportunity to work on a variety of fascinating and challenging topics. I appreciate the freedom I was granted to pursue my own way through all projects and to finalise them in a matter that was coherent with my personality. I have learned a great many things throughout the last four years, about different domains of particle physics, about analysis techniques and detector operation, and about working with and amongst collaborators.

I would like to thank all members of the University of Bristol particle physics group. I thank Dr. Jack Benton for introducing me to the analysis subject and the necessary tools and software. I thank Dr. Andrew Cook for explaining the methodology of the RICH alignment to me. I thank Dr. Sam Harnew for his huge part in the measurement of the hadronic parameters of the $D^0 \to 2\pi^+2\pi^-$ decay, his patience for all my many questions and for his feedback on my thesis. I thank Dr. Paras Naik and Dr. Jeremy Dalseno for always being available for questions and discussions.

I thank the LHCb Oxford group, especially Dr. Guy Wilkinson, Dr. Chris Thomas and Dr. Sneha Malde for including me in the measurement of the *CP*-even fraction of the $D^0 \to 2\pi^+2\pi^-$ decay. I am grateful for all the tutoring that I received about the CLEO-c experiment and analysis techniques used at CLEO-c.

I thank all the members of the LHCb RICH software group, most importantly Dr. Antonis Papanestis, Dr. Chris Jones and Dr. Mike McCann who took a great interest in the RICH mirror alignment and taught me much about the RICH detectors and the LHCb reconstruction software. I am very grateful for their willingness to discuss any raising subject and problems and to share their immense knowledge with me.

I thank the LHCb online computing team, namely Dr. Roel Aaij, Dr. Beat Jost and Dr. Clara Gaspar for their support during my work on the real-time alignment. I am grateful for the knowledge that was shared with me and for the many hours that were spent commissioning the RICH alignment software. Without their precious support it would not have been possible to execute this project.

I am grateful to the alignment conveners Dr. Silvia Borghi and Dr. Lucia Grillo for introducing me to the LHCb alignment group and for promoting me within the collaboration.

I also thank Dr. Marie-Hélène Schune for being my mentor and friend throughout the four years. I very much appreciate her encouragement, support and guidance.

I thank Dr. Sascha Stahl for his endless patience, for sharing his comprehensive knowledge about the reconstruction and the trigger with me, for proof reading my thesis and for his overall tireless help.

Finally, I want to thank all my family and friends who supported me in many different ways during the Ph.D.

Author's Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: DATE:

Contents

Introduction

The Standard Model of particle physics is a very successful theory in that it explains a vast range of phenomena observed in nature. Its predictive powers exceed those of any previous theory and its mathematical structure is well defined and organised. The Standard Model describes the fundamental aspects of matter and energy by expressing matter particles (fermions) and particles that mediate interactions (bosons) in a unified theory. The latest achievement of particle physics was the discovery of the Higgs Boson, whose existence had been predicted over forty years ago as an inherent part of the Standard Model, by CMS [2] and ATLAS [3].

Despite these achievements there are still questions unanswered by the Standard Model. One open issue is the inclusion of gravity and the explanation of the hierarchy of the fundamental interactions. Another unexplained phenomenon is the discrepancy between the amount of matter and anti-matter observed in the universe. While the Standard Model does implement the mechanism of charge-parity (CP) violation, the generated effects within the Standard Model are not large enough to explain today's observations.

For this reason, many other theories – such as supersymmetry (SUSY) [4] or the Left-Right Symmetric Model (LRSM) [5] – have been developed as extensions of the Standard Model. These theories are designed so that their predictions agree with the Standard Model at low energies but differ at higher energies. Usually new particles are introduced with heavier masses, changing the physics at high energy scales with respect to the Standard Model.

All high energy physics experiments aim at performing precision measurements of Standard Model parameters or at discovering physics beyond the Standard Model. The LHC has introduced a promising environment with an unprecedented center of mass energy and luminosity that provides the allocated experiments with a high amount of statistics. The experiments ATLAS and CMS – the so-called 'general purpose' detectors – mainly aim at finding new particles through direct searches. The LHCb experiment however, performs indirect searches for physics beyond the Standard Model, for example by identifying deviations from Standard Model predictions in loop-diagram and box-diagram processes. The LHCb experiment focuses on rare decays of beauty and charm hadrons.

A precise determination of the CP-violating quantities is one of the primary objectives of heavy flavour physics. This thesis focuses on the measurement of the CKM angle γ which is one of the angles of the Unitarity Triangle and the CP-violating phase between $b \to u$ and $b \to c$ quark transitions. The CKM angle γ is the only angle of the Unitarity Triangle that can be measured without significant contributions from physics beyond the Standard Model, making it a Standard Model key measurement. The comparison between values of the angle γ obtained trough Standard Model measurement and measurements where physics beyond the

Standard Model can contribute is a powerful way to verify the consistency of the Standard Model and possibly identify sources of physics beyond the Standard Model. From the three angles of the Unitarity Triangle the angle γ is the least constraint angle. The best direct measurement of γ has an uncertainty of approximately $15°$ [1], whereas the combination of measurements reduces the uncertainty to approximately $7°$. To obtain a precise measurement of the CKM angle γ it is important to combine many different measurements, often involving studies of hadronic B meson decays with low rates.

To first order, $B^{\pm} \to DK^{\pm}$ decays are mediated by tree-level processes and offer a theoretically clean way to measure the CKM angle γ, unlikely to be affected by physics beyond the Standard Model. Since γ is also the *CP*-violating phase between the $b \to u$ and $b \to c$ quark transitions, it is observable in the interference between the $B^{\pm} \to D^0 K^{\pm}$ and the $B^{\pm} \to \overline{D}^0 K^{\pm}$ transition when the D meson is reconstructed in a final state accessible to both D^0 and $\overline{D}^0$ mesons. The precise measurement of γ requires the combination of many analyses in each of which the D meson is reconstructed in a different final state. The content of this thesis is the work towards the measurement of γ using $B^{\pm} \to D(\to 2\pi^+2\pi^-)K^{\pm}$ decays where the D meson is reconstructed in the self-conjugate multi-body final state $D \to 2\pi^+2\pi^-$. Sensitivity to γ can be gained both by integrating over the entire D meson decay phase space or by evaluating the interference pattern over the D meson decay phase space.

The parameters that quantify the interference between D^0 and $\overline{D}^0$ decays are a source of systematic uncertainty on γ. To avoid the significant systematic uncertainty associated with modelling the strong phase of the D decay amplitude across the five-dimensional phase space of the four body decay, model-independent parameters are defined. These parameters can be measured at the CLEO-c experiment at Cornell University which has collected a sample of quantum-correlated $\psi(3770) \to D^0\overline{D}^0$ decays. In this thesis, the model-independent *CP*-even fraction of the $D^0 \to 2\pi^+2\pi^-$ decay is determined with CLEO-c data. This measurement has already been used as input to the measurement of *CP* observables with $B^{\pm} \to D(\to 2\pi^+2\pi^-)K^{\pm}$ decays at LHCb [6].

An important part of all precision measurements, but especially analyses with many hadrons in the final state, is very efficient particle identification. For the measurement of the *CP*-violating phase γ with $B^{\pm} \to DK^{\pm}$ decays it is particularly important to distinguish kaons from pions in order to cleanly separate $B^{\pm} \to DK^{\pm}$ events from $B^{\pm} \to D\pi^{\pm}$ events. Both decays have almost identical kinematics but very different *CP* asymmetries and the $B^{\pm} \to D\pi^{\pm}$ decay is Cabibbo favoured compared to the $B^{\pm} \to DK^{\pm}$ decay. In LHCb, the main source of particle identification information for charged kaons and pions are the two ring imaging Cherenkov (RICH) detectors. Both detectors possess intricate optical systems for detecting Cherenkov photons. In order for the particle identification to function optimally, the position of all optical components has to be known to the best precision. This is achieved by a data-driven alignment procedure. For Run II of the LHC, LHCb has implemented a novel approach: it is the first ever High Energy Physics detector which is aligned, calibrated, and fully reconstructed in real-time. The implementation of the RICH mirror alignment procedure into the real-time alignment framework was essential to achieving the full alignment and calibration of the LHCb detector, and a substantial part of this thesis.

This thesis is organised in seven chapters. The first chapter comprises the theoretical background needed to model-independently measure the CKM angle γ from $B^{\pm} \to DK^{\pm}$ decays. This chapter includes the introduction of the model-independent parameters of the D decay and their measurement using correlated $D^0\overline{D}^0$ decays. The second chapter introduces the CLEO-c experiment located at the CESR accelerator. The different subdetectors are presented as well as the trigger and the reconstruction software. The third chapter contains a detailed documentation of the measurement of the CP-even fraction $F_{4\pi}^{+}$ of the $D^0 \to 2\pi^+2\pi^-$ decay using quantum correlated $D^0\overline{D}^0$ meson pairs. The CP-even fraction is measured by reconstructing one D meson as $D^0 \to 2\pi^+2\pi^-$ and the other D meson as $D \to K^0_{\mathrm{S,L}}\pi^+\pi^-$. The fourth chapter presents the LHCb experiment located at the LHC and introduces the detector as well as the different trigger stages and the LHCb software projects. The fifth chapter outlines the implementation of the alignment of the LHCb RICH mirror systems into the real-time alignment framework. The sixth chapter contains the study of the sensitivity to the CKM angle γ that can be obtained using $B^{\pm} \to D(\to 2\pi^+2\pi^-)K^{\pm}$ decays collected by LHCb.

1. Theory

This chapter introduces the theoretical understanding of direct *CP* violation in $B^{\pm} \to DK^{\pm}$ decays and how the *CP*-violating phase γ becomes an observable. In the first section, a brief introduction to the Standard Model of particle physics is given. In the second section the direct *CP* violation in the weak interaction is discussed. In the third section the CKM matrix is presented which facilitates the direct *CP* violation in the Standard Model. In the fourth section the method for a model-independent measurement of the CKM angle γ using $B^{\pm} \to DK^{\pm}$ decays is described. In the sixth section the measurement of the model-independent, hadronic parameters of the D meson decay is explained. In the seventh section the binning scheme for the D meson decay phase space is illustrated. The last section provides a summary of the before mentioned sections.

1.1 The Standard Model of Particle Physics

The Standard Model of particle physics is a comprehensive theory of the microscopic structure of the universe[1]. It contains the fundamental particles that build all matter and describes the interaction between these fundamental particles.

The matter fields described by the Standard Model are spin 1/2 particles (*fermions*) grouped into two categories: leptons and quarks. The particles can be organised in doublets, connected via the $SU(2)$ symmetry of the weak interaction. All fundamental fermions are listed in Table 1.1. Each fermion has a corresponding antiparticle which has the same mass but opposite additive quantum numbers.

There are three interactions in the Standard Model which are mediated by force carriers of spin 1 (*bosons*). The three interactions and their properties are listed in Table 1.2. The first interaction is the electromagnetic interaction. It is mediated by chargeless and massless *photons*. All charges particles participate in the electromagnetic interaction.

The second interaction is the weak interaction. It is mediated by the neutral Z^0 boson and the charged W^+ and W^- bosons. Unlike the massless photons, the bosons of the weak interaction have masses of $80\,\mathrm{GeV}/c^2$ and $91\,\mathrm{GeV}/c^2$, for the $W^{\pm}$ bosons and the Z^0 bosons, respectively. The mass of the bosons is responsible for the relatively small strength of the weak interaction compared to the other two interactions. All fermions carry the weak charge and participate in the weak interaction. The charged weak bosons $W^{\pm}$ couple exclusively to left-handed chiral states of fermions

[1]The introduction to the Standard Model is based on References [7–9].

Fermions

	Leptons			**Quarks**		
Generation	Type	Charge	Mass	Type	Charge	Mass
1	ν_e	0	$< 2\,\mathrm{eV}/c^2$	u	$+2/3$	$2.3\,\mathrm{MeV}/c^2$
	e^-	-1	$511.0\,\mathrm{eV}/c^2$	d	$-1/3$	$4.8\,\mathrm{MeV}/c^2$
2	ν_μ	0	$< 2\,\mathrm{eV}/c^2$	c	$+2/3$	$1.28\,\mathrm{GeV}/c^2$
	μ^-	-1	$105.7\,\mathrm{MeV}/c^2$	s	$-1/3$	$95\,\mathrm{MeV}/c^2$
3	ν_τ	0	$< 2\,\mathrm{eV}/c^2$	t	$+2/3$	$173.5\,\mathrm{GeV}/c^2$
	τ^-	-1	$1.78\,\mathrm{GeV}/c^2$	b	$-1/3$	$4.18\,\mathrm{GeV}/c^2$

Table 1.1: *Fundamental fermions of the Standard Model that build all matter. There are three generations of doublets for leptons and quarks respectively.*

and right-handed chiral states of antifermions. The neutral weak boson Z^0 couples to the chirality left-handed part of fermions and the chirality right-handed part with different strengths. The weak interaction is discussed in more detail in Section 1.2 in the context of *CP* violation. The electromagnetic and the weak interaction can be unified to the *electroweak interaction* generated by the electroweak symmetry group $SU(2)_L \otimes U(1)_Y$.

The third interaction is the *strong interaction* described by *quantum chromodynamics*. The strong interaction is mediated by eight *gluons* which couple exclusively to quarks (and gluons themselves) since only quarks and gluons carry colour charge. Unlike in the electroweak interaction, the coupling constant of the strong interaction gets smaller at high energies. This principle is called *asymptotic freedom* and describes that quarks are quasi-free at small distances. Additionally, the phenomenon of *colour confinement* has the consequence that only colour-neutral objects exist. This means that colour charged objects such as quarks and gluons are only found bound together as hadrons. The symmetry group of the strong interaction is the $SU(3)_C$.

All three interactions together give the symmetry group of the Standard Model

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \ .$$

All interactions of the Standard Model have to obey laws of conservation, such as charge conservation and energy conservation. The electromagnetic and the strong interactions are also invariant under parity transformation P and charge conjugation C while the weak interaction maximally violates both parity and charge conjugation. The parity and charge conjugation violation as well as the composite *CP* violation of the weak interaction is explored in detail in the next section.

The latest confirmed particle of the Standard Model is the *Higgs boson* which was discovered in 2012 by CMS [2] and ATLAS [3]. The Higgs boson has a spin of 0 and a mass of 125 GeV/c^2. The Higgs originates from the spontaneous symmetry breaking of the electroweak symmetry group which leads to the masses of the weak bosons Z^0, W^+ and W^-. The Higgs field is also responsible for the masses of the fermions which are generated by Yukawa interactions between the Higgs field and the fermions.

interaction	Relative Strength	Typical Lifetime	Typical Cross Section	Source	Force Carrier	Mass
Strong	1	10^{-23} s	10 mb	colour charge	8 gluons	0
Electro-magnetic	1/137	10^{-20} s	10 μb	electric charge	photon	0
Weak	10^{-6}	10^{-10} s	10 pb	weak charge	$W^\pm$ Z^0	80 GeV/c^2 91 GeV/c^2

Table 1.2: *Description of the three fundamental interactions incorporated in the Standard Model. The typical lifetime denotes the order of the lifetime of particles that decay through the respective interaction. The typical cross section represents the order of the cross section for processes that proceed through the respective interaction.*

While the Standard Model is widely successful and extensively tested, it cannot be a complete theory of the universe. The Standard Model leaves a number of questions unanswered and does not describe the entire universe as it is observed today. Amongst those questions is the baryon asymmetry of the universe, suggesting a strong effect of *CP*-violation while the *CP*-violation in the Standard Model is comparably small. Furthermore, the Standard Model only describes a small part of the universe, namely the matter. There are no descriptions of dark matter or dark energy in the Standard Model. Additionally, it does not incorporate the gravitational force. The Standard Model does also not explain why there are three generations of leptons and quarks nor does it predict the values for the fermion masses or the coupling constants.

1.2 *CP* Violation in the Charged Weak Interaction

The *CP* transformation is the subsequent application of the parity transformation P and the charge conjugation C. Both the parity transformation and the charge conjugation are discrete transformations, making the *CP* transformation a discrete transformation as well. While the electromagnetic interaction and the strong interaction are invariant under parity transformation and charge conjugation, the charged weak interaction violates both maximally. Additionally, *CP* violation is possible in the charged weak interaction due to a complex phase in the matrix that transforms the mass eigenstates of quarks into their weak eigenstates.

The following sections introduce the principle of discrete transformation and explore the parity transformation and the charge conjugation with respect to the charged weak interaction. Then the *CP* violation in the weak interaction is explored.

1.2.1 Discrete Transformations and Symmetries

A symmetry is present in a system if the system is invariant under a particular transformation. The Standard Model contains both *continuous* symmetries and *discrete* symmetries. A symmetry is continuous if a macroscopic transformation can be expressed as a series of infinitesimal transformations. For a discrete symmetry, the system can only be in a finite number of configurations which thus cannot be reached by microscopic transformations. Continuous symmetries give additive quantum numbers while discrete symmetries give multiplicative quantum numbers.

Noether's first theorem states that there is a conservation law for each continuous (differentiable) symmetry of the action of a physical system [10]. An example is the symmetry of a system under translations in space which yields the conservation of momentum. If the physics of a system is invariant under a discrete symmetry g and g can be represented by a hermitian operator $U(g)$, then U itself is an observable conserved quantity. This means that if a system is in an eigenstate of U, then transitions of the system can only occur to eigenstates with the same eigenvalue. Two discrete transformations, namely the parity transformation and the charge conjugation are discussed in Section 1.2.2 and Section 1.2.3 respectively.

1.2.2 Parity Transformation

The parity transformation, represented by the operator P, corresponds to the inversion of the spacial coordinates of a physical system with respect to the origin, given by

$$(x, y, z) \xrightarrow{P} (-x, -y, -z) \ . \tag{1.1}$$

All other quantities such as time or spin are unchanged.

The wave functions $\Psi(\mathbf{p}, t, s, a)$ describing fermions with three momentum $\mathbf{p}$, time coordinate, t, spin, s and additive quantum numbers, a, are eigenvectors to P with eigenvalue n_P

$$P \ \Psi(\mathbf{p}, t, s, a) = \Psi(-\mathbf{p}, t, s, a) = \ n_P \ \Psi(\mathbf{p}, t, s, a) \ . \tag{1.2}$$

The twofold application of the parity transformation of a system has to yield the same system. For a fermion the twofold application of the parity transformation is

$$\begin{aligned} P \ \left(P \ \Psi(\mathbf{p}, t, s, a)\right) &= P \ \Psi(-\mathbf{p}, t, s, a) \\ &= P \ n_P \ \Psi(\mathbf{p}, t, s, a) \\ &= n_P^2 \ \Psi(\mathbf{p}, t, s, a) \\ &\stackrel{!}{=} \Psi(\mathbf{p}, t, s, a) \ . \end{aligned} \tag{1.3}$$

Thus, the possible eigenvalues of the parity transformation are $n_P = +1$ (*symmetric* or *even*) and $n_P = -1$ (*antisymmetric* or *odd*). By convention the parity of leptons and quarks is +1, while that of anti-leptons and anti-quarks is –1.

While the electromagnetic interaction and the strong interaction conserve parity, parity is observed to be maximally violated by the charged weak interaction. This is expressed in the $V - A$ structure (vector – axial-vector structure) of the flavour changing charged currents coupling to $W^{\pm}$ bosons. An example is illustrated in Figure 1.1 which shows the the charged weak interaction $d \to u\,W^-$. The charged weak current j^μ for the incoming d quark and the outgoing u quark is given by

$$j^\mu \propto \bar{u}\gamma^\mu \frac{1}{2}(1-\gamma^5)d \tag{1.4}$$

where d represents the spinor for the d quark, $\bar{u}$ the adjoint spinor for the u quark with $\bar{u} = u^\dagger\gamma^0$, $\gamma^\mu = \{\gamma^0, \gamma^1, \gamma^2, \gamma^3\}$ are the Dirac matrices and γ^5 is the product of all Dirac matrices given by $\gamma^5 := i\gamma^0\gamma^1\gamma^2\gamma^3$.

The expression $P_L = \frac{1}{2}(1-\gamma^5)$ is called the *chirality left-handed projection operator*. Applied to a particle spinor it projects out the chirality left-handed component while applied to an anti-particle spinor it gives the chirality right-handed component. This means that only the chirality left-handed component of fermions and the chirality right-handed component of anti-fermions participate in the charged weak interaction.

The action of the parity transformation on the spinor of a quark q, can also be represented by the matrix γ^0 as $Pq = \gamma^0 q$. Applying the parity transformation to the left-handed component of a quark yields

$$Pq_L = \gamma^0 q_L = \gamma^0 \frac{1}{2}(1-\gamma^5)q \tag{1.5}$$

$$= \frac{1}{2}(1+\gamma^5)\gamma^0 q = n_P P_R q \tag{1.6}$$

where $P_R = \frac{1}{2}(1+\gamma^5)$ is the *chirality right-handed projection operator*. Thus the parity transformation transforms a left-handed fermion into a right-handed fermion which means the $V - A$ structure of the weak interaction violates parity.

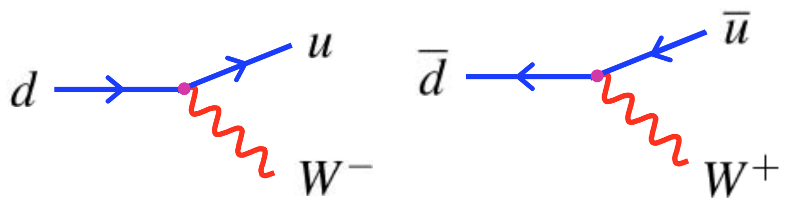


Figure 1.1: *Feynman diagram for the charged weak interaction $d \to u\,W^-$ (left) and its charge conjugate $\overline{d} \to \overline{u}\,W^+$ (right).*

1.2.3 Charge Conjugation

The charge conjugation, represented by the operator C, corresponds to the inversion of the sign of any additive quantum number (such as electric charge and baryon number). As for the parity transformation, the twofold application of the charge conjugation has to yield the same system meaning that the eigenvalues of the charge conjugation are $n_C = +1$ (*symmetric* or *even*) and $n_C = -1$ (*antisymmetric* or *odd*). Only neutral particles can be eigenstates of C.

The electromagnetic interaction and the strong interaction are invariant under charge conjugation, however the weak interaction is not. Under charge conjugation a chirality left-handed fermion — which participates in the charged weak interaction — is transformed into a chirality right-handed fermion — which does not participate in the charged weak interaction.

1.2.4 Direct *CP* Violation in the Standard Model

The *CP* transformation is the subsequent application of the parity transformation P and the charge conjugation C. Since the electromagnetic interaction and the strong interaction conserve both parity and charge conjugation separately, they also conserve *CP*. If the weak current was exactly as given in Equation 1.4, the weak interaction would also be invariant under *CP* transformation. The *CP* transformation transforms a chirality left-handed fermion into a chirality right-handed anti-fermion and would thus conserve the strength of the charged weak current as given in Equation 1.4.

The more accurate expression of the charged weak current for a transition of the down-type quark q to an u-type quark p and a W^- boson, given by $q \to pW^-$, is

$$j^\mu \propto \bar{p}\,\gamma^\mu \frac{1}{2}(1-\gamma^5)q' \tag{1.7}$$

where q' differs from the mass eigenstates q. The transformation between the spinors q and q' of the down-type quarks is given in terms of the complex, unitary CKM matrix (see Section 1.3 for more detail on the CKM matrix) as

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.8}$$

while the transformation between the adjoint spinors $\bar{q}$ and $\bar{q}'$ of the down-type quarks is given by

$$\begin{pmatrix} \bar{d}' & \bar{s}' & \bar{b}' \end{pmatrix} = \begin{pmatrix} \bar{d} & \bar{s} & \bar{b} \end{pmatrix} V^\dagger_{CKM} \ . \tag{1.9}$$

The charged weak current for the $d \to uW^-$ is thus given in terms of the mass eigenstates of the quarks, by

$$j^\mu_{d'u} = \bar{u}\left[-i\frac{g_W}{\sqrt{2}}\gamma^\mu \frac{1}{2}(1-\gamma^5)\right] V_{ud}\, d \tag{1.10}$$

where g_W is the dimensionless weak coupling constant. The current for the $\overline{d} \to \overline{u}W^+$ transition is given by

$$j^\mu_{\overline{d}'\overline{u}} = j^\mu_{ud'} = \bar{d}\, V^*_{ud}\left[-i\frac{g_W}{\sqrt{2}}\gamma^\mu \frac{1}{2}(1-\gamma^5)\right] u \ . \tag{1.11}$$

If the CKM matrix was real, the vertex factors $\mathcal{V}_{d'u}$ and $\mathcal{V}_{\overline{d}'\overline{u}}$ for both transitions given by

$$\mathcal{V}_{d'u} = \left[-i\frac{g_W}{\sqrt{2}}\gamma^\mu\frac{1}{2}(1-\gamma^5)\right] V_{ud} \tag{1.12}$$

and

$$\mathcal{V}_{\overline{d}'\overline{u}} = V^*_{ud}\left[-i\frac{g_W}{\sqrt{2}}\gamma^\mu\frac{1}{2}(1-\gamma^5)\right] \tag{1.13}$$

would be identical and the weak interaction would conserve CP. Since the CKM matrix has one complex phase, the weak interaction can be CP violating[2].

1.3 The CKM Matrix and the Unitarity Triangle

The CKM matrix V_{CKM} describes the connection between the flavour contents of a given initial state and a corresponding final state, whose transition occurs through a flavour-changing charged current mediated by charged $W^\pm$ bosons.

The CKM matrix is composed of two unitary matrices which represent the transformation of the mass eigenstates of the down-type quarks ($d\ s\ b$) and the up-type quarks ($u\ c\ t$) into the weak eigenstates ($d_w\ s_w\ b_w$) and ($u_w\ c_w\ t_w$) respectively, via

$$\begin{pmatrix} d_w \\ s_w \\ b_w \end{pmatrix} = A_{D_wD}\begin{pmatrix} d \\ s \\ b \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} u_w \\ c_w \\ t_w \end{pmatrix} = A_{U_wU}\begin{pmatrix} u \\ c \\ t \end{pmatrix} . \tag{1.14}$$

In order to persevere the norm and therefore the probability, the transformation matrices A_{D_wD} and A_{U_wU} have to be unitary

$$A^\dagger_{D_wD}A_{D_wD} = \mathbf{1} \qquad \text{and} \qquad A^\dagger_{U_wU}A_{U_wU} = \mathbf{1} . \tag{1.15}$$

The CKM matrix is given by

$$V_{CKM} \equiv A^\dagger_{U_wU}A_{D_wD} \tag{1.16}$$

and is therefore also unitary. The general form of the unitary matrix is written as

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

denoting that the amplitude of a transition of a quark q to a quark p is proportional to CKM matrix element V_{pq} while the transition from antiquark $\overline{q}$ to antiquark $\overline{p}$ is proportional to the complex conjugate matrix element V^*_{pq}.

As a complex, unitary matrix the CKM matrix has nine real free parameters. Due to the structure of the theory and how observables are calculated, five complex phases

[2]As can be seen in Section 1.4, further conditions have to be fulfilled to make CP violation observable.

of the CKM matrix can be absorbed into the quark fields. The CKM matrix has therefore four real parameters, three rotation angles in flavour space and one complex phase. The complex phase is the source of direct *CP* violation in the Standard Model.

1.3.1 Parametrisations of the CKM Matrix

The parametrisation of the CKM matrix is not unambiguous. A standard parametrisation [11] is

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

where $s_{ij} \equiv \sin(\theta_{ij})$, $c_{ij} \equiv \cos(\theta_{ij})$ and δ_{13} is the *CP*-violating phase. The parameters θ_{ij} are called the *quark mixing angles.*

Another convenient parametrisation is the Wolfenstein parametrisation [12] where the free parameters of the CKM matrix are chosen to be the four real parameters A, ρ, η and λ. The elements of the CKM matrix are then expanded in orders of the parameter λ which expresses the hierarchy of the size of the matrix elements. The Wolfenstein parametrisation of the CKM matrix up to $\mathcal{O}(\lambda^4)$ is given by

$$V_{CKM} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \quad . \tag{1.19}$$

This parametrisation illustrates that the transitions within the three generations are favoured while transitions between generations are suppressed. It also shows the *CP* violation is greatest in the CKM matrix elements V_{ub} and V_{td} since all other matrix elements are real to order $\mathcal{O}(\lambda^3)$.

1.3.2 The Unitarity Triangle

The unitarity requirement of the CKM matrix leads to nine equations. In particular, the equation

$$V_{ud}V_{ub}^* \quad + \quad V_{cd}V_{cb}^* \quad + \quad V_{td}V_{tb}^* = 0 \tag{1.20}$$

is of special interest for the *CP* violation in $B^\pm \to DK^\pm$ decays. Equation 1.20 can be represented as the *Unitarity Triangle* in the complex plane, which is depicted in Figure 1.2. Presuming *CPT* invariance[3], the three angles of the Unitarity Triangle, α, β and γ, have to add up to 180° for three generations of quarks. A deviation from 180° would indicate the existence of physics beyond the Standard Model. It is therefore important to over constrain the parameters of the Unitarity Triangle.

[3]Invariance under the subsequent application of time reversal T, parity transformation P and charge conjugation C.

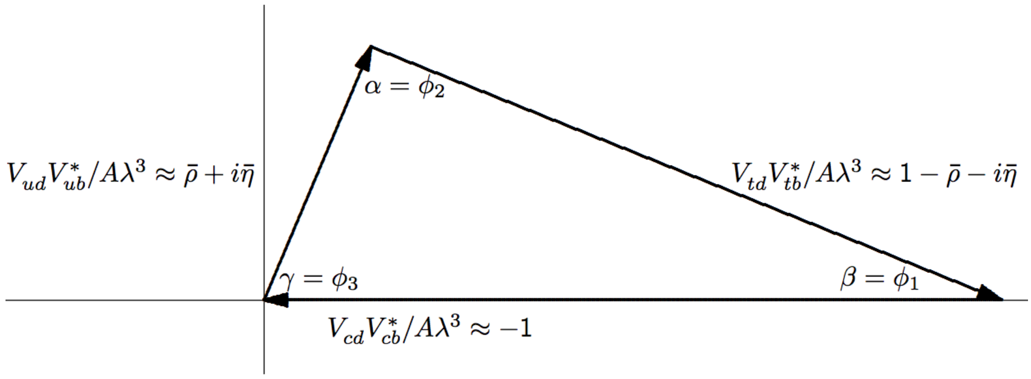


Figure 1.2: *Schematic illustration of the Unitarity Triangle in the complex plane. The angle γ corresponds to the complex phase in the CKM matrix in the standard parametrisation given in Equation 1.18.*

Figure 1.3 shows current constraints on the Unitarity Triangle and Figure 1.4 shows the constraints on all three angles. The PDG [1] cites the average values from the combination of different measurements as $\alpha = \left(87.6^{+3.5}_{-3.3}\right)^\circ$, $\beta \mathrel{\hat{=}} (21.85 \pm 0.49)^\circ$ and $\gamma = \left(73.2^{+6.3}_{-7.0}\right)^\circ$ which makes γ the least constrained angle of the CKM matrix.

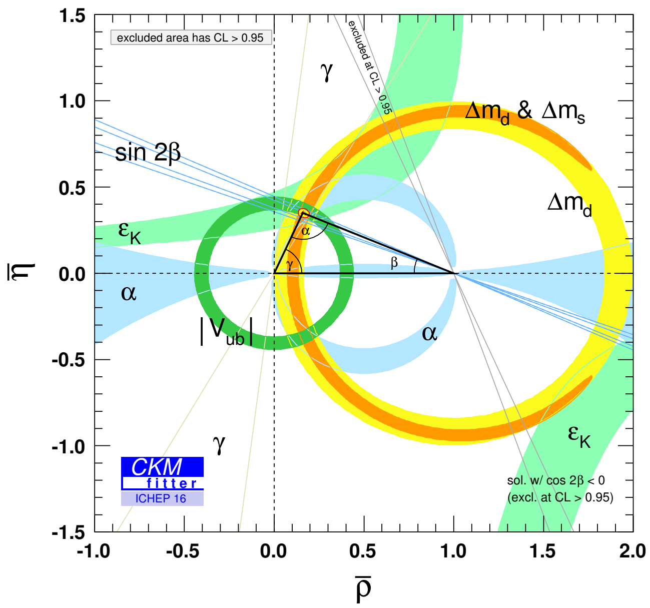


Figure 1.3: *Current state of the Unitarity Triangle as determined by the CKM fitter group [13]. The shaded areas correspond to constraints on the different parameters from a combination of different sources. The red area indicates the 68% confidence level region of the apex of the triangle.*

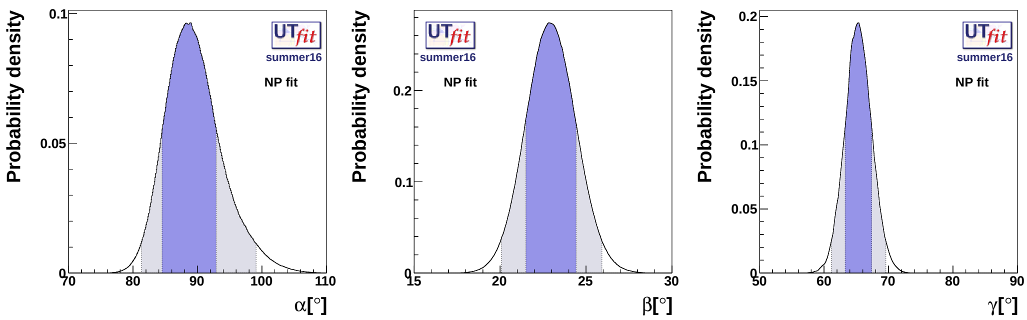


Figure 1.4: *Current predictions for the three angles of the Unitarity Triangle α (left), β (center) and γ (right) from the UT$_{fit}$ Collaborations in 2016 [14]. The constraints are obtained by performing a fit of many different measurements to the Unitarity Triangle. The effect of potential physics beyond the Standard Model is allowed in the fit.*

1.3.3 Current State of the Measurement of the CKM Angle γ at LHCb

As discussed in the last section, the PDG [1] published in 2017 lists the value for γ as $\left(73.2^{+6.3}_{-7.0}\right)^\circ$. This value is the result of the combination of direct measurements of γ perfromed by the CKM fitter group [13]. The CKM fitter group also lists the indirectly determined value of γ– obtained by a fit to the Standard Model excluding the direct measurements – as $\left(66.85^{+0.94}_{-3.44}\right)^\circ$. Thus, in order to identify a discrepancy between these two values and to probe the Standard Model, the uncertainty from the direct measurements has to be reduced.

The most current combination of γ measurements by the LHCb collaboration yields a value of $\left(76.8^{+5.1}_{-5.7}\right)$[4] [15]. This value comes from a combination of 13 different analyses of $B \to DK$ decays where the D meson is reconstructed in a number of final states, such as $B^\pm \to D(\to h^+\pi^-\pi^+\pi^-)K^\pm$, $B^\pm \to D^{*0}(\to h^+h^-)K^\pm$ and $B^0_s \to D^+_s(\to h^+h^-\pi^\pm)K^\pm$. Figure 1.5 shows the contribution of different decay modes to the overall constraint on the angle γ.

The best measurement of the CKM angle γ from a single decay channel comes from the use of the GGSZ method – which will be explained in detail in Section 1.4 – with decays at LHCb. This analysis used the 3.0 fb^{-1} collected by LHCb in Run I and measured γ to be $\gamma = \left(62^{+15}_{-14}\right)^\circ$ [16]. The difference between the uncertainty and the uncertainty on the full γ combination, in addition to the results illustrated in Figure 1.5 show that a precise measurement of γ heavily relies on the combination of many different decay modes. The $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ decay mode used in this thesis is expected to have a statistical significance comparable to the $B^\pm \to D(\to K^0_S\pi^+\pi^-)K^\pm$ decay[5] and therefore be able to significantly contribute to the overall constraint on the CKM angle γ.

[4] The uncertainty on this value is smaller than the uncertainty on the result listed in the 2017 PDG. This is because the LHCb combination is newer.

[5] While the branching fraction of the $D^0 \to K^0_S\pi^+\pi^-$ decay is greater than the fraction fraction of the $D^0 \to 2\pi^+2\pi^-$, the reconstruction and selection efficiency for the $D \to K^0_S\pi^+\pi^-$ decay is lower due to the additional complication in the reconstruction and selection of the K^0_S candidate.

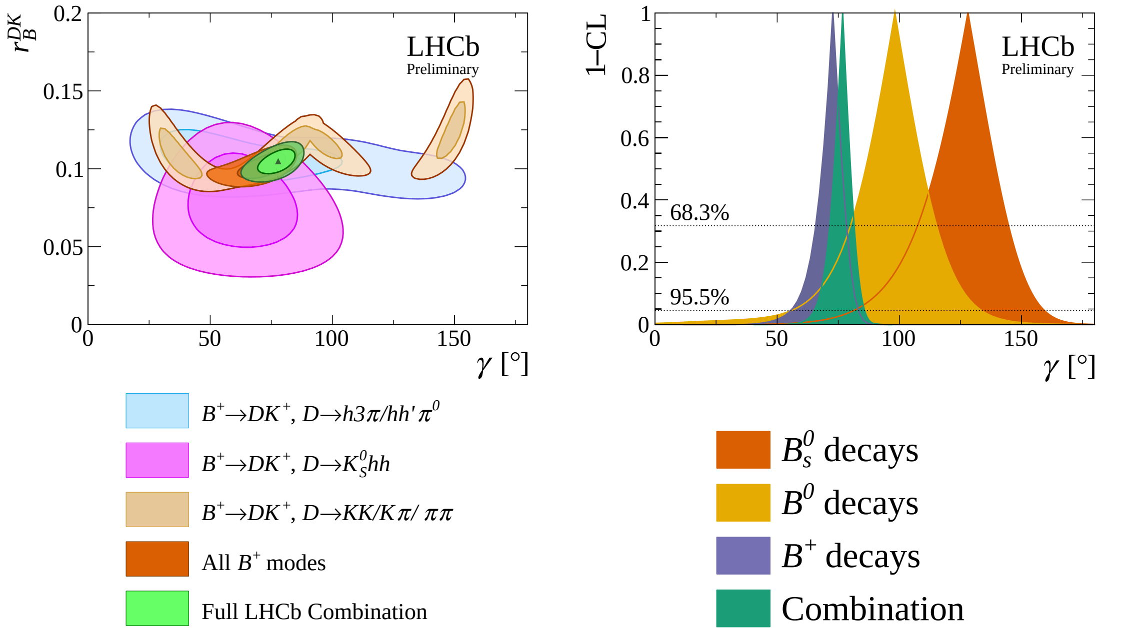


Figure 1.5: Left: *Profile likelihood contours of γ vs. r_B for various sub combinations of decays. Dark and light confidence regions corresond to 68.3% and 95.5% Confidence Levels (CL), respectively. [15]* Right: *1-CL plots of the measurements of the CKM angle γ split by the initial B meson flavour. [15]*

1.4 Model-independent Measurement of the CKM Angle γ with $B^{\pm} \to DK^{\pm}$ Decays

The $B^{\pm} \to DK^{\pm}$ decay offers a theoretically clean way to measure γ since they are predominately mediated by the tree-level processes shown in Figure 1.6. Physics beyond the Standard Model is expected to manifest itself as virtual particles in diagrams with loops. Therefore the value of γ measured with tree-level decays is unlikely to be affected by physics beyond the Standard Model and the measured value for γ corresponds to its definition in the Standard Model. The value of γ obtained using tree-level decays can be compared to alternative measurements of γ using decays that involve loop processes, and a discrepancy between the measurements will be a strong indication of physics beyond the Standard Model.

1.4.1 Principle of the measurement

This section contains the description of the principle of measuring the CKM angle γ using $B^{\pm} \to DK^{\pm}$ decays. The detailed description including the mathematical formalism is given in the following sections.

In the standard representation of the CKM matrix in Equation 1.18 the angle γ corresponds to the complex phase δ_{13}. For an amplitude $\mathcal{A} = Ae^{i\phi}$ where A is real and ϕ is the complex phase, observables – such as the decay width Γ – are proportional to $\mathcal{A}\mathcal{A}^* = AA$. In order to measure the complex phase of an amplitude directly, interference between two amplitudes has to occur. If two amplitudes $\mathcal{A}_1 =$

$A_1 e^{i\phi_1}$ and $\mathcal{A}_2 = A_2 e^{i\phi_2}$ are possible for the same process, interference between the amplitudes applies where

$$\Gamma \propto (\mathcal{A}_1 + \mathcal{A}_2)(\mathcal{A}_1 + \mathcal{A}_2)^* = A_1^2 + A_2^2 + 2A_1 A_2 \cos(\phi_1 - \phi_2) \tag{1.21}$$

and the complex phases ϕ_1 and ϕ_2 become measurable in the interference term of the decay width Γ.

The $B^\pm \to DK^\pm$ decays are an ideal way to measure the CKM angle γ since the $B^\pm$ meson can decay both via $B^\pm \to D^0 K^\pm$ and $B^\pm \to \overline{D}^0 K^\pm$. As can be seen in the Feynman diagrams shown in Figure 1.6, the $B^- \to \overline{D}^0 K^-$ decay contains the b quark to u quark transition – and thus the CKM matrix element $V_{ub} = |V_{ub}|e^{-i\gamma}$ – while the $B^- \to D^0 K^-$ decay contains the b quark to c quark transition.

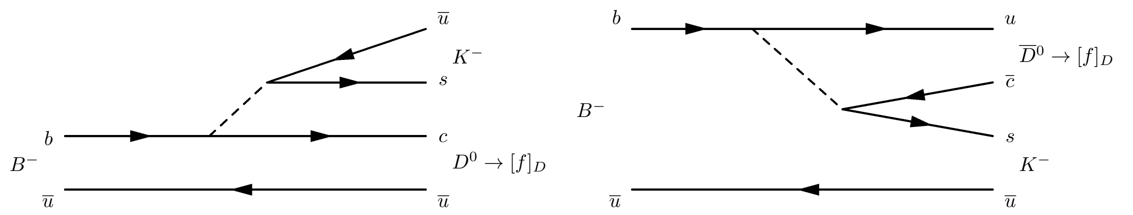


Figure 1.6: *Feynman diagrams for the $B^\pm \to D^0 K^\pm$ decays (left) and $B^\pm \to \overline{D}^0 K^\pm$ decays (right). The amplitude for the $B^\pm \to \overline{D}^0 K^\pm$ decays contains the CKM matrix element $V_{ub} = |V_{ub}|e^{-i\gamma}$.*

If the D meson is reconstructed in a final state $f_{\mathbf{p}}$ accessible to both D^0 mesons and $\overline{D}^0$ mesons, interference between the two amplitudes $B^\pm \to D^0(\to f_{\mathbf{p}})K^\pm$ and $B^\pm \to \overline{D}^0(\to f_{\mathbf{p}})K^\pm$ takes place. The principle of this interference is illustrated in Figure 1.7.

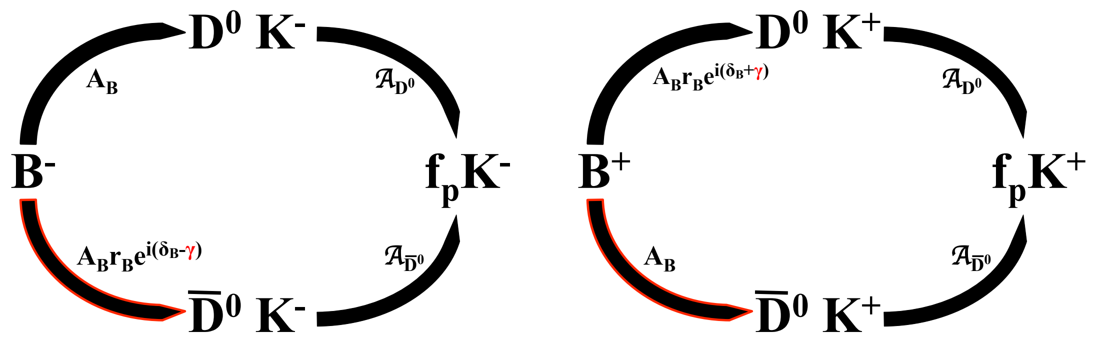


Figure 1.7: *Schematic illustration of interference in $B^\pm \to DK^\pm$ decays. The $B^\pm \to DK^\pm$ decay can proceed via $B^\pm \to D^0 K^\pm$ and $B^\pm \to \overline{D}^0 K^\pm$, thus if the D meson is reconstructed in a final state accessible to both D^0 and $\overline{D}^0$ mesons interference between the two amplitudes takes place.*

The phases of the B^+ and B^- decay amplitudes can then be determined by comparing the observed decay rate with the decay rates expected for the $B^\pm \to D^0 K^\pm$ transition and the $B^\pm \to \overline{D}^0 K^\pm$ transition.

In the analysis described in this document, the D meson is reconstructed in an multi-body final state, in particular as $D \to 2\pi^+ 2\pi^-$. The four charged pions span a five-dimensional phase space, over which the magnitude and the strong phase of the D^0 meson decay varies[6]. The interference pattern of the $B^{\pm} \to D^0K^{\pm}$ amplitude and the $B^{\pm} \to \overline{D}^0K^{\pm}$ amplitude can then be observed over the phase space of the D meson decay. Since this means that γ can be observed at all points of the D decay, the sensitivity to γ is enhanced with respect to two-body D decays and analyses where the observables are integrated over the entire D meson decay phase space.

In the GGSZ method used in this analysis (see Section 1.4.3), the D meson decay phase space is divided into bins and the observable decay width is integrated over these bins. This allows for an enhanced sensitivity to γ by observing γ at different points of the D meson decay phase space while also allowing for a model-independent approach where the quantities related to the D meson decay amplitude are expressed by amplitude-model independent quantities (see Section1.4.3).

1.4.2 The CKM Angle γ as an Observable in $B^{\pm} \to DK^{\pm}$ Decays

The amplitudes for the B^- meson decay can be defined as

$$\mathcal{A}(B^- \to D^0K^-) = A_B \tag{1.22}$$

$$\mathcal{A}(B^- \to \overline{D}^0K^-) = A_B r_B e^{i(\delta_B - \gamma)} \tag{1.23}$$

where A_B is real, r_B is the ratio between the amplitudes and δ_B is the *CP* invariant phase difference between the amplitudes, called the *the strong phase difference*. The corresponding amplitudes for the B^+ meson decay are given by

$$\mathcal{A}(B^+ \to D^0K^+) = A_B r_B e^{i(\delta_B + \gamma)} \tag{1.24}$$

$$\mathcal{A}(B^+ \to \overline{D}^0K^+) = A_B \tag{1.25}$$

where the strong phase δ_B appears with the same sign as in the *CP*-conjugated B^- amplitude while the *CP*-violating phase γ changes sign under the *CP* conjugation. The amplitudes of the subsequent D^0 and $\overline{D}^0$ meson decays to a final state $f_{\mathbf{p}}$ can be written as

$$\mathcal{A}(D^0 \to f(\mathbf{p})) = A^f_{\mathbf{p}} e^{i\delta^f_{\mathbf{p}}} \tag{1.26}$$

$$\mathcal{A}(\overline{D}^0 \to f(\mathbf{p})) = \bar{A}^f_{\mathbf{p}} e^{i\bar{\delta}^f_{\mathbf{p}}} \tag{1.27}$$

where $A^f_{\mathbf{p}}$ and $\bar{A}^f_{\mathbf{p}}$ are real and $\delta^f_{\mathbf{p}}$ and $\bar{\delta}^f_{\mathbf{p}}$ are the *CP* invariant strong phases. The parameter $\mathbf{p}$ describes a point in the phase space of the $D \to f$ decay, and has a dimensionality that depends on the number of final state particles and their spin. For two-, three- and four-body pseudo-scalar final states the phase space dimensionality is 0, 2 and 5, respectively. Since no signs of *CP* violation has been observed in charm decays [17], *CP* conservation is assumed in $D^0 \to 2\pi^+2\pi^-$ decays and the amplitudes for the D^0 meson and the $\overline{D}^0$ meson decay can be related via

$$\mathcal{A}(\overline{D}^0 \to f(\mathbf{p})) = \mathcal{A}(D^0 \to f(\overline{\mathbf{p}})) \tag{1.28}$$

[6] The assumption of negligible violation in D meson decays is made. The weak phase of the D mesons decay is thus set to zero.

where $\overline{\mathbf{p}}$ is the *CP* conjugate point of $\mathbf{p}$ which means that for all final state particles the charges are reversed (C) and three momenta are reversed (P).

The observable decay width, Γ, for $B^- \to D\,(\to f_{\mathbf{p}})\,K^-$ decays is build from the amplitudes according to

$$\begin{aligned}\Gamma(B^- \to DK^-, D \to f_{\mathbf{p}}) \propto |\, &\mathcal{A}(B^- \to D^0K^-) \cdot \mathcal{A}(D^0 \to f(\mathbf{p})) \\ &+ \mathcal{A}(B^- \to \overline{D}{}^0K^-) \cdot \mathcal{A}(\overline{D}{}^0 \to f(\mathbf{p}))\,|^2\end{aligned} \tag{1.29}$$

and equivalently for $B^+ \to D\,(\to f_{\mathbf{p}})\,K^+$ as

$$\begin{aligned}\Gamma(B^+ \to DK^+, D \to f_{\mathbf{p}}) \propto |\, &\mathcal{A}(B^+ \to D^0K^+) \cdot \mathcal{A}(D^0 \to f(\mathbf{p})) \\ &+ \mathcal{A}(B^+ \to \overline{D}{}^0K^+) \cdot \mathcal{A}(\overline{D}{}^0 \to f(\mathbf{p}))\,|^2\ .\end{aligned} \tag{1.30}$$

The decay widths can be expressed in terms of the amplitudes in Equations 1.22, 1.24 and 1.26 as

$$\begin{aligned}\Gamma(B^- \to DK^-, D \to f_{\mathbf{p}}) &\propto \left| A^f_{\mathbf{p}}\, e^{i\delta^f_{\mathbf{p}}} + r_B\, e^{i(\delta_B-\gamma)}\, \bar{A}^f_{\mathbf{p}}\, e^{i\bar{\delta}^f_{\mathbf{p}}} \right|^2 \\ &\propto A^{f\,2}_{\mathbf{p}} + r_B^2\, \bar{A}^{f\,2}_{\mathbf{p}} + 2\, A^f_{\mathbf{p}}\, \bar{A}^f_{\mathbf{p}} \left[x_- \cos(\Delta\delta^f_{\mathbf{p}}) + y_- \sin(\Delta\delta^f_{\mathbf{p}})\right],\end{aligned} \tag{1.31}$$

and equivalently for the $B^+ \to DK^+$ decay as

$$\begin{aligned}\Gamma(B^+ \to DK^+, D \to f_{\mathbf{p}}) &\propto \left| r_B\, e^{i(\delta_B+\gamma)} A^f_{\mathbf{p}}\, e^{i\delta^f_{\mathbf{p}}} + \bar{A}^f_{\mathbf{p}}\, e^{i\bar{\delta}^f_{\mathbf{p}}} \right|^2 \\ &\propto r_B^2\, A^{f\,2}_{\mathbf{p}} + \bar{A}^{f\,2}_{\mathbf{p}} + 2\, A^f_{\mathbf{p}}\, \bar{A}^f_{\mathbf{p}} \left[x_+ \cos(\Delta\delta^f_{\mathbf{p}}) + y_+ \sin(\Delta\delta^f_{\mathbf{p}})\right],\end{aligned} \tag{1.32}$$

where $\Delta\delta^f_{\mathbf{p}} = \delta^f_{\mathbf{p}} - \bar{\delta}^f_{\mathbf{p}}$ is the strong phase difference of the D decays and $x_\pm = r_B \cos(\delta_B \pm \gamma)$ and $y_\pm = r_B \sin(\delta_B \pm \gamma)$ contain the CKM angle γ. The angle γ is thus observable in the difference of the decay widths of B^- mesons and B^+ mesons in $B^\pm \to DK^\pm$ decays where the D mesons decays to a final state accessible to both D^0 and $\overline{D}{}^0$ mesons.

Equations 1.31 and 1.32 also show that apart from a complex phase in the CKM matrix, two further conditions need to be fulfilled for *CP* violation to be observable. The first condition is the existence of at least two paths from an initial state to a final state which interfere with each other. The second condition is that the strong phase difference δ_B differs from zero.

While measurements of the CKM angle γ with $B^\pm \to DK^\pm$ decays usually determine γ, δ_B and r_B simultaneously, the real amplitude $A^f_{\mathbf{p}}$ and the strong phase difference $\Delta\delta^f_{\mathbf{p}}$ of the D meson decay at the point $\mathbf{p}$ in phase space have to be determined independently. These quantities can either be obtained from a D meson decay amplitude model (see Section 1.6.1) or measured in a model-independent way. While a D meson decay amplitude model has the advantage of describing every infinitesimal point in phase space, it also introduces a significant systematic uncertainty associated with modelling the strong phase of the D decay amplitude across the multi-dimensional phase. To avoid this systematic uncertainty the D meson decay phase space can be divided into bins and integrating over each bin. The decay widths from Equations 1.31 and 1.32 can then be expressed in terms of model-independent quantities of the D meson decay called the hadronic parameters. This method will be explained in the next section.

1.4.3 The GGSZ Method for $D \to 2\pi^+2\pi^-$

The neutral D meson can be reconstructed in a variety of final states. The method that reconstructs the D meson in a *CP* eigenstate is called the GLW method [18]. Alternatively, the D meson can be reconstructed in a final state which is Cabibbo suppressed for either D^0 or $\overline{D}^0$ and Cabibbo favoured for the other D meson. This method is called the ADS method [19]. The current best single measurement of the angle γ comes from an analysis using the GGSZ method [20]. In the GGSZ method the D meson is reconstructed in a self-conjugate multi-body final state and the measurement for γ is evaluated at different points of the D decay phase space.

By reconstructing the D meson in a multi-body final state the sensitivity to γ is enhanced through observation of the interference pattern over the phase space of the D meson decay. The original proposal of the GGSZ method – named after its proponents A. Giri, Y. Grossman, A. Soffer and J. Zupan – uses $D \to K^0_S\pi^+\pi^-$, $D \to K^0_S K^+K^-$ and $D \to K^0_S\pi^+\pi^-\pi^0$ decays. In the analysis described in this document the D mesons are reconstructed as $D \to 2\pi^+2\pi^-$.

The GGSZ method is a model-independent approach. This means that the quantities in Equations 1.31 and 1.32 that are related to the D meson decay amplitudes, namely $A^f_{\mathbf{p}}$, $\bar{A}^f_{\mathbf{p}}$ and $\Delta\delta^f_{\mathbf{p}}$, are not determined from a D^0 decay amplitude model but rather replaced with model-independent quantities. The decay amplitude of a multi-body decay is a phenomenological object and a significant uncertainty is associated with the modelling of the complex strong phase over the multi-body phase space. This systematic uncertainty can be avoided by constructing model-independent quantities called *hadronic parameters* for the D decay, which are described in the following.

The *flavour-tagged yields* K^f_i and $\bar{K}^f_i$ in bin i of $D^0 \to f$ and $\overline{D}^0 \to f$ respectively, are defined as

$$K^f_i = \int_i |A^f_{\mathbf{p}}|^2 \phi(\mathbf{p}) \mathrm{d}\mathbf{p} \qquad \bar{K}^f_i = \int_i |\bar{A}^f_{\mathbf{p}}|^2 \phi(\mathbf{p}) \mathrm{d}\mathbf{p}, \tag{1.33}$$

where $\phi(\mathbf{p})$ gives the density of states at $\mathbf{p}$. The flavour-tagged yields can be normalised to give the *flavour-tagged fraction* of yields in bin i

$$T^f_i = \frac{K^f_i}{\sum_i K^f_i} \qquad \bar{T}^f_i = \frac{\bar{K}^f_i}{\sum_i \bar{K}^f_i} \tag{1.34}$$

with $\sum_i T^f_i = 1$ and $\sum_i \bar{T}^f_i = 1$.

The *amplitude weighted cosine*, c_i^f, and *amplitude weighted sine*, s_i^f, averaged over bin i are defined as

$$c_i^f = \frac{1}{\sqrt{K_i^f \bar{K}_i^f}} \int_i |A_{\mathbf{p}}^f||\bar{A}_{\mathbf{p}}^f| \cos\left(\Delta\delta_{\mathbf{p}}^f\right) \phi(\mathbf{p}) \mathrm{d}\mathbf{p}, \tag{1.35}$$

$$s_i^f = \frac{1}{\sqrt{K_i^f \bar{K}_i^f}} \int_i |A_{\mathbf{p}}^f||\bar{A}_{\mathbf{p}}^f| \sin\left(\Delta\delta_{\mathbf{p}}^f\right) \phi(\mathbf{p}) \mathrm{d}\mathbf{p} \ . \tag{1.36}$$

Integrating the decay widths in Equations 1.31 and 1.32 over the phase space of bin i gives,

$$\Gamma(B^- \to DK^-, D \to f_i) \propto \bar{T}_i^f r_B^2 + T_i^f + 2\sqrt{T_i^f \bar{T}_i^f}(c_i^f x_- + s_i^f y_-) \tag{1.37}$$

$$\Gamma(B^+ \to DK^+, D \to f_i) \propto T_i^f r_B^2 + \bar{T}_i^f + 2\sqrt{T_i^f \bar{T}_i^f}(c_i^f x_+ - s_i^f y_+) \tag{1.38}$$

where the CKM angle γ is found in $x_\pm = r_B \cos(\delta_B \pm \gamma)$ and $y_\pm = r_B \sin(\delta_B \pm \gamma)$.

Furthermore, the symmetry of the self-conjugate D meson final state $2\pi^+2\pi^-$ can be exploited by defining bins in pairs which map onto each other under *CP* transformation. The bins are labelled such that bin i maps onto bin $-i$. For any point $\mathbf{p}$ that falls in bin i the *CP* conjugate point $\overline{\mathbf{p}}$ falls into bin $-i$. This choice of binning and Equation 1.28 leads to the relations between the hadronic parameters of bin i and $-i$

$$\begin{aligned} K_{-i}^f &= \bar{K}_i^f \qquad & \bar{K}_{-i}^f &= K_i^f \\ T_{-i}^f &= \bar{T}_i^f \qquad & \bar{T}_{-i}^f &= T_i^f \\ c_{-i}^f &= c_i^f \qquad & s_{-i}^f &= -s_i^f \ . \end{aligned} \tag{1.39}$$

This choice of binning scheme is adapted throughout the rest of this chapter. More details on the exact binning scheme used in later analyses is given in Section 1.6.

With the relations in Equation 1.39 the decay widths from Equation 1.31 and Equation 1.32 can be expressed as

$$\Gamma(B^- \to DK^-, D \to f_i) \propto T_{-i}^f r_B^2 + T_i^f + 2\sqrt{T_i^f T_{-i}^f}(c_i^f x_- + s_i^f y_-) \tag{1.40}$$

$$\Gamma(B^+ \to DK^+, D \to f_i) \propto T_i^f r_B^2 + T_{-i}^f + 2\sqrt{T_i^f T_{-i}^f}(c_i^f x_+ - s_i^f y_+) \ . \tag{1.41}$$

This shows that the CKM angle γ can be measured with $B^\pm \to DK^\pm$ decays in a way that is independent of the D meson decay model. Additional sensitivity is gained by dividing the D meson decay phase space into bins. The comparison between the distribution of $B^- \to DK^-$ events over the bins and the distribution of $B^+ \to DK^+$ events over the bins enables the decoupling of the strong phase δ_B and the *CP*-violating weak phase γ. The measurement of the hadronic parameters is explained in Section 1.5.

1.4.4 Sensitivity to the CKM Angle γ with the *CP*-even Fraction F_+

While an increased sensitivity to the angle γ is obtained by dividing the D meson decay phase space into bins, γ can also be measured using the simpler approach of integrating over the entire D meson decay phase space. The $B^\pm \to DK^\pm$ decay widths can then be expressed in terms of the model-independent *CP*-even fraction of the D^0 decay.

The *CP*-even fraction F_+^f of a D meson decaying to a final state f is defined as

$$F_+^f \equiv \frac{\int_{\mathbf{p}} |\mathcal{A}(D_{CP+} \to f(\mathbf{p}))|^2 \phi(\mathbf{p})\mathrm{d}\mathbf{p}}{\int_{\mathbf{p}} |\mathcal{A}(D_{CP+} \to f(\mathbf{p}))|^2 + |\mathcal{A}(D_{CP-} \to f(\mathbf{p}))|^2 \ \phi(\mathbf{p})\mathrm{d}\mathbf{p}} \ . \tag{1.42}$$

where *CP* eigenstates $|D_{CP+}\rangle$ and $|D_{CP-}\rangle$ of the neutral D meson can be expressed in terms of the mass eigenstates $|D^0\rangle$ and $|\overline{D}^0\rangle$ through

$$\begin{aligned} |D_{CP+}\rangle &= \frac{1}{\sqrt{2}}\left(|D^0\rangle + |\overline{D}^0\rangle\right) \\ |D_{CP-}\rangle &= \frac{1}{\sqrt{2}}\left(|D^0\rangle - |\overline{D}^0\rangle\right) \ . \end{aligned} \tag{1.43}$$

This relates the decay amplitudes of the *CP* eigenstates to the amplitudes of the mass eigenstates via

$$\begin{aligned} \mathcal{A}(D_{CP+} \to f(\mathbf{p})) =& \frac{1}{\sqrt{2}}\left(\mathcal{A}(D^0 \to f(\mathbf{p})) + \mathcal{A}(\overline{D}^0 \to f(\mathbf{p}))\right) \\ \mathcal{A}(D_{CP-} \to f(\mathbf{p})) =& \frac{1}{\sqrt{2}}\left(\mathcal{A}(D^0 \to f(\mathbf{p})) - \mathcal{A}(\overline{D}^0 \to f(\mathbf{p}))\right) \ . \end{aligned} \tag{1.44}$$

The *CP*-even fraction F_+^f is thus given by

$$F_+^f = \frac{\int_{\mathbf{p}} A_{\mathbf{p}}^{f\,2} + \bar{A}_{\mathbf{p}}^{f\,2} + A_{\mathbf{p}}^f \bar{A}_{\mathbf{p}}^f \cos\left(\Delta\delta_{\mathbf{p}}^f\right) \phi(\mathbf{p})\mathrm{d}\mathbf{p}}{\int_{\mathbf{p}} A_{\mathbf{p}}^{f\,2} + \bar{A}_{\mathbf{p}}^{f\,2} \ \phi(\mathbf{p})\mathrm{d}\mathbf{p}} \ . \tag{1.45}$$

With the definitions for the D decay amplitudes from Equation 1.26, $\Delta\delta_{\mathbf{p}}^f = \delta_{\mathbf{p}}^f - \bar{\delta}_{\mathbf{p}}^f$ and a binning scheme for which the relations from Equation 1.39 hold, the *CP*-even fraction transforms to

$$F_+^f = \frac{1}{2}\sum_{i>0}\left(T_i^f + T_{-i}^f + 2c_i^f\sqrt{T_i^f T_{-i}^f}\right) \ . \tag{1.46}$$

Summing the decay width for $B^- \to DK^-$ decays in Equation 1.40 over all bins and using $\sum_i T_i^f = 1$, $c_i^f = c_{-i}^f$ and $s_i^f = -s_{-i}^f$ gives

$$\begin{aligned} \Gamma(B^- \to DK^-, D \to f) &\propto r_B^2 + 1 + 4x_- \sum_{i>0}\left(c_i^f\sqrt{T_i^f T_{-i}^f}\right) \\ &\propto r_B^2 + 1 + 2x_-\left(2F_+^f - 1\right) \end{aligned} \tag{1.47}$$

and equivalently for $B^+ \to DK^+$ decays in Equation 1.41

$$\begin{aligned}\Gamma(B^+ \to DK^+, D \to f) &\propto 1 + r_B^2 + 4x_+ \sum_{i>0} \left(c_i^f \sqrt{T_i^f T_{-i}^f} \right) \\ &\propto r_B^2 + 1 + 2x_+ \left(2F_+^f - 1 \right) \ . \end{aligned} \tag{1.48}$$

Thus, a model-independent measurement of the CKM angle γ can be performed with a multi-body D decay final state even when integrating over the entire D decay phase space. The comparison between the decay width of $B^- \to DK^-$ events and the decay width of $B^+ \to DK^+$ events enables the decoupling of the strong phase δ_B and the *CP*-violating weak phase γ. The measurement of the *CP*-even fraction $F_{4\pi}^+$ is explained in Section 1.5.4.

1.5 Measurement of the Hadronic Parameters of the Multi-body D Decay

The hadronic parameters of the D meson decay can be measured using pairs of correlated D^0 and $\overline{D}^0$ mesons. A good source for these correlated pairs is the $\psi(3770)$ resonance which can directly and efficiently be created in e^+e^- collisions. The $\psi(3770)$ resonance then decays into the $D^0\overline{D}^0$ meson pair with a branching fraction of 93% [1]. The $\psi(3770)$ resonance has defined parity and charge conjugation quantum numbers. Since the decay of the $\psi(3770)$ resonance to the $D^0\overline{D}^0$ meson pair proceeds via the strong interaction which is invariant under parity transformation and charge conjugation, the $D^0\overline{D}^0$ meson pair has the same, defined quantum numbers. This also means that the flavour and *CP* content of the $D^0\overline{D}^0$ meson pair is known.

1.5.1 Principle of the measurement

This section contains the description of the principle of measuring the hadronic parameters of a multi-body D meson decay using $\psi(3770) \to D\overline{D}$ events. The detailed description including the mathematical formalism is given in the following sections.

The amplitude-model independent hadronic parameters of the multi-body D decay can be measured using correlated pairs of $D^0\overline{D}^0$ mesons [21]. Correlated $D^0\overline{D}^0$ mesons are created in the strong decays of the $\psi(3770) \to D\overline{D}$ resonance. The $\psi(3770)$ resonance has a charge conjugation eigenvalue of $n_C = -1$ [1] which results in an asymmetric wave function for the resulting pair of D^0 meson and $\overline{D}^0$ meson, given by

$$|\psi(3770)\rangle = |D^0\overline{D}^0\rangle - |\overline{D}^0 D^0\rangle \ . \tag{1.49}$$

The antisymmetry of the $\psi(3770)$ wave function is thus transferred to the $D\overline{D}$ state and induces quantum correlations between the two D mesons. In particular, if one D meson is in a flavour eigenstate such as D^0 the other meson has to be in the opposite flavour eigenstate $\overline{D}^0$. Equivalently, if one D meson is in a *CP* eigenstate, the other D meson is required to be in a *CP* eigenstate with opposite eigenvalue. More generally, due to the specific quantum numbers of the $\psi(3770)$ resonance ($n_C = -1$ and $n_P = -1$), the flavour and *CP* content of the $D\overline{D}$ state is fixed.

This means that the flavour or *CP* content of one D meson can be *tagged* by reconstructing the other D meson in a flavour or *CP* eigenstate. This can directly be used to determine the flavour-tagged fractions T_i as well as the *CP*-even fraction F_+ of a decay of interest. The combination of the knowledge of the flavour-tagged fractions T_i with the analysis of the events where one meson is reconstructed in a *CP* eigenstate and the other meson is reconstructed in the decay of interest, gives access to the amplitude weighted cosine and sine of the strong phase differences c_i and s_i.

The sensitivity to the *CP*-even fraction of the D meson decay of interest can be further enhanced by also reconstructing events where the other D meson is not in a *CP* or flavour eigenstate – such as $D \to K^0_{\mathrm{S,L}}\pi^+\pi^-$. The hadronic parameters – and therefore the *CP* content – of $D \to K^0_{\mathrm{S}}\pi^+\pi^-$ is known and since the *CP* content of the $D^0\overline{D}^0$ meson pair is known, the *CP* content of the decay of interest can be inferred. Additional sensitivity can be gained by observing the variation of $D^0\overline{D}^0$ events over the $D \to K^0_{\mathrm{S,L}}\pi^+\pi^-$ phase space. Different regions of the $D \to K^0_{\mathrm{S,L}}\pi^+\pi^-$ decay phase space have different *CP* contents and are thus effected in a different way by being reconstructed against the decay of interest.

1.5.2 Measurement of the Flavour-tagged Fractions T_i

The quantum correlation of the $D^0\overline{D}^0$ meson pair can be used to measure the flavour-tagged fractions T_i^f by reconstructing one D meson in the decay of interest – such as $2\pi^+2\pi^-$ – and the other D meson in a flavour-eigenstate – such as $K^\pm e^\mp \nu$. The flavour-tagged fraction T_i^f is then the fraction of $D \to 2\pi^+2\pi^-$ events in bin i.

1.5.3 Measurement of the Amplitude-Weighted Cosine and Sine of the Strong Phase Differences c_i and s_i

If one D meson is reconstructed in a final state $f(\mathbf{p})$ and the other meson in a final state $g(\mathbf{q})$, the decay amplitude is given by[7]

$$\begin{aligned}\mathcal{A}\left(\psi(3770) \to D^0\overline{D}^0 \to f(\mathbf{p})g(\mathbf{q})\right) \propto\, & \mathcal{A}(D^0 \to f(\mathbf{p}))\,\mathcal{A}(\overline{D}^0 \to g(\mathbf{q})) \\ & - \mathcal{A}(D^0 \to g(\mathbf{q}))\,\mathcal{A}(\overline{D}^0 \to f(\mathbf{p}))\ . \end{aligned} \tag{1.50}$$

[7]The effects from D meson mixing are neglected throughout. Since the $D^0\overline{D}^0$ mesons are correlated and evolve coherently, the effects of mixing only become observable once one D meson decays. As can be seen in Reference [22] the D meson mixing adds a term to the decay width which is scaled by $5.6 \cdot 10^{-5}$ with respect to the decay width given by Equation 1.52. This is completely negligible.

which is expressed with the D decay amplitudes defined in Equation 1.26 as

$$\begin{aligned}\mathcal{A}\left(\psi(3770) \to D^0\overline{D}^0 \to f(\mathbf{p})g(\mathbf{q})\right) &\propto A^f_{\mathbf{p}}\bar{A}^g_{\mathbf{q}}e^{i\delta_f}e^{i\bar{\delta}_g} - \bar{A}^f_{\mathbf{p}}A^g_{\mathbf{q}}e^{i\bar{\delta}_f}e^{i\delta_g} \\ &\propto A^{f\,2}_{\mathbf{p}}\,\bar{A}^{g\,2}_{\mathbf{q}} + \bar{A}^{f\,2}_{\mathbf{p}}\,A^{g\,2}_{\mathbf{q}} \\ &+ 2A^f_{\mathbf{p}}\,\bar{A}^f_{\mathbf{p}}\,\cos\left(\Delta\delta^f_{\mathbf{p}}\right)A^g_{\mathbf{q}}\,\bar{A}^g_{\mathbf{q}}\,\cos\left(\Delta\delta^g_{\mathbf{q}}\right) \\ &+ 2A^f_{\mathbf{p}}\,\bar{A}^f_{\mathbf{p}}\,\sin\left(\Delta\delta^f_{\mathbf{p}}\right)A^g_{\mathbf{q}}\,\bar{A}^g_{\mathbf{q}}\,\sin\left(\Delta\delta^g_{\mathbf{q}}\right)\ . \qquad (1.51)\end{aligned}$$

The decay width of $\psi(3770) \to D^0\overline{D}^0 \to f_i g_j$ integrated over bin[8] i of final state f and bin j of final state g is therefore given by

$$\Gamma[\psi(3770) \to D^0\overline{D}^0 \to f_i g_j] \propto T^f_i T^g_{-j} + T^f_{-i}T^g_j - 2\sqrt{T^f_i T^g_{-j}T^f_{-i}T^g_j}\left(c^f_i c^g_j + s^f_i s^g_j\right)\ . \qquad (1.52)$$

If state g is a *CP* eigenstate, then $\mathcal{A}(\mathbf{p}) = \mathcal{A}(\bar{\mathbf{p}})$ and thus $T^g_j = T^g_{-j}$, $s^g_j = 0$ and $c^g_j = n_{CP}$. This results in a simplified equation for the decay widths of

$$\Gamma[\psi(3770) \to D^0\overline{D}^0 \to f(\mathbf{p})g(\mathbf{q})] \propto T^f_i + T^f_{-i} + 2n^g_{CP}\sqrt{T^f_i T^f_{-i}}c^f_i\ . \qquad (1.53)$$

By reconstructing one D meson in the channel of interest – such as $2\pi^+2\pi^-$ – and the other D meson in different *CP* eigenstates, the values of c^f_i can be measured. The values for s^f_i can subsequently be measured using Equation 1.52 by either reconstructing both D mesons as the decay of interest or reconstructing one D meson as the decay of interest and the other D mesons as a decay for which the values of c^f_i and s^f_i are known – such as $K^0_{\mathrm{S,L}}\pi^+\pi^-$.

1.5.4 Measurement of the *CP*-even Fraction F_+

Direct sensitivity to F^f_+ can be obtained by summing Equation 1.52 over all bins i of final state f. This yields the relation for the decay widths in bin j of final state g of

$$\begin{aligned}\Gamma[\psi(3770) \to D^0\overline{D}^0 \to fg(\mathbf{q})] &\propto T^g_{-j}\sum_i T^f_i + T^g_j\sum_i T^f_{-i} \\ &+ 2\sqrt{T^g_{-j}T^g_j}\,c^g_j\,\sum_i\sqrt{T^f_i T^f_{-i}}\,c^f_i \\ &+ 2\sqrt{T^g_{-j}T^g_j}\,s^g_j\,\sum_i\sqrt{T^f_i T^f_{-i}}\,s^f_i\ . \qquad (1.54)\end{aligned}$$

By using $\sum_i T^f_i = \sum_i T^f_{-i} = 1$, $c^f_i = c^f_{-i}$ and $s^f_i = -s^f_{-i}$ the expression can be simplified to

$$\begin{aligned}\Gamma[\psi(3770) \to D^0\overline{D}^0 \to fg(\mathbf{q})] &\propto T^g_{-j} + T^g_j \\ &+ 2\sqrt{T^g_{-j}T^g_j}\,c^g_j\left(2F^f_+ - 1\right)\ . \qquad (1.55)\end{aligned}$$

[8]The binning scheme introduced in Section 1.4.3 is used in which bins are defined in pairs which map onto each other under *CP* transformation.

Therefore the *CP*-even fraction can be measured using correlated $D^0\overline{D}^0$ decays by reconstructing one D meson in the channel of interest – such as $2\pi^+2\pi^-$ – and the other D meson in a channel for which the hadronic parameters are known – such as $K^0_{\mathrm{S,L}}\pi^+\pi^-$. The measurement of the *CP*-even fraction $F^+_{4\pi}$ of $D \to 2\pi^+2\pi^-$ is described in Chapter 3.

Since the *CP*-even fraction is defined as given in Equation 1.42, it can also be measured reconstructing one D meson in the channel of interest and the other D meson in a *CP* eigenstate. Due to the correlation of the $D^0\overline{D}^0$ pair the first D meson has to be in a *CP* eigenstate with quantum number opposite to the D meson reconstructed in the *CP* eigenstate (neglecting effects from D meson mixing). This measurement is not part of this thesis.

1.6 The Binning Scheme for the *D* Meson Decay Phase Space

Sensitivity to the CKM angle γ is obtained through the interference term

$$2\sqrt{T_i^f\bar{T}_i^f}\left(c_i^f x_\pm \mp s_i^f y_\pm\right) \tag{1.56}$$

where γ appears in $x_\pm = r_B\cos(\delta_B \pm \gamma)$ and $y_\pm = r_B\sin(\delta_B \pm \gamma)$. The factors $x_\pm$ and $y_\pm$ are enhanced by c_i^f and s_i^f respectively.

Since the D^0 and $\overline{D}^0$ decay amplitudes vary over the multi-body phase space, so do the amplitude weighted cosine c_i^f and sine c_i^f of the strong phase difference. To reach an optimal precision the bins should be chosen in such a way that the integration over the phase space does not dilute the strength of the amplitude weighted (co)sine of the strong phase difference. One such binning scheme, the *equal* $\Delta\delta^f_{\mathbf{p}}$, is presented in Section 1.6.2. In this binning scheme the D decay phase space is divided according to the strong phase difference $\Delta\delta^f_{\mathbf{p}}$.

In order to find the strong phase difference at every point of the phase space a model of the D decay is used. Although that model is used to define the bins, the measured values of the hadronic parameters are still model independent. The model influences only the statistical significance of the measurement.

1.6.1 Amplitude Model

The amplitude model $\mathcal{A}^0_D(\mathbf{p})$ for a multi body D decay can be constructed by forming the coherent sum over all intermediate amplitudes $\mathcal{A}_i(\mathbf{p})$, each weighted by a complex coefficient a_i such that

$$\mathcal{A}^0_D(\mathbf{p}) = \sum_i a_i\mathcal{A}_i(\mathbf{p}) \ . \tag{1.57}$$

The intermediate amplitudes are constructed using the isobar approach which is based on the assumption that the decay process can be factorised into subsequent two-body decay amplitudes. For three body decays the decay chain looks like $D^0 \to (R \to h_1h_2)\,h_3$. For four body modes two topologies are possible where the first topology is the *twin resonance decay* $D^0 \to (R_1 \to h_1h_2)\,(R_2 \to h_3h_4)$ and the second topology is that of a *cascade decay* with $D^0 \to [R_1 \to (R_2 \to h_1h_2)\,h_3]\,h_4$. For both topologies each intermediate amplitude $\mathcal{A}_i(\mathbf{p})$ is expressed as

$$\mathcal{A}_i(\mathbf{p}) = B_{L_D}(\mathbf{p})\left[B_{L_{R_1}}(\mathbf{p})T_{R_1}(\mathbf{p})\right]\left[B_{L_{R_2}}(\mathbf{p})T_{R_2}(\mathbf{p})\right]S_iB(\mathbf{p}) \tag{1.58}$$

where the B_L are the form factors for each vertex of the decay tree, the T_R are the Breit-Wigner propagators for each resonance and S_i is the spin factor representing the overall angular distribution.

Additional constraints can be applied to the amplitude models to account for the negligible *CP* violation in neutral D decays and the two pairs of indistinguishable pions in the $D^0 \to 2\pi^+2\pi^-$ decay. The complex coefficients a_i in Equation 1.57 are determined experimentally by fitting the amplitude model to data. More detail on the $D^0 \to K^0_\mathrm{S}\pi^+\pi^-$ amplitude model and the $D^0 \to 2\pi^+2\pi^-$ amplitude model which are used in Chapter 3 and Chapter 6 can be found in Reference [23] and Reference [24], respectively.

1.6.2 Equal $\Delta\delta$ Binning

The binning scheme used to bin the $D \to K^0_{\mathrm{S,L}}\pi^+\pi^-$ decays in the measurement of the *CP*-even fraction of $D^0 \to 2\pi^+2\pi^-$ in Chapter 3 and the $D \to 2\pi^+2\pi^-$ decays in the preparation of the γ measurement in Chapter 6 is called the *equal* $\Delta\delta$ *binning*. In this binning scheme the D decay phase space is divided according to the strong phase difference $\Delta\delta$. Using a model for the D decay amplitude, a value for $\Delta\delta$ is appointed to each point $\mathbf{p}$ of the phase space. A bin number is then assigned using

$$\begin{aligned} +i:&\quad \delta_{i-1} < \Delta\delta^{4\pi}_{\mathbf{p}} < \delta_i \\ -i:&\ -\delta_{i-1} > \Delta\delta^{4\pi}_{\mathbf{p}} > -\delta_i \end{aligned} \tag{1.59}$$

where $\delta_0 \equiv 0$, $\delta_{\mathcal{N}} \equiv \pi$ and $\delta_i < \delta_{i+1}$. This automatically fulfils the requirement mentioned in Section 1.4.3 that bin $+i$ maps to bin $-i$ under *CP*, since $\Delta\delta^{4\pi}_{\mathbf{p}} \equiv -\Delta\delta^{4\pi}_{\overline{\mathbf{p}}}$. The values of δ_i are chosen according to $\delta_i = i\pi/\mathcal{N}$ where $\mathcal{N}$ is the number of pairs of bins.

1.7 Summary

The CKM angle γ is the least well measured angle of the Unitarity Triangle. In the Standard parametrisation of the CKM matrix, γ also represents the *CP*-violating complex phase and appears in the weak transition of a b quark to a u quark. Since γ is a complex phase it can only be measured through interference. $B^\pm \to DK^\pm$ decays offer a theoretically clean way to measure γ through the interference of $B^\pm \to D^0K^\pm$ and $B^\pm \to \overline{D}{}^0K^\pm$ when the D meson is reconstructed in a final state accessible to both D^0 and $\overline{D}{}^0$ mesons. Since both decays $B^\pm \to D^0K^\pm$ and $B^\pm \to \overline{D}{}^0K^\pm$ are tree level processes the measured value of γ corresponds to the Standard Model parameter without any influence of possible physics beyond the Standard Model. The value of γ obtained through the analysis of $B^\pm \to DK^\pm$ decays can be compared to alternative measurements of γ with decays that involve loop processes, and a discrepancy between the measurements will be a strong indication of physics beyond the Standard Model.

By reconstructing the D meson in a multi-body, self-conjugate final state – such as $2\pi^+2\pi^-$ – the sensitivity to γ is increased through observation of the interference pattern over the five-dimensional phase space of the D meson decay. The significant systematic uncertainty which is associated with modelling the complex phase of the D decay amplitude across the five-dimensional phase space of the four body decay can be avoided by adapting a model-independent approach. This is realised by dividing the D decay phase space into bins and integrating over each bin. The resulting D decay related quantities are called hadronic parameters and can be measured model independently.

The hadronic parameters, namely the flavour-tagged fractions T_i^f, the amplitude weighted cosine and sine of the strong phase difference c_i^f and s_i^f, can be measured using quantum correlated $D^0\overline{D}^0$ pairs created in decays of the $\psi(3770)$ resonance. One D meson is reconstructed in the decay of interest – here $D \to 2\pi^+2\pi^-$– while the other D meson is reconstructed as a flavour eigenstate for the T_i^f measurement, a CP eigenstate for the c_i^f measurement and either the final state of interest itself or a final state whose hadronic parameters are known – such as $K^0_{\mathrm{S,L}}\pi^+\pi^-$.

Sensitivity to the CKM angle γ can also be obtained by integrating over the full phase space of the multi-body D decay. The decay width can then be expressed in terms of the CP-even fraction F_+ of the D decay. Like the hadronic parameters, the CP-even fraction can be measured using quantum correlated $D\overline{D}$ pairs created in decays of the $\psi(3770)$ resonance.

The measurement of the CP-even fraction of the $D^0 \to 2\pi^+2\pi^-$ decay is documented in Chapter 3. The study of the sensitivity to the CKM angle γ that can be obtained using the GGSZ method for $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ decays at LHCb is performed in Chapter 6.

2. The CLEO-c Experiment

The CLEO experiment was the only large particle physics experiment located at the Cornell Electron Storage Ring (CESR) at the Laboratory for Elementary Particle Physics (LEPP) in Ithaca (NY). It recorded data from electron-positron-collisions at a center of mass energy between 3 and 11 GeV. While B physics was its main focus, many analyses on a variety of topics — such as charm, quarkonium and τ physics — have been performed. The collaboration's most cited result is the first measurement of the flavour-changing neutral current in the $b \to s\gamma$ transition [25], which was measured to be in accordance with the Standard Model expectation and placed constrains on possible contributions from physics beyond the Standard Model, for example from charged Higgs bosons.

In its run period from 1979 to 2008 the CLEO detector underwent several upgrades, the last of which was the CLEO-c detector. While the operation time before the CLEO-c detector was mainly spent at the $\Upsilon(4S)$ resonance to study B mesons, the operation time of CLEO-c was spent at three lower center of mass energies of $\sqrt{s} = 3770\,\mathrm{MeV}$, $\sqrt{s} = 4140\,\mathrm{MeV}$ and $\sqrt{s} = 3100\,\mathrm{MeV}$. Those energies corresponded to the $D_s^+\,D_s^-$ threshold, the $\psi(3770)$ resonance and the J/ψ resonance, respectively. The first two center of mass energies allowed CLEO-c to study D and D_s^+ mesons that were created just above threshold and in the case of the neutral D mesons in correlated pairs. This means that the events were very clean with low multiplicity, resulting in high efficiencies and low systematic errors.

The CLEO III and therefore its successor the CLEO-c detector were much more advanced with respect to previous charm experiments — such as BES and Mark III— having e.g. substantially superior particle identification, mass resolution and photon energy resolution. The CLEO-c detector also had an increased solid angle coverage of 25% relative to BES giving it an advantage in any measurements that required the reconstruction of both D mesons. Additionally, CLEO-c worked at a much higher luminosity, resulting in a data sample three magnitudes bigger than the Mark III datasets and 270 times as much D and D_s^+ data compared to BES.

This chapter is dedicated to the description of the CLEO-c detector. After a general overview over the detector the different subdetectors are presented and a summary of the detector performance is given. Then the CLEO-c trigger system is introduced as well as the reconstruction of D meson candidates within the CLEO-c reconstruction framework.

2.1 Overview of the CLEO-c Detector

The series of CLEO detectors were designed as general purpose detectors and offer a 93% coverage of the solid angle.

The small kinetic energy of the D mesons created during CLEO-c collision periods made a vertex detector obsolete. Furthermore the final state particles from collisions in CLEO-c had a smaller average momentum than in the previous CLEO experiments. Therefore, the CLEO-c detector was built with a minimal amount of material budget, even replacing its predecessor's silicon vertex detector with a wire drift chamber. Additionally, the magnetic field was lowered. The muon chambers were obsolete in CLEO-c since not even the muons could penetrate them.

The CLEO-c detector was composed of several subdetectors shown in Figure 2.1 which are presented in the following sections.

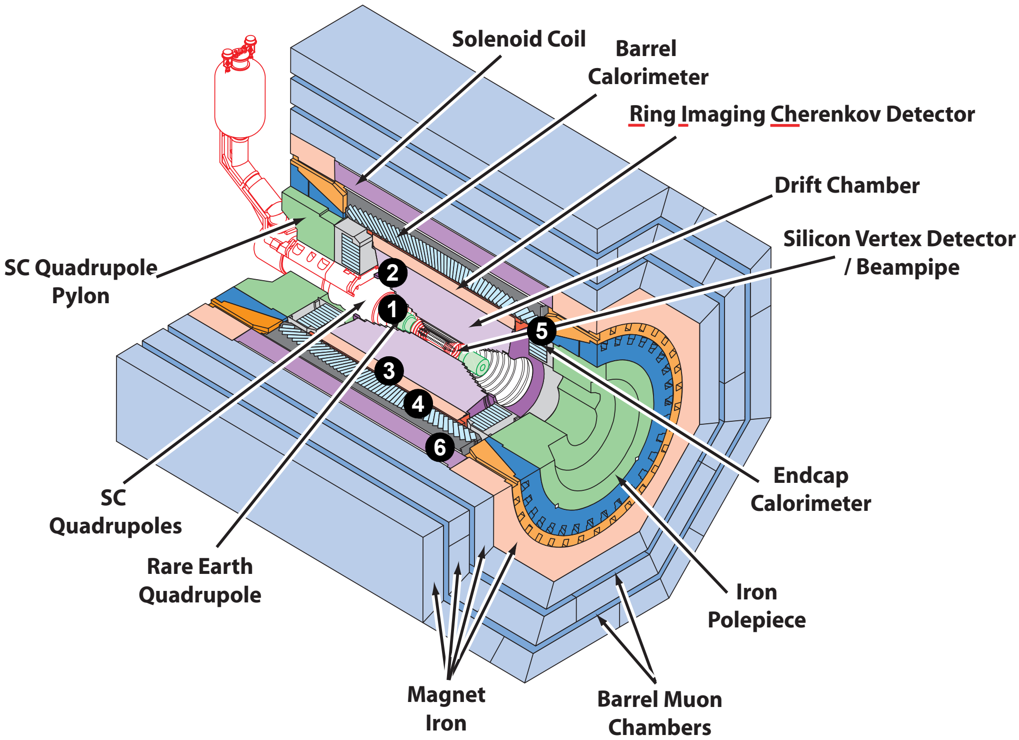


Figure 2.1: *The CLEO-c detector. 1: Inner wire drift chamber ZD, 2: Wire drift chamber, 3: Rich imaging Cherenkov detector (RICH), 4: calorimeter barrel, 5: endcap calorimeter, 6: the magnet.* [26]

2.1.1 The Inner Wire Chamber (ZD)

The ZD was the detector closest to the beam pipe. It replaced the silicon vertex detector of CLEO III with a six layer wire drift chamber, since it featured a lower material budget than the silicon detector. The ZD was used for pattern recognition in the trigger and momentum measurement of charged particle tracks. It did not provide vertexing, which was not needed since the D mesons were created barely above threshold and resulting in a decay length of about 20 to 40 μm.

The ZD consisted of 300 square cells, each measuring 10 mm across. The cells were made of one sense wire surrounded by eight parallel field wires with a potential of 1900 V between the sense and the field wires. The gas of the drift chamber was a mixture of 60% helium and 40% propane gas, chosen for its long radiation length.

The wires were approximately parallel to the beam pipe but arranged at a small stereo angle, with the innermost layer having a stereo angle of 4.4° to the beam axis and the outermost layer a stereo angle of 5.8°. The stereo angle allowed for the reconstruction of the z component of particle trajectories.

A schematic illustration as well as a close up photo of the ZD is shown in Figure 2.2.

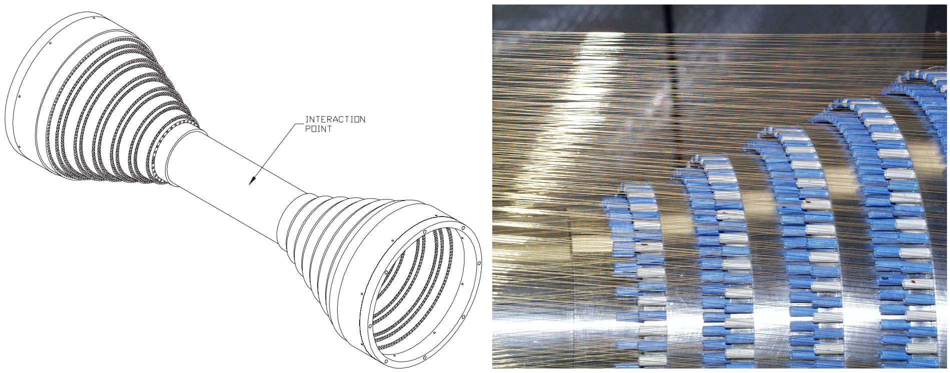


Figure 2.2: Left: *Schematic illustration of the ZD.* Right: *Close up photo of one side of the ZD of the CLEO-c detector. The photo shows the different layers of wires arranged at slightly different stereo angles.*

2.1.2 The Wire Drift Chamber

The wire drift chamber surrounded the ZD. It had the same gas mixture as the ZD and was used for tracking as well as particle identification (see Section 2.1.6).
The wire drift chamber was made up of square cells measuring 14 mm across. As for the ZD, the cells of the wire drift chamber consisted of one sense wire surrounded by eight field wires. The potential between the sense and the field wires was 2100 V [27].

The wire drift chamber had 47 layers in total. The 16 innermost layers were parallel to the beam pipe while the following layers were grouped into superlayers of 4 layers. These superlayers were arranged at alternating stereo angles between 1.2° and 1.6°. Additionally, the outer wall of the wire drift chamber was lined with cathode pads with 1 cm segmentation in z direction.

2.1.3 The Ring Imaging CHerenkov (RICH) Detector

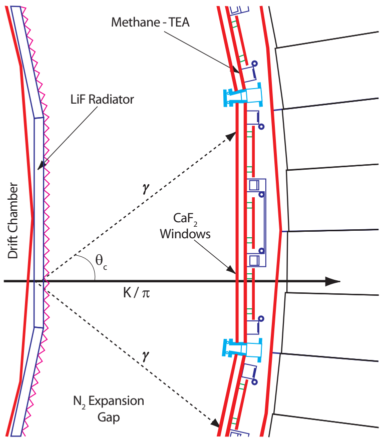
The Ring Imaging CHerenkov (RICH) detector was located outside the wire drift chamber and had a solid angle coverage of 83%. It offered particle identification for charged particles with momentum greater than 0.7 GeV [28]. The RICH detector relied on the principle of *Cherenkov radiation* which occurs when charged particles pass through a dielectric medium at a velocity greater than the speed of light in that medium [29]. The Cherenkov radiation is emitted by the charged particle at an angle that depends on the mass and momentum of the charged particle as well as the refractive index of the RICH radiator. By measuring the particle momentum p with the tracking system and the Cherenkov angle θ_C the particle mass m can be determined using

$$\cos\theta_C = \frac{1}{n\beta} \tag{2.1}$$

$$= \frac{1}{n}\frac{\sqrt{m^2+p^2}}{p} \tag{2.2}$$

where n is the refractive index of the RICH radiator and β is the ratio of the speed of the particle to the speed of light.

The RICH radiator was located at the inner boundary of the RICH detector and consisted of 14 rows of Lithium fluoride (LiF) crystals. The four rows closest to the collision point were equipped with a ”sawtooth” edge in order to avoid total internal reflection of the Cherenkov photons which can occur when the charged particles enter the crystal at a right angle as can be seen in Figure 2.3. The main volume of the RICH was an expansion gap filled with nitrogen gas. Travelling through this volume allowed the different Cherenkov photons from the same track to separate further and thus to be detected as separate photons. The outer boundary of the RICH was made of Multi Wire Proportional Chambers (MWPC) filled with a methane-triethylamine mixture in which the Cherenkov photons converted into photo-electrons. The multiplied signal of the photo-electrons was then measured via charge induced on an array of cathode pads of size 75 mm × 8 mm.



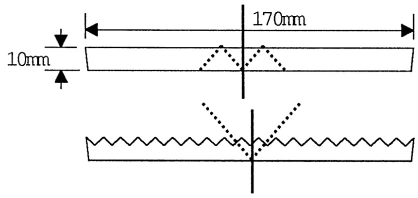


Figure 2.3: Left: *Schematic illustration of the part of the CLEO-c RICH detector. The charged particles travel from the left to the right side. The LiF radiators can be seen on the left side and the nitrogen-filled expansion gap makes up the biggest part of the detector.* Right: *Illustration of the internal reflection of Cherenkov photons in square radiators (top) and the avoidance of internal reflection from using a ”saw-tooth” edge on the relectors outer side.* [28]

2.1.4 The Calorimeter

The CLEO-c calorimeter was used for the detection and energy measurement of neutral particles such as photons, neutral pions and η mesons, for electron identification and luminosity studies [26].

The calorimeter covered about 93% of the solid angle and consisted of 7800 scintillating crystals of dimension 5 cm $\times$ 5 cm $\times$ 30 cm. The crystals were located in the central barrel region and in the two endcaps as shown in Figure 2.4. In order to reduce the probability of particles flying into the gap between the crystal and thus being undetected, the crystals in the barrel pointed towards a point slightly displaced from the collision point.

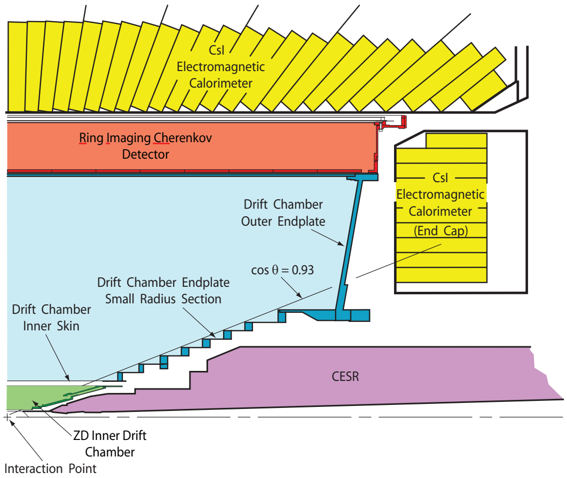


Figure 2.4: *Cross section of a quadrant to the CLEO-c detector. The orientation of the calorimeter crystals can be seen at the top.*

2.1.5 The Magnet

The CLEO-c magnet was a superconducting solenoid coil which surrounded all subdetectors, except for the obsolete muon chambers. The magnet created a uniform field with a strength of 1 T. The field lines were parallel to the beam axis (z direction), meaning that charged particles are curved in the $x - y$ plane.

2.1.6 Particle Identification

Three subdetectors were used for the particle identification (PID) in CLEO-c, namely the drift chamber, the RICH detector and the calorimeter.

The PID with the drift-chambers used the dE/dx principle of specific ionisation. This principle describes that the energy loss per distance travelled in a medium depends on the charged particle species. The mean dE/dx for charged particles is described by the Bethe-Bloch formula [1], except for electrons where it is described by the Berger-Seltzer formula [30]. Example distributions obtained by CLEO-c can be seen in Figure 2.5. In order to identify the particle species, h, the number of standard deviations from the expected energy loss for the particle hypothesis, $\chi_{\mathrm{d}E/\mathrm{d}x}(h)$, was calculated for each of the particle hypotheses, i.e. for electron, muon, pion and kaon:

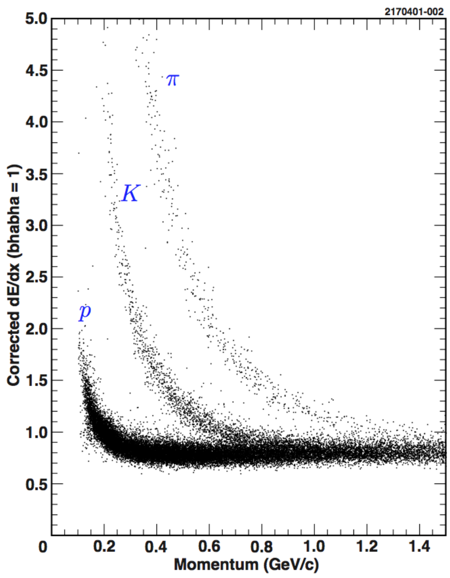
$$\chi^2_{\mathrm{d}E/\mathrm{d}x}(h) = \frac{(\mathrm{d}E/\mathrm{d}x(h) - < \mathrm{d}E/\mathrm{d}x >)^2}{\sigma^2} \tag{2.3}$$

To complement the dE/dx method, CLEO-c also used the principle of Cherenkov radiation with the RICH detector. The RICH information could only be evaluated for particle trajectories within the detector acceptance, meaning the angle θ between the track and the beam axis had to fulfil the requirement of $|\cos\theta| < 0.80$. Additionally, the RICH information was only used for tracks with a momentum above 0.7 GeV which guaranteed a decent separation of pions and kaons as can be seen in Figure 2.5. For a given particle hypothesis all photon hits within 5 standard deviations of the expected ring size for Cherenkov photons are combined to produce a likelihood value, χ_{RICH}. These likelihood are compatible with the $\chi^2_{\mathrm{d}E/\mathrm{d}x}(h)$ values from the dE/dx measurements and can be combined.

For all pion and kaon candidates a delta log likelihood $\Delta\mathcal{L}_{K-\pi}$ was constructed from the individual dE/dx and RICH particle hypothesis.

$$\Delta\mathcal{L}_{K-\pi} = \chi^2_{\mathrm{d}E/\mathrm{d}x}(K^\pm) - \chi^2_{\mathrm{d}E/\mathrm{d}x}(\pi^\pm) + \chi_{RICH}(K^\pm) - \chi_{RICH}(\pi^\pm) \tag{2.4}$$

If either the dE/dx information or the RICH information was not available for a given track, the respective χ value was set to zero.



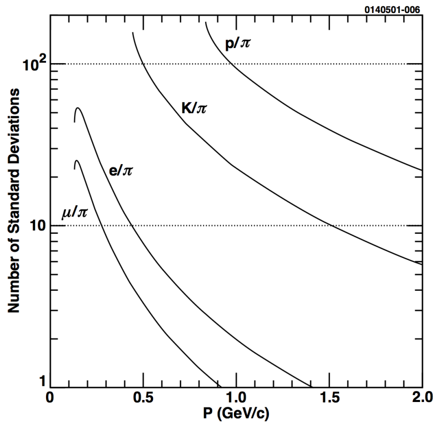


Figure 2.5: Left: dE/dx *curves for different charged particle species in the CLEO-c drift chambers. The separation between protons and kaons ceases to work above 0.5* GeV *and between kaons and pions above 0.9* GeV. *[26]* Right: *Separation in terms of the number of standard deviations between various particle hypotheses as a function of momentum as achieved by the CLEO-c detector. [31]*

Since the principles of specific ionisation and Cherenkov radiation only work for charged particles the identification of neutral particles was provided by the calorimeter. The neutral particles were identified by the shape of their shower in the calorimeter. For electron candidates, additional info was provided by the calorimeter in form of the ratio of the energy deposited in the calorimeter to the measured momentum [32].

2.1.7 CLEO-c Performance Summary

Even though the first CLEO detectors were designed for physics at a higher center of mass energy, due to the lower magnetic field and the low material budget of the inner drift chamber, the CLEO-c detector matched the performance of its predecessors.

The average drift distance resolution for the drift chambers was 85 µm [33] which yielded a momentum resolution of less than 0.5% for tracks with a momentum below 1 GeV [26]. The calorimeter attained an energy resolution of 1.5% for 5 GeV photons, 4% for 100 MeV photons and 7% for 30 MeV photons [26]. The resulting mass resolution of neutral pions from $\pi^0 \to \gamma\gamma$ is shown in Figure 2.6. Figure 2.6 also shows that the efficiency for identifying charged pions with a momentum below 0.7 GeV was stable and above 97% while the rate of wrongly identifying kaons as pions lay below 1% [34].

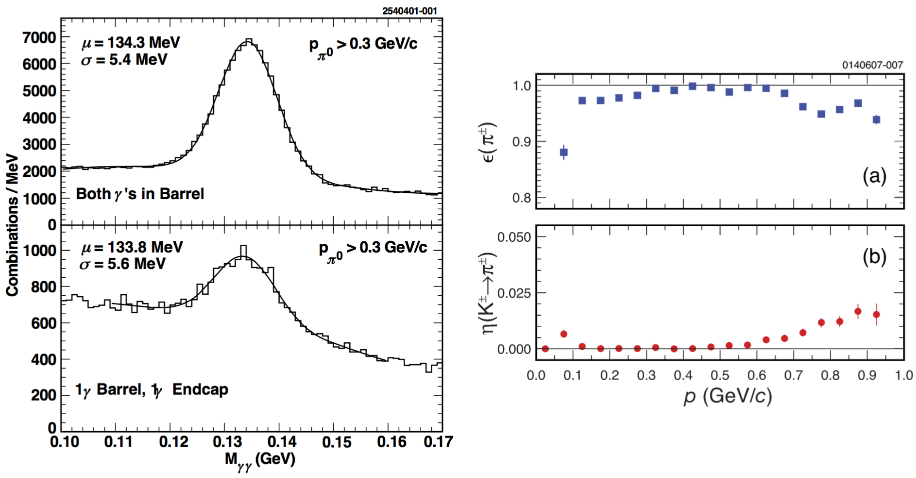


Figure 2.6: Left: *Resolution of the π^0 mass in the CLEO-c calorimeter. The π^0 candidate is reconstructed in its decay channel $\pi^0 \to \gamma\gamma$ with both γ reconstructed in the main calorimeter barrel (top) and with one γ reconstructed in the main barrel and the other γ reconstructed in the endcap (bottom).* Right: *Pion-identification efficiency as a function of the momentum p (top) and the probability of misidentifying a kaon as a pion [34].*

2.2 The CLEO-c Trigger System

The CESR accelerator had a peak luminosity of $\mathcal{L} \approx 10^{-33} \frac{1}{\mathrm{cm}^2\,\mathrm{s}}$ [35]. The events of interest in the CLEO-c physics program where predominantly $e^+e^- \to D\overline{D}$ events. At a center-of-mass energy 3770 MeV the cross-section for $e^+e^- \to D\overline{D}$ is $\sigma(D\overline{D}) = 6.57\,\mathrm{nb}$ [34]. Thus, the rate for decays of interests was approximately 7 Hz. The data acquisition system required that the data were collected and written to disk at a rate of 80 Hz, which was an order of magnitude greater than the rate by which the events of interest occurred. This means that the trigger could accept almost all events of interest, rejecting only background events such as interactions of the beam with the beam pipe. Additionally, events consistent with Bhabha scattering into the CLEO-c barrel were retained for the luminosity measurement. The trigger also provided the timing information that was used later in the event reconstruction.

The CLEO-c trigger was fully hardware based and ran on field-programmable gate arrays (FPGA). Trigger decisions were based on information from the calorimeter and the drift chamber. In a first step *trigger primitives* were calculated by individual customised circuit boards. These trigger primitives were quantities such as the shower count and topology in the calorimeter, and track count and topology in the drift chamber. Track-finding algorithms were applied to the axial and stereo components of the drift chambers individually to build all possible track patterns whose closest approach to the beam line was 5 mm or less. The showers in the calorimeter were placed in three categories according to their energy.

The information from the trigger primitives was correlated by a global trigger circuit to generate a global trigger decision for the event. An example for a positive trigger decision would be an event with more than 2 tracks and one low-energy shower. The trigger had an efficiency greater than 99% for hadronic events.

2.3 The CLEO-c D Meson Tagging (`DTag`)

The CLEO-c framework offers an efficient and clean preselection for final state particles as well as the construction of intermediate particles, such as K_S^0 mesons, and D meson candidates. Within this framework final state particles are reconstructed from tracks in the drift chambers or showers in the calorimeters. Different selection criteria are placed on the track quantities, shower energies and particle identification. These selection criteria for the final state particles were optimised for the kinematic distributions of the decays of interest as well as for the detector performance. The specific selection criteria for the final state particles used in the measurement of the CP even fraction of $D^0 \to 2\pi^+2\pi^-$ are listed in Section 3.3.2.

The final state particles are combined by the CLEO-c framework to intermediate candidates which are further combined to form a specific D meson decay. This construction and selection of D meson candidates is called `DTag` and ensures that no physics object such as tracks and showers are used more than once per event. Two individual `DTag` candidates such as $D \to 2\pi^+2\pi^-$ and $D \to K_S^0\pi^+\pi^-$ can be combined to form a `DDoubleTag`.

3. Measurement of the CP-even Fraction F_+ of the $D^0 \to 2\pi^+ 2\pi^-$ Decay using quantum correlated $D\bar{D}$ Pairs at CLEO-c

This chapter describes the measurement of the *CP*-even fraction $F_{4\pi}^+$ of the $D^0 \to 2\pi^+ 2\pi^-$ decay using $D \to K^0_{\rm S}\pi^+\pi^-$ and $D \to K^0_{\rm L}\pi^+\pi^-$ decays to tag the signal mode. In the first section of this chapter the strategy used to perform the measurement is outlined. In the second section the different data and simulation samples used in the analysis are introduced. In the third section the reconstruction and selection of the $D^0 \to 2\pi^+ 2\pi^-$ *vs.* $D \to K^0_{\rm S}\pi^+\pi^-$ and $D \to 2\pi^+ 2\pi^-$ *vs.* $D \to K^0_{\rm L}\pi^+\pi^-$ candidates is presented. In the fourth and fifth section the procedure used to extract the number of $D \to 2\pi^+ 2\pi^-$ *vs.* $D \to K^0_{\rm S}\pi^+\pi^-$ and $D \to 2\pi^+ 2\pi^-$ *vs.* $D \to K^0_{\rm L}\pi^+\pi^-$ signal events, respectively, is explained. In the sixth section the fit to extract the *CP* even fraction $F_{4\pi}^+$ is described. In the seventh section all considered systematic uncertainties are summarised. The eighth and last section contains a conclusion that includes the comparison of the results of this analysis to other measurements of $F_{4\pi}^+$.

3.1 Strategy

As outlined in Chapter 1, $F_{4\pi}^+$ can be determined using correlated $D^0\overline{D}^0$ decays where one D meson decays to $2\pi^+ 2\pi^-$ and the other D meson decays to a *CP*-mixed final state such as $K^0_{\rm S}\pi^+\pi^-$ or $K^0_{\rm L}\pi^+\pi^-$. These *double tags* are denoted as $D \to 2\pi^+ 2\pi^-$ *vs.* $D \to K^0_{\rm S}\pi^+\pi^-$ and $D \to 2\pi^+ 2\pi^-$ *vs.* $D \to K^0_{\rm L}\pi^+\pi^-$, respectively.

The sensitivity to $F_{4\pi}^+$ lies in the variation of density of events over the Dalitz plot of the $D \to K^0_{\rm S}\pi^+\pi^-$ or $D \to K^0_{\rm L}\pi^+\pi^-$ candidate. If the Dalitz plot is divided into bins, the number of $D \to 2\pi^+ 2\pi^-$ *vs.* $D \to K^0_{\rm S}\pi^+\pi^-$ events $N_i^{K^0_{\rm S}\pi^+\pi^-}$ in bin i of the $D \to K^0_{\rm S}\pi^+\pi^-$ Dalitz plot is given by

$$N_i^{K^0_{\rm S}\pi^+\pi^-} = h\left(T_i + T_{-i} - 2\,\sqrt{T_i T_{-i}}\;c_i\;\left(2\;F_+^{4\pi} - 1\right)\right) \tag{3.1}$$

and the number of $D \to 2\pi^+ 2\pi^-$ *vs.* $D \to K^0_{\rm L}\pi^+\pi^-$ events $N_i^{K^0_{\rm L}\pi^+\pi^-}$ in bin i of the $D \to K^0_{\rm L}\pi^+\pi^-$ Dalitz plot is given by

$$N_i^{K^0_{\rm L}\pi^+\pi^-} = h'\left(T'_i + T'_{-i} - 2\,\sqrt{T'_i T'_{-i}}\;c'_i\;\left(2\;F_+^{4\pi} - 1\right)\right) \tag{3.2}$$

where h and h' are normalisation factors specific to the double tag. The factors T_i and T'_i are the flavour-tagged fractions of the $D \to K^0_{\rm S}\pi^+\pi^-$ and $D \to K^0_{\rm L}\pi^+\pi^-$ candidate, respectively. The c_i and c'_i are the amplitude weighted cosine of the strong-phase difference between $D^0 \to K^0_{\rm S,L}\pi^+\pi^-$ and $\overline{D}^0 \to K^0_{\rm S,L}\pi^+\pi^-$ decays averaged over bin i. The values for T_i, T'_i, c_i and c'_i have been previously determined [36] [22] and are used as external inputs in this analysis. The *CP*-even fraction $F^+_{4\pi}$ and both the normalisation factors h and h' are determined in this analysis.

The binning scheme for the $K^0_{\rm S,L}\pi^+\pi^-$ Dalitz plot is chosen to be the $\Delta\delta$ *BaBar 2008* binning described in Section 1.6 of Chapter 1, in which the Dalitz plot is divided according to the strong-phase difference between the $D^0 \to K^0_{\rm S,L}\pi^+\pi^-$ decay and the $\overline{D}^0 \to K^0_{\rm S,L}\pi^+\pi^-$ decay, as predicted by a model developed by the BaBar collaboration [23]. The shape of the bins in the $D \to K^0_{\rm S,L}\pi^+\pi^-$ Dalitz plot can be seen in Figure 3.1. The Dalitz plot of the $D \to K^0_{\rm S,L}\pi^+\pi^-$ mode is divided into eight pairs of symmetric bins by the line $m^2_{K^0_{\rm S,L}\pi^+} = m^2_{K^0_{\rm S,L}\pi^-}$ where $m^2_{K^0_{\rm S,L}\pi^\pm}$ is the invariant-mass squared of the $K^0_{\rm S,L}\,\pi^\pm$ pair. The bins lying on one side of this line ($m^2_{K^0_{\rm S,L}\pi^+} > m^2_{K^0_{\rm S,L}\pi^-}$) are labelled $-1 \to -8$, and those on the other side $1 \to 8$.

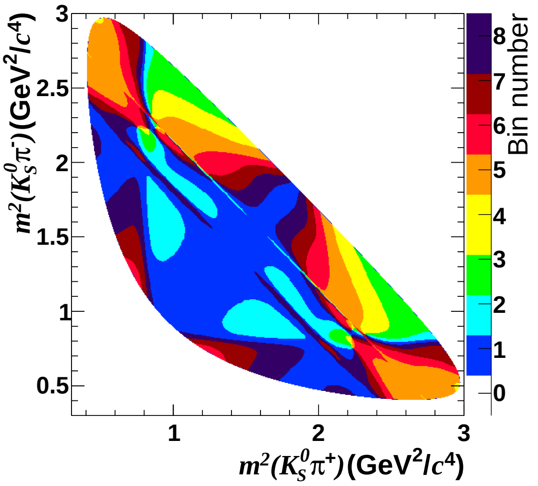


Figure 3.1: *Shape of the bins in the $D \to K^0_{\rm S,L}\pi^+\pi^-$ Dalitz plot in the $\Delta\delta$ BaBar 2008 binning scheme. The Dalitz plot is divided into bins according to the strong-phase difference between $D^0 \to K^0_{\rm S,L}\pi^+\pi^-$ and $\overline{D}^0 \to K^0_{\rm S,L}\pi^+\pi^-$ decays as predicted by the model developed by the BaBar collaboration [23].*

The yields given by Equations 3.1 and 3.2 for bin i and bin $-i$ are identical. They can therefore be added and the yields in the bin with the absolute bin number $|i|$ given by

$$N^{K^0_{\rm S}\pi^+\pi^-}_{|i|} = N^{K^0_{\rm S}\pi^+\pi^-}_{i} + N^{K^0_{\rm S}\pi^+\pi^-}_{-i} \qquad N^{K^0_{\rm L}\pi^+\pi^-}_{|i|} = N^{K^0_{\rm L}\pi^+\pi^-}_{i} + N^{K^0_{\rm L}\pi^+\pi^-}_{-i} \tag{3.3}$$

are studied in this analysis. In the following, bins denoted by i refer to the union of bin i and bin $-i$ as defined in the above binning scheme.

The first part of this analysis consists of the measurement of the distribution of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_{\rm S}\pi^+\pi^-$ events and $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_{\rm L}\pi^+\pi^-$ events over the respective $K^0_{\rm S,L}\pi^+\pi^-$ Dalitz plot. The measurement is carried out on the full CLEO-c dataset recorded in $e^+e^- \to \psi(3770)$ collisions. The $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ candidate events and $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ candidate events are reconstructed and selected using different selection criteria. Data-driven and simulation-driven techniques are used to determine the remaining contribution from different background sources. The background is divided into peaking background, and combinatorial and continuum background. After the background contributions are subtracted from the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ yields and the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ yields, the reconstruction and selection efficiency for both signal channels is determined using signal Monte Carlo. The distribution of background-subtracted signal events over the $K^0_{S,L}\pi^+\pi^-$ bins is then corrected for the efficiency to account for relative bin-to-bin efficiency variations. Note that while $F^+_{4\pi}$ is only sensitive to the relative variation of the signal events over the $K^0_{S,L}\pi^+\pi^-$ Dalitz plot, the yields and efficiencies in this analysis are still calculated in absolute terms. This has no impact on the result of $F^+_{4\pi}$.

The second part of this analysis is the fit to extract the *CP*-even fraction $F^+_{4\pi}$ of $D^0 \to 2\pi^+2\pi^-$. The background-subtracted and efficiency-corrected distributions of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ events and $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ events are simultaneously used in a least square fit. The fit parameters are the *CP*-even fraction $F^+_{4\pi}$ and the overall normalisations h and h'. The values of T_i, T'_i , c_i and c'_i are also fitted, but with their measurement uncertainties and correlations imposed with Gaussian constraints.

The last part of the analysis is the consideration of several different sources of potential bias. The effect of the possible biases is estimated either with pseudo-experiments or by using an independent technique to determine a certain quantity and recalculating $F^+_{4\pi}$.

3.2 Data and Simulation Samples

In this section the data samples and the different simulation samples used in the measurement of the *CP*-even fraction of $D^0 \to 2\pi^+2\pi^-$ are presented.

3.2.1 Recorded Data Sample

The dataset used in this analysis is the full CLEO-c dataset recorded in $e^+e^- \to \psi(3770)$ collisions[1], consisting of a total of $(818 \pm 8)\,\mathrm{pb}^{-1}$ in ten sets. The integrated luminosity of each dataset is given in Table 3.1.

3.2.2 Simulated Data Samples

All samples of simulated data (*Monte Carlo* data) are produced within the CLEO-c framework. Different Monte Carlo samples are used for different purposes, such as the determination of the reconstruction and selection efficiencies and the identification of different sources of background.

[1] Dataset 35 is not used in the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ data samples.

Dataset	Integrated Luminosity	Dataset	Integrated Luminosity
31	19.104 pb^{-1}	37	109.349 pb^{-1}
32	30.487 pb^{-1}	43	116.648 pb^{-1}
33	6.184 pb^{-1}	44	173.987 pb^{-1}
35	47.677 pb^{-1}	45	108.192 pb^{-1}
36	68.575 pb^{-1}	46	137.084 pb^{-1}

Table 3.1: *The integrated luminosity of each dataset used in this analysis. The total integrated luminosity is 818.3 ± 8.2 pb*$^{-1}$.

The CLEO-c framework employs the EvtGen software package [37] as a generator to simulate e^+e^- collisions and propagate the decay of their products. The final state particles are passed to the Geant4 software package [38] which is used to describe the CLEO-c detector, the passage of particles through the detector including material interactions, possible noise and bremsstrahlung. The simulated detector responses are recorded in the same format as the real responses obtained from the data acquisition thus enabling an identical reconstruction of the Monte Carlo data and in real data.

3.2.2.1 Signal Monte Carlo

Dedicated signal Monte Carlo samples are generated within the CLEO-c framework in order to study selection criteria and estimate reconstruction efficiencies.

Signal Monte Carlo samples are generated for $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ events, $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ events, $D \to K^0_S\pi^+\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ events and $D \to K^0_S\pi^+\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ events. All samples are reconstructed under the signal decay hypotheses. The signal Monte Carlo samples are generated with non-resonant amplitude models in the CLEO-c framework where correlations between the D mesons are neglected. After reconstruction, the samples are *reweighted* to take into account amplitude models with resonant structures and the correlation between the two D mesons in the event. The reweighting procedure uses the amplitude models for the $D \to 2\pi^+2\pi^-$ and $D \to K^0_S\pi^+\pi^-$ decays from amplitude analyses in Reference [24] and Reference [39], respectively. Since there is no amplitude model for the $D \to K^0_L\pi^+\pi^-$ decay, its amplitude is approximated using the $D \to K^0_S\pi^+\pi^-$ amplitude, i.e.

$$A(D^0 \to K^0_L\pi^+\pi^-) = A(D^0 \to K^0_S\pi^+\pi^-) \tag{3.4}$$

and

$$A(\overline{D}^0 \to K^0_L\pi^+\pi^-) = -A(\overline{D}^0 \to K^0_S\pi^+\pi^-)\ . \tag{3.5}$$

As shown in Appendix A, this approximation holds when the doubly Cabibbo suppressed amplitudes contributing to the decay are negligible with respect to the Cabibbo favoured amplitudes.

The amplitude model for $\psi(3770) \to D^0\overline{D}^0 \to f_1 f_2$ decays is given by (see Chapter 1)

$$A(\psi(3770) \to D^0\overline{D}^0 \to f_1 f_2) \propto A(D^0 \to f_1) \cdot A(\overline{D}^0 \to f_2) \quad (3.6)$$
$$- A(D^0 \to f_2) \cdot A(\overline{D}^0 \to f_1) \quad (3.7)$$

which takes into account the correlation between the two D mesons.

An example for the distribution of events in the $D \to K^0_S\pi^+\pi^-$ Dalitz plane before and after the reweighting procedure is shown in Figure 3.2. The resonant structures from the $K^{*\pm}$ and the ρ are clearly visible after the reweighting.

A systematic uncertainty is applied in Section 3.7.4 to account for any model dependence in the efficiency determination.

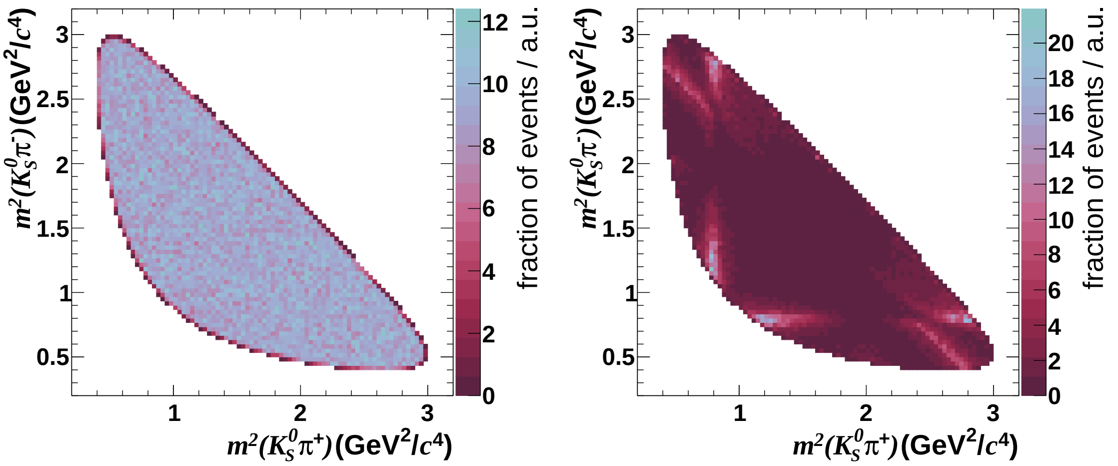


Figure 3.2: *Distribution of events over the $D \to K^0_S\pi^+\pi^-$ Dalitz plane for $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$ events at generator level.* Left: *Distribution of events generated with non-resonant amplitude model.* Right: *Distribution of events after the reweighting procedure. The K^{*+} and K^{*-} resonance can be seen as vertical and horizontal lines at 0.96* GeV^2/c^4*. The diagonal structure is the ρ resonance.*

3.2.2.2 Generic Monte Carlo

The CLEO-c collaboration provides a sample of *generic Monte Carlo* used to identify and estimate background contributions. The generic Monte Carlo contains both $e^+e^- \to \psi(3770) \to D^0\overline{D}^0$ and $e^+e^- \to \psi(3770) \to D^+D^-$ events. Each D meson decays to a final state f with a probability equal to its measured branching fraction $\mathcal{B}(D \to f)$ in the 2004 PDG [40]. Decay modes not listed in the PDG – such as $D \to K^0_L\pi^+\pi^-$ – are added with estimated branching ratios. The integrated luminosity of the generic Monte Carlo samples corresponds to approximately 10 times the luminosity in data for the datasets 31 - 33 and 35 - 37 (abbreviated as datasets 31 - 37) and 20 times the luminosity for the datasets 43 - 46. The generic Monte Carlo uses non-resonant decay models for the D mesons.

The generic Monte Carlo can be used to estimate the background contribution from specific decays. Therefore, the yields in the generic Monte Carlo samples have to be scaled to match the yield expected in the data sample. Two different scaling factors need to be applied to the generic Monte Carlo yields. The first scaling factor accounts for the different number of $D^0\,\overline{D}^0$ events between the generic Monte Carlo and the data. The ratio of $\psi(3770) \to D^0\overline{D}^0$ events to $\psi(3770) \to D^+D^-$ events for generic Monte Carlo samples 31 - 37 was assumed to be 1.33. With a production cross section $\sigma\left(e^+e^- \to \psi(3770)\right) = 7.25\,\mathrm{nb}$ [34] and the branching ration $\mathcal{BR}\left(\psi(3770) \to D\overline{D}\right) = 0.83$ [34] this results in a total number of $9.789 \cdot 10^6$ $D^0\,\overline{D}^0$ events. The number of $D^0\,\overline{D}^0$ events in the corresponding data sample was measured to be $(1.031 \pm 0.015) \cdot 10^6$ [34] which means a factor of $f^{31-37}_{\mathrm{lumi}} = 0.105$ needs to be applied to the yields in the generic Monte Carlo sample 31-37. For the samples 43-46 the factor is taken from the ratio of the luminosities in data and generic Monte Carlo and is $f^{43-46}_{\mathrm{lumi}} = 0.05$.

The generic Monte Carlo was generated without taking into account quantum correlations between the two D mesons. For two correlated decays $D \to f_i$ and $D \to f_j$ the scaling factor f^{ij}_{quan} applied to the generic Monte Carlo is

$$f^{ij}_{\mathrm{quan}} = \frac{\kappa}{4} \cdot \left(1 - \frac{\delta_{ij}}{2}\right) \left[(1 - F^+_i)F^+_j + (1 - F^+_j)F^+_i\right] \tag{3.8}$$

where F^+_i and F^+_j are the *CP*-even fractions of $D \to f_i$ and $D \to f_j$, respectively, and $\kappa \approx 4.0004$ [41] is a factor that ensures that the sum over all combinations of final states i and j yields the total number of quantum-correlated $D^0\,\overline{D}^0$ events.

3.2.2.3 Continuum Monte Carlo

The CLEO-c collaboration also provides a sample of *continuum Monte Carlo* used to identify and estimate background contributions. Continuum events are created by off-resonant $e^+e^- \to q\overline{q} \quad \{q = u, d, s\}$ interactions. The model used to simulate the continuum events is the Lund string model [42] which is very successful at describing the creation of additional hadrons from two quarks moving apart.

The continuum Monte Carlo was generated for the data samples 31 - 37 with a luminosity 5 times higher than the data. This continuum Monte Carlo sample is used to represent the full data sample used in this analysis.

3.3 Reconstruction and Selection of the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_{S,L}\pi^+\pi^-$ Candidates

In this section the procedure to select the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_{\mathrm{S}}\pi^+\pi^-$ candidates and $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_{\mathrm{L}}\pi^+\pi^-$ candidates is described. This builds the basis for the measurement of $F^+_{4\pi}$. First, the topology of the signal decays and what quantities can be used to distinguish it from background is introduced. Second, the selection of the D decay candidates within the CLEO-c `DTag` framework is described. Finally, the additional selection — including a veto for $D \to K^0_{\mathrm{S}}\pi^+\pi^-$ candidates within the $D \to 2\pi^+2\pi^-$ candidates — is explained.

3.3.1 Topology of the Signal Events

The electron and positron beams at CESR have the same energy and are collided at a crossing angle of 3.3 mrad. In the $e^+e^- \to \psi(3770) \to D\overline{D}$ reaction the total energy of the colliding electron-positron pair is transferred to the $\psi(3770)$ and subsequently to the $D\overline{D}$ pair. This means that the four momentum of the $D\overline{D}$ pair is fully known. If the electron and the positron each have an energy E_{beam} and the crossing angle is α, the center of mass energy of the event is given by

$$s = 2\,E^2_{beam}\,(1 - \cos\alpha) \approx 4\,E^2_{beam}\ . \tag{3.9}$$

The approximation holds for small angles such as the crossing angle at CLEO-c of $\alpha = 3.3$ mrad. Thus, each D meson in $e^+e^- \to \psi(3770) \to D\overline{D}$ has an energy of E_{beam}. This is used to construct the *beam constrained D mass*, m^D_{BC}, where

$$m^D_{BC} \equiv \sqrt{E^2_{beam} - p^2_D}\ . \tag{3.10}$$

Since the energy of the electron beam is known to a much higher precision than the energies of the D decay products, the beam constrained D mass is more precise than the D meson mass calculated from the energies of D decay products.
Additionally the *energy difference*, ΔE,

$$\Delta E \equiv E^D - E_{beam} \tag{3.11}$$

is calculated for each D meson in the event. For events that truly come from the $e^+e^- \to \psi(3770) \to D\overline{D}$ reaction the value of ΔE is 0 within the energy resolution of the CLEO-c detector. Figure 3.3 shows the distribution of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ signal Monte Carlo events in the plane spanned by $\Delta E^{2\pi^+2\pi^-}$ *vs.* $m^{2\pi^+2\pi^-}_{BC}$. The two variables are nearly uncorrelated.

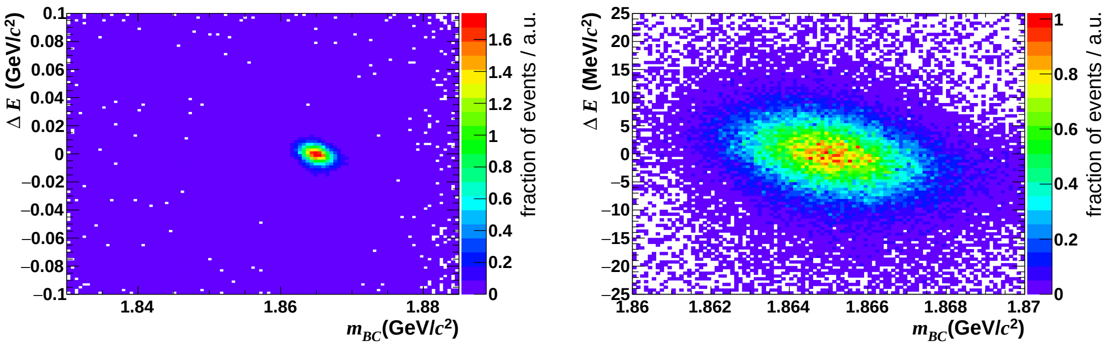


Figure 3.3: *Distribution $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$ signal Monte Carlo events in the plane spanned by $\Delta E^{2\pi^+2\pi^-}$ vs. $m^{2\pi^+2\pi^-}_{BC}$ for all events that pass the CLEO-c preselection (left) and a zoom around the signal region (right). The correlation factors are -0.05 and -0.23, respectively.*

The D mesons in $e^+e^- \to \psi(3770) \to D\overline{D}$ events are created just above threshold, meaning that they have very little kinetic energy and are created almost at rest. It is therefore not possible to distinguish any secondary vertices and all final state particles originate very close to the primary vertex. This information can be used to reject background events by selecting only events where all final state particles have small impact parameters.

Even though the kinetic energy of the D mesons is small, in the center-of-mass frame of the $D\,\overline{D}$ pair their three-momenta point in opposite directions. In the case that only one of the two D mesons is fully reconstructed – as is the case for $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ events where the K^0_L meson is not reconstructed – the four-momentum of the other D meson is calculated using the assumption that in the center-of-mass frame of the $\psi(3770)$ the D^0 and the $\overline{D}^0$ mesons have equal energy and opposite three momenta.

3.3.2 The CLEO-c Reconstruction and Preselection

The preselections for the fully reconstructible decays $D \to 2\pi^+2\pi^-$ and $D \to K^0_S\pi^+\pi^-$ and the partially reconstructible decay $D \to K^0_L\pi^+\pi^-$ are performed by the CLEO-c framework that was introduced in Section 2.3. Within the CLEO-c framework, charged pions are identified as charged tracks in the detector that fulfil a number of selection requirements on the track parameters and the particle identification variables for these tracks. First, the pion candidates have to lay within the fiducial volume of the CLEO-c detector meaning that the angle θ between the particle track and the beam line has to have a $|\cos(\theta)| \leq 0.9$. Additional selections are placed on the momentum, p, the impact parameters, z_0 (impact parameter in z direction) and, d_b (impact parameter in the $x-y$ plane), the track fit quality, χ^2_{track}, and the fraction of hits in the drift chambers, f_{hit}. In order to distinguish the charged pions from kaons, selections are executed on the combined likelihood difference $\Delta\mathcal{L}_{K-\pi}$ for particle identification (see Section 2.1.6).

The K^0_S candidates are constructed by combining two pions of opposite charge. The selection criteria for K^0_S candidates demands that the two opposite sign pions can be successfully constrained to a common vertex (indicated by a $\chi^2_{vtx} >= 0$ of the fit).

The `DTag` candidates for $D \to 2\pi^+2\pi^-$ decays are constructed by combining two positively-charged pions and two negatively-charged pions. The `DTag` candidates for $D \to K^0_S\pi^+\pi^-$ decays are constructed through the combination of a K^0_S candidate, one negative and one positive pion. Selection criteria are placed on the beam constrained D mass m^D_{BC} and the energy difference ΔE of the D meson candidate.

Due to the long lifetime of the K^0_L meson, the $D \to K^0_L\pi^+\pi^-$ decay cannot be fully reconstructed and therefore has no `DTag` candidate. It is build within the CLEO-c framework by combining the `DTag` candidates for the $D \to 2\pi^+2\pi^-$ decays with two pions of opposite charge. The selection requires that there are no additional tracks or neutral pion candidates in the events. The four-momentum of the K^0_L candidates is calculated using the fact that in the center-of-mass frame of the $\psi(3770)$ the D^0 and the $\overline{D}^0$ mesons have equal energy and opposite three momenta. Using the momenta of the reconstructed pions in $D \to K^0_L\pi^+\pi^-$, the missing momentum from the K^0_L meson can be inferred.

All selection criteria from the CLEO-c framework are listed in Table 3.2. The reconstruction and selection efficiency from the CLEO-c framework as determined on the signal Monte Carlo is 45% for $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ decays and 62% for $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ decays.

Particle	Parameter	Selection Criterion
D	m^D_{BC}	$> 1.83\,\mathrm{GeV}$
	ΔE	$< 0.1\,\mathrm{GeV}$
K^0_{S}	Vertex Quality χ^2_{vtx}	> 0
$\pi^{\pm}$	Track Momentum p	$\geq 50\,\mathrm{MeV}$ and $\leq 2\,\mathrm{GeV}$
	Impact Parameter z_0	$\leq 5\,\mathrm{cm}$
	Impact Parameter d_b	$\leq 5\,\mathrm{mm}$
	Track Angle $\lvert\cos\theta\rvert$	≤ 0.9
	Track Fit Quality χ^2_{track}	≤ 100000
	Hit Fraction f_{hit}	≥ 0.5
	Particle ID $\Delta\mathcal{L}_{K-\pi}$	> 0

Table 3.2: *Selection criteria for all particles for $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_{\mathrm{S,L}}\pi^+\pi^-$ in the `DTag` framework. The K^0_{S} vertex quality cut does not apply to the $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_{\mathrm{L}}\pi^+\pi^-$ candidates.*

3.3.2.1 Kinematic Fitting

The CLEO-c reconstruction framework provides the option of performing fits on kinematic objects like tracks and showers. Kinematic parameters like the momenta of the final-state particles can be determined more accurately by imposing constraints on for example the invariant mass of a group of particles or the net momentum in the laboratory frame. The kinematic fit takes the magnetic field, the beam spot, the beam energy and the beam crossing-angle into account. These fits are performed by the FitEvt package [43] and have a variety of options.

In this analysis the kinematic fitting is applied to the $D \to K^0_{\mathrm{S,L}}\pi^+\pi^-$ candidates, allowing the position of the event on the Dalitz plane to be determined more accurately. A kinematic fit is also applied to the $D \to 2\pi^+2\pi^-$ candidates since a selection on the quality of the fit can be used to reject background events. The kinematic fits use the method of least squares.

For the $D \to K^0_{\mathrm{S}}\pi^+\pi^-$ candidates a first fit of two opposite sign pion tracks forming the K^0_{S} candidate is performed. The two pions are required to originate from the same vertex and the invariant mass of the $\pi^+\,\pi^-$ pair is fixed to the nominal K^0_{S} mass. If the fit is successful (indicated by a positive χ^2), the K^0_{S} candidate and the two additional pions are constrained to the same vertex and the nominal D mass in a second fit.

For the $D \to K^0_{\mathrm{L}}\pi^+\pi^-$ candidates the K^0_{L} candidates information is given as the four-momentum determined previously. The error matrix for the K^0_{L} candidates are set to the identity matrix. The tracks of the two pions and the K^0_{L} candidate are required to originate from the same vertex and are constrained to the nominal mass of the

D meson. For the $D \to 2\pi^+2\pi^-$ candidates the tracks of the four pions are required to originate from the same vertex and their combined invariant mass is constrained to the nominal mass of the D meson.

The effect of the kinematic fit on the reconstructed position of events in phase space is quantified with signal Monte Carlo. Without the kinematic fit an average of 76% of the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ candidates are reconstructed in the correct bin while with the kinematic fit an average of 95% of the candidates are reconstructed in the correct bin. Without the kinematic fit an average of 74% of the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ candidates are reconstructed in the correct bin while with the kinematic fit an average of 76% of the candidates are reconstructed in the correct bin. Figure 3.4 shows the purity of events per bin, i.e. the number of events that were produced in a given bin divided by the total number of events reconstructed in that bin.

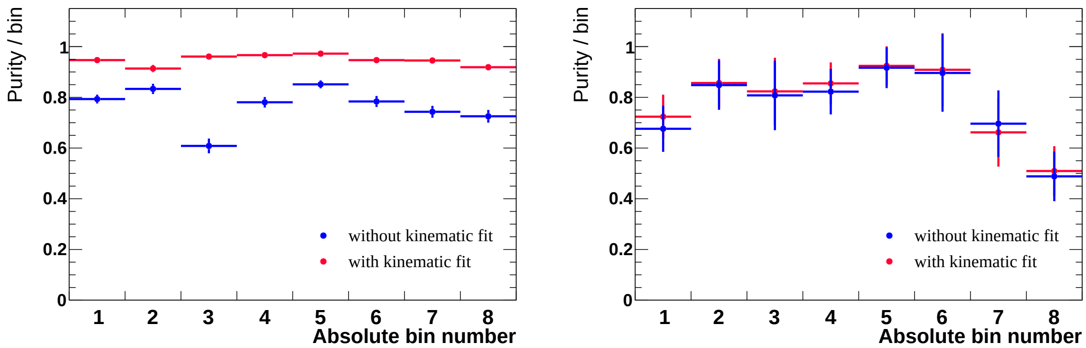


Figure 3.4: *Distribution of purity of events per bin as determined on Monte Carlo for* $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ *candidates (left) and* $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ *candidates (right) without the kinematic fitting (blue) and with the kinematic fitting (red).*

3.3.3 Additional Selection Criteria

Additional selection criteria are placed on the signal candidates to further reduce background and increase the signal purity.

3.3.3.1 $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ Candidate Selection

Tighter selections are placed on the energy difference ΔE for both the $D \to 2\pi^+2\pi^-$ and $D \to K^0_S\pi^+\pi^-$ candidates. Additionally, all kinematic fits have to successfully converge. In order to reject background from $D \to 2\pi^+2\pi^-$ decays reconstructed as $D \to K^0_S\pi^+\pi^-$ a selection is placed on the flight distance significance χ^2_{FS} of the K^0_S candidate – defined as the separation of the K^0_S decay vertex from the interaction region divided by the uncertainty assigned to the K^0_S flight distance. Another cut is placed on the invariant mass of the K^0_S candidate.

The additional selection criteria for $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ candidates are listed in Table 3.3. The selection efficiency from the additional selection as determined on the signal Monte Carlo is 45% for $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ decays and the background rejection as determined from the sidebands in data is 99%.

Parameter	Selection Criterion
Energy Difference for $D \to 2\pi^+2\pi^-$	$\Delta E < 0.025\,\mathrm{GeV}$
Energy Difference for $D \to K^0_S\pi^+\pi^-$	$\Delta E < 0.020\,\mathrm{GeV}$
Kinematic Fits	all converged
K^0_S invariant mass	$\left\|m_{K^0_S} - m^{PDG}_{K^0_S}\right\| < 7.5\,\mathrm{MeV}$
K^0_S flight distance significance	$\chi^2_{FS} > 2$

Table 3.3: *Selection criteria for $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$ candidates.*

3.3.3.2 $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_L\pi^+\pi^-$ Candidate Selection

A tighter selection is placed on the beam constrained D mass $m_{BC}^{2\pi^+2\pi^-}$ and the energy difference ΔE of the $D \to 2\pi^+2\pi^-$ candidate. Additionally, the kinematic fits for both decays have to be successful. In order to reject background from decays that mimic $D \to K^0_L\pi^+\pi^-$ candidates because of particles that were not reconstructed, selection criteria are placed on calorimeter showers that are not directly associated to $D \to K^0_L\pi^+\pi^-$. Each shower that was not associated by the CLEO-c reconstruction can be classified according to two variables: the energy of the shower E_S and the angle θ_S between the shower and the inferred missing momentum in the event. A two dimensional cut is performed on all not-associated calorimeter showers in the event and if one shower in the event fails the selection, the event is rejected. This selection is parameterised by

$$\begin{gathered} -1 < \cos\theta_S < 0.9 \text{ and } E_S < 0.1\,\mathrm{GeV} \\ \text{or} \\ 0.9 < \cos\theta_S < 0.98 \text{ and } E_S < 2.5\,\cos\theta_S + 2.15 \end{gathered} \tag{3.12}$$

The distribution of signal and background events in the E_S *vs.* θ plane and the selection criteria is shown in Figure 3.5.

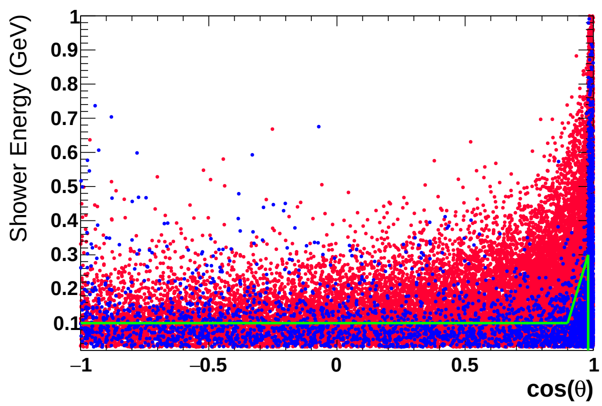


Figure 3.5: *The distribution of signal (red) and background (blue) events in the E_S vs $\cos\theta$ plane of the not-associated shower in the events. The green line represents the selection criteria.*

All additional selection criteria for $D \to 2\pi^+2\pi^-$ *vs.* $D \to K_L^0\pi^+\pi^-$ candidates are listed in Table 3.4. The selection efficiency from the additional selection as determined on the signal Monte Carlo is 33% for $D \to 2\pi^+2\pi^-$ *vs.* $D \to K_L^0\pi^+\pi^-$ decays and the background rejection as determined from the sidebands in data is 97%.

Parameter	Selection Criterion
$m_{BC}^{2\pi^+2\pi^-}$	$1.86\,\mathrm{GeV}/c^2 < m_{BC}^{2\pi^+2\pi^-} < 1.87\,\mathrm{GeV}/c^2$
$\Delta E^{2\pi^+2\pi^-}$	$\Delta E < 0.025\,\mathrm{GeV}$
Kinematic Fits	all converged
Not-associated Showers	$-1 < \cos\theta_S < 0.9$ and $E_S < 0.1\,\mathrm{GeV}$
	or
	$0.9 < \cos\theta_S < 0.98$ and $E_S < 2.5\,\cos\theta_S + 2.15$

Table 3.4: *Selection criteria for $D \to 2\pi^+2\pi^-$ vs. $D \to K_L^0\pi^+\pi^-$ candidates.*

3.3.3.3 K_S^0 Veto Selection

A selection is developed to reject $D \to K_S^0\pi^+\pi^-$ decays reconstructed as $D \to 2\pi^+2\pi^-$ candidates. The signal Monte Carlo sample of $D \to K_S^0\pi^+\pi^-$ *vs.* $D \to K_S^0\pi^+\pi^-$ events is reconstructed as $D \to 2\pi^+2\pi^-$ *vs.* $D \to K_S^0\pi^+\pi^-$ and compared to the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K_S^0\pi^+\pi^-$ signal Monte Carlo. Within the CLEO-c framework all combinations of pions of opposite charge from the $D \to 2\pi^+2\pi^-$ candidates are combined to test if they can form a K_S^0 candidate (as defined in the CLEO-c preselection in Section 3.3.2). One or more K_S^0 candidates are found in 28% of the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K_S^0\pi^+\pi^-$ decays and in 99% of the $D \to K_S^0\pi^+\pi^-$ *vs.* $D \to K_S^0\pi^+\pi^-$ decays. All events with no K_S^0 candidate are accepted by the selection. If a K_S^0 candidate is found, a selection is applied on its flight distance significance.

This selection retains 85% of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K_S^0\pi^+\pi^-$ events while rejecting 93% of $D \to K_S^0\pi^+\pi^-$ *vs.* $D \to K_S^0\pi^+\pi^-$ events. The distribution of the flight distance significance of all found K_S^0 candidates for $D \to 2\pi^+2\pi^-$ *vs.* $D \to K_S^0\pi^+\pi^-$ events and $D \to K_S^0\pi^+\pi^-$ *vs.* $D \to K_S^0\pi^+\pi^-$ events and the selection criteria can be seen in Figure 3.6.

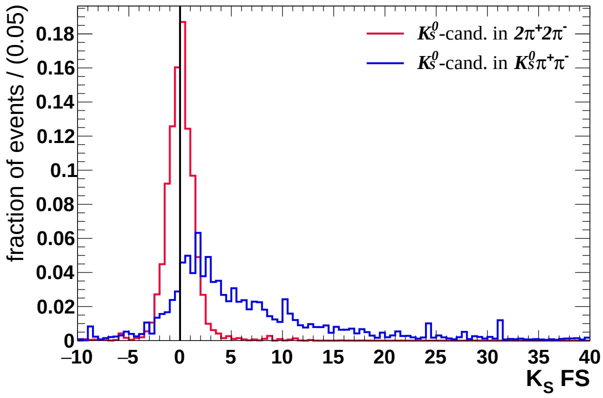


Figure 3.6: *Distribution of the flight distance significance of K^0_S candidates found in the $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$ events (red) and $D \to K^0_S\pi^+\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$ candidates (blue) reconstructed as $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$ events. The black lines indicates the K^0_S veto selection cut.*

3.3.3.4 Multiple Candidate Selection

A selection is applied to events where multiple candidates in the same event have been reconstructed and have passed the above selection. The amount of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ events that have multiple candidates after the selection is 2% while the amount of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ events with multiple candidates after the selection is 1%.

For the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ events the candidate with the smallest figure of merit δ_F is chosen, where

$$\delta_F = \left| \frac{m_{BC}^{2\pi^+2\pi^-} + m_{BC}^{K^0_S\pi^+\pi^-}}{2} - m_{D^0} \right| \tag{3.13}$$

m_{D^0} the nominal D^0 meson mass taken from the PDG [1]. The figure of merit for the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ candidates is defined as

$$\delta_P = \left| m_{BC}^{2\pi^+2\pi^-} - m_{D^0} \right| \tag{3.14}$$

and the candidate with the smallest δ_P is selected.

In order to account for a potential bias arising from the multiple candidate selection, a systematic uncertainty is evaluated in Section 3.7.3.

3.4 Determination of the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ Signal Event Yields

In this section the distribution of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ events over the bins of the $D \to K^0_S\pi^+\pi^-$ Dalitz plot is calculated. The data in each bin is treated independently of the other bins. First the signal and background regions are defined. Second, the reconstructed number of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ candidates is counted for each bin. Third, the background contribution is estimated for each bin.

The selection and reconstruction efficiency per bin is determined. The reconstructed number of signal candidates in each bin is corrected for the background yields and corrected for the efficiency.

3.4.1 Signal and Background Regions

The two-dimensional space spanned by the beam constrained D masses $m_{BC}^{K_S^0\pi^+\pi^-}$ *vs.* $m_{BC}^{2\pi^+2\pi^-}$ is divided into five regions as illustrated in Figure 3.7. The region labelled **S** is the signal region while the regions labelled **A** - **D** are backgrounds regions used to estimate the contribution of combinatorial background. The regions are defined in Table 3.5. Each region corresponds to a different kind of background. Events in region **A** are events where the $D \to K_S^0\pi^+\pi^-$ candidate is correctly reconstructed while the $D \to 2\pi^+2\pi^-$ candidate is misidentified while the opposite is the case for events in region **B**. Events in region **C** are events where tracks between the D meson candidates are swapped. Since the total momentum of all tracks in the event is zero, this results in both D candidate masses being either lower or higher than the nominal D meson mass. The region above the signal region **S** is omitted here since the source of background events is the same as in region **C** and the density of events is expected to be the same as in region **C**. The region **D** is taken to identify the density of events from a random combination of tracks that mimic the signal decay. This background is uniformly distributed over the $m_{BC}^{K_S^0\pi^+\pi^-}$ *vs.* $m_{BC}^{2\pi^+2\pi^-}$ space.

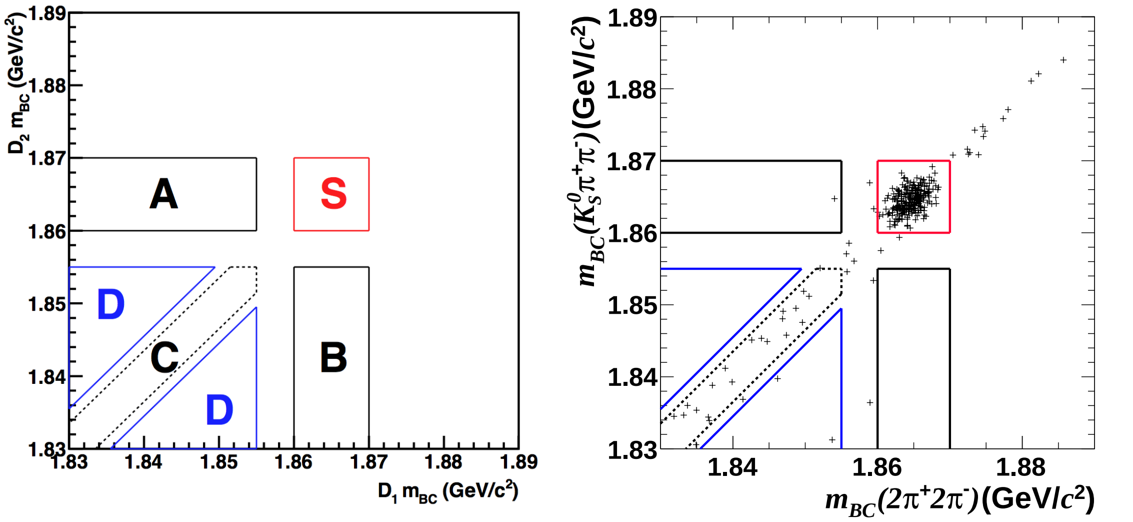


Figure 3.7: Left: *Definition of the signal region **S** and the sidebands **A** - **D** for the case where both D mesons are fully reconstructible, here for $D \to 2\pi^+2\pi^-$ vs. $D \to K_S^0\pi^+\pi^-$.* Right: *Distribution of $D \to 2\pi^+2\pi^-$ vs. $D \to K_S^0\pi^+\pi^-$ candidates over the $m_{BC}^{K_S^0\pi^+\pi^-}$ vs. $m_{BC}^{2\pi^+2\pi^-}$ space after reconstruction and selection.*

Region	$m^{4\pi}_{BC}$(GeV/c^2)		$m^{K^0_S\pi\pi}_{BC}$(GeV/c^2)		δm_{BC}(GeV/c^2)
	Min.	Max.	Min.	Max.	
S	1.860	1.870	1.860	1.870	
A	1.830	1.855	1.860	1.870	
B	1.860	1.870	1.830	1.855	
C	1.830	1.855	1.830	1.855	≤ 0.0035
D	1.830	1.855	1.830	1.855	≥ 0.0055

Table 3.5: *Definition of the signal region **S** and the sidebands **A** - **D** for the case where both D mesons are fully reconstructible, here for $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$.*

3.4.2 Reconstructed $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ Candidates

The distribution of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ candidates after reconstruction and selection over the $m^{K^0_S\pi^+\pi^-}_{BC}$ *vs.* $m^{2\pi^+2\pi^-}_{BC}$ space is shown in Figure 3.7. The total number of candidates in the signal region is 248. There are 23 events in all background regions combined. The number of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ candidates per region is listed in Table 3.6. The one dimensional distributions of the beam constrained masses of both D meson candidates in the signal region **S** are shown in Figure 3.8.

Region	Number of Events	Region	Number of Events
A	1	C	25
B	0	D	2
S	249		

Table 3.6: *Number of $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$ candidates after reconstruction and selection for the different regions in $m^{K^0_S\pi^+\pi^-}_{BC}$ vs. $m^{2\pi^+2\pi^-}_{BC}$ space.*

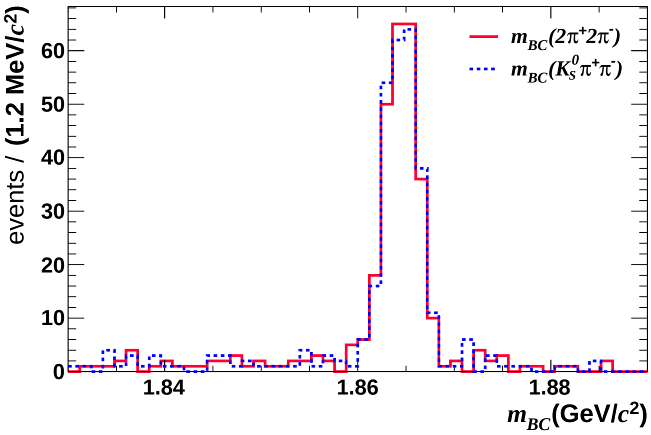


Figure 3.8: *Distribution of the beam constrained mass of the $D \to 2\pi^+2\pi^-$ candidates (red line) and the $D \to K^0_S\pi^+\pi^-$ candidates(blue dashed line) in the signal region **S**.*

The distribution of reconstructed and selected candidates per $K^0_S\pi^+\pi^-$ bin in the signal region over the $D \to K^0_S\pi^+\pi^-$ Dalitz plane is illustrated in Figure 3.9 and the number of events per bin is listed in Table 3.7.

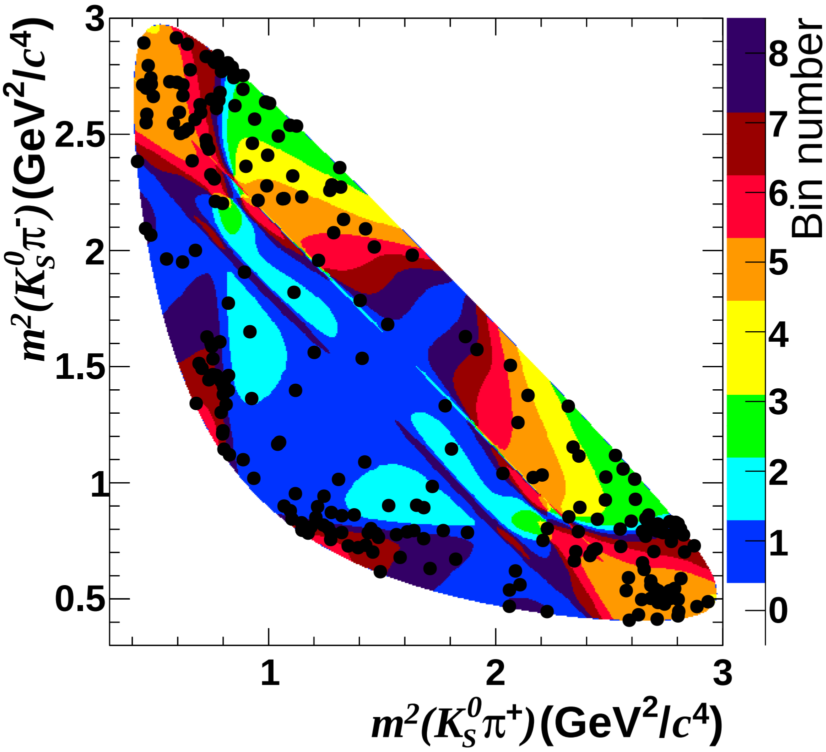


Figure 3.9: *Distribution of the $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$ events over the Dalitz plot of the $D \to K^0_S\pi^+\pi^-$ candidate in the full data sample after reconstruction and selection.*

Bin	Number of Events	Bin	Number of Events
1	38 ± 6.16	5	59 ± 7.68
2	22 ± 4.69	6	24 ± 4.90
3	21 ± 4.58	7	30 ± 5.48
4	13 ± 3.61	8	42 ± 6.48

Table 3.7: *Number of $D\to 2\pi^+2\pi^-$ vs. $D\to K^0_S\pi^+\pi^-$ candidates after reconstruction and selection for the different $K^0_S\pi^+\pi^-$ bins in the signal region **S**. The uncertainty is the Poisson uncertainty on the yields.*

3.4.3 Background Estimation

The background for $D\to 2\pi^+2\pi^-$ *vs.* $D\to K^0_S\pi^+\pi^-$ events is divided into two categories: background from specific misidentified decays and background from the combination of tracks to mimic one or both sides of the signal decay. The background from specific misidentified decays is called *peaking* background since it peaks in $m_{BC}^{2\pi^+2\pi^-}$ *vs.* $m_{BC}^{K^0_S\pi\pi}$ space like the signal. It is expected to have a non-uniform structure in the $m^2_{K^0_S\pi^-}$ *vs.* $m^2_{K^0_S\pi^+}$ space of the $D\to K^0_S\pi^+\pi^-$ candidate.

The background from the random combination of tracks is called *combinatorial* background[2]. It is assumed to be distributed uniformly over the $m^2_{K^0_S\pi^-}$ *vs.* $m^2_{K^0_S\pi^+}$ space of the $D\to K^0_S\pi^+\pi^-$ candidate[3].

For each category, the background yields per $K^0_S\pi^+\pi^-$ bin are determined in two steps. In the first step the total number of background events in the signal region **S** from a certain source is estimated. In the second step these total yields are distributed over the $K^0_S\pi^+\pi^-$ bins.

3.4.3.1 Interlude: Statistical Uncertainty on Very Small Event Yields

The estimated background yields for individual bins can be very small. In order to give a statistical uncertainty on this yield a Poisson distribution is generated for each yield where the Poisson parameter λ is the respective yield. The 68% Confidence Level interval is identified by finding the inner 68% of the distribution. The asymmetric uncertainty is then taken to be the differences between the yield and the lower and upper boundaries of the confidence interval. An example is illustrated in Figure 3.10.

[2]The contribution from continuum events is expected to be very small due to the high track multiplicity in $D\to 2\pi^+2\pi^-$ *vs.* $D\to K^0_S\pi^+\pi^-$ events. It is not considered separately but is implicitly included in the estimation of combinatorial background events.

[3]There is no significant peaking structure from continuum events reconstructed as $D\to 2\pi^+2\pi^-$ *vs.* $D\to K^0_S\pi^+\pi^-$ decays [44].

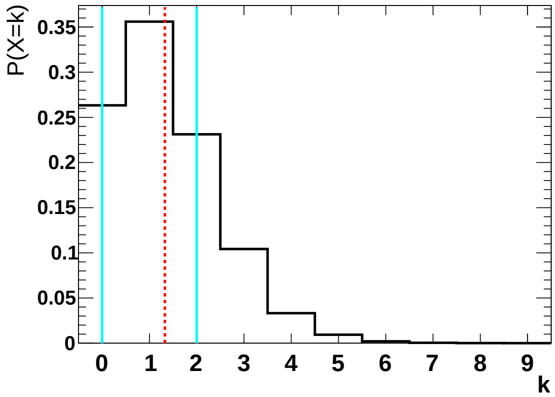


Figure 3.10: *Example to illustrate the determination of the statistical uncertainty on very small yields. The histogram shows the Poisson distribution for $\lambda = 1.33$ (dashed red line). The lower and upper boundaries of the inner 68% Confidence Level interval are shown as blue lines. The differences between λ and the lower and upper boundaries are taken to be the asymmetric statistical uncertainties.*

3.4.3.2 Peaking Background Contribution

The decays for the peaking background contribution are identified using the generic Monte Carlo sample. All considered contributions are extrapolated from the yields in the generic Monte Carlo sample using the formalism outlined in Section 3.2.2.2. All results are listed in Table 3.8. The uncertainties are calculated by propagating the Poisson uncertainties on the yields in the generic Monte Carlo samples. The only significant contribution is found to be from $D \to K^0_S\pi^+\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ decays with an estimated total number of $D \to K^0_S\pi^+\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ events in the signal region of 19.52 ± 1.15.

The value for the *CP*-even fraction $F_+^{4\pi}$ of $D^0 \to 2\pi^+2\pi^-$ used in Equation 3.8 to estimate the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S K^0_S$ contribution is taken from an independent amplitude analysis of $D^0 \to 2\pi^+2\pi^-$ decays [24]. To evaluate a potential bias from using $F_{4\pi}^+$ as input to the analysis, the number of expected $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S K^0_S$ events in the data is calculated for different values of $F_{4\pi}^+$. The greatest contribution is achieved with $F_{4\pi}^+ = 0$ and yields 1.41 ± 0.41 events. This accounts for less than 0.5% of events in the signal region and can be neglected.

Decay	$F_+^{D_1}$	$F_+^{D_2}$	N_{exp}^{data}
$D \to K^0_S\pi^+\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$	0.556 [41]	0.556 [41]	19.52 ± 1.15
$D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S K^0_S$	0.729 [24]	1 [41]	0.38 ± 0.11
$D \to K^0_S\pi^+\pi^-$ *vs.* $D \to K^0_S K^0_S$	0.556 [41]	1 [41]	0.14 ± 0.10

Table 3.8: *Contribution of different possible peaking backgrounds to the reconstructed event yield estimated from the full generic Monte Carlo sample. The uncertainties are calculated by propagating the Poisson uncertainties on the yields in the generic Monte Carlo samples.*

The distribution of $D \to K^0_S\pi^+\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ events over the $K^0_S\pi^+\pi^-$ bins is determined using the $D \to K^0_S\pi^+\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ Monte Carlo sample described in Section 3.2.2.1. The results are listed in Table 3.9. The first uncertainty is the statistical uncertainty determined as described in Section 3.4.3.1. The second uncertainty is the combination from the uncertainty on the total number of $D \to K^0_S\pi^+\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ events and the uncertainty from the limited $D \to K^0_S\pi^+\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ Monte Carlo sample used to determine the distribution of the $D \to K^0_S\pi^+\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ events over the bins. For the purpose of simplicity the second uncertainty is called the systematic uncertainty.

A systematic uncertainty is assigned to the total number of $D \to K^0_S\pi^+\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ events in Section 3.7.5.1 as well as to the amplitude model used to reweight the Monte Carlo samples in Section 3.7.4.

Bin	Number of Events	Bin	Number of Events
1	$2.17\ ^{+1.83}_{-1.17} \pm 0.37$	5	$2.88\ ^{+2.12}_{-1.88} \pm 0.42$
2	$3.94\ ^{+2.06}_{-1.94} \pm 0.49$	6	$2.34\ ^{+1.65}_{-1.34} \pm 0.38$
3	$2.24\ ^{+1.76}_{-1.24} \pm 0.37$	7	$1.68\ ^{+1.32}_{-1.68} \pm 0.32$
4	$1.28\ ^{+0.72}_{-1.28} \pm 0.28$	8	$2.98\ ^{+2.02}_{-1.98} \pm 0.43$

Table 3.9: *Estimated distribution of $D \to K^0_S\pi^+\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$ events over the $K^0_S\pi^+\pi^-$ bins in the signal sample. The first uncertainty is the statistical uncertainty of the yields per $K^0_S\pi^+\pi^-$ bin, determined as described in Section 3.4.3.1. The second uncertainty is the combination of the uncertainty on the total number of $D \to K^0_S\pi^+\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$ events and the uncertainty from the limited $D \to K^0_S\pi^+\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$ Monte Carlo sample used to determine the distribution of the $D \to K^0_S\pi^+\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$ events over the bins. For the purpose of simplicity the second uncertainty is called the systematic uncertainty.*

3.4.3.3 Combinatorial Background Contribution

The total contribution of combinatorial background events is determined using a fully data-driven method. The events in the sidebands **A - D** (defined in Section 3.4.1) are assumed to contain no peaking background. The number of events in the sideband is extrapolated to give the total number of combinatorial background events in the signal region **S**. The number of expected combinatorial background events N_S^{comb} in the signal region is calculated using

$$N_S^{comb} = \frac{a_S}{a_D} \cdot N_D + \sum_{X=A,B,C} \frac{a_S}{a_X} \left(N_X - \frac{a_X}{a_D} \cdot N_D \right) \tag{3.15}$$

where a_i is the area of region i ($i = S,\ A,\ B,\ C,\ D$) and N_i is the number of events in region i. This formulation ensures that there is no double counting of events. The formalism predicts 14.71 ± 3.19 combinatorial background events in the signal region **S**. The uncertainty is the propagated Poisson uncertainty on the yields in the sidebands.

Since the combinatorial background events are expected to be uniformly distributed over the $m^2_{K^0_S\pi^-}$ *vs.* $m^2_{K^0_S\pi^+}$ space of the $D\to K^0_S\pi^+\pi^-$ candidate, the fraction of combinatorial background events per $K^0_S\pi^+\pi^-$ bin is allotted according to the bins' area in the $K^0_S\pi^+\pi^-$ Dalitz plot. The resulting combinatorial background yields for each $K^0_S\pi^+\pi^-$ bin are listed in Table 3.10. The first uncertainty is the statistical uncertainty determined as described in Section 3.4.3.1 and the second uncertainty is the propagated uncertainty from the total number of combinatorial background events. For the purpose of simplicity the second uncertainty is called the systematic uncertainty.

A systematic uncertainty related to the method used to estimate the total number of combinatorial background events is evaluated in Section 3.7.7.1. A systematic uncertainty is assigned to the distribution of the events over the $K^0_S\pi^+\pi^-$ bins in Section 3.7.7.2.

Bin	Number of Events	Bin	Number of Events
1	$4.86\ ^{+2.14}_{-1.86} \pm 1.05$	5	$1.96\ ^{+1.04}_{-0.96} \pm 0.42$
2	$1.68\ ^{+1.32}_{-1.68} \pm 0.36$	6	$1.19\ ^{+0.81}_{-1.19} \pm 0.26$
3	$0.94\ ^{+1.06}_{-0.94} \pm 0.20$	7	$1.24\ ^{+0.76}_{-1.24} \pm 0.27$
4	$0.87\ ^{+1.13}_{-0.87} \pm 0.19$	8	$1.97\ ^{+1.03}_{-0.97} \pm 0.43$

Table 3.10: *Estimated distribution of combinatorial background events over the K^0_S $\pi^+\pi^-$ bins in the signal sample. The first uncertainty is the statistical uncertainty of the yields per $K^0_S\pi^+\pi^-$ bin, determined as described in Section 3.4.3.1 and the second uncertainty is the propagated uncertainty from the total number of combinatorial background events. For the purpose of simplicity the second uncertainty is called the systematic uncertainty.*

3.4.4 The Reconstruction and Selection Efficiency

The reconstruction and selection efficiency of the $D\to 2\pi^+2\pi^-$ *vs.* $D\to K^0_S\pi^+\pi^-$ signal candidates is determined using the signal Monte Carlo described in Section 3.2.2.1. The absolute efficiency for each $K^0_S\pi^+\pi^-$ bin is listed in Table 3.11. The uncertainty on the efficiency is taken to be the binomial uncertainty.

A systematic uncertainty is assigned in Section 3.7.4 to account for a possible bias in the amplitude models used in the reweighting procedure.

Bin	Efficiency [%]	Bin	Efficiency [%]
1	23.34 ± 0.73	5	23.18 ± 0.65
2	22.80 ± 1.03	6	24.77 ± 1.06
3	27.29 ± 0.89	7	21.43 ± 1.02
4	26.23 ± 0.79	8	23.66 ± 0.97

Table 3.11: *Reconstruction and selection efficiency of the $D\to 2\pi^+2\pi^-$ vs. $D\to K^0_S\pi^+\pi^-$ signal events for each $K^0_S\pi^+\pi^-$ bin obtained from the signal Monte Carlo sample. The uncertainty on the efficiency is the binomial uncertainty.*

3.4.5 $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$ Signal Event Yields

The number of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ signal events in the data sample is determined by subtracting the peaking background yields (Section 3.4.3.2) and the combinatorial background yields (Section 3.4.3.3) from the reconstructed and selected $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ candidates (Section 3.4.2) and then correcting for the reconstruction and selection efficiency (Section 3.4.4).

The distribution of the reconstructed $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ candidates over the $K^0_S\pi^+\pi^-$ bins is shown in Figure 3.11. The figure shows the estimated contributions from the peaking and the combinatorial background. The background-subtracted and efficiency-corrected number of signal events in the data sample for each $K^0_S\pi^+\pi^-$ bin is listed in Table 3.12. The first uncertainty is the statistical uncertainty propagated from the Poisson uncertainties on the reconstructed $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ candidates and the individual background contributions. The second uncertainty is the combination of the systematic uncertainties on the individual background yields. The third uncertainty is the propagated binomial uncertainty from the reconstruction and selection efficiency.

Bin	Number of Signal Events	Bin	Number of Signal Events
1	$132.7\ ^{+29.03}_{-28.04}\ \pm 4.78\ \pm 0.08$	5	$233.6\ ^{+34.66}_{-34.36}\ \pm 2.58\ \pm 0.09$
2	$71.89\ ^{+23.22}_{-23.44}\ \pm 2.69\ \pm 0.14$	6	$82.63\ ^{+21.13}_{-21.06}\ \pm 1.86\ \pm 0.14$
3	$65.28\ ^{+18.40}_{-17.73}\ \pm 1.56\ \pm 0.09$	7	$126.4\ ^{+26.54}_{-27.36}\ \pm 1.96\ \pm 0.19$
4	$41.35\ ^{+14.66}_{-14.96}\ \pm 1.30\ \pm 0.10$	8	$156.6\ ^{+29.02}_{-28.93}\ \pm 2.56\ \pm 0.13$

Table 3.12: *Background-subtracted and efficiency-corrected number of $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$ signal events in the full CLEO-c data sample for each $K^0_S\pi^+\pi^-$ bin. The first uncertainty is the purely statistical uncertainty propagated from the statistical uncertainties on the reconstructed yields and the individual background yields. The second uncertainty is the combination of the systematic uncertainties on the individual background yields. The third uncertainty is the propagated binomial uncertainty from the reconstruction and selection efficiency.*

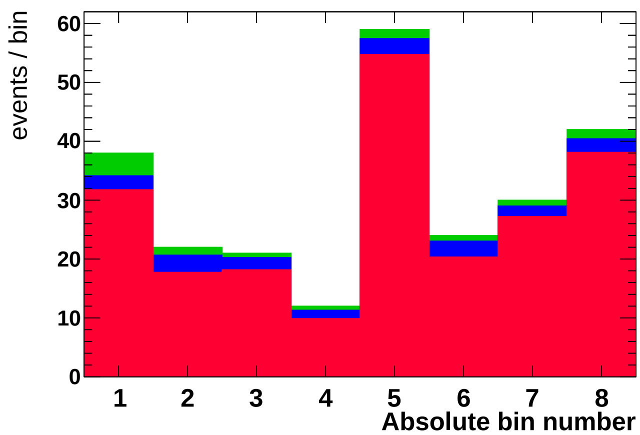


Figure 3.11: *Distribution of reconstructed and selected $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$ candidates over the $K^0_S\pi^+\pi^-$ bins. The figure shows the contribution of signal events (red), peaking background events from $D \to K^0_S\pi^+\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$ decays (blue) and combinatorial background events (green).*

3.5 Determination of the $D \to 2\pi^+\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ Signal Event Yields

In this section the distribution of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ events over the bins of $D \to K^0_L\pi^+\pi^-$ Dalitz plot is calculated. The data in each bin is treated independently of the other bins. First, the signal and background regions are defined. Second, the reconstructed number of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ candidates is counted for each bin. Third, the background contribution is estimated for each bin. The selection and reconstruction efficiency per bin is determined. The reconstructed number of signal candidates in each bin is corrected for the background yields and corrected for the efficiency.

3.5.1 Signal and Background Regions

The signal region and sidebands for the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ candidates are defined in terms of the missing mass squared m^2_{miss} in the events. The signal region **S** is centred around the nominal K^0_S mass. All regions are illustrated in Figure 3.12 and defined in Table 3.13.

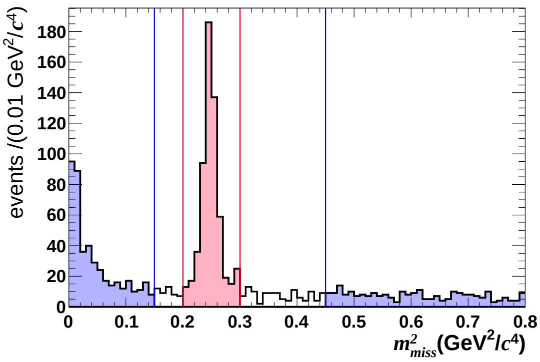


Figure 3.12: *Distribution of the m^2_{miss} of the reconstructed $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_L\pi^+\pi^-$ candidates. The Signal region is shaded in red and the lower and upper sidebands are shaded in blue.*

Region	m^2_{miss}(GeV/c^2)	
	Min.	Max.
LS	0.0	0.15
S	0.2	0.3
US	0.45	0.8

Table 3.13: *Signal region and sidebands as defined for $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_L\pi^+\pi^-$.*

3.5.2 Reconstructed $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ Candidates

The total number of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_{\rm L}\pi^+\pi^-$ candidates after reconstruction and selection in the signal region is 592 ± 24.33. There are 411 events in both background regions combined. The number of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_{\rm L}\pi^+\pi^-$ candidates per region is listed in Table 3.14 and their distribution over the m^2_{miss} range shown in Figure 3.12. The distribution of the events in the signal region over the $D \to K^0_{\rm L}\pi^+\pi^-$ Dalitz plane is illustrated in Figure 3.13 and the number of candidates per $K^0_{\rm L}\pi^+\pi^-$ bin in the signal region is listed in Table 3.15.

Region	Number of Events
LS	213 ± 14.59
S	592 ± 24.33
US	198 ± 14.07

Table 3.14: *Number of $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_{\rm L}\pi^+\pi^-$ candidates after reconstruction and selection for the different m^2_{miss} regions. The uncertainty is the Poisson uncertainty on the yields.*

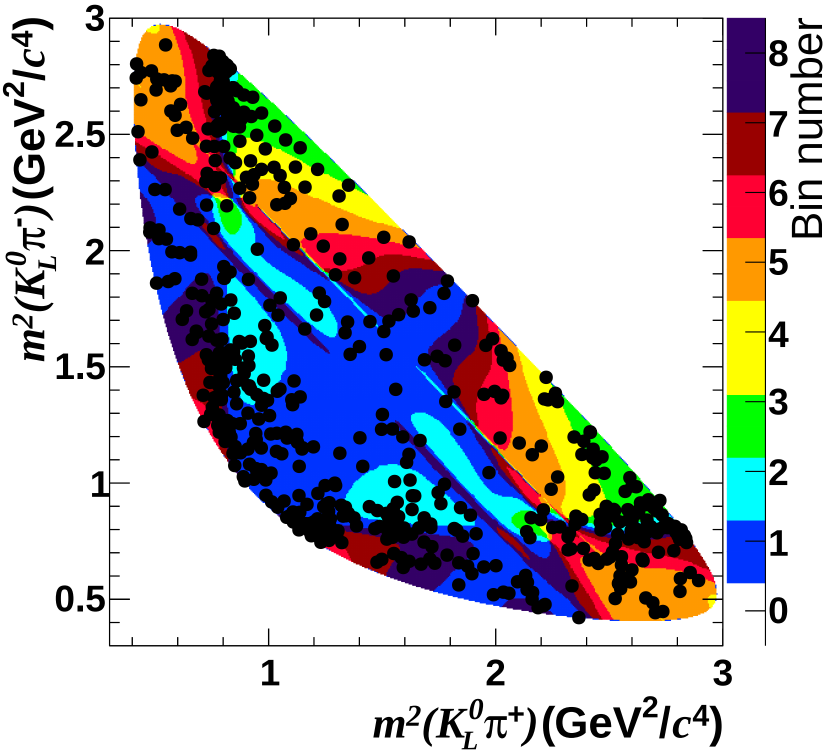


Figure 3.13: *Distribution of the $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_{\rm L}\pi^+\pi^-$ events in the data sample after reconstruction and selection the Dalitz plot of the $D \to K^0_{\rm L}\pi^+\pi^-$ candidate.*

Bin	Number of Events	Bin	Number of Events
1	172 ± 13.11	5	58 ± 7.62
2	72 ± 8.49	6	33 ± 5.74
3	62 ± 7.87	7	72 ± 8.49
4	24 ± 4.90	8	99 ± 9.95

Table 3.15: *Number of $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_{\rm L}\pi^+\pi^-$ candidates after reconstruction and selection for the different $K^0_{\rm L}\pi^+\pi^-$ bins in the signal region $\boldsymbol{S}$. The uncertainty is the Poisson uncertainty on the yields.*

3.5.3 Background Estimation

The background for $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_{\rm L}\pi^+\pi^-$ events is divided into three categories: background from specific misidentified decays, background from the combination of tracks from $\psi(3770) \to D^+_{(s)}D^-_{(s)}/D^0\overline{D}^0$ which mimic one or both sides of the signal decay and background from continuum events. The background from specific misidentified decays is called *peaking* background since it *peaks* in $m_{BC}^{2\pi^+2\pi^-}$ *vs.* $m_{BC}^{K^0_{\rm L}\pi\pi}$ space like the signal. It is also expected to have a non-uniform structure in the $m^2_{K^0_{\rm L}\pi^-}$ *vs.* $m^2_{K^0_{\rm L}\pi^+}$ space of the $D \to K^0_{\rm L}\pi^+\pi^-$ candidate.

The background from the combination of tracks from $\psi(3770) \to D^+_{(s)}D^-_{(s)}/D^0\overline{D}^0$ events is called *combinatorial* background and the background from the combination of tracks from $e^+e^- \to q\overline{q} \quad \{q = u, d, s\}$ interactions is called *continuum* background. Both the combinatorial background and the continuum background [4]are assumed to be distributed uniformly over the $m^2_{K^0_{\rm L}\pi^-}$ *vs.* $m^2_{K^0_{\rm L}\pi^+}$ space of the $D \to K^0_{\rm L}\pi^+\pi^-$ candidate.

For each category, the background yields per $K^0_{\rm L}\pi^+\pi^-$ bin are determined in two steps. In the first step the total number of background yields from a certain source are estimated. In the second step these yields are distributed over the $K^0_{\rm L}\pi^+\pi^-$ bins.

3.5.3.1 Peaking Background

The peaking background contributions are identified using the generic Monte Carlo sample. All considered contributions are listed in Table 3.16 and their distributions over the squared missing mass m^2_{miss} range are illustrated in Figure 3.14. The uncertainties are calculated by propagating the Poisson uncertainties on the yields in the generic Monte Carlo samples. The only significant contributions were found to be from $D \to K^0_{\rm S}\pi^+\pi^-$ *vs.* $D \to K^0_{\rm L}\pi^+\pi^-$ decays and $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_{\rm S}\pi^+\pi^-$ decays.

The distribution of the peaking background events over the $K^0_{\rm S}\pi^+\pi^-$ bins is determined using the respective Monte Carlo samples described in Section 3.2.2.1.

[4]There is no significant peaking structure from continuum events for $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_{\rm L}\pi^+\pi^-$ decays [44].

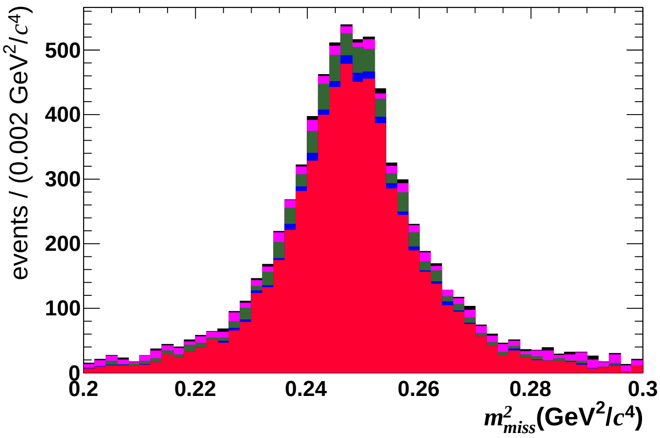


Figure 3.14: *Distribution of the reconstructed and selected events in the full generic Monte Carlo sample, signal events (red), $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$ (blue), $D \to K^0_S\pi^+\pi^-$ vs. $D \to K^0_L\pi^+\pi^-$ (green), $D \to K^0_S\pi^+\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$ (magenta) and other events (black).*

Decay	$F_+^{D_1}$	$F_+^{D_2}$	N_{exp}^{data}
$D \to 2\pi^+2\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$	0.729 [24]	0.556 [41]	8.50 ± 0.72
$D \to K^0_S\pi^+\pi^-$ vs. $D \to K^0_L\pi^+\pi^-$	0.556 [41]	0.630 [22]	35.71 ± 1.52
$D \to K^0_S\pi^+\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$	0.556 [41]	0.556 [41]	1.10 ± 0.27

Table 3.16: *Contribution of different possible peaking backgrounds to the reconstructed event yield estimated from the full generic Monte Carlo sample. The uncertainties are calculated by propagating the Poisson uncertainties on the yields in the generic Monte Carlo samples.*

The results are listed in Table 3.17. The first uncertainty is the statistical uncertainty determined as described in Section 3.4.3.1. The second uncertainty is the combination from the uncertainty on the total number of events of the respective peaking background and the uncertainty from the limited Monte Carlo sample used to determine the distribution of the respective peaking background events over the bins. For the purpose of simplicity the second uncertainty is called the systematic uncertainty.

A systematic uncertainty is assigned to the total number of peaking background events from $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$ events in Section 3.7.6. This systematic uncertainty accounts for a possible bias from using $F^+_{4\pi}$ to calculate the contribution of $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_S\pi^+\pi^-$ events. A systematic uncertainty is also assigned to the total number of $D \to K^0_S\pi^+\pi^-$ vs. $D \to K^0_L\pi^+\pi^-$ events in Section 3.7.5.1. As mentioned in Section 3.2.2.2, an estimation was used for the branching ratio of $D^0 \to K^0_L\pi^+\pi^-$ in the generic Monte Carlo. The latter systematic also takes into account the uncertainty on this branching ratio. Another systematic uncertainty is assigned in Section 3.7.4, taking into account the uncertainty on the amplitude model used to reweight the Monte Carlo samples.

Bin	$D\to K_S^0\pi^+\pi^-$ *vs.* $D\to K_L^0\pi^+\pi^-$ Events	$D\to 2\pi^+2\pi^-$ *vs.* $D\to K_L^0\pi^+\pi^-$ Events
1	$10.44^{+3.56}_{-3.44} \pm 0.69$	$2.93\ ^{+2.07}_{-1.93} \pm 0.50$
2	$3.53\ ^{+1.47}_{-1.53} \pm 0.41$	$0.95\ ^{+1.05}_{-0.95} \pm 0.28$
3	$6.47\ ^{+2.53}_{-2.47} \pm 0.55$	$0.67\ ^{+0.33}_{-0.67} \pm 0.24$
4	$1.32\ ^{+0.68}_{-1.32} \pm 0.25$	$0.26\ ^{+0.74}_{-0.26} \pm 0.15$
5	$3.83\ ^{+2.17}_{-1.83} \pm 0.43$	$0.62\ ^{+0.38}_{-0.62} \pm 0.23$
6	$2.23\ ^{+1.77}_{-1.23} \pm 0.33$	$0.45\ ^{+0.55}_{-0.45} \pm 0.20$
7	$2.95\ ^{+2.05}_{-1.95} \pm 0.38$	$1.15\ ^{+0.85}_{-1.15} \pm 0.31$
8	$4.94\ ^{+2.06}_{-1.94} \pm 0.48$	$1.47\ ^{+1.53}_{-1.47} \pm 0.35$

Table 3.17: *Estimated distribution of peaking background events over the $K_L^0\pi^+\pi^-$ bins in the signal sample. The first uncertainty is the statistical uncertainty of the yields per $K_L^0\pi^+\pi^-$ bin, determined as described in Section 3.4.3.1. The second uncertainty is the combination of the uncertainty on the total number of peaking background events and the uncertainty from the limited Monte Carlo sample used to determine the distribution of the peaking background events over the bins. For the purpose of simplicity the second uncertainty is called the systematic uncertainty.*

3.5.3.2 Combinatorial and Continuum Background

The total contribution of combinatorial and continuum background events is determined simultaneously using different Monte Carlo samples and information from the data sample.

The total event yield N_X in data for each region $\mathbf{X}$ (as defined in Section 3.5.1) is expressed as the the sum of the signal event yields N_X^{sig}, the peaking background yields N_X^{peak}, the combinatorial background yields N_X^{comb} and the continuum background yields N_X^{cont}

$$N_X = N_X^{sig} + N_X^{peak} + N_X^{comb} + N_X^{cont} \tag{3.16}$$

$$\Rightarrow N_X - N_X^{peak} = N_X^{sig} + N_X^{comb} + N_X^{cont} \quad . \tag{3.17}$$

Under the assumption that the ratio of events between the signal region and the sidebands is accurately modelled in the different Monte Carlo samples, the yield for e.g. the combinatorial background in data in region X can be expressed in terms of the combinatorial background yield in data in region Y with

$$N_X^{comb} = \frac{M_X^{comb}}{M_Y^{comb}} N_Y^{comb} \tag{3.18}$$

where M_X^{comb} and M_Y^{comb} represent the combinatorial background yields in the generic Monte Carlo samples in region $\mathbf{X}$ and $\mathbf{Y}$. Analogous, the signal event yields in data in region **X** can be calculated from the signal event yield in data in region **Y** using the ratio of yields from the signal Monte Carlo sample and the continuum background yields can be calculated using the continuum Monte Carlo sample. Using these substitutions, the three possible equations from Equation 3.16 are rewritten as

$$N_{LS} - N_{LS}^{peak} = N_{LS}^{cont} + \frac{M_{LS}^{sig}}{M_S^{sig}} N_S^{sig} + \frac{M_{LS}^{comb}}{M_{US}^{comb}} N_{US}^{comb} \tag{3.19}$$

$$N_S - N_S^{peak} = \frac{M_S^{cont}}{M_{LS}^{cont}} N_{LS}^{cont} + N_S^{sig} + \frac{M_S^{comb}}{M_{US}^{comb}} N_{US}^{comb} \tag{3.20}$$

$$N_{US} - N_{US}^{peak} = \frac{M_{US}^{cont}}{M_{LS}^{cont}} N_{LS}^{cont} + \frac{M_{US}^{S}}{M_S^{S}} N_S^{sig} + N_{US}^{comb} \; . \tag{3.21}$$

This is expressed in matrix formalism as $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$, where

$$\boldsymbol{A} = \begin{pmatrix} 1 & \frac{M_{LS}^{sig}}{M_S^{sig}} & \frac{M_{LS}^{comb}}{M_{US}^{comb}} \\ \frac{M_S^{cont}}{M_{LS}^{cont}} & 1 & \frac{M_S^{comb}}{M_{US}^{comb}} \\ \frac{M_{US}^{cont}}{M_{LS}^{cont}} & \frac{M_{US}^{S}}{M_S^{S}} & 1 \end{pmatrix} \tag{3.22}$$

and

$$\boldsymbol{y} = \begin{pmatrix} N_{LS} - N_{LS}^{peak} \\ N_S - N_S^{peak} \\ N_{US} - N_{US}^{peak} \end{pmatrix} \qquad \boldsymbol{x} = \begin{pmatrix} N_{LS}^{cont} \\ N_S^{sig} \\ N_{US}^{comb} \end{pmatrix} . \tag{3.23}$$

This equation can be solved to obtain the total number of combinatorial and continuum background events in the signal region. The results are 42.36 ± 4.60 combinatorial background events and 18.52 ± 7.99 continuum background events in the signal sample. The uncertainties are the propagated Poisson uncertainties on the individual data and Monte Carlo yields.

Since the combinatorial and continuum background events are expected to be uniformly distributed over the $m^2_{K^0_{\rm L}\pi^-}$ *vs.* $m^2_{K^0_{\rm L}\pi^+}$ space of the $D \to K^0_{\rm L}\pi^+\pi^-$ candidate, the fraction of combinatorial background events per $K^0_{\rm L}\pi^+\pi^-$ bin is allotted according to the bins' area in $m^2_{K^0_{\rm L}\pi^-}$ *vs.* $m^2_{K^0_{\rm L}\pi^+}$ space. The resulting background yields for each $K^0_{\rm L}\pi^+\pi^-$ bin are listed in Table 3.18. The first uncertainty is the statistical uncertainty determined as described in Section 3.4.3.1. The second uncertainty is the propagated uncertainty from the total number of combinatorial background events. For the purpose of simplicity the second uncertainty is called the systematic uncertainty.

A systematic uncertainty from the method used to calculate the total number of combinatorial and continuum background events is evaluated in Section 3.7.8.1. A systematic uncertainty is assigned to the distribution of the events over the $K^0_{\rm L}\pi^+\pi^-$ bins in Section 3.7.8.2.

Bin	Combinatorial Events	Continuum Events
1	$14.42\,^{+3.58}_{-3.42} \pm 1.25$	$6.30\,^{+2.70}_{-2.30} \pm 1.65$
2	$4.84\,^{+2.16}_{-1.84} \pm 0.72$	$2.11\,^{+1.89}_{-1.11} \pm 0.96$
3	$2.61\,^{+1.39}_{-1.61} \pm 0.53$	$1.14\,^{+0.86}_{-1.14} \pm 0.70$
4	$2.41\,^{+1.59}_{-1.41} \pm 0.51$	$1.05\,^{+0.95}_{-1.05} \pm 0.67$
5	$5.49\,^{+2.51}_{-2.49} \pm 0.77$	$2.40\,^{+1.60}_{-1.40} \pm 1.02$
6	$3.43\,^{+1.57}_{-1.43} \pm 0.61$	$1.50\,^{+1.50}_{-1.50} \pm 0.81$
7	$3.51\,^{+1.49}_{-1.51} \pm 0.62$	$1.54\,^{+1.46}_{-1.54} \pm 0.81$
8	$5.66\,^{+2.34}_{-2.66} \pm 0.78$	$2.47\,^{+1.53}_{-1.47} \pm 1.03$

Table 3.18: *Estimated distribution of flat background events over the $K^0_{\mathrm{L}}\pi^+\pi^-$ bins in the signal sample. The first uncertainty covers the statistical uncertainty of the yields per $K^0_{\mathrm{L}}\pi^+\pi^-$ bin while the second uncertainty stems from the method used to determine the total number of combinatorial and continuum background events in the signal sample.*

3.5.4 The Reconstruction and Selection Efficiency

The efficiency for the reconstruction of the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_{\mathrm{L}}\pi^+\pi^-$ signal events is determined from the signal Monte Carlo described in Section 3.2.2.1. The absolute efficiency for each $K^0_{\mathrm{L}}\pi^+\pi^-$ bin is listed in Table 3.19. The uncertainty on the efficiency is taken to be the binomial uncertainty.

A systematic uncertainty is assigned in Section 3.7.4 to account for a possible bias in the amplitude models used in the reweighting procedure.

Bin	Efficiency [%]	Bin	Efficiency [%]
1	27.01 ± 0.71	5	27.01 ± 0.72
2	26.64 ± 0.78	6	25.43 ± 1.75
3	25.92 ± 1.01	7	26.39 ± 1.06
4	29.3 ± 1.01	8	26.94 ± 1.01

Table 3.19: *Reconstruction and selection efficiency of the $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_{\mathrm{L}}\pi^+\pi^-$ signal events for each $K^0_{\mathrm{L}}\pi^+\pi^-$ bin obtained from the signal Monte Carlo sample. The uncertainty on the efficiency is the binomial uncertainty.*

3.5.5 $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ Signal Event Yields

The number of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_{\mathrm{L}}\pi^+\pi^-$ signal events in the data sample is determined by subtracting the peaking background yields (Section 3.5.3.1), and combinatorial and continuum background yields (Section 3.5.3.2) from the reconstructed and selected $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ candidates (Section 3.5.2) and subsequently correcting for the reconstruction and selection efficiency (Section 3.5.4). The distribution of the reconstructed $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ candidates over the $K^0_L\pi^+\pi^-$ bins is shown in Figure 3.15. The figure also shows the contributions from the different background sources. The background-subtracted and efficiency-corrected number of signal events in the data sample for each $K^0_L\pi^+\pi^-$ bin is listed in Table 3.20. The first uncertainty is the statistical uncertainty propagated from the Poisson uncertainties on the reconstructed $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ candidates and the individual background contributions. The second uncertainty is the combination of the systematic uncertainties on the individual background yields. The third uncertainty is the propagated binomial uncertainty from the reconstruction and selection efficiency.

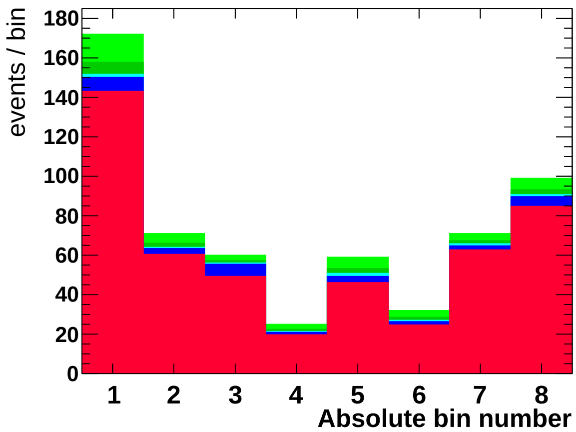


Figure 3.15: *Distribution of reconstructed and selected $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_L\pi^+\pi^-$ candidates over the $K^0_L\pi^+\pi^-$ bins. The figure shows the contribution of signal events (red), peaking background events from $D \to K^0_S\pi^+\pi^-$ vs. $D \to K^0_L\pi^+\pi^-$ decays (dark blue) and $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_L\pi^+\pi^-$ decays (light blue), combinatorial background events (dark green) and continuum events (light green).*

Bin	Number of Signal Events	Bin	Number of Signal Events
1	$510.70\,^{+53.55}_{-52.95}\ \pm 8.29\ \pm 0.14$	5	$169.10\,^{+31.35}_{-30.96}\ \pm 5.06\ \pm 0.01$
2	$227.30\,^{+34.3}_{-33.54}\ \pm 4.87\ \pm 0.02$	6	$99.79\,^{+25.21}_{-24.56}\ \pm 4.25\ \pm 0.14$
3	$197.20\,^{+32.55}_{-32.83}\ \pm 4.11\ \pm 0.04$	7	$238.20\,^{+34.16}_{-34.27}\ \pm 4.29\ \pm 0.06$
4	$64.72\,^{+18.2}_{-18.35}\ \pm 3.06\ \pm 0.07$	8	$313.50\,^{+39.53}_{-39.66}\ \pm 5.30\ \pm 0.01$

Table 3.20: *Background-subtracted and efficiency-corrected number of $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_L\pi^+\pi^-$ signal events expected in the full CLEO-c data sample for each $K^0_L\pi^+\pi^-$ bin. The first uncertainty is the purely statistical uncertainty propagated from the statistical uncertainties on the reconstructed yields and the individual background yields. The second uncertainty is the combination of the systematic uncertainties on the individual background yields. The third uncertainty is the propagated binomial uncertainty from the reconstruction and selection efficiency.*

3.6 Extraction of the CP-even Fraction $F_+^{4\pi}$

The method used to compute $F_{4\pi}^+$ from the signal event yields determined in Section 3.4 and Section 3.5 is a least square fit. As shown in Section 1.5.4 and Section 3.1, the expected number of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K_S^0\pi^+\pi^-$ events $N_{|i|}^{K_S^0\pi^+\pi^-}$ in the
$K_S^0\pi^+\pi^-$ bin i[5], for a given value of $F_{4\pi}^+$ is

$$N_{|i|}^{K_S^0\pi^+\pi^-} = h\left(T_i + T_{-i} - 2\sqrt{T_i T_{-i}}\; c_i\;\left(2\;F_+^{4\pi} - 1\right)\right) \tag{3.24}$$

and the expected number of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K_L^0\pi^+\pi^-$ events $N_{|i|}^{K_L^0\pi^+\pi^-}$ in the $K_L^0\pi^+\pi^-$ bin i is given by

$$N_{|i|}^{K_L^0\pi^+\pi^-} = h'\left(T'_i + T'_{-i} - 2\sqrt{T'_i T'_{-i}}\; c'_i\;\left(2\;F_+^{4\pi} - 1\right)\right) \tag{3.25}$$

where h and h' are normalisation factors specific to the double tag.

The fit parameters are the CP-even fraction $F_{4\pi}^+$ as well as the normalisation factors h and h'. The values for T_i, T'_i , c_i and c'_i are let to vary in the fit under Gaussian constraints. The means of the respective Gaussians are the measured values from Reference [36] for T_i and T'_i and from Reference [22] for c_i and c'_i. The widths of the Gaussians are determined by the respective measurement uncertainties. Correlations between the $c_i^{(\prime)}$ are taken into account. The resulting expression that is minimised during the fitting procedure is given by

$$\begin{aligned}\chi^2 = &\sum_i^{bins} \frac{\left(N_{|i|}^{K_S^0\pi^+\pi^-} - N_{|i|}^{K_S^0\pi^+\pi^-\,fit}\right)^2}{\sigma^2(N_{|i|}^{K_S^0\pi^+\pi^-})} \\ &+\sum_i^{bins} \frac{\left(N_{|i|}^{K_L^0\pi^+\pi^-} - N_{|i|}^{K_L^0\pi^+\pi^-\,fit}\right)^2}{\sigma^2(N_{|i|}^{K_L^0\pi^+\pi^-})} \\ &+\sum_i^{bins} \frac{\left(T_i - T_i^{fit}\right)^2}{\sigma^2(T_i)} \\ &+\sum_i^{bins} \frac{\left(T'_i - T_i'^{fit}\right)^2}{\sigma^2(T'_i)} \\ &+\boldsymbol{\Delta}_{c,c'}^T \boldsymbol{V}^{-1} \boldsymbol{\Delta}_{c,c'}\end{aligned} \tag{3.26}$$

where $\boldsymbol{\Delta}_{c,c'}$ is the vector consisting of the $c_i - c_i^{fit}$ and $c'_i - c_i'^{fit}$ and $\boldsymbol{V}$ is the covariance matrix that contains the correlation between the individual c_i's and the c'_i's. The asymmetric uncertainties on the number of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K_S^0\pi^+\pi^-$ events and $D \to 2\pi^+2\pi^-$ *vs.* $D \to K_L^0\pi^+\pi^-$ events, listed in Table 3.12 and Table 3.20 respectively, are averaged in the fitting procedure.

Individual fits are also performed as a cross check to the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K_S^0\pi^+\pi^-$ events and $D \to 2\pi^+2\pi^-$ *vs.* $D \to K_L^0\pi^+\pi^-$ events separately.

[5] Here bin i refers to the union of bin i and bin $-i$ as defined in the $\Delta\delta$ *BaBar 2008* binning scheme described in Section 1.6 of Chapter 1.

3.6.1 Validation of the Fitting Procedure

The fitting procedure is validated using pseudo-experiments. A *true* distribution of the number of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K_S^0\pi^+\pi^-$ and $D \to 2\pi^+2\pi^-$ *vs.* $D \to K_S^0\pi^+\pi^-$ events over the $K_{S,L}^0\pi^+\pi^-$ bins is generated for three different values of $F_{4\pi}^{+}$ respectively. These *true* distributions are varied 20 000 times according to a Gaussian with the width corresponding to the uncertainties on the events yields per bin listed in Table 3.12 and Table 3.20. A fit to extract $F_{4\pi}^{+}$ is performed on each of the 20 000 distributions. The pulls of these pseudo-experiments — i.e. the difference of $F_{4\pi}^{+}$ used to generate the *true* distribution and the fitted $F_{4\pi}^{+}$ divided by the uncertainty on the fitted $F_{4\pi}^{+}$— are shown in Figure 3.16. The means and widths of these distributions are listed in Table 3.21. The means of all three distributions are consistent with 0 and their standard deviations consistent with 1. This indicates that there is no bias in the fitting procedure and that the uncertainty on $F_{4\pi}^{+}$ is calculated correctly by the fitter.

$F_{4\pi}^{+}$	Mean	Standard Deviation
0.69	0.00	0.99
0.74	0.00	0.98
0.79	0.02	0.98

Table 3.21: *Mean and widths of the pull distributions of the pseudo-experiments used to validate the fitting procedure.*

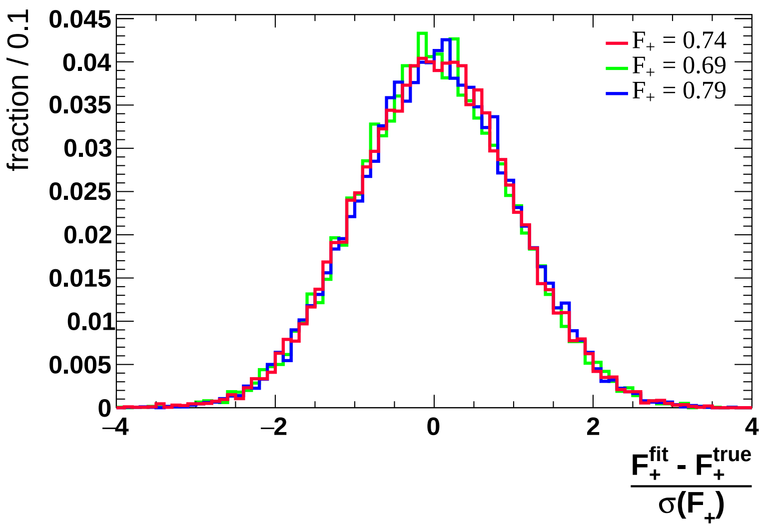


Figure 3.16: *Distribution of the pulls — i.e. the difference of $F_{4\pi}^{+}$ used to generate the true distribution and the fitted $F_{4\pi}^{+}$ divided by the uncertainty on the fitted $F_{4\pi}^{+}$— of the pseudo-experiments used to evaluate the fitting procure.*

3.6.2 Fit Result for the *CP*-even Fraction $F_{+}^{4\pi}$

The results of the individual fits to the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_\mathrm{S}\pi^+\pi^-$ distribution and the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_\mathrm{L}\pi^+\pi^-$ distribution are $F_{4\pi}^{+} = 0.858 \pm 0.077$ and $F_{4\pi}^{+} = 0.682 \pm 0.063$, respectively. The result of the simultaneous fit is $F_{4\pi}^{+} = 0.755 \pm 0.053$. The uncertainty is the combined statistical uncertainty and the uncertainty from the hadronic parameters of the $D^0 \to K^0_\mathrm{S,L}\pi^+\pi^-$ decays. The results on $F_{4\pi}^{+}$ from $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_\mathrm{S}\pi^+\pi^-$ events and $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_\mathrm{L}\pi^+\pi^-$ events are in agreement within $1.8\,\sigma$.

The χ^2_{ndof} of the individual fit results are 0.76 and 0.42 for the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_\mathrm{S}\pi^+\pi^-$ distribution and the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_\mathrm{L}\pi^+\pi^-$ distribution, respectively. The χ^2_{ndof} for the simultaneous fit is 0.49.

All results of the fitting procedure are listed in Table 3.22.

	$F_{4\pi}^{+}$	h	h'	χ^2
$D \to K^0_\mathrm{S}\pi^+\pi^-$	0.858 ± 0.077	479.49 ± 45.16		0.76
$D \to K^0_\mathrm{L}\pi^+\pi^-$	0.682 ± 0.063		826.66 ± 58.27	0.42
simultaneous fit	0.755 ± 0.053	468.29 ± 42.77	794.81 ± 58.83	0.49

Table 3.22: *Results of the individual fits to the $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_\mathrm{S}\pi^+\pi^-$ distribution and the $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_\mathrm{L}\pi^+\pi^-$ distribution and the simultaneous fit.*

3.7 Systematic Uncertainties

A comprehensive list of systematic uncertainties is studied. Sources of potential bias are identified and their effect on the value of $F_{4\pi}^{+}$ is estimated using either pseudo-experiments or an alternative method of estimating a certain quantity and recalculating $F_{4\pi}^{+}$. The contribution of all individual sources to the uncertainty on $F_{4\pi}^{+}$ is summarised at the end of this section in Table 3.24.

3.7.1 Uncertainty on the $D^0 \to K^0_S\pi^+\pi^-$ Hadronic Parameters

The hadronic parameters of the $D \to K^0_\mathrm{S}\pi^+\pi^-$ decay and the $D \to K^0_\mathrm{L}\pi^+\pi^-$ decay are constrained in the nominal fit to extract $F_{4\pi}^{+}$. In order to estimate the contribution of the uncertainties of the hadronic parameter to the uncertainty to $F_{4\pi}^{+}$, the fit is performed once with the hadronic parameters constrained as shown in Equation 3.26 and once with the hadronic parameters fixed to their measured values. The contribution of the uncertainty on the hadronic parameters σ_had to the uncertainty on $F_{4\pi}^{+}$ is then calculated as

$$\sigma^2_\mathrm{had} = \sigma^2 - \sigma^2_\mathrm{stat} \qquad (3.27)$$

where σ_{stat} is the uncertainty on $F_{4\pi}^{+}$ given by the fit when the hadronic parameters are fixed to their measured values and σ is the uncertainty on $F_{4\pi}^{+}$ given by the fit when the hadronic parameters are constrained as described by Equation 3.26. This procedure yields a systematic uncertainty of 0.02 on $F_{4\pi}^{+}$ in the simultaneous fit coming from the uncertainty on the hadronic parameters used as input to the fit. This makes it the largest systematic uncertainty to $F_{4\pi}^{+}$.

3.7.2 Bin-to-bin Migration

Due to the finite momentum resolution of the CLEO-c detector, an event can be reconstructed in a different $K^0_{\mathrm{S,L}}\pi^+\pi^-$ bin from its *true* bin. The effect of this bin migration on the value of $F_{4\pi}^{+}$ is assessed with pseudo-experiments.

The bin-migration matrices for the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_{\mathrm{S}}\pi^+\pi^-$ and $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_{\mathrm{L}}\pi^+\pi^-$ events are determined on the respective signal Monte Carlo samples. Uncertainties are assigned to the matrix elements due to the limited number of events in the Monte Carlo samples. The elements of the bin-migration matrices are varied 20 000 times according to a Gaussian with the mean corresponding to the matrix element itself and the width corresponding to the matrix elements' uncertainties. These 20 000 bin-migration matrices are each applied to the *true* distribution of number of events per $K^0_{\mathrm{S,L}}\pi^+\pi^-$ bin. The true distributions of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_{\mathrm{S}}\pi^+\pi^-$ events and $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_{\mathrm{L}}\pi^+\pi^-$ events were generated according to Equation 3.24 and Equation 3.25, respectively. Three different input values of $F_{4\pi}^{+}$ are used, namely 0.69, 0.74 and 0.79.

Figure 3.17 shows the resulting distribution of the difference between the $F_{4\pi}^{+}$ used to generate the *true* distributions and the $F_{4\pi}^{+}$ from the fit to the bin-migrated distributions. Table 3.23 lists the means and widths of the distributions. The average of the means of the three distributions is taken as systematic uncertainty associated with the bin migration. This results in a systematic uncertainty of 0.011 for the simultaneous fit result which makes it one of the three biggest systematic uncertainties.

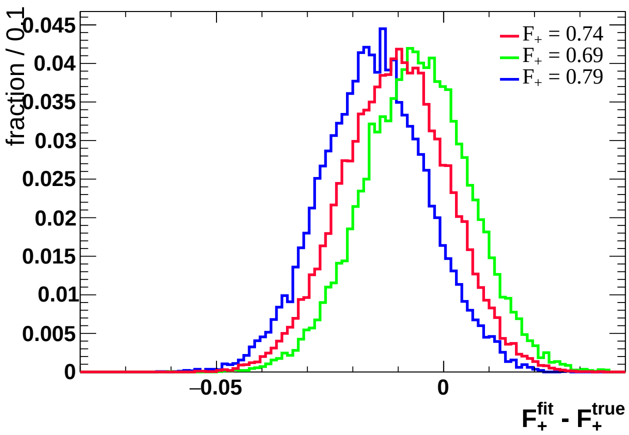


Figure 3.17: *Distribution of the difference between the true $F_{4\pi}^{+}$ and the fitted $F_{4\pi}^{+}$ from the pseudo-experiments used to evaluate the systematic uncertainty associated to the migration of events between the $K^0_{\mathrm{S,L}}\pi^+\pi^-$ bins.*

$F_{4\pi}^{+}$	Mean	Standard Deviation
0.69	-0.011	0.012
0.74	-0.006	0.012
0.79	-0.015	0.012

Table 3.23: *Mean and widths of the distributions of the pseudo-experiments to determine the effect of bin-to-bin migration of events.*

3.7.3 Multiple Candidate Selection

In order to evaluate a possible bias on $F_{4\pi}^{+}$ from the multiple candidate selection in Section 3.3.3.4, the analysis procedure is redone using an alternative multiple candidate selection. In this selection one candidate in the events is chosen at random. The value for $F_{4\pi}^{+}$ is recalculated and the systematic uncertainty assigned to the multiple candidate selection is the difference between the resulting value of $F_{4\pi}^{+}$ and the nominal $F_{4\pi}^{+}$. This yields a systematic uncertainty of 0.004 for the simultaneous fit result.

3.7.4 Amplitude Model used in Reweighting Procedure

To account for a possible bias from an incorrect model used to reweight the signal Monte Carlo samples in Section 3.2.2.1 the samples are reweighted with an alternative model corresponding to the density of states, e.g. neglecting any correlation between the two D mesons in the event. The PDF for this reweighting is given by

$$\left|A(\psi(3770) \to D^0\overline{D}^0 \to f_1 f_2)\right|^2 \propto \left|A(D^0 \to f_1) \cdot A(\overline{D}^0 \to f_2)\right|^2 \tag{3.28}$$

$$+\left|A(D^0 \to f_2) \cdot A(\overline{D}^0 \to f_1)\right|^2 \tag{3.29}$$

The different reweighting impacts the selection and reconstruction efficiencies of the signal decays as well as the relative distribution of peaking background events over the $K^0_{\mathrm{S,L}}\pi^+\pi^-$ bins. The value for $F_{4\pi}^{+}$ is recalculated and the systematic uncertainty is the difference between the resulting value of $F_{4\pi}^{+}$ and the nominal $F_{4\pi}^{+}$. This yields a systematic uncertainty of 0.013 for the simultaneous fit result which makes this the second biggest systematic uncertainty.

3.7.5 Peaking Background from $D \to K^0_S\pi^+\pi^-$ *vs.* $D \to K^0_{S,L}\pi^+\pi^-$ Events

3.7.5.1 Absolute Number of $D \to K^0_S\pi^+\pi^-$ *vs.* $D \to K^0_{S,L}\pi^+\pi^-$ Background Events

The nominal way of determining the absolute peaking background yields from $D \to K^0_{\mathrm{S}}\pi^+\pi^-$ *vs.* $D \to K^0_{\mathrm{S,L}}\pi^+\pi^-$ events is based on the generic Monte Carlo sample. To estimate a systematic uncertainty associated to this method the result is compared to a data-driven method. Instead of the K^0_{S} veto described in Section 3.3.3.3

a K^0_S selection is applied to K^0_S candidates within the $D \to 2\pi^+2\pi^-$ candidates. Events with no K^0_S candidate in the $D \to 2\pi^+2\pi^-$ events are rejected. This selection gives a data sample that consists of 96% of $D \to K^0_S\pi^+\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ events for the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ candidates and 91% of $D \to K^0_S\pi^+\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ events for the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ candidates[6]. The ratio of events in the K^0_S selection over the events in the K^0_S veto can be determined from the signal Monte Carlo sample. This ratio is used to extrapolate the number of $D \to K^0_S\pi^+\pi^-$ *vs.* $D \to K^0_{S,L}\pi^+\pi^-$ events in the K^0_S selection in data to the number of $D \to K^0_S\pi^+\pi^-$ *vs.* $D \to K^0_{S,L}\pi^+\pi^-$ events in the K^0_S veto in the data.

The value for $F^+_{4\pi}$ is recalculated and the systematic uncertainty is the difference between the resulting value of $F^+_{4\pi}$ and the nominal $F^+_{4\pi}$. This yields a systematic uncertainty of 0.001 for the simultaneous fit result.

This systematic uncertainty also covers uncertainty from the unknown branching fraction of $D^0 \to K^0_L\pi^+\pi^-$ decays used in the generic Monte Carlo sample.

3.7.5.2 Distribution of $D \to K^0_S\pi^+\pi^-$ *vs.* $D \to K^0_{S,L}\pi^+\pi^-$ Background Events over the $K^0_{S,L}\pi^+\pi^-$ Bins

The distribution of the $D \to K^0_S\pi^+\pi^-$ *vs.* $D \to K^0_{S,L}\pi^+\pi^-$ events over the $K^0_{S,L}\pi^+\pi^-$ bins is determined according to the distribution in the signal Monte Carlo samples as described in Section 3.4.3.2 and Section 3.5.3.1. In order to evaluate the systematic uncertainty from an incorrect model used in the reweighting procedure of the Monte Carlos samples (see Section 3.2.2.1), the Monte Carlo samples are reweighted according to the density of states without taking into account correlation between the D mesons. This systematic uncertainty is incorporated in the systematic uncertainty associated with the amplitude model used to reweight the Monte Carlo described in Section 3.7.4.

3.7.6 Peaking Background from $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ Events

3.7.6.1 Absolute Number of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ Background Events

The impact of a possible bias from using $F^+_{4\pi}$ in the estimation of the number of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ events in the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ candidates is studied by calculating the total number of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ events for a value of $F^+_{4\pi} = 0$ and $F^+_{4\pi} = 1$ which will yield the greatest and smallest possible contributions respectively. The values for $F^+_{4\pi}$ are recalculated for both cases and systematic uncertainty is calculated as the difference between the two resulting values for $F^+_{4\pi}$ and the nominal value for $F^+_{4\pi}$. This yields a systematic uncertainty smaller than $^{+0.001}_{-0.001}$ for the simultaneous fit result which is less than 0.1% relative uncertainty and considered negligible.

[6] The percentages were determined on the generic Monte Carlo samples.

3.7.6.2 Distribution of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ Background Events over the $K^0_L\pi^+\pi^-$ Bins

The distribution of the $D\to 2\pi^+2\pi^-$ *vs.* $D\to K^0_{\rm S}\pi^+\pi^-$ events over the $K^0_{\rm L}\pi^+\pi^-$ bins is determined according to the distribution in the signal Monte Carlo samples as described in Section 3.5.3.1. In order to evaluate the systematic uncertainty from an incorrect model used in the reweighting procedure of the Monte Carlos samples (see Section 3.2.2.1), the Monte Carlo samples are reweighted according to the density of states without taking into account correlation between the D mesons. This systematic uncertainty is incorporated in the systematic uncertainty associated with the amplitude model used to reweight the Monte Carlo described in Section 3.7.4.

3.7.7 Combinatorial Background to $D\to 2\pi^+2\pi^-$ *vs.* $D\to K^0_S\pi^+\pi^-$ Candidates

3.7.7.1 Absolute Number of Combinatorial Background Events

In the determination of the absolute number of combinatorial background events in the $D\to 2\pi^+2\pi^-$ *vs.* $D\to K^0_{\rm S}\pi^+\pi^-$ candidates (see Section 3.4.3.3), the assumption was made that the density of combinatorial background events in the sidebands is the same as the density of combinatorial background events in the signal region. In order to estimate a systematic uncertainty from the total number of combinatorial background events being incorrect, the value of $F^+_{4\pi}$ is determined for 50% and 200% of the nominal number of combinatorial background events. The systematic uncertainty is the difference between the resulting values of $F^+_{4\pi}$ and the nominal value of $F^+_{4\pi}$. It yields an uncertainty of $^{+0.001}_{-0.004}$ for the simultaneous fit result.

3.7.7.2 Non-uniform Distribution of Combinatorial Background Events

In Section 3.4.3.3 the assumption was made that the combinatorial background events are distributed uniformly over the $m^2_{K^0_{\rm S}\pi^-}$ *vs.* $m^2_{K^0_{\rm S}\pi^+}$ space of the $D\to K^0_{\rm S}\pi^+\pi^-$ candidate. To account for a potential bias introduced by this assumption, the distribution of the combinatorial background events over the $K^0_{\rm S}\pi^+\pi^-$ bins is determined using the distribution of the combinatorial background events as represented in the generic Monte Carlo samples. The value of $F^+_{4\pi}$ is recalculated under the new assumption and the uncertainty assigned to a potentially non-uniform distribution of combinatorial background events is the difference between the resulting $F^+_{4\pi}$ and the nominal $F^+_{4\pi}$. This yields a systematic uncertainty of 0.006 for the simultaneous fit result.

3.7.8 Combinatorial and Continuum Background to $D\to 2\pi^+2\pi^-$ *vs.* $D\to K^0_L\pi^+\pi^-$ Candidates

3.7.8.1 Absolute Number of Combinatorial and Continuum Background Events

The absolute number of combinatorial and continuum background events events in the signal sample is calculated in Section 3.5.3.2 under the assumption that the ratio of events between the signal region and the sidebands is accurately modelled in the different Monte Carlo samples. In order to estimate a potential bias from this assumption the number of combinatorial and continuum background events in the signal sample is extrapolated from the sidebands defined by $0.075\,\mathrm{GeV}^2/c^4 \leq m^2_{miss} \leq 0.15\,\mathrm{GeV}^2/c^4$ or $0.45\,\mathrm{GeV}^2/c^4 \leq m^2_{miss} \leq 0.8\,\mathrm{GeV}^2/c^4$. The value for $F^+_{4\pi}$ is calculated with the new number of combinatorial and continuum events. The systematic uncertainty is the difference between the resulting $F^+_{4\pi}$ and the nominal $F^+_{4\pi}$ and is 0.006 for the simultaneous fit result.

3.7.8.2 Non-uniform Distribution of Combinatorial and Continuum Background Events

In Section 3.5.3.2 the assumption was made that the combinatorial background and the continuum background events are distributed uniformly over the $m^2_{K^0_{\mathrm{L}}\pi^-}$ *vs.* $m^2_{K^0_{\mathrm{L}}\pi^+}$ space of the $D \rightarrow K^0_{\mathrm{L}}\pi^+\pi^-$ candidate. To account for a potential bias introduced by this assumption, the distribution of the combinatorial and continuum background events over the $K^0_{\mathrm{L}}\pi^+\pi^-$ bins is determined according to the distribution of events in a sideband in data. In order to exclude any contributions from peaking background events, the sideband is defined as $1.83\,\mathrm{MeV}/c^2 \leq m^D_{BC} \leq 1.85\,\mathrm{MeV}/c^2$ for the beam constrained D mass of the $D \rightarrow 2\pi^+2\pi^-$ candidate and $0.075\,\mathrm{GeV}^2/c^4 \leq m^2_{miss} \leq 0.15\,\mathrm{GeV}^2/c^4$ or $0.45\,\mathrm{GeV}^2/c^4 \leq m^2_{miss} \leq 0.8\,\mathrm{GeV}^2/c^4$ for the square of the missing mass in the event. The fit for $F^+_{4\pi}$ is rerun and the systematic uncertainty is taken to be the difference between this $F^+_{4\pi}$ value and the nominal $F^+_{4\pi}$ value. This results in a systematic uncertainty of 0.005 for the simultaneous fit result.

3.7.9 Summary of Systematic Uncertainties

All systematic uncertainties are listed in Table 3.24. The sources of greatest uncertainties are the uncertainties on the $D \rightarrow K^0_{\mathrm{S,L}}\pi^+\pi^-$ hadronic parameters, the bin-to-bin migration and the use of a different amplitude model in the reweighting of the signal Monte Carlo. This is expected since they have the most direct impact on the relative distribution of events over the $K^0_{\mathrm{S,L}}\pi^+\pi^-$ bins.

The resulting value for $F^+_{4\pi}$ from the simultaneous fit is $F^{4\pi}_+ = 0.755 \pm 0.050\,(\mathrm{stat}) \pm 0.029\,(\mathrm{sys})$. Even with the complete CLEO-c $e^+e^- \rightarrow \psi(3770)$ data sample, the statistical uncertainty dominates over the systematic uncertainty.

	$\sigma^{K^0_S\pi^+\pi^-}$	$\sigma^{K^0_L\pi^+\pi^-}$	$\sigma^{\text{simultaneous}}$
hadronic parameters	$\pm$ 0.021	$\pm$ 0.030	$\pm$ 0.021
bin migration	$\pm$ 0.001	$\pm$ 0.020	$\pm$ 0.011
multiple candidate selection	$\pm$ 0.007	$\pm$ 0.006	$\pm$ 0.004
amplitude model	$\pm$ 0.014	$\pm$ 0.0101	$\pm$ 0.013
$D \to K^0_S\pi^+\pi^-$ *vs.* $D \to K^0_{S,L}\pi^+\pi^-$ bkg	$\pm$ 0.002	$\pm$ 0.001	$\pm$ 0.001
$D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ bkg		${}^{+0.000}_{-0.000}$	${}^{+0.000}_{-0.000}$
$D \to K^0_S\pi^+\pi^-$ absolute combinatorial bkg	${}^{+0.030}_{-0.016}$		${}^{+0.001}_{-0.004}$
$D \to K^0_S\pi^+\pi^-$ relative combinatorial bkg	$\pm$ 0.009		$\pm$ 0.006
$D \to K^0_L\pi^+\pi^-$ absolute combinatorial bkg		$\pm$ 0.006	$\pm$ 0.006
$D \to K^0_L\pi^+\pi^-$ relative combinatorial bkg		$\pm$ 0.011	$\pm$ 0.005
total	$\pm$ 0.036	$\pm$ 0.040	$\pm$0.029

Table 3.24: *Summary of systematic uncertainties on $F^+_{4\pi}$ from the measurement using $D \to 2\pi^+2\pi^-$ vs. $D \to K^0_{S,L}\pi^+\pi^-$ decays in the complete CLEO-c $e^+e^- \to \psi(3770)$ data sample.*

3.8 Results

The results of the individual fits to the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ distribution and the $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ distribution including the systematic uncertainties are $F^+_{4\pi} = 0.858 \pm 0.075\,(\text{stat}) \pm 0.036\,(\text{sys})$ and $F^+_{4\pi} = 0.682 \pm 0.056\,(\text{stat}) \pm 0.040\,(\text{sys})$, respectively. The result of the simultaneous fit including the systematic uncertainty is $F^+_{4\pi} = 0.755 \pm 0.050\,(\text{stat}) \pm 0.029\,(\text{sys})$. The uncertainty from the limited signal Monte Carlo sample used to calculate the reconstruction and selection efficiency is less than 0.1% and therefore completely negligible. The results on $F^+_{4\pi}$ from $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ events and $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ events are in agreement within $\approx 1.4\,\sigma$.

The resulting distributions of $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_S\pi^+\pi^-$ and $D \to 2\pi^+2\pi^-$ *vs.* $D \to K^0_L\pi^+\pi^-$ events after the fit are compared to the data in Figure 3.18 for the simultaneous fit.

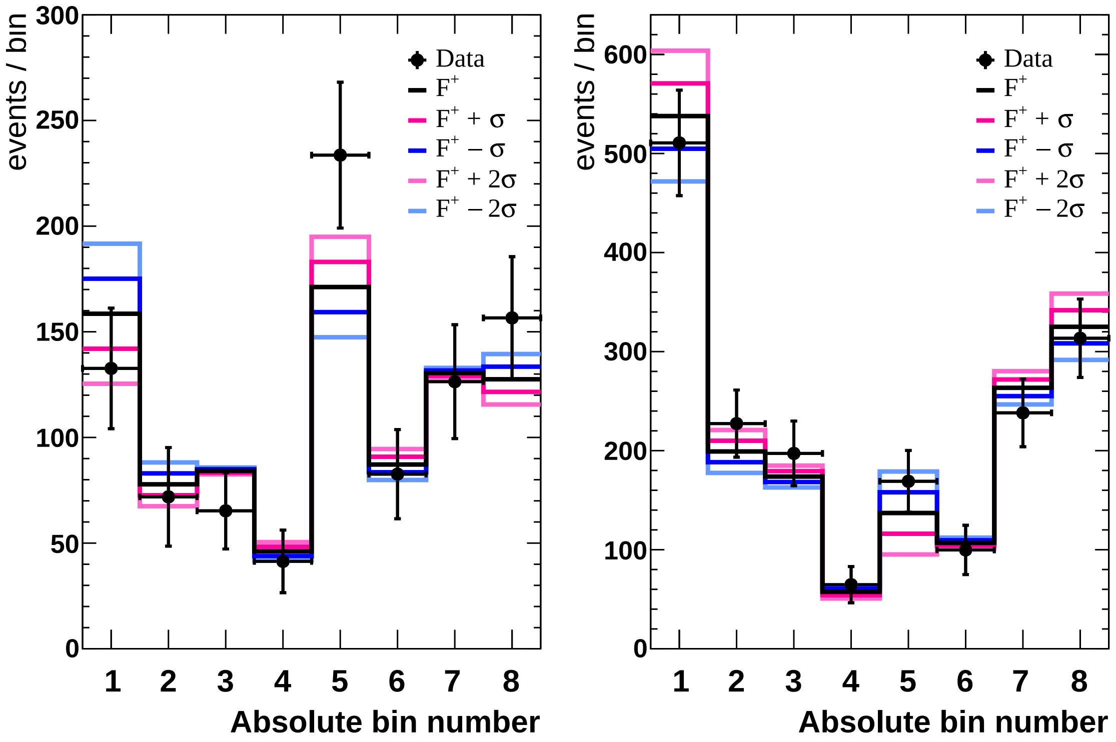


Figure 3.18: *Distribution of events over the $K^0_{\mathrm{S,L}}\pi^+\pi^-$ bins for $D\to 2\pi^+2\pi^-$ vs. $D\to K^0_{\mathrm{S}}\pi^+\pi^-$ events (left) and $D\to 2\pi^+2\pi^-$ vs. $D\to K^0_{\mathrm{L}}\pi^+\pi^-$ events (right). The black markers represent the distributions of signal events from data and the black histograms the distributions determined by the fitted valued of $F^+_{4\pi}$. The blue/red histograms represent the distributions for $F^+_{4\pi}\pm\sigma$ while the light blue/light red histograms represent the distributions for $F^+_{4\pi}\pm 2\sigma$.*

3.9 Conclusion

In this chapter the measurement of the *CP*-even fraction of the $D^0\to 2\pi^+2\pi^-$ decay was performed using a dataset corresponding to 818 pb^{-1} of quantum correlated $D\overline{D}$ decays produced in electron-positron collisions at the $\psi(3770)$ resonance collected by the CLEO-c experiment. In the analysis, one of the correlated D mesons is reconstructed as $D\to 2\pi^+2\pi^-$ while the other D meson is reconstructed as $D\to K^0_{\mathrm{S,L}}\pi^+\pi^-$. The phase space of the $D\to K^0_{\mathrm{S,L}}\pi^+\pi^-$ decays is divided into bins and sensitivity to the *CP*-even fraction of $D^0\to 2\pi^+2\pi^-$ is obtained by determining the variation of signal yields over these bins. The *CP*-even fraction is measured to be $F^+_{4\pi} = 0.755\pm 0.050\,(\mathrm{stat})\pm 0.029\,(\mathrm{syst})$.

This analysis is part of the first measurement of the *CP*-even fraction $F^+_{4\pi}$ of the $D\to 2\pi^+2\pi^-$ decay and was published in Reference [45][7]. This analysis was published alongside two other analyses which also used correlated D mesons. The first analysis reconstructed the other D meson in various *CP* eigenstates while the second

[7] The value for $F^+_{4\pi}$ from $K^0_{\mathrm{S,L}}\pi^+\pi^-$ tags in Reference [45] differs from the value determined in the analysis in this document. The difference is due to improved analysis techniques since 2015 as well as to a different treatment of the systematic uncertainties.

analysis used a technique analogous to this analysis with the other D mesons reconstructed as $\pi^+\pi^-\pi^0$. These analyses yield results of $F_+^{4\pi} = 0.754 \pm 0.031 \pm 0.021$ and $F_+^{4\pi} = 0.695 \pm 0.050 \pm 0.021$, respectively. The CP-even fraction was also determined from an amplitude model of the $D^0 \to 2\pi^+2\pi^-$ decay and found to be $F_+^{4\pi} = 0.729 \pm 0.009 \pm 0.015 \pm 0.01$ [24]. All these results are in good agreement with each other and with the result of this analysis of $F_+^{4\pi} = 0.755 \pm 0.050 \pm 0.029$.

The precisely measured value of $F_+^{4\pi} = 0.755 \pm 0.050 \pm 0.029$ and the relatively high branching fraction of $D \to 2\pi^+2\pi^-$ makes this channel is a valuable addition to the number of D decays that can be used for the measurement of the unitarity-triangle angle γ through the process $B^\pm \to DK^\pm$. A measurement of the CP observables in $B^\pm \to DK^\pm$ and $B^\pm \to D\pi^\pm$ decays with two-and four-body D decays has already been performed at LHCb using the result of this thesis [6].

The measurement of the CP-even fraction of $D^0 \to 2\pi^+2\pi^-$ also laid the groundwork for the measurement of the hadronic parameters of the $D \to 2\pi^+2\pi^-$ decay. The measurement of the hadronic parameters was performed using the same CLEO-c data sample of correlated $D\overline{D}$ pairs and has just been published in Reference [46].

4. The LHCb Experiment

LHCb is one of the four large experiments allocated at the Large Hadron Collider (LHC). It is dedicated to the study of *CP* violation and the search for new physics in rare decays of B and D mesons. As much focus is put on the precision measurement of SM parameters as on the indirect detection of physics beyond the SM that can be identified through its contribution to loop-mediated processes.

The LHC provides an ideal environment for new research in the sector of heavy flavour physics. The nominal center of mass energy of the LHC in Run II is $\sqrt{s} = 13\,\text{TeV}$. The according $b\bar{b}$ production cross sections within the LHCb acceptance is $\sigma_{b\bar{b}X}(13\,\text{TeV}) = 515\,\mu b$ [47] resulting in $6 \cdot 10^{11}$ B hadrons produced in the LHCb detector acceptance 2015 and 2016 combined ($2\,\text{fb}^{-1}$ integrated luminosity collected in 2015 + 2016). Additionally to providing high statistics, the LHC gives access to the system of the B_s meson – which was mostly beyond the reach of previous b factories such as BaBar and Belle – allowing for the investigation of *CP* violation and the search for new physics in the B_s^0 system.
The LHCb collaboration has already published several new results constraining SM parameters and possible contributions from new physics, for example the restraining of SUSY parameters with the decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ [48].

This chapter is dedicated to the description of the LHCb detector in Run II. After a general overview of the detector, the different subdetectors of the tracking system and the particle identification system are presented and a summary of the detector performance is given. Then the LHCb trigger system and different trigger lines are introduced. At the end of this chapter the LHCb software is described briefly.

4.1 The LHCb Detector

The LHCb detector was designed as a single arm forward spectrometer [49] to match the kinematics of $b\bar{b}$ production in pp collisions shown in Figure 4.1. Corresponding to this distribution, the LHCb detector covers an angular range of $10\,\text{mrad} - 300\,\text{mrad}$ in the horizontal, bending and $10\,\text{mrad} - 250\,\text{mrad}$ in the vertical, non-bending plane. At a center of mass energy of $14\,\text{TeV}$ $24\,\%$ of the produced $b\bar{b}$ pairs are in the detector acceptance.

To enable precision measurements, the LHCb detector was built with a minimal amount of material budget. Additionally, the LHC conducts luminosity levelling for LHCb, keeping the number of proton-proton interactions per bunch-crossing adjusted to an average of 1.5 throughout the entire run. This is achieved by displacing the proton-beams with respect to each other at the LHCb collision point, giving the bunches only a small overlap while colliding.

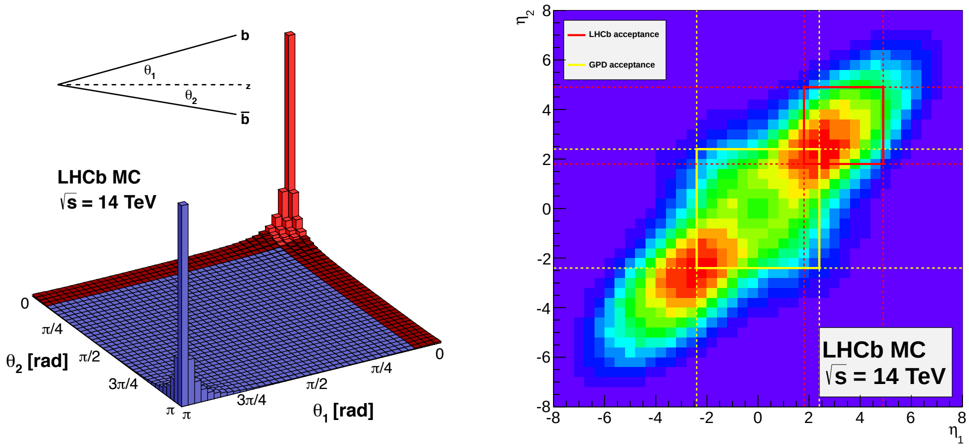


Figure 4.1: *At hadron colliders the dominant mechanism for $b\bar{b}$ production is through gluon gluon fusion. Due to the statistical partition of the energy inside the protons the $b\bar{b}$ pairs originated from the gluon gluon interactions are boosted in the forward or backward direction.* Left: *Density of $b\bar{b}$ events produced in proton-proton collisions as a function of the angular distribution of the b quark and the $\bar{b}$ quark.* Right: *Density of $b\bar{b}$ events produced in proton-proton collisions as a function of the pseudo rapidity of the b quark and the $\bar{b}$ quark. The yellow solid box shows the geometrical acceoptance of the general purpose detectors while the red solid box shows the LHCb geometrical acceptance.* [50]

The LHCb detector is composed of several subdetectors shown in Figure 4.2 which are presented in the following sections. Altogether the LHCb detector incorporates precision vertexing and tracking, worldwide leading particle identification and efficient triggering through a system of dynamic triggers.

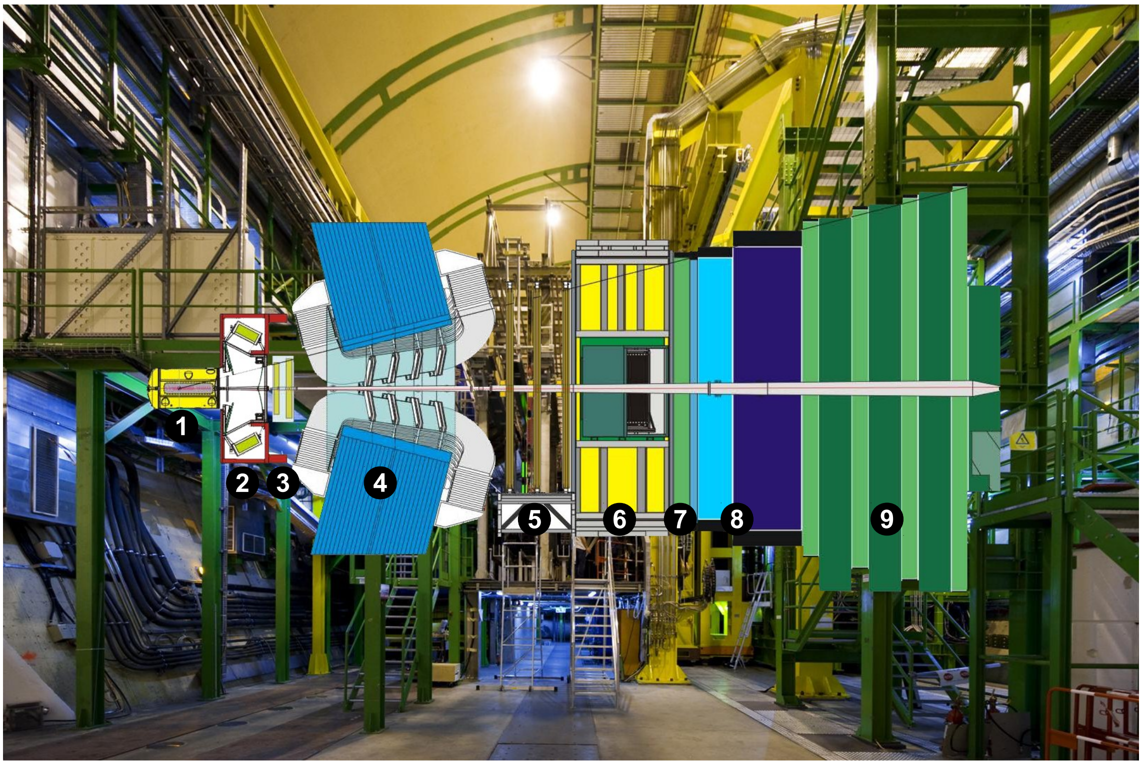


Figure 4.2: *The LHCb detector. 1: Vertex Locator (VELO), 2: Ring Imaging Cherenkov detector 1 (RICH1), 3: Tracker Turicensis (TT), 4: the magnet, 5: the tracking stations, 6: Ring Imaging Cherenkov detector 2 (RICH2), 7: Muon chamber 1 (M1), 8: Scintillating Pad Detector (SPD), PreShower Detector (PS), electromagnetic calorimeter (ECAL) and hadronic calorimeter (HCAL), 9: Muon chambers 2 to 5 (M2 to M5).* [51]

4.1.1 The Tracking System

The LHCb tracking system's purpose is the detection of charged particle tracks and the measurement of their momenta. It is also responsible for the reconstruction of production and decay vertices of B and D mesons. This is essential not only for lifetime measurements but also for an efficient background rejection in the analysis of rare decays.
The tracking system consists of the Vertex Locator (VELO), the magnet, the Silicon Tracker (ST) and the Outer Tracker (OT). The Silicon Tracker makes up for the entire Tracker Turicensis (TT) and the Inner Tracker (IT) which is the inner part of the three tracking stations (T1, T2 and T3) downstream of the magnet. The Outer Tracker is the outer part of the tracking stations (see 5. in Figure 4.2).

4.1.1.1 The Vertex Locator (VELO)

The VELO [52] accurately measures the positions of tracks close to the interaction point and allows for a very precise reconstruction of the primary vertex and the impact parameters of all tracks.

As shown in Figure 4.3, the vertex locator is approximately 1 m long and consists of 21 disc-like modules arranged around the beam pipe. Each module is composed of two halves, each having two sides with silicon strips arranged to measure the R and Φ coordinate respectively. The modules can be approached as close as 8 mm to the beam and have a total radius of 4.2 cm. The strip pitch varies from 38 µm wide strips nearest to the beam pipe to 102 µm wide strips for the R sensors and 79 µm wide strips for the Φ sensors at the outer margin, matching the particle density in the detector.

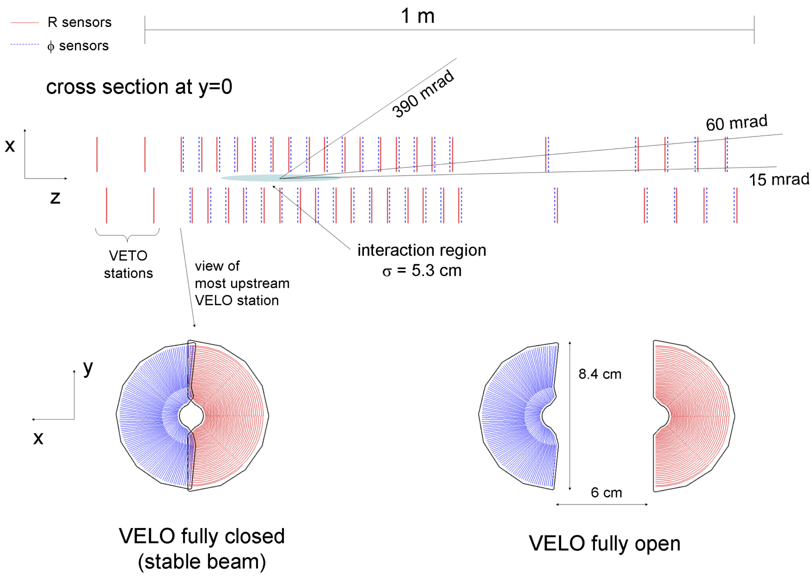


Figure 4.3: *Schematic illustration of the VELO. Top: top view of the 21 modules. Bottom: Two VELO modules shown with both halves to measure Φ (blue) and R (red). Module in closed position (left) and in open position (right) which is engaged while unstable beam conditions.* [52]

4.1.1.2 The Magnet

The LHCb magnet [53] is a warm dipole magnet with saddle shaped coils. The field lines of the magnetic field are parallel to the y axis, making the $(x-z)$ plane the bending plane. Since the relative momentum resolution, $\frac{\sigma_p}{p}$, varies with magnetic field, B, as

$$\frac{\sigma_p}{p} \propto \frac{1}{B} \,, \tag{4.1}$$

the integrated magnetic field was chosen to be 4 Tm. Additionally, the magnet polarisation is regularly changed to reduce systematic effects on measurements due to geometrical acceptances.

4.1.1.3 The Silicon Tracker (ST)

The ST [54] implements silicon microstrip technology for the Tracker Turicensis (TT) and the Inner Tracker (IT) with a strip pitch of about 200 µm. Combining the TT and the IT, the ST has four stations. Each of these four stations consists of four layers that are arranged in a $(x-u-v-x)$ geometry with vertical strips in the x layers and strips rotated by a stereo angle of $-5°$ and $+5°$ in the u and v layer respectively.

The TT is located upstream of the magnet and covers the total LHCb angular acceptance. It allows reconstruction of low momentum particles which are swept out of the detector acceptance after entering the magnetic field and additionally provides information for the trigger by performing transverse momentum measurements for tracks with a large impact parameter. The TT consists of sensors of size $9.4\,\mathrm{cm} \times 9.6\,\mathrm{cm}$ as is illustrated in Figure 4.4. In the inner region the sensors are read out individually, while in regions with lower occupancy the sensors are bonded together in groups of two, three or four sensors.

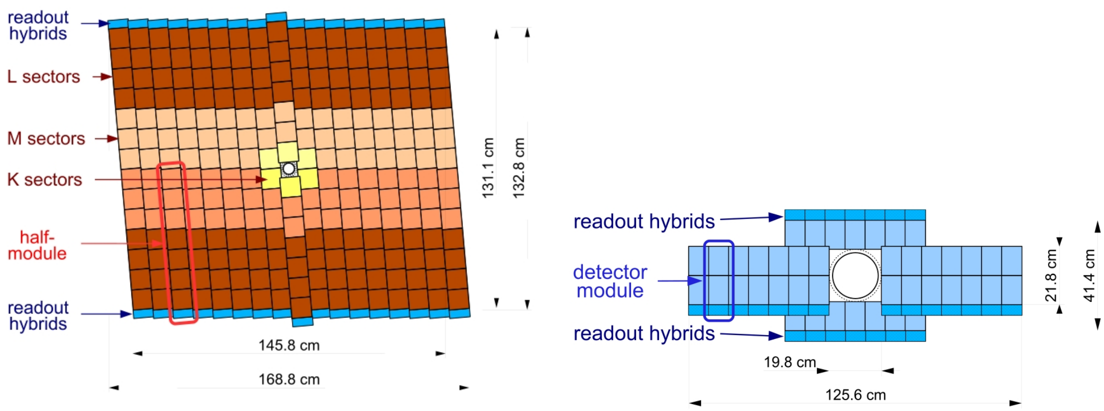


Figure 4.4: *The two parts of the Silicon Tracker (ST).* Left: *Layer of the Tracker Turicensis (TT). The sensors are read out individually (yellow) or bonded together in groups of two, three (light brown) or four (dark brown). The blue boxes at the bottom and the top represent the readout hybrids.* Right: *Layer of the Inner Tracker (IT) which is the inner part of the downstream tracking stations. The blue boxes at the bottom and the top of the modules represent the readout hybrids.* [54]

The IT covers the inner region of the downstream tracking stations where the particle flux is highest. One layer of the IT can be seen in Figure 4.4 on the right. It consists of four pieces arranged in a criss-cross pattern around the beam pipe that cover a total area of $0.35\,\mathrm{m}^2$.

4.1.1.4 The Outer Tracker (OT)

The OT [55] covers the outer region of the downstream tracking stations. Since the particle flux is lower in this area the OT implements the technology of drift tubes. The inner boundaries filled by the IT were determined by the requirement that occupancies in the straw-tubes should not exceed 10% at the nominal running luminosity of $2 \cdot 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

As for the ST, each OT tracking station has four layers arranged in a $(x-u-v-x)$ geometry with vertical tubes in the x layers and tubes rotated by a stereo angle of $-5°$ and $+5°$ in the u and v layer respectively. Each of these layers are composed of two staggered layers (monolayers) of straw-tubes. All straw-tubes are 5 mm thick and filled with a gas mixture of $Argon$ and CO_2 (70 : 30) which combines fast response time (< 50 ns) with a high resolution of the drift coordinate (200 µm).

The structure of the OT layers is shown in Figure 4.5. Each layer is built of two different types of modules. The F modules are 4.85 m long and their monolayers are split along the y axis, at different positions for each monolayer to avoid insensitive regions in the middle of the modules. The S modules are shorter than the F modules and located above and below the beam pipe.

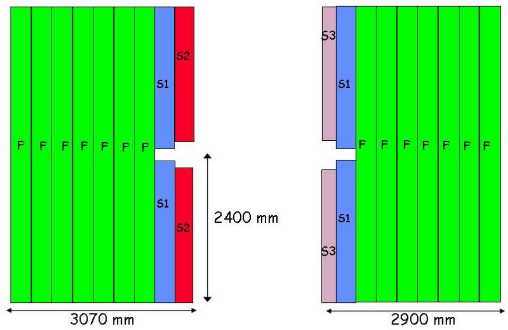


Figure 4.5: *One layer of the Outer Tracker (OT) with the long F modules and the shorter S modules combined out of* 256 *and* 128 *drift tubes respectively.* [55]

4.1.2 The Particle Identification System

The LHCb particle identification system consists of two Ring Imaging CHerenkov (RICH) detectors, the calorimeter system and the muon system. The information of these subdetectors is combined and evaluated with the use of a likelihood function.

4.1.2.1 The RICH1 and RICH2

The LHCb detector includes two RICH detectors [56] to allow for a precise separation of charged pions and kaons over the entire momentum spectrum. When the ultra-relativistic charged particles traverse the gas of the RICH detectors they emit Cherenkov radiation [29]. As can be seen in Figure 4.6, the angle under which the Cherenkov light is emitted depends on the mass and momentum of the charged particle as well as the refractive index of the RICH radiators. Since the particle momentum is established by the tracking system, the particle mass and therefore the particle type can be determined by the RICH detectors.

Both LHCb RICH detectors use a system of primary and secondary mirrors to reflect the Cherenkov photons emitted by the charged particles onto the Hybrid Photon Detectors (HPD) located outside the detector acceptance. This is illustrated in Figure 4.7.

RICH1 is placed between the VELO and the TT and covers the full angular acceptance of the LHCb detector. It uses C_4F_{10} gas to distinguish charged particles with a momentum between 1 GeV/c − 60 GeV/c.

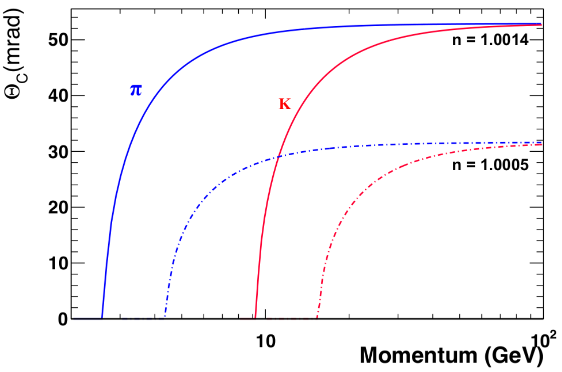


Figure 4.6: *Cherenkov angle versus the particle momentum for pions (blue lines) and kaons (red lines) in the radiators of RICH1 (solid lines) and RICH2 (dashed lines).*

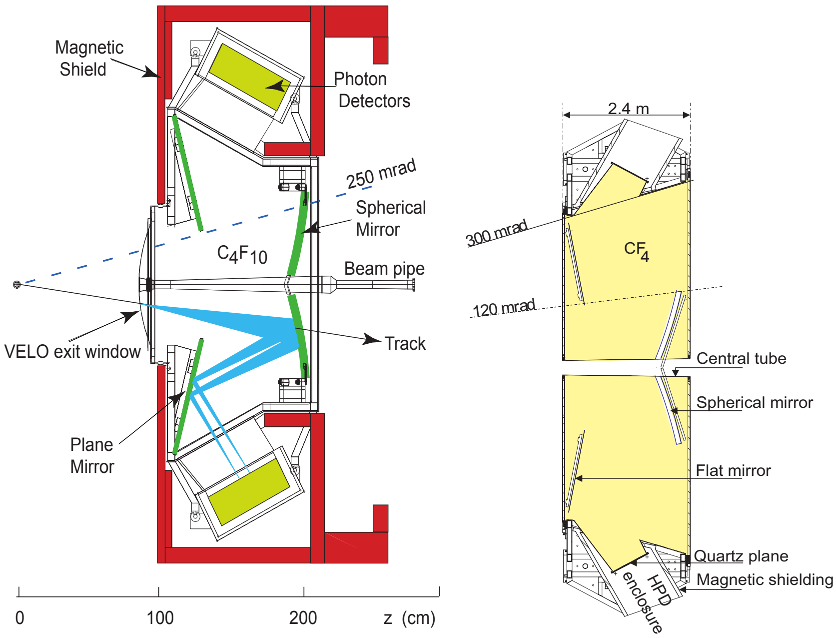


Figure 4.7: *Schematic illustration of the RICH1 (left) and the RICH2 (right) detectors. The Cherenkov photons are reflected from spherical mirrors onto plane mirrors and then onto the photon detector planes which are located outside of the LHCb acceptance.*

RICH2 is located downstream of the magnet behind the tracking stations. It is designed to cover the momentum range from 15 GeV/c to 100 GeV/c using CF_4 radiator gas. Corresponding to the region where high momentum particles are produced, the RICH2 covers an angular acceptance of 15 mrad to 120 mrad (bending plane) and 100 mrad (non-bending plane). More detailed information about the RICH detectors, particularly about the mirror systems, can be found in Chapter 5.

4.1.2.2 The Calorimeter System

The LHCb calorimeter system [57] consists of a Preshower detector (PS), a Scintillating Pad Detector (SPD), an electromagnetic calorimeter (ECAL) and a hadronic calorimeter (HCAL). The main purpose of the calorimeter system is the identification of light particles such as electrons and neutral particles such as photons or π^0 as illustrated in Figure 4.8 [58] and the measurement of energy and position of neutral particles that can't be detected by the tracking system. The calorimeters also provide information for the hardware based L0 trigger (see Section 4.2.1).

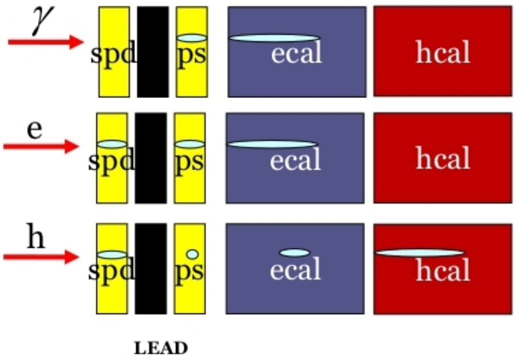


Figure 4.8: *Schematic illustration of the particle separation using the calorimeter system. The lead layer between the Scintillation Pad Detector (SPD) and the PreShower detector (PS) is thick enough to induce electromagnetic showers.* [57]

Each subdetector of the calorimeter system is composed of square cells varying in size according to the particle flux. For the SPD, PS and the ECAL three different zones were chosen with cell sizes of 4 cm, 6 cm and 12 cm as shown in Figure 4.9. The HCAL is segmented into two zones with larger cells (13.13 cm in the inner section and 26.26 cm in the outer section) because of the size of hadronic showers. All calorimeters implement the technology of scintillation light that is transmitted to Photo-Multipliers (PMT) by wavelength-shifting (WLS) fibres.

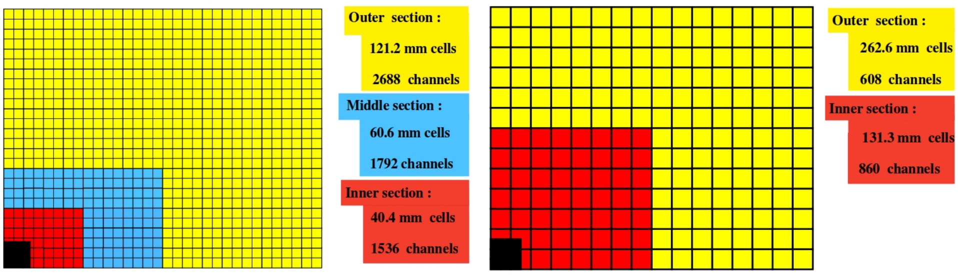


Figure 4.9: *Schematic illustration of the segmentation of the Scintillating Pad Detector (SPD), PreShower detector (PS) and the ECAL (left) and the HCAL (right). The pictures show one quarter of the calorimeter surfaces. The black area on the bottom left of both images is the area close to the beam pipe where the particle flux is too high for any calorimeter performance.* [57]

The PS and the SPD are located before the ECAL and separated by lead sheet of 2.5 X_0 thickness allowing for a separation of electrons from a high background of pions. Both PS and SPD are made of one single layer of 15 mm thick plastic-scintillator tiles.

The ECAL is composed of cells that are built after the shashlik principle, one cell is composed of alternating layers of 2 mm thick lead and 4 mm thick polystyrene scintillator tiles. One cell, made out of 66 lead and scintillator layers, is illustrated in Figure 4.10. The HCAL employs a non-typical structure where the scintillating tiles are arranged parallel to the beam pipe as shown in Figure 4.10. The absorber for the HCAL was chosen to be 1 cm iron tiles and the scintillator layers are made out of 3 mm thick doped polystyrene tiles.

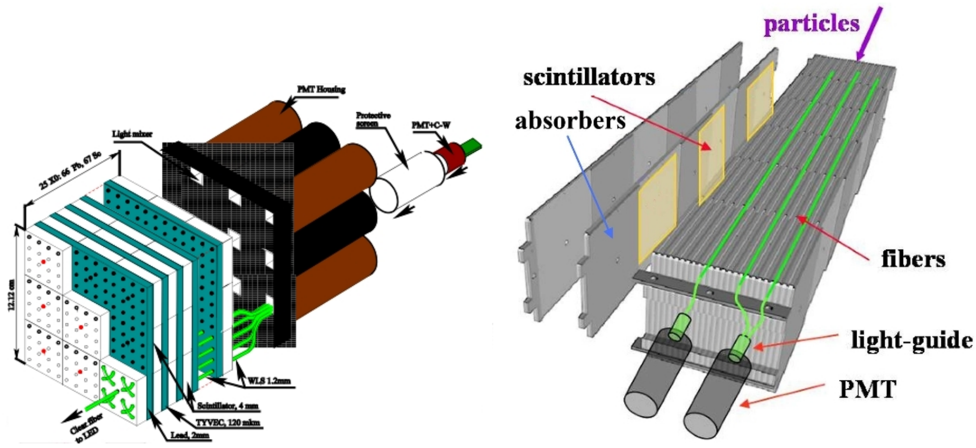


Figure 4.10: *One cell of the ECAL (left) and the HCAL (right). The scintillating tiles of the ECAL are arranged perpendicular to the beam while they are parallel to the beam for the HCAL. The absorber is made of lead for the ECAL and iron for the HCAL.* [57]

4.1.2.3 The Muon Chambers

The LHCb detector contains five muon stations (M1- M5) dedicated to muon identification and muon triggering [59].

The first muon station M1 is placed upstream of the calorimeters to provide a precise transverse momentum measurement. The other muon stations are placed downstream of the calorimeters and are interleaved with 80 cm thick iron absorbers. The stations M2 and M3 yield a very good spacial resolution for the tracks while the last two stations only deliver information for the particle identification. The side view of the muon system is shown in Figure 4.11.

As can be seen on the right in Figure 4.11, each muon chamber is divided into four regions R1 to R4. The granularity in the regions decreases radially which provides an isotropic channel occupancy for each muon station. Except for the inner region R1 of M1, all muon stations are composed of Multi Wire Proportional Chambers (MWPC). R1 of M1 implements the Triple GEM [60] technology because the particle flux in this region would overburden the MWPC.

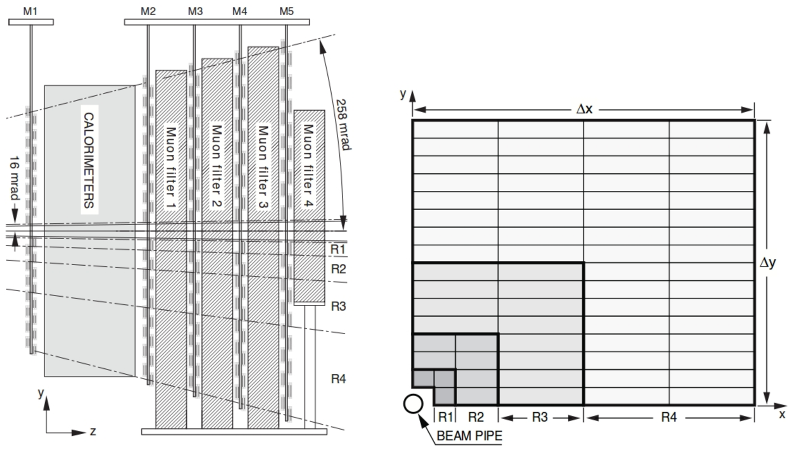


Figure 4.11: *Side view of the muon stations (left) and the schematic front view of the upper right quarter of one muon station with the segmentation into four regions R1 to R4. All modules are composed of MWPC except for R1 of M1 which uses the Triple GEM technology due to extremely high particle flux.* [59]

4.1.2.4 The Discriminant Particle Identification Variables

The information from all the detectors of the particle identification system is combined to calculate two different sets of discriminative variables. These variables are calculated for all final state tracks of an event.

The first set of variables (DLL) are based on a combined log likelihood. It is computed for each subsystem by testing the hypothesis that a track is a certain particle with respect to the hypothesis that the track comes from another particle. The reference particle is usually chosen to be a pion. The likelihoods from the different subdetectors are then summed to obtain the global discriminative variable.
The second sets of variables (ProbNN) are the output of a multivariate data analysis for each particle hypothesis. The multivariate tool used is a multilayer perceptron (MLP)[1] that takes various output of the PID detectors as input as well as information from the tracking system. In contrast to the DLL variables, the ProbNN variables are limited to a range between 0 and 1 and represents the probability for certain hypothesis [61].

Even though the input from all subdetectors is combined, the information on the particle identification variable for kaons and pions comes mainly from the RICH1 and RICH2. For neutral particles and electrons the calorimeter system provides the significant information while the particle identification variable for muons is calculated by relying especially on the muons system.

[1] A multilayer perceptron is a deep neural network that consists of a series of single perceptrons. The linear predictor functions of the single perceptons is used to identify small linearly separable sections of the inputs (see Figure 6.5 for an example of a linear predictor funtion). The combination of layers of single percetrons to a multilayer perceptron allows the solution of nonlinearly separable problems by producing arbitrarily shaped decision regions. Multilayer perceptrons are capable of separating any classes (*Kolmogorov theorem*).

4.1.3 LHCb Performance Summary

The LHCb detector has shown a high and stable performance over the entire range of data taking [62].

The resolution of the x and y coordinates of the primary vertex and the resolution on the impact parameter's x coordinate are shown in Figure 4.12. For 25 tracks in the event, a resolution of 13 µm on the x and y coordinates and a resolution of 69 µm on the z coordinate of the primary vertex can be reached [63]. Additionally an equally excellent impact parameter resolution of < 35 µm for tracks with a transverse momentum $p_T > 1$ GeV can be obtained. The tracking system has an average relative momentum resolution of

$$\frac{\Delta p}{p} = 0.5 \tag{4.2}$$

The relative energy resolution for the ECAL is

$$\frac{\Delta E}{E} = \frac{10\%}{\sqrt{E[\,\mathrm{GeV}/c]}} \oplus 1\% \tag{4.3}$$

and for the HCAL

$$\frac{\Delta E}{E} = \frac{80\%}{\sqrt{E[\,\mathrm{GeV}/c]}} \oplus 10\% \tag{4.4}$$

Using $Bs \to J/\psi\,\phi$ decays, the decay time resolution of LHCb has been determined to be about 50 fs [64]. As can be seen in Figure 4.12 the decay time resolution is essentially independent of the B momentum.

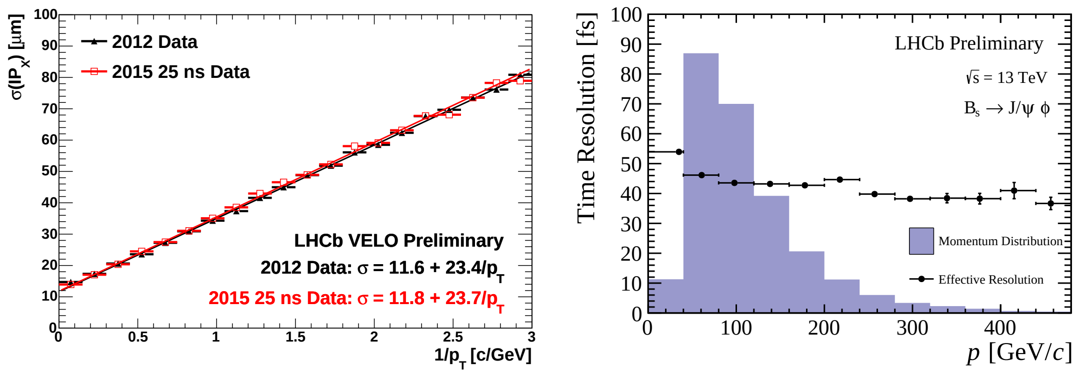


Figure 4.12: *Summary of the LHCb VELO performance obtained from 2015 data. The resolution of the x coordinate of the impact parameter as a function of the transverse momentum p_T (left) and the decay time resolution as a function of the B momentum p (right).*

4.2 The LHCb trigger system

The LHC collides bunches at a maximal rate of 40 MHz. Due to the LHCb's luminosity levelling the visible[2] rate of interaction is about 25 MHz which has to be

[2]To be visible, events must have at least two charged particles producing enough hits in the tracking system to be reconstructed.

reduced to about 12.5 kHz to be permanently stored for offline analysis. At LHCb this is realised by a two-stage trigger system [65] [66]: the Level-0 (L0), a purely hardware based trigger, followed by the High Level Trigger (HLT) which is executed on a CPU farm.

Since the beginning of Run II, LHCb has implemented a stage of real-time alignment and calibration of the subdetectors within the HLT. This allows for a new concept - called Turbo stream - where physics measurements are performed on the output of the HLT. A schematic presentation of the trigger system for Run II can be seen in in Figure 4.13.

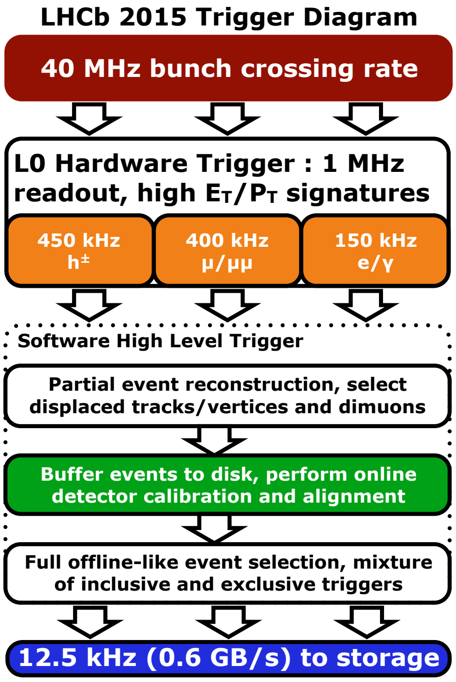


Figure 4.13: *A schematic representation of the LHCb trigger system. The first state is the hardware based Level-0 trigger. Afterwards the High Level Triggers are executed on a processor farm. Each subtrigger uses several trigger lines for dynamic and specific triggering.*

4.2.1 The Level-0 Trigger

The LHCb Level-0 (L0) trigger is made from custom electronics and reduces the rate from 25 MHz to 1 MHz. To identify events from B hadrons the L0 trigger uses the fact that the large mass of B hadrons provides a significant amount of transverse kinetic energy to the decay particles. Therefore it selects events with a high amount of transverse energy deposited in the calorimeter or the muon system. Additionally, the L0 accesses pileup information from the VELO and the SPD to reject events with too many tracks. There are different L0 trigger lines for different final state particles.

The most important L0 line for the analysis is the **L0 Hadron**. The L0 Hadron trigger line is activated by one cluster in the HCAL. The threshold transverse energy

E_T for the L0 Hadron is the sum of the transverse energy in the HCAL with the transverse energy in the corresponding ECAL cells.

4.2.2 The High Level Trigger

All events that pass the L0 trigger are processed by the High Level Trigger (HLT). Since the beginning of Run II, the HLT consist of three stages. The first stage is the HLT1 stage after which the accepted events are buffered to disk. The HLT1 performs a partial event reconstruction by reconstructing tracks, primary vertices and potential secondary vertices. The second stage is the full alignment and calibration of the LHCb detector which is performed on a subset of the buffered events. The third stage is the HLT2 which performs a full reconstruction of the buffered events using the newly determined alignment and calibration constants. Due to an upgrade of the CPU farm for the HLT and the increased performance of the reconstruction algorithms, the reconstruction of the events in HLT2 is identical to the offline reconstruction. By performing the alignment and calibration of the detector before HLT2 it is ensured that there are no more differences between the *online* and offline reconstruction.

The HLT1 has access to information from the VELO, the tracking system and the muon stations as well as the information available to the L0. It reconstructs particle tracks in the VELO and determines the position of the primary vertices in each event. The HLT1 then makes a decision based upon impact parameter, momentum, transverse momentum and/or track quality of one or more tracks in the event. For each L0 line several HLT1 lines are executed. The HLT1 reduces the rate to approximately 1 MHz.

After HLT1 the events are buffered to disk while the calibrations and alignments are performed. The calibrations are the RICH refractive index calibration, the RICH hybrid photon detector image calibration, the OT time calibration and the ECAL calibration. These can be performed as simple fits on monitoring histograms. The alignments are the spacial alignment of the VELO modules, the tracking stations, the RICH mirror systems and the muon chambers. The alignments require more complex algorithms that run on the same CPU farm as the HLT. The first three calibrations are run every hour during data taking while the ECAL calibration and the alignments are performed every couple of hours. The newly determined calibration and alignment constants are applied in the subsequent HLT2 processing of the events.

The HLT2 performs a full event reconstruction identical to the offline analysis. Starting by reconstructing all tracks in the event using the VELO tracks as seeds, the HLT2 reconstructs intermediate particles and resonances and identifies displaced vertices. Afterwards, depending on the HLT2 line, different sets of selections are applied which are dominantly designed to identify decays of B and D hadrons. The HLT2 finally reduces the rate to about 12.5 kHz which is either stored for the offline analysis or directly used for physics analysis in the Turbo Stream. The use of real-time alignment and calibration of the detector also enables the HLT2 to use the particle identification variables and to perform at optimal efficiency at all times.

Two HLT1 line and several HLT2 lines will be used in the analysis of the $B^{\pm} \to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ decays. The HLT1 lines are the *Hlt1TrackMVA* line and the *Hlt1TwoTrackMVA* [67]. The former line selects events with at least one track with a high transverse momentum p_T and a great impact parameter (IP) with respect to the primary vertex, since this track is likely to come from a decaying B hadron. The latter line searches for two tracks that form a secondary vertex that is displaced with respect to the primary vertex.

The HLT2 lines that are used are the *HLT2 topological lines* [67]. They were designed to trigger inclusive n-body ($n = 2, 3, 4$) B decays. The HLT2 lines implement a Boosted Decision Tree (see Section 6.5.2) to determine if n tracks show the topology of a B decay. Missed tracks are compensated for by taking into account the difference between the total momentum of the n tracks and the momentum of the hypothetical B hadron. More information about the HLT lines used in the analysis of the $B^{\pm} \to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ decay are given in Section 6.4.1 of Chapter 6.

4.2.3 The Turbo Stream

Since HLT2 performs a reconstruction that is identical to the offline reconstruction, the output of HLT2 can be directly used for physics analysis. This new concept is called the Turbo stream [68]. Currently about 30% of events go through the Turbo stream in Run II.

The Turbo stream consists of a collection of Turbo lines which are equivalent to HLT2 lines. While after the HLT2 decision its reconstruction is discarded, all physics objects (such as tracks, primary vertices etc.) reconstructed in HLT2 are saved for the Turbo lines. For all Turbo candidates the raw detector information is discarded and only the objects reconstructed by HLT2 are kept, halving the size of the event.

The events in the Turbo stream are then processed by the software TESLA (see Section 4.4) which extracts the objects reconstructed in HLT2 and converts them for physics analysis. Performing analyses on the output of the online reconstruction is only possible since there is a real-time alignment and calibration of the LHCb detector.

The Turbo stream was designed for use during Run III (starting in 2020). In Run III the center-of-mass energy of the proton proton collisions will be increased to 14 TeV while the instantaneous luminosity seen at LHCb will be increased by a factor of five. In order to maintain signal efficiencies close to those obtained during Run I the hardware trigger stage has to be removed and higher output rates are necessary. This will be archived by using only the Turbo stream and dispensing with an offline reconstruction. Run II is used as commissioning period for the Turbo stream.

4.3 The LHCb Preselection (Stripping)

In order to reduce computing time, all events are preselected by a certain *stripping line*. Each analysis performed at LHCb has its own dedicated stripping line. Most stripping lines aim at identifying a specific decay and consist of a loose selection for rejecting as much background as possible while keeping as much signal as possible. The stripping is executed by the physics analysis software DaVinci (see Section 4.4).

During the stripping procedure every event is scanned for particle candidates that can be combined to make the signal decay. Events with signal candidates are stored for later use.

The stripping line used in the analysis of the $B^{\pm} \to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ decay is the `B2D0KD2HHHH` line that selects $B^{\pm} \to DK^{\pm}$ decays with the D meson decaying to a hadronic four-body final state (see Section 6.4.2).

4.4 The LHCb Software

The data processing at LHCb goes through a series of steps which are executed by different software frameworks. To ensure consistency between these frameworks and the way that data and Monte Carlo are treated, all LHCb core software is embedded in the Gaudi framework [69]. There are six main applications, each responsible for a different stage of event processing:

- **Gauss** is used for the generation of Monte Carlo. Therein, proton-proton collisions are simulated with Pythia [50], the decay of B hadrons is made with EvtGen [37] and the detector simulation is implemented in Geant4 [38].

- **Boole** takes the output from Gauss and simulates the digitisation of data to give it the same format as the LHCb data obtained by the electronics and data acquisition systems.

- **Moore** performs online subdetector and global reconstruction in HLT1 and HLT2 during data taking. For Run II the reconstruction in Moore is identical to the offline reconstruction performed in Brunel.

- **Brunel** performs offline subdetector and global reconstruction using pattern recognition for both Monte Carlo and data.

- **TESLA** processes the events in the Turbo stream by extracting the physics objects constructed by Moore in HLT 2.

- **DaVinci** executes the last step, namely the reconstruction of the final signal events and running the stripping [70].

5. Implementation and Commissioning of the Real-time Alignment of the LHCb RICH Mirror Systems

The alignment procedure of the LHCb RICH mirror systems in Run II of the LHC is performed in real-time. This chapter describes the alignment procedure of the RICH mirrors in Run II using proton-proton collision data, in particular the implementation of the alignment procedure into the online framework and the commissioning and improvement of the alignment method during the first two years of Run II.

The first section of this chapter is an overview of the importance of particle identification and the role of the RICH mirror alignment. The second section provides an introduction to Cherenkov radiation and its use in obtaining the particle identification information in LHCb. The third section is a presentation of the LHCb RICH optical systems. The fourth section explains the irreducible limitations to the precision of the particle identification. The fifth section describes the method used to align the RICH mirror systems. The sixth section outlines the alignment strategy in Run I of the LHC while the seventh section gives a in-depth explanation of the alignment strategy and its implementation in Run II. The eighth section summarises the results of several studies aimed at improving the RICH mirror alignment procedure. The ninth section presents the first alignment performed in 2016 and the tenth section summarises the information collected during the automatic running of the alignment procedure throughout 2016. The last section gives the conclusion and provides an outlook on the future of the alignment for the LHCb upgrade.

The author implemented the RICH mirror alignment in the online system and commissioned the alignment throughout 2015 and 2016. The implementation of the RICH mirror alignment in the online system is described in Section 5.7.4. The author also facilitated the monitoring of the procedure as described in Section 5.7.5 and performed the studies presented in Sections 5.8, 5.9 and 5.10.

5.1 Introduction

The primary role of the RICH system is the identification of charged hadrons such as pions, kaons and protons. One of the major requirements for charged hadron identification in flavour-physics experiments is for the exact reconstruction of the invariant mass of a decaying particle – such as $B^0 \to \pi^+\pi^-$ — to reduce combinatorial background. The most precise invariant mass can only be obtained if the masses of all final state particles — and therefore their species — is known. Another crucial use of particle identification is the discrimination between final states which are topologically identical but have different final state particles, such as $B^{\pm} \rightarrow DK^{\pm}$ and $B^{\pm} \rightarrow D\pi^{\pm}$. These decays have almost identical kinematics but very different CP asymmetries. Extracting a signal using only kinematic and vertex-related selection criteria would mean summing over all $B^{\pm} \rightarrow Dh^{\pm}$ decay modes and diluting CP violation effects. It is therefore necessary to identify all final state particles. An example of the impact of the particle identification information is shown in Figure 5.3 and in Section 6.5.5 of Chapter 6.

The LHCb RICH system consists of two RICH detectors which together cover the particle identification in the momentum range from 2 to 100 GeV/c. Both RICH detectors have two sets of mirrors, amounting to a total of 20 mirrors in RICH1 and 96 mirrors in RICH2 (see Section 5.3). The mirrors reflect the Cherenkov photons emitted by charged particles travelling through the RICH radiators onto the detector plane. Consequently, the particle identification requires the accurate knowledge of the position and orientation of the RICH mirrors, more precisely their position and orientation with respect to the LHCb tracking system. The orientation of the mirrors is also called *the alignment* of the mirrors.

The RICH mirror alignment procedure determines the difference between the position of each the mirror in the *LHCb conditions database* and its actual position. The LHCb conditions database stores the non-event time-varying data pertaining to detector conditions including the orientation of all sub-detectors. The difference between the orientation of a mirror in the conditions database and its actual position is called *the misalignment* of the mirror. The orientations in the conditions database can then be corrected for any misalignments and the LHCb conditions database updated. The alignment procedure corrects for rotations around the mirrors individual y- and z-axes (see Section 5.3).

The LHCb detector performed well during Run I [62] [64] but the data-taking conditions changed significantly for Run II. Firstly, the centre-of-mass energy of the LHC was increased from 8 TeV to 13 TeV with a reduced bunch spacing of 25 ns instead of 50 ns. This means an increase in instantaneous luminosity. In order to keep the selection efficiencies in the high level trigger as high as in Run I or even higher, requirements on the particle identification have to be used in the high level trigger in Run II. This requires the full alignment and calibration of the RICH detectors within the high level trigger sequence (see Section 4.2.2). Secondly, the Turbo stream was implemented, which allows the direct use of the output of HLT2 for physics analyses (see Section 4.2.3). Hence, the alignment and calibration of the detector has to be performed before HLT2. An additional advantage of the real-time alignment and calibration of the detector is that there are no more differences between the reconstruction during the high level trigger stage and the offline reconstruction. This enables much more precise understanding of the reconstruction efficiencies.

In Run II the full LHCb detector is aligned and calibrated either each *run* or each *fill*[1]. The RICH refractive index calibration, the calibration of the RICH hybrid

[1]A *run* takes usually an hour while a *fill* takes on average about 12 hours.

photon detectors and the global time calibration of the outer tracker are performed per run while the spacial alignment of the full tracking system, the alignment of the RICH mirror systems and the calibration of the calorimeter is performed once per fill.

5.2 Cherenkov Radiation and Particle Identification at LHCb

When a charged particle traverses a medium at a speed higher than the phase velocity of light in that medium, the particle emits photons. This radiation is called *Cherenkov radiation* and is emitted at a polar angle θ_C - called the *Cherenkov angle* - with respect to the particle's direction and uniformly in azimuthal angle ϕ. For a charged particle with momentum p and mass m travelling through a medium with refractive index n the Cherenkov angle is

$$\cos\theta_C = \frac{1}{n\beta} \tag{5.1}$$

$$= \frac{1}{n}\frac{\sqrt{m^2+p^2}}{p}. \tag{5.2}$$

The curves of the Cherenkov angle over the particle momentum can be seen in Figure 5.1 for protons, pions and kaons in the radiators of RICH1 and RICH2. The graph shows that for higher momentum the Cherenkov angles converge to the same value. The Cherenkov angle is said to be *saturated*. The momentum at which the Cherenkov angle saturates and the value of the saturated angle depend on the refractive index of the material.

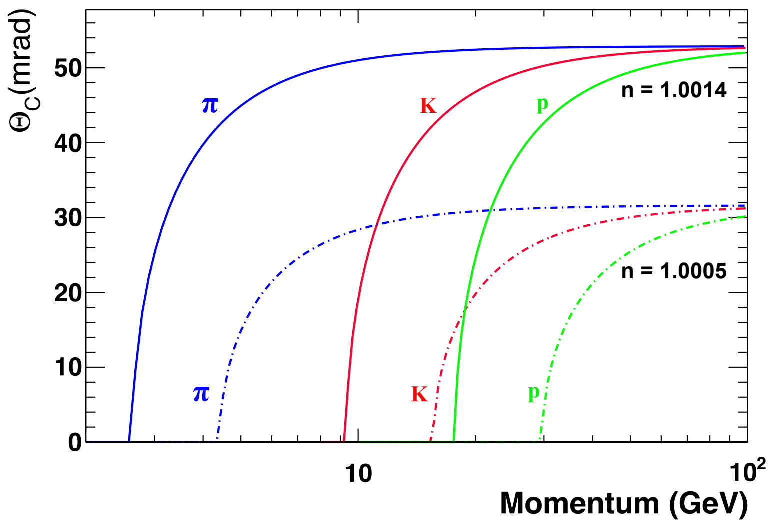


Figure 5.1: *Cherenkov angle as a function of particle momentum for pions, kaons and protons in the C_4F_{10} radiator (n = 1.0014) of RICH1 and the CF_4 radiator (n = 1.0005) of RICH2.*

Figure 5.2 shows two examples of the images on the HPD planes for two selected events during nominal LHCb run conditions. While in one example the Cherenkov rings from different tracks are distinguishable, these events are very rare. Typical events have high occupancy and several overlapping Cherenkov cones, making it difficult to distinguish individual Cherenkov rings. Therefore, particle identification at LHCb using the RICH systems is performed using an overall event log-likelihood minimisation algorithm, where all tracks in the event and in both RICH detectors are considered simultaneously [71] [72]. The process can be described in two steps. The first step is called *ray tracing* and consists of taking a photon hit on the detector plane and matching it to a charged track under the assumption that the hit came from a Cherenkov photon emitted by the track. Since the emission point of the Cherenkov photon cannot be known it is taken to be in the middle of the track in the respective RICH detector volume. From this the Cherenkov angle is computed. Each track is matched to all hits on the detector plane that lie within an expected region around the track. This region is defined in terms of the Cherenkov angle resolution σ_θ.

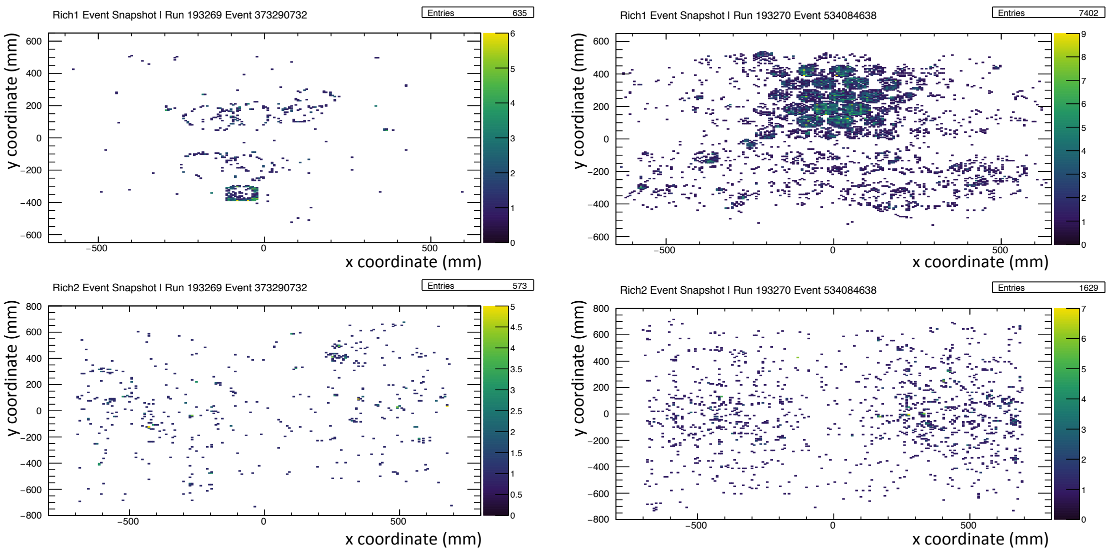


Figure 5.2: *Examples of the images on the HPD planes of RICH1 (top) and RICH2 (bottom) for two nominal LHCb events. The left figures show an event with distinguishable, individual Cherenkov rings in RICH1. The right figure shows a typical events with high occupancy and indistinguishable Cherenkov rings.*

The second step of the procedure is the construction of the likelihood function. The probability distribution for Cherenkov photons from tracks (i.e. the signal) is taken to be Gaussian in θ and uniform in ϕ and given by

$$f_{h_j}(\theta, \phi) = \frac{1}{(2\pi)^{3/2}\sigma_\theta} e^{-\frac{(\theta - \theta_C(h_j))^2}{2\sigma_\theta^2}}, \tag{5.3}$$

where $\theta_C(h_j)$ is the expected Cherenkov angle for the track under the mass hypothesis h_j. The background contribution to the pattern observed on the detector plane is also modelled in the likelihood function. It is estimated using a data-driven technique where the observed signal in each HPD is compared to the expected signal in that HPD that would be caused by only the tracks in the event. Additionally, detector inefficiencies are taken into account.

The log likelihood function is calculated for different particle hypotheses for each track for the entire event. Since pions are the most common particles in proton-proton collisions, the likelihood minimisation procedure starts by assuming all charged particles are pions. The overall event likelihood, computed from the observed distribution of photon hits on the detector plane, the associated tracks and their errors, is then calculated for this set of hypotheses. Then, for each track in turn, the likelihood is recomputed changing the mass hypothesis to $e^{\pm}$, $\mu^{\pm}$, $\pi^{\pm}$, $K^{\pm}$ and proton, whilst leaving all other hypotheses unchanged. The change in mass hypothesis amongst all tracks that gives the largest increase in the event likelihood is identified, and the mass hypothesis for that track is set to its preferred value. This procedure is repeated until all tracks have been set to their optimal hypotheses. To lower the CPU usage of this algorithm, tracks are sorted according to the size of their likelihood change from the previous step, and the search starts with the track most likely to change its hypothesis. If the improvement in the likelihood for the first track is above a given threshold, this iteration is stopped and the hypothesis for that track is immediately changed. Additionally, if a track shows a clear preference for one mass hypothesis, it is set to that hypothesis and is removed from the minimisation procedure.

The final results of the log-likelihood minimisation algorithm are differences in the log-likelihood values (DLL) for each track individually. The DLL describe the difference in the overall event log-likelihood between the pion hypothesis and each of the $e^{\pm}$, $\mu^{\pm}$, $K^{\pm}$ and proton hypotheses. These values can then be used to apply selection criteria on particle types. An example of the effect of the particle identification information can be seen in Figure 5.3 for invariant mass distribution in $B^0 \to \pi^+\pi^-$.

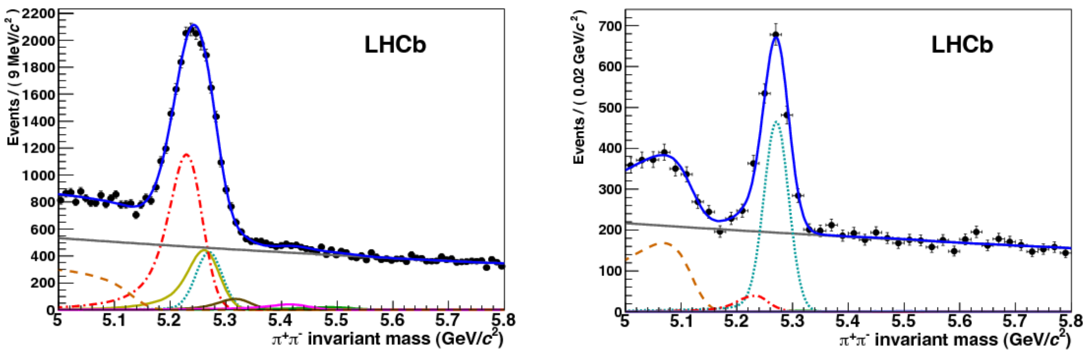


Figure 5.3: *Demonstration of the effect of the particle identification on the invariant mass distribution of $B^0 \to \pi^+\pi^-$. [73]* Left: *Distribution before the use of the particle identification information. The signal decay $B^0 \to \pi^+\pi^-$ (turquoise line) has a relatively small contribution. Contributions from different b-hadron decay modes, such as $B^0 \to K^+\pi^-$ decays (red line), $B^0 \to$ 3-body decays (orange line), $B_s^0 \to K^+K^-$ decays (yellow line), $B_s^0 \to K^+\pi^-$ decays (brown line), $\lambda_B \to pK$ decays (purple line) and $\lambda_B \to p\pi$ decays (green line), are clearly visible.* Right:*Distribution after the use of the particle identification information. Most contribution from other decay modes have been eliminated. The remaining two background contributions are much suppressed with respect to the left plot.*

The exact knowledge of the physical position of all components of the RICH optical systems, i.e. the alignment of the RICH optical system, is of crucial importance at several different points of the particle identification procedure. The ray tracing algorithm relies on a precise description of the RICH systems in the LHCb software. Furthermore, an incorrect alignment of the optical systems results in a greater Cherenkov angle resolution σ_θ which has a direct impact on the probability distribution function used in the log likelihood function. Additionally, the Cherenkov angle resolution defines the regions on the detector plane from which hits are matched to a given track. A smaller Cherenkov angle resolution would result in a smaller region and therefore fewer hits. This would reduce the amount of background as well as the time needed to run the particle identification algorithm.

5.3 The RICH Optical Systems

The designs of the RICH optical systems are shown in Figure 5.4 [56, 74]. Both RICH1 and RICH2 have similar optical systems with a set of focusing spherical *primary* mirrors, and the *secondary*, much flatter mirrors. Each optical system is divided into two halves, with RICH1 being divided vertically above and below the beam pipe and RICH2 horizontally on either side of the beam pipe. The four primary mirrors of RICH1 are made from carbon-fibre while its 16 secondary mirrors and the 56 primary mirrors of RICH2 and its 40 secondary mirrors employ a thin glass substrate. All mirrors are lined with an aluminium magnesium fluoride ($Al + MgF_2$) coating which was chosen for the mirrors to obtain a good reflectivity in the UV range.

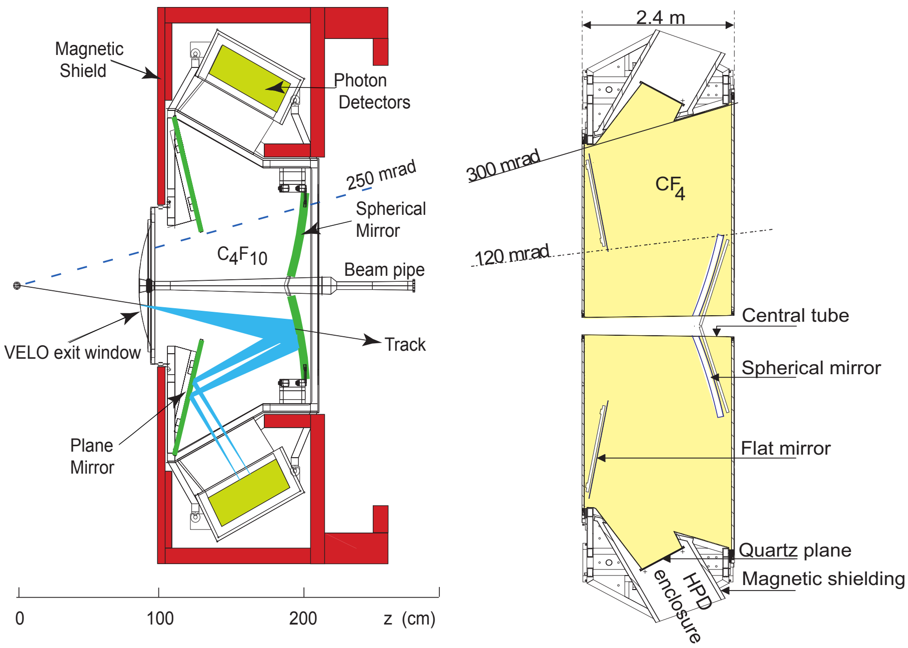


Figure 5.4: *Schematic representation of the LHCb RICH detectors in Run II.* Left: *Side view of the RICH1 detector. The formation of a Cherenkov ring and its propagation through the detector is illustrated in blue.* Right: *Top view of the RICH2 detector.*

Cherenkov photons emitted by a charged track are reflected off a primary mirror onto a secondary mirror, and from there out of the LHCb acceptance onto the detector plane. The detector plane consists of hybrid photon detectors (HPDs) and coincides with the focal plane of the given part of the optical systems. The HPDs are vacuum tubes with a 75 mm active diameter. Photoelectrons generated in the photo cathode are focused onto silicon pixel arrays. The pixel arrays consist of 23×23 pixel of size 2.5×2.5 mm^2. A total of 196 HPDs are tightly packed to cover both detector planes in RICH1 and 288 HPDs cover the detector planes in RICH2.

Each mirror has its own coordinate system in which potential misalignments are described as rotations around the y- and z-axes. The right-handed coordinate system of each mirror is defined by placing the origin at the centre of its curvature with the x-axis pointing towards the mirror, the y-axis pointing upwards and the corresponding z-axis being horizontal. Figure 5.5 shows the arrangement and the numbering of the primary and secondary mirrors of RICH1. Each secondary mirror only receives photons from one primary mirror, making each primary mirror and its secondary mirrors an independent system. Figure 5.6 shows the arrangement and numbering of the primary and secondary mirrors of RICH2. Both halves of the system are independent of each other. The primary and secondary mirrors of each half are interconnected.

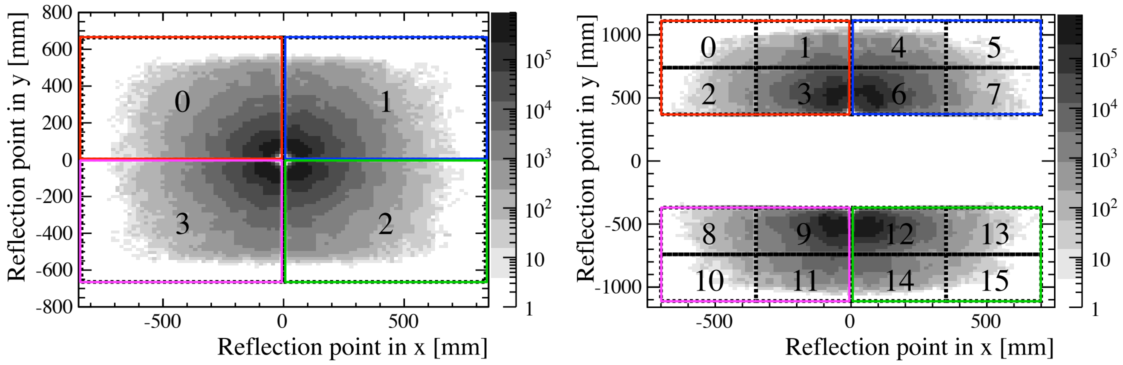


Figure 5.5: *Arrangement and numbering of the 4 primary (left) and 16 secondary (right) mirrors of RICH1. Each secondary mirror only receives photons from one primary mirror, as indicated by the colours.*

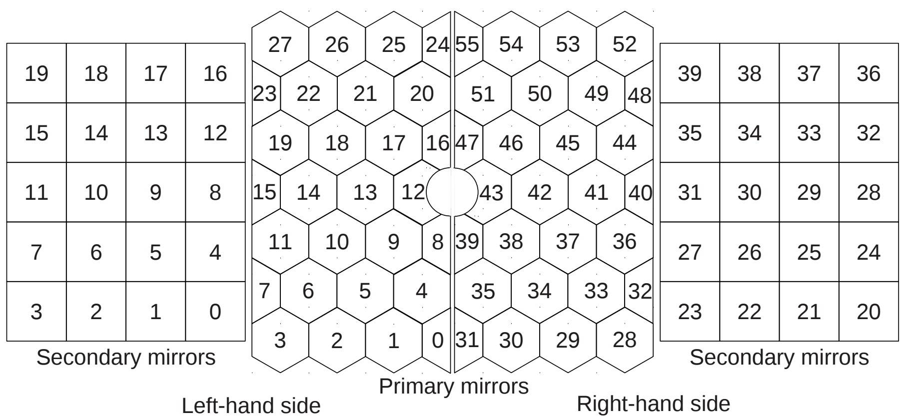


Figure 5.6: *Arrangement and numbering of the 56 primary mirrors (inner region) and 40 secondary mirrors (left and right) of RICH2.* [75]

5.4 Limitations on the Precision of the Single Photon Cherenkov Angle Resolution

The performance of the RICH detector can be quantified in terms of the Cherenkov angle resolution. The resolution is limited by four irreducible sources of uncertainty:

1. **Emission point**
   The emission point uncertainty of the Cherenkov angle comes from the fact that the real point of emission of the Cherenkov photon along the charged particle track is unknown. The particle identification algorithm assumes the emission point to be at the center of the charged particle track in the RICH detector volume.

2. **Chromatic dispersion**
   The Cherenkov angle depends on the refractive index n of the medium, as shown in Equation 5.1. The refractive index depends on the wavelength of the medium and therefore so does the Cherenkov angle. A charged particle with momentum p can emit Cherenkov photons under different Cherenkov angles.

3. **Pixel size** The pixels of the hybrid photon-detectors have a finite size which adds an uncertainty to the measured Cherenkov angle.

4. **Tracking** The uncertainty on the direction of the charged particle track determined by the tracking system contributes to the uncertainty on the Cherenkov angle.

The size of each uncertainty is presented in Table 5.1. Adding the uncertainties in quadrature gives the minimal total uncertainty, which can theoretically be obtained by having an optimal alignment of all optical components of the RICH detectors. The minimal single photon Cherenkov angle resolution for RICH1 is about 1.6 mrad and for RICH2 0.7 mrad. The alignment procedure aims at reaching this resolution.

	σ_θ [mrad]	
	RICH1	RICH2
Emission point	0.9	0.2
Chromatic dispersion	0.9	0.5
Pixel size	0.6	0.2
Tracking	0.4	0.4
Total	1.6	0.7

Table 5.1: *Sources of uncertainty on the measurement of a single photon Cherenkov angle for the LHCb RICH detectors. The total uncertainty is the minimal uncertainty that can be achieved with an optimal alignment of all optical components of the RICH detectors [75]. Due to the removal of the aerogel in RICH1 the emission point error in Run II is slightly larger than in Run I.*

There is an additional uncertainty for RICH1 which has not yet been quantified. The field of the LHCb magnet reaches into the RICH1 detector and distorts the image in the HPDs by deflecting photoelectrons on their way from the photocathode to the silicon anode. A system is devised to reduce this effect – called the *magnetic distortion correction system* – which measures the resulting distortion of the images. The correction is applied when reconstructing the Cherenkov angle resolution. The residual uncertainty of the distortion of the image adds to the overall Cherenkov angle uncertainty of RICH1.

5.5 Alignment Method for the RICH Mirror Systems

This section provides a description of the method used to align the RICH mirror systems. First, the identification of a misalignment of the mirrors is described. Second, an overview of the entire alignment procedure is given. Subsequently, two intricate parts of the procedure are explained in detail.

5.5.1 Identification of a Misalignment in the RICH Mirror Systems

In order to align all mirrors of the RICH mirror system, misalignments in the system have to be identified and corrected for. The misalignment of a mirror is defined as the difference between the position of the mirror as described in the LHCb software and the real position of that mirror.

The effect of misalignments in the mirror system can be explained by considering the image on the detector plane. This principle is illustrated in Figure 5.7. For a charged track with a sufficient number of Cherenkov photons, a Cherenkov ring is formed which is visible as a circle on the detector plane. In the case of perfect alignment, the projection point of the charged particle track onto the detector plane is right in the center of the corresponding Cherenkov ring. However, if a misalignment is present for one or several mirrors, the projected track position on the detector plane is shifted with respect to the center of the circle. This shift is determined by analysing the difference $\delta\theta$ between the measured Cherenkov angle θ and the expected angle θ_C

$$\delta\theta(\phi) = \theta(\phi) - \theta_C \tag{5.4}$$

as a function of the azimuthal angle ϕ. For a well aligned detector the Cherenkov angle is independent of the azimuthal angle, whereas a sinusoidal dependence occurs for a misaligned detector. The expression for $\delta\theta$ is then

$$\delta\theta(\phi) = \Theta^y \sin\phi + \Theta^z \cos\phi \tag{5.5}$$

where Θ^y and Θ^z are called the *misalignments on the detector plane*. An example of the $\delta\theta$ *vs.* ϕ distributions of a misaligned detector and an aligned detector can be seen in Figure 5.8. The projection of the two dimensional histogram onto the y axis is shown in Figure 5.9. The distributions of $\delta\theta$ are fitted with a Gaussian function for the signal component and a second order polynomial for the background component. The width of the Cherenkov angle distribution with the misaligned mirror combination is greater than the width for the aligned combination. Additionally, the mean of the Gaussian for the misaligned mirrors is shifted by -0.65 mrad with respect to zero. Both these factors will negatively impact the performance of the particle identification.

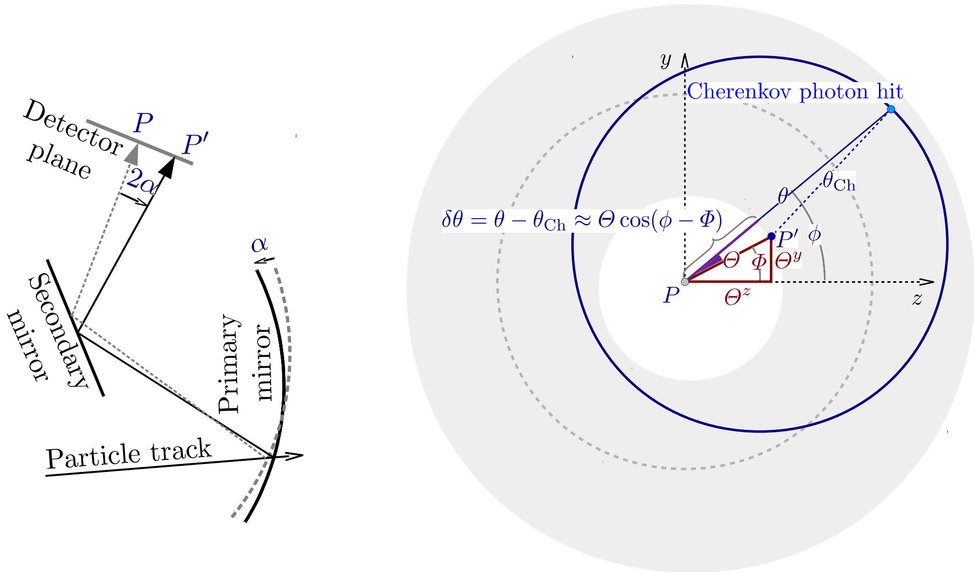


Figure 5.7: *Schematic illustration of the effect of a rotational misalignment of a RICH mirror on the image observed on the detector plane. Left: A rotational misalignment of a RICH mirror causes a shift of the projected track position on the detector plane from P' to P. Right: The Cherenkov ring on the detector plane is illustrated as a blue circle. Due to the misalignment, the projected track position on the detector plane is shifted from the center of the Cherenkov ring at P' to P. The Cherenkov angle θ shows a sinusoidal dependency of the azimuthal angle ϕ.* [75]

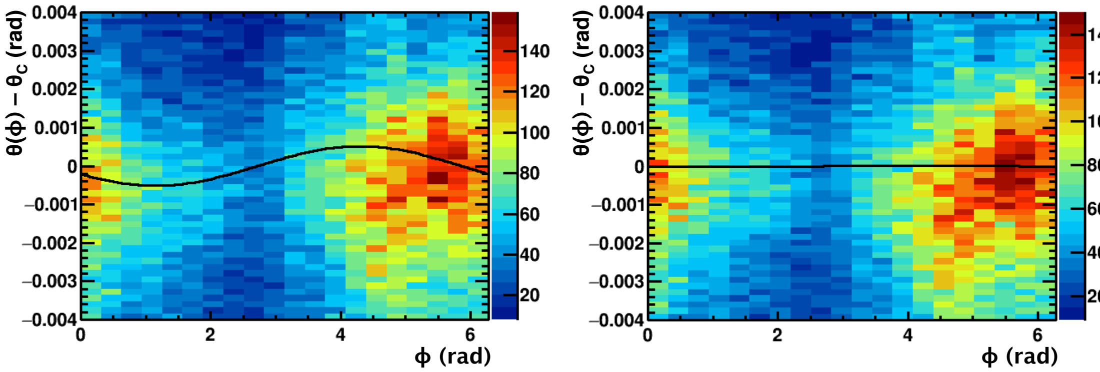


Figure 5.8: *Distribution of $\delta\theta$ vs. ϕ for the combination of primary mirror 22 and secondary mirror 14 of RICH2 with a misalignment (left) and after correction of the misalignment (right). The black line represents the sinusoidal function given in Equation 5.5 fitted to the histograms. The full fitting procedure is described in Section 5.5.3.*

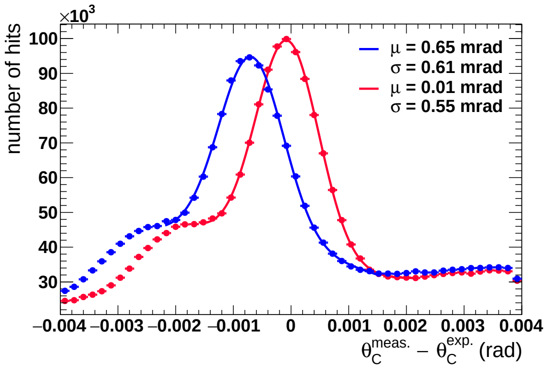


Figure 5.9: *Distribution of $\delta\theta$ for the misaligned mirror combination (blue) and the aligned mirror combination(red). The distributions are fitted with a Gaussian function for the signal component and a second order polynomial for the background component. The signal component of the misaligned mirror combination is shifted with respect to zero and has a greater widths.*

5.5.2 The RICH Mirror Alignment Procedure

The alignment procedure is iterative and completely data-driven. An flow chart of the procedure is shown in Figure 5.10 and an overview is given below. A detailed description of the steps for the alignment is given later in Section 5.5.3 and Section 5.5.4 of this chapter.

1. The alignment starts with a sample of preselected events and a given database entry, usually the latest version of the LHCb conditions database.

2. The events are reconstructed with the current database entry and tracks with a momentum above 20 GeV for RICH1 and 40 GeV for RICH2 are selected. These tracks have a momentum high enough for the Cherenkov angle to be saturated and therefore not to depend on the particle species.

3. For each possible combination of primary and secondary mirror a $\delta\theta$ vs. ϕ histogram (introduced in Section 5.5.1) is created and filled with photon candidates that can be associated to the high momentum tracks from 2. A sinusoidal function is fitted to each histogram to extract the misalignments in the detector plane. The filling and the fitting of the $\delta\theta$ vs. ϕ histograms is described in further detail in Section 5.5.3.

4. The misalignments on the detector plane are disentangled to find the misalignments of the individual primary and secondary mirrors. The image on the detector plane, and therefore the $\delta\theta$ vs. ϕ histograms, are created by a combination of a primary mirror with a secondary mirror. Obtaining the individual primary mirror and secondary mirror misalignments requires the solution of a system of linear equations.

5. A new database entry is produced containing the new mirror orientations. The convergence criterion — which is formulated in terms of the maximal rotation of any mirror — is verified.

(a) If the convergence criterion is fulfilled the alignment procedure has converged and the LHCb conditions database can be updated with the newly produced database entry.

(b) If the convergence criterion is not met, another iteration of the alignment procedure is set up. The alignment parameters calculated in the previous iteration are used in the reconstruction of events for this iteration. The alignment procedure iterates until the convergence criterion is met.

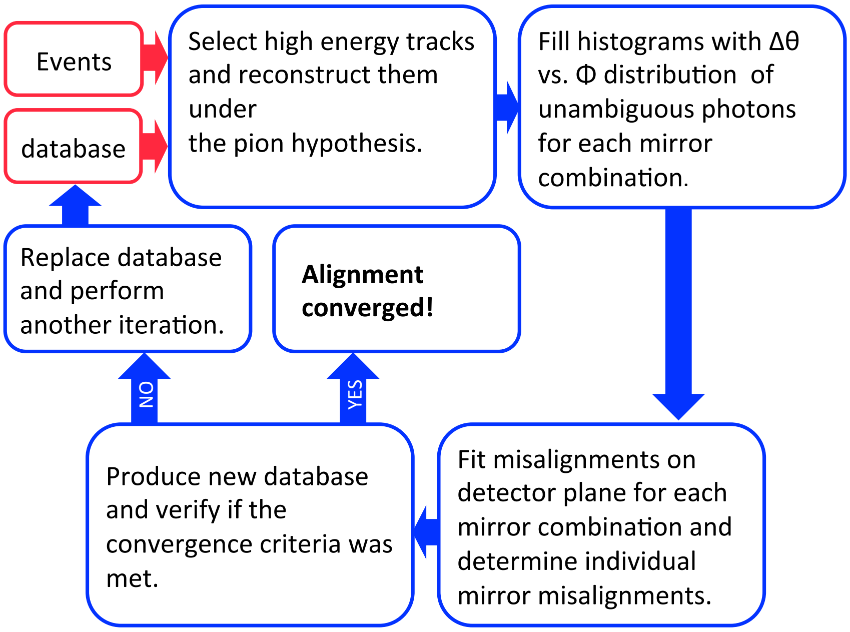


Figure 5.10: *Overview of the iterative, data-driven procedure to align the RICH mirrors. The procedure is explained in detail in the text.*

5.5.3 Filling and Fitting the $\delta\theta$ vs. ϕ Histograms

For each possible combination of primary mirror and secondary mirror a $\delta\theta$ vs. ϕ histogram is filled. Each track that passes the selection is projected onto the detector plane. The hits in the detector in a region around the track position are taken as Cherenkov photon candidates. For all photon candidates the Cherenkov angle with respect to the track is determined. Since the exact emission point of the photon is unknown it is assumed to be emitted at the middle point of the track in the detector volume. For the mirror alignment procedure the noise from incorrectly associated photons can be reduced by selecting only *unambiguous* photons. These are photon candidates that, regardless of their emission point along the track, will be reflected by the same pair of primary and secondary mirrors. The expected Cherenkov angle θ_C can be calculated under the pion hypothesis and the difference $\delta\theta$ between the measured Cherenkov angle θ and the expected angle is computed as $\delta\theta(\phi) = \theta(\phi) - \theta_C$.

Each $\delta\theta$ vs. ϕ histogram is divided into 20 bins in ϕ. For each bin the $\delta\theta$ distribution is fitted with a Gaussian function plus a second order polynomial. The means of the Gaussian functions are connected by the sinusoidal function given as

$$\delta\theta(\phi) = \Theta^z \cos\phi + \Theta^z \sin\phi \tag{5.6}$$

where Θ^z and Θ^y are called the *misalignment on the detector plane* in y and z respectively.

Two options have been tested for the fit.

1. The functions fitted to each of the ϕ bins within one $\delta\theta$ vs. ϕ histogram are completely independent of each other apart from the means of the Gaussians that are constrained to the sinusoidal function.

2. The widths of the Gaussians are constrained to be the same in each bin. The widths of the Gaussians correspond to the Cherenkov angle resolution.

Both approaches are evaluated and compared in Section 5.8.1.

5.5.4 Determination of Individual Mirror Misalignments

The misalignments on the detector plane Θ^y, Θ^z are caused by the rotations of the primary and/or secondary mirror with respect to their orientation in the LHCb conditions database. For rotations α_p^y, α_p^z of the primary mirror p around y, z respectively, and rotations β_s^y, β_s^z of the secondary mirror s around y, z respectively, the misalignments on the detector plane are expressed as

$$\Theta^y = A_{ps}^y \alpha_p^y + B_{ps}^y \beta_s^y + a_{ps}^z \alpha_p^z + b_{ps}^z \beta_s^z \tag{5.7}$$

$$\Theta^y = A_{ps}^z \alpha_p^z + B_{ps}^z \beta_s^z + a_{ps}^z \alpha_p^y + b_{ps}^y \beta_s^y \tag{5.8}$$

where A_{ps}^y, B_{ps}^y, A_{ps}^z and B_{ps}^z are the *major magnification coefficients* and a_{ps}^y, b_{ps}^y a_{ps}^z and b_{ps}^z are the *minor magnification coefficients*. The magnification coefficients translate the effect of a rotation of a mirror onto the detector plane. They depend on the path the Cherenkov photons travel through the detector. The magnification coefficients for different combinations of primary mirror p and secondary mirror s are very similar. The average values of the major magnification coefficients are listed in Table 5.2. The absolute values of the minor magnification coefficients vary approximately between 0.001 and 0.250.

	RICH1	RICH2
$\langle A^y \rangle$	1.86	2.05
$\langle A^z \rangle$	2.02	1.83
$\langle B^y \rangle$	-0.55	-1.04
$\langle B^z \rangle$	0.81	0.61

Table 5.2: *Average of the major magnification coefficients for both RICH detectors.*

Each possible combination of primary and secondary mirrors has the two Equations 5.7 and 5.8. For RICH1 this results in 32 equations — each secondary mirror receives photons from exactly one primary mirror — for 40 unknowns. This means the system is under-constrained. This can be illustrated by considering each quadrant — a primary mirror and its four associated secondary mirrors — independently. A rotation of the primary mirror in one direction can be compensated by an opposite rotation of all secondary mirrors, causing the same image on the detector plane. While many more combinations of primary and secondary mirrors are possible in RICH2 and therefore more equations, this principle still holds in the approximation that the magnification coefficients for all primary and secondary mirrors are the same. It can be shown that the system of equations for RICH2 is degenerate (see Appendix B for proof). Since the magnification coefficients are not completely identical a unique solution to the system of equations exists. As the difference between the magnification coefficients is very small, this solution is unstable.

Two methods have been used to solve the system of equations.

1. **The Algebraic Method**
   For the mirrors of RICH1 the algebraic method consists of two steps. In the first step the rotations of the primary mirrors are determined by assuming the secondary mirrors to be fully aligned. Since there are four secondary mirrors of each primary mirror, four values are obtained for each α_p^y and α_p^z. The mirror rotations for the primary mirrors are taken to be the average of the four values. In the second step the secondary mirrors are aligned with respect to the primary mirrors whose rotations α_p^y and α_p^z have been determined in the first step.

   For each of the two halves of RICH2 all mirrors are aligned with respect to one selected mirror, primary mirrors 12 and 43. All primary and secondary mirrors are linked in a chain as shown in Figure 5.11. The alignment of the mirrors of RICH2 happens along this chain by assuming no misalignment in primary mirror 12 and aligning secondary mirror 9. From there the next primary mirror in the chain, e.g. primary mirror 17 is aligned and so on until the entire system is aligned. There is more than one way to build the chain linking all mirrors. The chain for this method is chosen in such a way that it contains the mirror combinations with the highest population in the histograms.

   This method solves the problem of degeneracy in the system of equations in a simple manner. The disadvantage of this method is that it singles out the fixed mirrors. The solution from this method is also unstable.

2. **The L2 regularisation** The principle of the L2 regularisation method is a least square fit with an additional minimisation term. This term is the sum of the squares of all individual mirror rotations. The term minimised in the least square fit is thus given by

$$\sum_{p,s}\Bigg[\frac{(\Theta^y_{p,s}-A^y_{p,s}\alpha^y_p-B^y_{p,s}\beta^y_s-a^y_{p,s}\alpha^z_p-b^y_{p,s}\beta^z_s)^2}{\sigma^2(\Theta^y_{p,s})+\sigma^2(A^y_{p,s})(\alpha^y_p)^2+\sigma^2(B^y_{p,s})(\beta^y_s)^2+\sigma^2(a^y_{p,s})(\alpha^z_p)^2+\sigma^2(b^y_{p,s})(\beta^z_s)^2}$$
$$+\frac{(\Theta^z_{p,s}-A^z_{p,s}\alpha^z_p-B^z_{p,s}\beta^z_s-a^z_{p,s}\alpha^y_p-b^z_{p,s}\beta^y_s)^2}{\sigma^2(\Theta^z_{p,s})+\sigma^2(A^z_{p,s})(\alpha^z_p)^2+\sigma^2(B^z_{p,s})(\beta^z_s)^2+\sigma^2(a^z_{p,s})(\alpha^y_p)^2+\sigma^2(b^z_{p,s})(\beta^y_s)^2}$$
$$+(\alpha^y_p)^2+(\alpha^z_p)^2+(\beta^y_s)^2+(\beta^z_s)^2\Bigg]. \tag{5.9}$$

About 1 000 combinations of primary and secondary mirrors are physically possible for RICH2. To solve the system of equations 96 combinations are chosen in such a way that all mirrors are included and that the $\delta\theta$ vs. ϕ histograms of the chosen combination have the highest population.

The L2 regularisation always provides one, stable solution. Additionally, unlike the algebraic method, the L2 regularisation treats all mirrors equally.

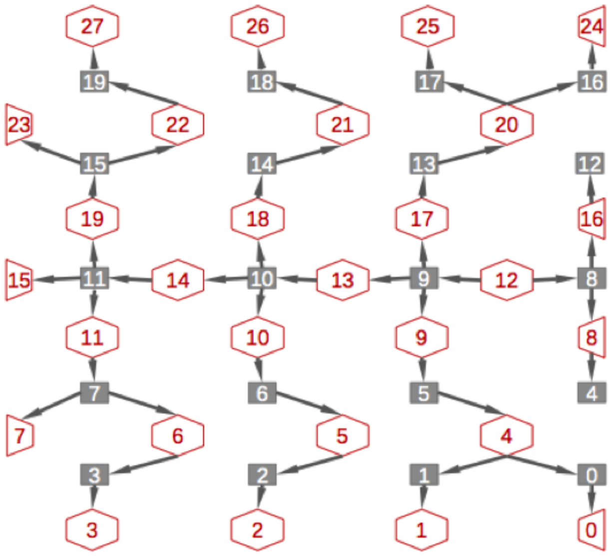


Figure 5.11: *Chain linking all primary and secondary mirrors of the left half of RICH2. The chain is not unique and here chosen in such a way that it contains the mirror combinations with the highest population in the histograms.*

Both methods for determining the individual mirror misalignments are evaluated and compared in Section 5.8.2.

The Magnification Coefficients

The magnification coefficients depend on the paths the photons take through the RICH detector to reach the HPD plane. They may therefore depend on the orientation of the mirror themselves since a rotation of a mirror may shorten or lengthen the path of the photon.

The magnification coefficients can be determined on data with a method similar to the determination of the misalignments of the individual mirrors. For each mirror combination the magnification coefficients are evaluated by introducing 8 independent calibrational rotations — positive or negative rotations about y or z for the primary or secondary mirrors — and by measuring the resulting misalignments on the detector plane. The final value of each factor is the arithmetical mean of the two corresponding values obtained with the calibrating rotations in opposite directions. The magnification coefficients are re-evaluated only after big changes of the mirror orientations. In stable periods a predetermined set of magnification coefficients is used for the alignment procedure.

5.6 Alignment Strategy in Run I

The data processing strategy of LHCb in LHC Run I (2010-2012) is shown in Figure 5.12. During the HLT stage the particle trajectories of the events are reconstructed in real-time. This event reconstruction is called the *online reconstruction*. All events accepted by the HLT were sent offline for permanent storage. The data sent offline included all raw information from the detector. An additional event reconstruction was performed on the LHC Computing Grid [76], the *offline reconstruction*, recreating particles and decays from the raw data using improved detector calibration. In Run I the online reconstruction was a simplified, faster version of the offline reconstruction.

The RICH mirror alignment was also performed *offline* and only once a year. The alignment procedure was run after the data-taking period and reapplied to the data in the reprocessing of the entire dataset for the year. One alignment was provided for each magnet polarity (*magnet up* and *magnet down*).

The reconstruction of the events, and therefore the filling of the $\delta\theta$ vs. ϕ histograms, was executed *offline* on the LHCb Computing Grid. The dataset was divided into subsets and the reconstruction submitted as individual jobs. The outputs of these jobs were merged and the individual mirror misalignments determined. An instability was introduced to the alignment procedure by requesting that in each iteration only 80% of the jobs were finished. A full alignment procedure required several days to run. The reprocessing of the entire dataset with the new alignment constants took several months.

5.7 Alignment Strategy in Run II

The dataflow for Run II is shown in Figure 5.12. As in Run I a rate of 1 MHz of events passes the level-0 trigger (L0) and is passed on to first high level trigger stage (HLT1). In HLT1 the events are partially reconstructed and accepted events are written to disk. At this stage the different alignment procedures are run on a dedicated part of the buffered data. In case of a change in alignment constants the new constants are propagated to the LHCb conditions database and used in the subsequent reconstruction of the events by the second high level trigger stage (HLT2). Due to an upgrade of the HLT computing facilities, the online reconstruction is identical to the offline reconstruction in Run II. About two thirds of the events accepted by the HLT are send to permanent offline storage, as in Run I. A third of the events are part of the Turbo stream which stores the offline-quality reconstruction from HLT2. See Section 4.2 of Chapter 4 for a more detailed description of the data-taking strategy in Run II.

The alignment tasks being performed between HLT1 and HLT2 are — in the order they are being run — VELO alignment, tracker alignment, RICH alignment and muon chamber alignment. Each alignment has its own dedicated HLT1 line which collects a given number of events at the beginning of each fill. It was found that $\sim 1\,$M events for RICH1 and $\sim 2\,$M events for RICH2 is sufficient to produce stable results. Once enough events have been collected the alignment procedure is automatically started.

The first year of Run II (2015) was dedicated to the implementation of the RICH mirror alignment into the online framework. All alignments in 2015 were started manually. In the second year of Run II (2016) the alignment was started automatically in each fill when enough events were collected. While the LHCb conditions database was only updated once at the beginning of the year, the frequent running of the alignment procedure enabled the monitoring of a potential movement of the RICH mirrors over a long period of time. The information gathered during 2016 were used to tune sensible thresholds for the automatic update of the alignment constants for 2017.

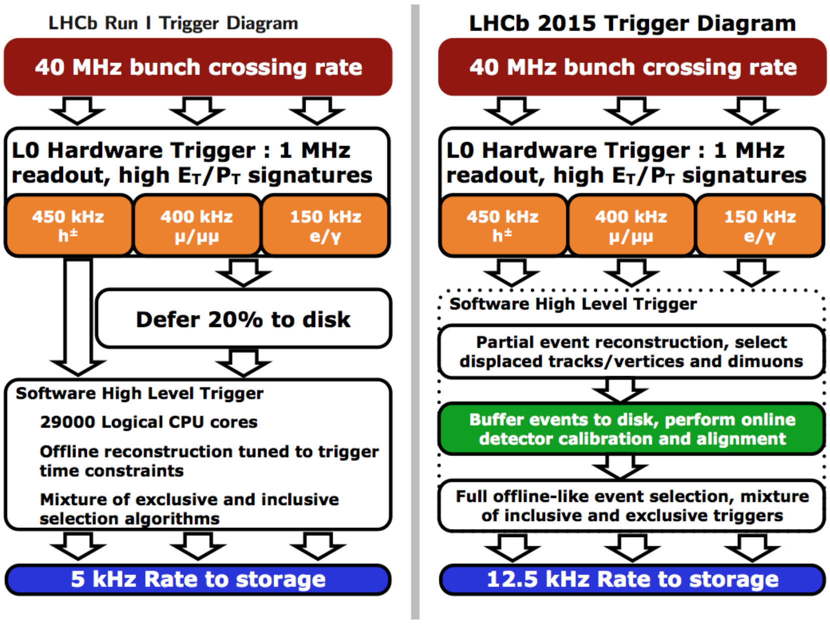


Figure 5.12: *LHCb dataflow for Run I (left) and Run II (right). In Run II the data is buffered after HLT1 and an alignment is performed for each fill. The HLT2 then processes the buffered events with the updated alignment constants.*

5.7.1 HLT1 Selection for the RICH Mirror Alignment

In order to perform a successful alignment, the $\delta\theta$ vs. ϕ histograms for each mirror combination have to contain enough entries for the fits described in Section 5.5.3 to converge. The minimum condition for a successful fit has been found to be that 16 of the 20 bins in ϕ contain at least 300 entries.

This is accomplished by having two dedicated HLT1 selections, one for each RICH detector. The lines trigger on tracks of high momentum particles whose Cherenkov photons would populate the mirror combinations containing the fewest photons. The other mirror combinations are then populated by the rest of the tracks in the events.

The variables used in the selection are the track momentum p, the transverse track momentum p_T, the pseudorapidity η, the goodness of fit for the track χ^2 and the polar angle of the track ϕ. The selection criteria of tracks that are triggered upon are listed in Table 5.3. The HLT1 lines for the RICH alignment can be optimised and updated at any time.

	RICH1	RICH2
momentum p	$p > 20$ GeV	$p > 40$ GeV
transverse momentum p_T	$p_T > 0.5$ GeV	$p_T > 0.5$ GeV
pseudorapidity η	$1.6 < \eta < 2.04$	$2.65 < \eta < 2.80$
track χ^2	$\chi^2 < 2$	$\chi^2 < 2$
polar angle ϕ	$-2.65 < \phi < -2.30$	$-2.59 < \phi < -2.49$
	$-0.80 < \phi < -0.50$	$-0.65 < \phi < -0.55$
	$0.50 < \phi < 0.80$	$0.55 < \phi < 0.65$
	$2.30 < \phi < 2.65$	$2.49 < \phi < 2.59$

Table 5.3: *Selection criteria for the HLT1 line for RICH1 and RICH2 in 2016. Events that are accepted by these trigger lines need to have at least one track that satisfies the selection criteria.*

5.7.2 The LHCb Event Filter Farm

The RICH mirror alignments are run on the LHCb Event Filter Farm (EFF) which also executes both HLT1 and HLT2. The EFF consists of 62 sub-farms, which are largely, but not completely, homogeneous. The sub-farms make up a total of approximately 1700 independent units. Each unit has a harddisk space between 4 Tbytes and 12 Tbytes and between 24 and 40 logical cores. This means that each unit can run between 24 and 40 processes in parallel making it possible to run about 50 000 processes at the same time. The data that passed the HLT1 selection is stored evenly distributed over the units until all alignments and calibrations have been performed. It can then be processed by HLT2.

For the purpose of the alignment, one central unit is singled out and called *iterator* while all other units are called *analysers*. The analysers work independently of each other on every node and reconstruct the events selected by the HLT1 line and thus produce the individual $\delta\theta$ vs. ϕ histograms. Each analyser produces the histograms for the data that is stored on it and when all analysers are finished the histograms are merged and evaluated by the iterator. The iterator also computes the individual mirror misalignments and verifies the convergence criteria. The tasks of the iterator and the analysers are explained in more detail in Section 5.7.4.

All units are completely independent of each other. The parallel processing of the units is asynchronous and has to be coordinated between the individual analysers and the iterator. This is described in the next section.

5.7.3 The Control Flow

The execution of the alignment tasks is under the control of the LHCb Experiment Control System (ECS), and is implemented as a *finite state machine*, which is illustrated in Figure 5.13. The principle of a finite state machine means that each component of the system (here every individual analyser and the iterator) has to be in one of a finite number of states at all times. The states used for the alignment procedure are also shown in Figure 5.13; **Ready**, **Running**, **Paused** and **Stopped**.

The alignment procedure is steered by the ECS which knows the state of all parts of the system. The ECS also has the ability to send commands — such as *configure*, *start*, *pause* and *stop* — to all individual parts of the system. If a command is received by a unit, it will go from its current state into the state declared by Figure 5.13. When in a new state, the unit will usually perform a task and once the task is finished set itself into another state.

For a given state only a certain number of commands are possible - for example if the component is in state **Paused** it can only receive the commands *continue* and *stop*.

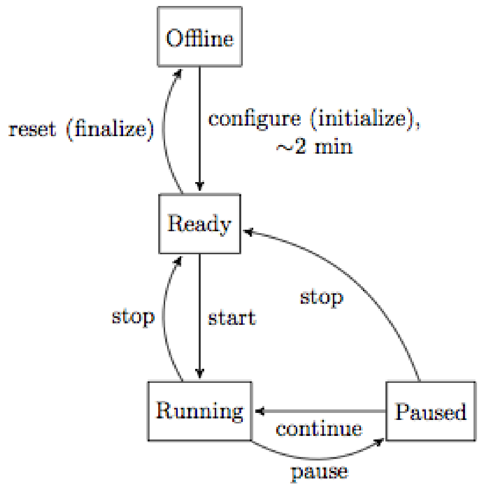


Figure 5.13: *Example of one unit within a system functioning under the principle of a finite state machine. The boxes show the states the unit is in while the arrows show the commands the unit gets from the run control.*

5.7.4 Implementation of the RICH Mirror Alignment for Run II

The interplay between the iterator, one example analyser and the run control during the course of the alignment procedure of a RICH detector is shown in Figure 5.14.

The individual analysers and the iterator all follow the same sequence of states, namely the one shown in Figure 5.13. When the alignment is being started the run control sends the command to *configure* to both the iterator and all analysers. All units will go into state **Configuring** while setting up to run the alignment. For the analysers this means that they read in the configuration for the reconstruction of the events, while the iterator sets up a directory in which all files for this alignment are saved. It also retrieves the current RICH mirror orientations (usually from the LHCb conditions database) and makes this information available to the analysers.

When finished configuring, each unit changes into state **Ready**. When all individual units are in the **Ready** the run control sends the command to *start* which makes all units change their status into *Running*. This prompts the analysers to start the reconstruction of the events and the production of the $\delta\theta$ vs. ϕ histograms. During this time the iterator is idle.

Each analyser that has completed processing its data updates its state to **Paused**. Once all analysers have reached this state, the run control sends the *stop* command and they update their states to **Ready**.

The iterator is then sent the *pause* command which changes its state into **Paused**. During this time the iterator retrieves the histograms produced by the analysers and performs the fits to the $\delta\theta$ vs. ϕ histograms. It then calculates the individual mirror misalignments and creates from them a new database entry. The iterator evaluates the convergence criterion and then either indicates that conversion has been reached by changing its state to **Ready**, or that another iteration is required by changing its state to **Running**.

In the latter case the iterator will provide the new database entry to the analysers before changing its state and another iteration is started.

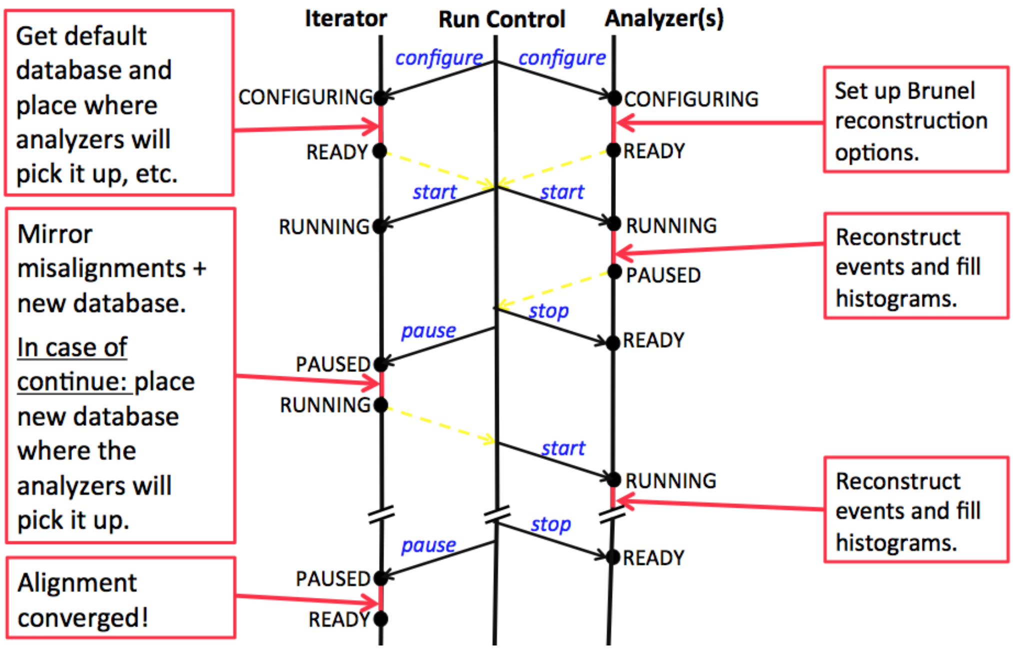


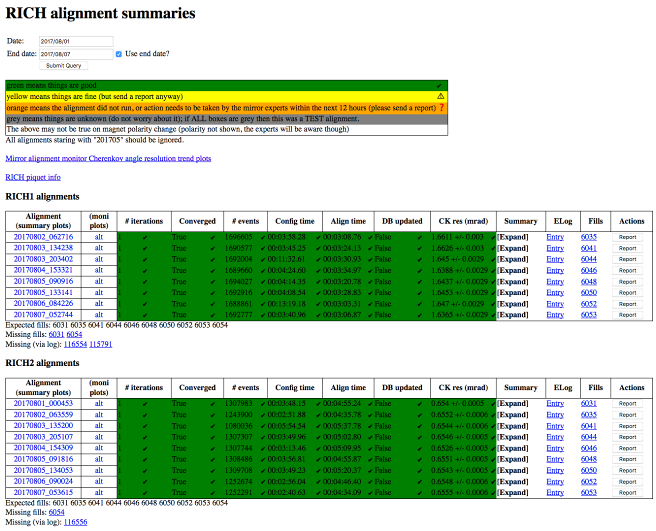
Figure 5.14: *Interplay between the iterator, one example analyser and the run control during the RICH alignment procedure. The analysers reconstruct the data and produce the $\delta\theta$ vs. ϕ histograms that the iterator evaluates. The run control sends commands to the iterator and the analysers to ensure that the alignment procedure happens in the necessary sequence.*

5.7.5 Monitoring of the Automatic Procedure

When performing frequent alignments, it is important to monitor the procedure. For this purpose a set of monitoring plots is produced after each alignment. The monitoring plots are accessible through the website of the LHCb RICH group and offer the possibility to look at every alignment ever performed.

Figure 5.15 is a screenshot of the layout of the monitoring website[2]. A time period can be chosen in the upper left corner of the site. All alignments that were performed in this time period will then be listed. The left most links in the table lead to a collection of plots that summarise the alignment procedure.

Figure 5.16 and Figure 5.17 show these plots for an example alignment of the RICH1 mirrors. The first plots show the distribution of $\delta\theta$ for each iteration. A fit is performed to extract the widths of the distributions which corresponds to the Cherenkov angle resolution. The subsequent plot shows the development of the Cherenkov angle resolution over the iterations of the alignment. The final plots show the determined total misalignments for the individual mirrors in y and z with respect to the LHCb conditions database.



RICH alignment summaries

Date: 2017/08/01
End date: 2017/08/07 Use end date?
Submit Query

green means things are good ✔
yellow means things are fine (but send a report anyway) ⚠
orange means the alignment did not run, or action needs to be taken by the mirror experts within the next 12 hours (please send a report) ?
grey means things are unknown (do not worry about it); if ALL boxes are grey then this was a TEST alignment.
The above may not be true on magnet polarity change (polarity not shown, the experts will be aware though)

All alignments staring with "201705" should be ignored.

Mirror alignment monitor Cherenkov angle resolution trend plots

RICH piquet info

RICH1 alignments

Alignment (summary plots)	(moni plots)	# iterations	Converged	# events	Config time	Align time	DB updated	CK res (mrad)	Summary	ELog	Fills	Actions
20170802_062716	alt	1 ✔	True ✔	1696605 ✔	00:03:58.28 ✔	00:03:08.76 ✔	False ✔	1.6611 +/- 0.003 ✔	[Expand]	Entry	6035	Report
20170803_134238	alt	1 ✔	True ✔	1690577 ✔	00:03:45.25 ✔	00:03:24.13 ✔	False ✔	1.6626 +/- 0.003 ✔	[Expand]	Entry	6041	Report
20170803_203402	alt	1 ✔	True ✔	1692004 ✔	00:11:32.61 ✔	00:03:30.93 ✔	False ✔	1.645 +/- 0.0029 ✔	[Expand]	Entry	6044	Report
20170804_153321	alt	1 ✔	True ✔	1689660 ✔	00:04:24.60 ✔	00:03:34.97 ✔	False ✔	1.6388 +/- 0.0029 ✔	[Expand]	Entry	6046	Report
20170805_090916	alt	1 ✔	True ✔	1694027 ✔	00:04:14.35 ✔	00:03:20.78 ✔	False ✔	1.6437 +/- 0.0029 ✔	[Expand]	Entry	6048	Report
20170805_133141	alt	1 ✔	True ✔	1692916 ✔	00:04:08.54 ✔	00:03:28.83 ✔	False ✔	1.6453 +/- 0.0029 ✔	[Expand]	Entry	6050	Report
20170806_084226	alt	1 ✔	True ✔	1688861 ✔	00:13:19.18 ✔	00:03:03.31 ✔	False ✔	1.647 +/- 0.0029 ✔	[Expand]	Entry	6052	Report
20170807_052744	alt	1 ✔	True ✔	1692777 ✔	00:03:40.96 ✔	00:03:06.87 ✔	False ✔	1.6365 +/- 0.0029 ✔	[Expand]	Entry	6053	Report

Expected fills: 6031 6035 6041 6044 6046 6048 6050 6052 6053 6054
Missing fills: 6031 6054
Missing (via log): 116554 115791

RICH2 alignments

Alignment (summary plots)	(moni plots)	# iterations	Converged	# events	Config time	Align time	DB updated	CK res (mrad)	Summary	ELog	Fills	Actions
20170801_000453	alt	1 ✔	True ✔	1307983 ✔	00:03:48.15 ✔	00:04:55.24 ✔	False ✔	0.654 +/- 0.0005 ✔	[Expand]	Entry	6031	Report
20170802_063559	alt	1 ✔	True ✔	1243900 ✔	00:02:51.88 ✔	00:04:35.78 ✔	False ✔	0.6552 +/- 0.0006 ✔	[Expand]	Entry	6035	Report
20170803_135200	alt	1 ✔	True ✔	1080036 ✔	00:05:54.54 ✔	00:05:37.78 ✔	False ✔	0.6544 +/- 0.0006 ✔	[Expand]	Entry	6041	Report
20170803_205107	alt	1 ✔	True ✔	1307307 ✔	00:03:49.96 ✔	00:05:02.80 ✔	False ✔	0.6546 +/- 0.0005 ✔	[Expand]	Entry	6044	Report
20170804_154309	alt	1 ✔	True ✔	1307744 ✔	00:03:13.46 ✔	00:05:09.95 ✔	False ✔	0.6526 +/- 0.0005 ✔	[Expand]	Entry	6046	Report
20170805_091816	alt	1 ✔	True ✔	1308486 ✔	00:03:56.81 ✔	00:04:55.87 ✔	False ✔	0.6551 +/- 0.0005 ✔	[Expand]	Entry	6048	Report
20170805_134053	alt	1 ✔	True ✔	1309708 ✔	00:03:49.23 ✔	00:05:20.37 ✔	False ✔	0.6543 +/- 0.0005 ✔	[Expand]	Entry	6050	Report
20170806_090024	alt	1 ✔	True ✔	1252674 ✔	00:02:56.04 ✔	00:04:46.40 ✔	False ✔	0.6548 +/- 0.0006 ✔	[Expand]	Entry	6052	Report
20170807_053615	alt	1 ✔	True ✔	1252291 ✔	00:02:40.63 ✔	00:04:34.09 ✔	False ✔	0.6555 +/- 0.0006 ✔	[Expand]	Entry	6053	Report

Expected fills: 6031 6035 6041 6044 6046 6048 6050 6052 6053 6054
Missing fills: 6054
Missing (via log): 116556

Figure 5.15: *Screenshot of the layout of the monitoring website. A time period can be chosen in the upper left corner of the site and all alignments performed in this time period will be listed. The links on the very left of the table lead to a collection of plots that summaries the alignment procedure.*

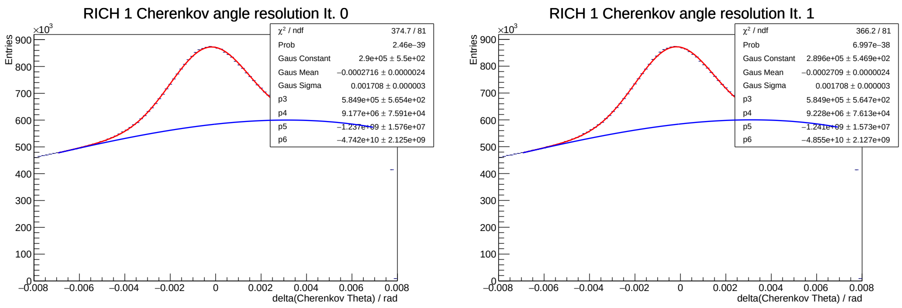


Figure 5.16: *Example of the monitoring plots produced after each alignment procedure. This example is a RICH1 alignment that took two iterations to converge. Distribution of the expected Cherenkov angle minus the measured Cherenkov angle for the first iteration (left) and the second iteration (right). Fits are performed to extract the width of the distribution which corresponds to the Cherenkov angle resolution.*

[2]The layout of the monitoring website for the RICH alignment has changed from its initial implementation but its functionality is similar

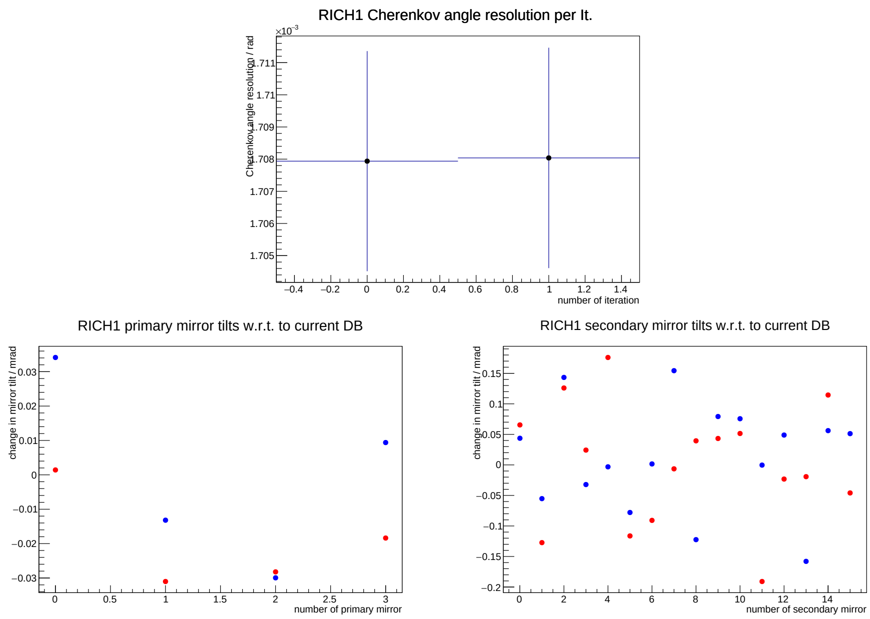


Figure 5.17: *Example of the monitoring plots produced after each alignment procedure. This example is a RICH1 alignment that took two iterations to converge.* Top: *The development of the Cherenkov angle resolution over the iterations of the alignment.* Bottom: *Total misalignments for the individual mirrors in y (red markers) and z (blue markers) for the primary mirrors (left) and the secondary mirrors(right). The mirror numbering scheme is explained in Section 5.3. The misalignments are determined with respect to the current LHCb conditions database.*

5.8 Improvements to the Alignment Procedure

This section presents two studies aimed at improving the RICH mirror alignment procedure. During the periodic running of the real-time alignments in Run II, speed is of crucial importance. Therefore two different methods to perform the fit to the $\delta\theta$ vs. ϕ histograms introduced in Section 5.5.3 are evaluated in the first study. In the second study the performance of the two methods used for the determination of the individual mirror misalignments, shown in Section 5.5.4, is compared.

5.8.1 Fitting of the $\delta\theta$ vs. ϕ Histograms

As introduced in Section 5.5.3, two different fit functions can be used to fit the $\delta\theta$ vs. ϕ histograms to extract the misalignments on the detector plane Θ^y and Θ^z for each combination of primary and secondary mirror. In the first method the widths of the Gaussians describing the signal contribution in each ϕ bin are independent of each other while in the second method they are fixed to a shared value.

Both methods have been applied to a set of histograms for each RICH detector. The histograms have varying degrees of misalignments. The performance of the two fit methods is compared by evaluating the fit results as well as the duration of the fits.

RICH1

Figure 5.18 shows the difference of the fit results of the method with the free Gaussian widths and the results of the methods with the unified Gaussian widths for RICH1. The means of the distributions lie at 0.00 mrad with a standard deviation of 0.02 mrad and 0.01 mrad for Θ^y and Θ^z, respectively.

Figure 5.19 shows the fit results of the method with the free Gaussian widths over the results of the methods with the unified Gaussian widths for RICH1. The points lie on a diagonal and show no significant bias between the two methods.

Figure 5.20 shows the duration of each fit for both methods. The method with the free Gaussian widths takes an average of 79 s while the method with the unified Gaussian widths takes an average of 15 s. The relative time difference $\Delta t/t^{free} = \frac{t^{free}-t^{unif.}}{t^{free}}$ is also shown in Figure 5.20. The method with the unified Gaussian width is on average 82% faster than the method with the free Gaussian widths.

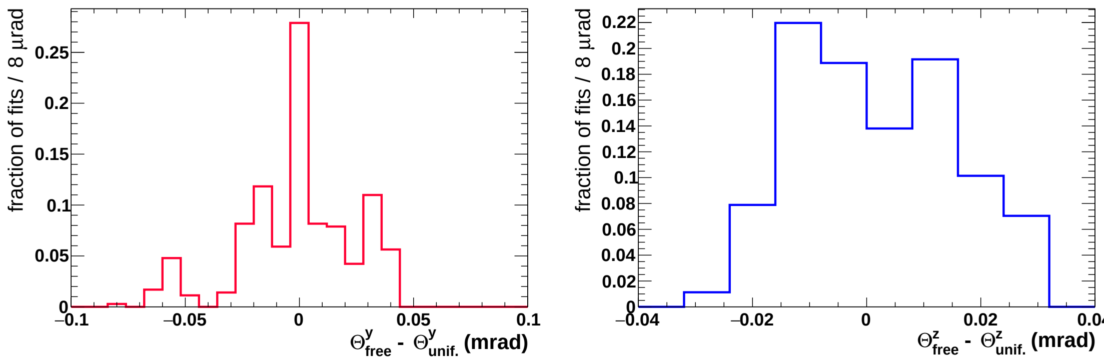


Figure 5.18: *Distribution of the difference of the fit results of the method with the free Gaussian widths and the results of the methods with the unified Gaussian widths for RICH1 for the misalignments on the detector plane in y (left) and in z (right). The means of the distributions are 0.00* mrad *and the standard deviation is 0.02* mrad *and 0.01* mrad *for* Θ^y *and* Θ^z*, respectively.*

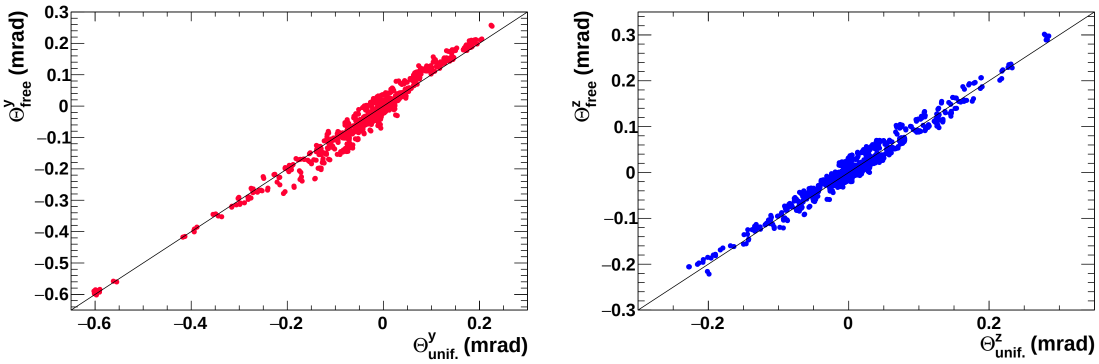


Figure 5.19: *Distribution the fit results of the method with the free Gaussian widths over the results of the methods with the unified Gaussian widths for RICH1 for the misalignments on the detector plane in y (left) and in z (right).*

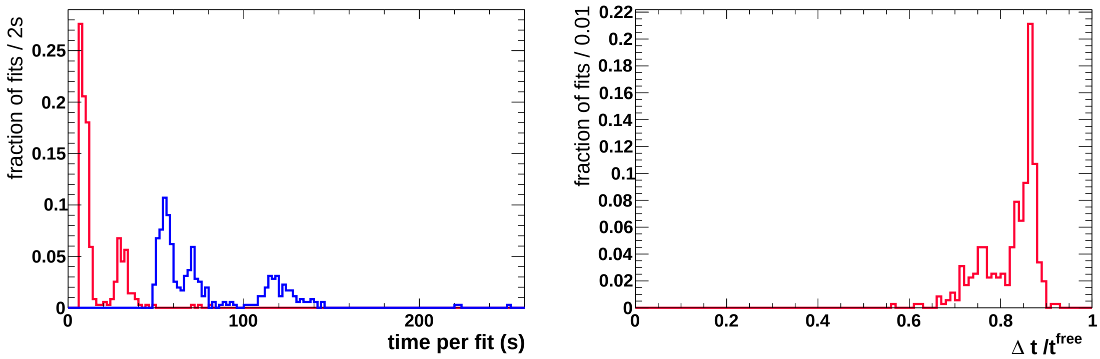


Figure 5.20: Right: *Distributions of the time taken to perform one fit for the method with the free Gaussian widths (blue) and the method with the unified Gaussian widths (red) for RICH1.* Left: *Distribution of the relative time difference* $\Delta t/t^{free} = \frac{t^{free} - t^{unif.}}{t^{free}}$ *between the two methods for RICH1.*

RICH2

Figure 5.21 shows the difference of the fit results of the method with the free Gaussian widths and the results of the methods with the unified Gaussian widths for RICH2. The means of the distributions lie at 0.00 mrad with a standard deviation of 0.006 mrad and 0.001 mrad for Θ^y and Θ^z, respectively.

Figure 5.22 shows the fit results of the method with the free Gaussian widths over the results of the methods with the unified Gaussian widths for RICH2. The points lie on a diagonal and show no significant bias between the two methods.

Figure 5.23 shows the duration of each fit for both methods. The method with the free Gaussian widths takes an average of 128 s while the method with the unified Gaussian widths takes an average of 27 s. The relative time difference $\Delta t/t^{free} = \frac{t^{free} - t^{unif.}}{t^{free}}$ is also shown in Figure 5.23. The method with the unified Gaussian width is on average 76% faster than the method with the free Gaussian widths.

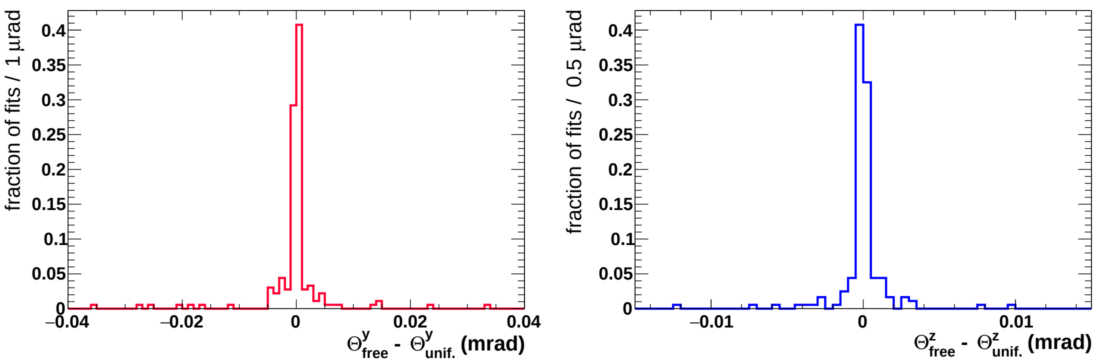


Figure 5.21: *Distribution of the difference of the fit results of the method with the free Gaussian widths and the results of the methods with the unified Gaussian widths for RICH2 for the misalignments on the detector plane in y (left) and in z (right). The means of the distributions are 0.00* mrad *and the standard deviation is 0.006* mrad *and 0.001* mrad *for* Θ^y *and* Θ^z*, respectively.*

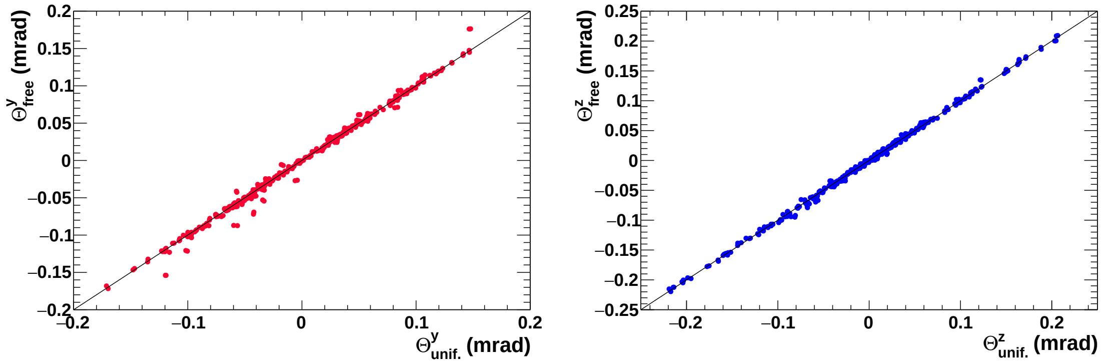


Figure 5.22: *Distribution the fit results of the method with the free Gaussian widths over the results of the methods with the unified Gaussian widths for RICH2 for the misalignments on the detector plane in y (left) and in z (right).*

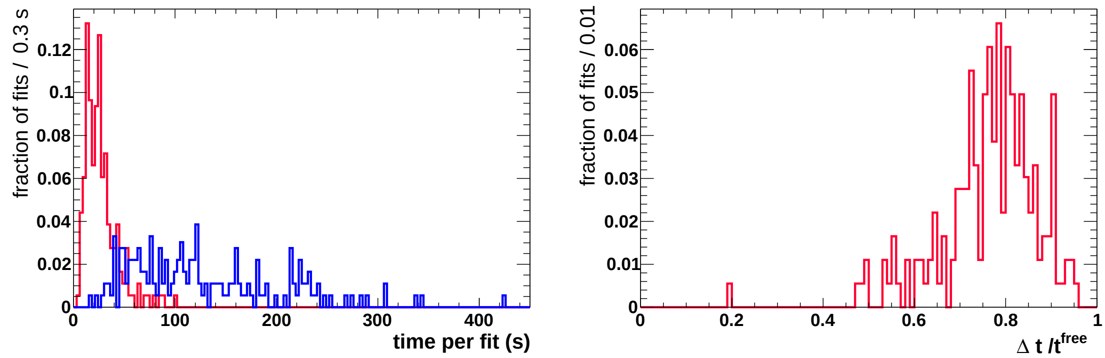


Figure 5.23: Right: *Distributions of the time taken to perform one fit for the method with the free Gaussian widths (blue) and the method with the unified Gaussian widths (red) for RICH2.* Left: *Distribution of the relative time difference* $\Delta t/t^{free} = \frac{t^{free} - t^{unif.}}{t^{free}}$ *between the two methods for RICH2.*

Conclusion

Both fit methods give consistent results for the misalignments on the detector plane Θ^y and Θ^z and no significant bias is observed between the results. The method using the unified Gaussian widths is much faster, taking on average 82% less time for RICH1 and 76% less time for RICH2. Since time is of crucial importance during the real-time alignment, the method with the unified Gaussian widths is chosen for the automatic alignment procedure.

5.8.2 Algebraic Method vs. L2 Regularisation

In this section the two methods used to compute the individual mirror misalignments are compared. The methods introduced in Section 5.5.4 are evaluated on a number of fills by comparing their performance in terms of Cherenkov angle resolution and speed. Each fill is used to perform the full alignment procedure, once with the algebraic method and once with the L2 regularisation method.

RICH1

The RICH1 alignments were performed on five different fills. All alignments procedures converged within one iteration as is shown in Figure 5.24. Figure 5.24 also shows for each fill the Cherenkov angle resolution obtained with the L2 regularisation method against the resolution obtained with the algebraic method. The corresponding resolutions are almost identical.

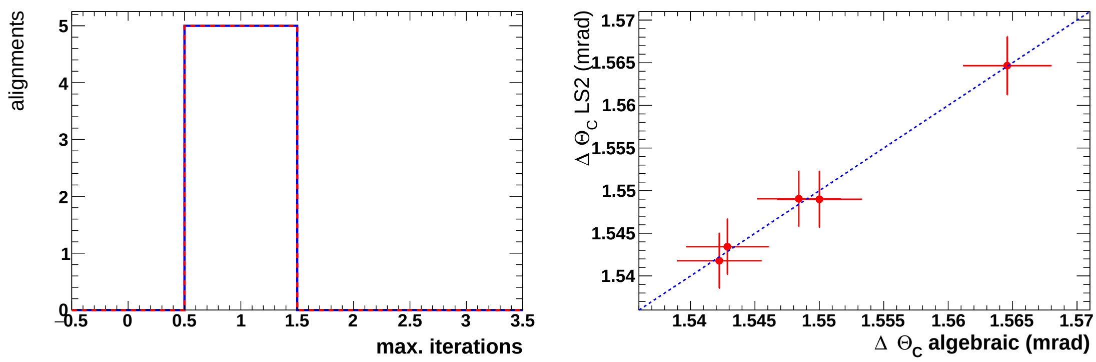


Figure 5.24: Right: *Distributions of the iterations needed to reach convergence for the alignments using the algebraic method (red) and the alignments using the L2 regularisation method (blue) for RICH1.* Left: *The Cherenkov angle resolution obtained with the L2 regularisation method against the resolution obtained with the algebraic method for each fill for RICH1.*

RICH2

The RICH2 alignments were performed on four different fills. As shown in Figure 5.25, the alignment procedures using the algebraic method took between two and 5 iterations to converge. All alignment procedures using the L2 regularisation method converged within two iterations. The reason these alignment procedures converge within two iterations rather than one iteration like all RICH2 alignment procedures performed automatically during Run II (see Section 5.10) is that the starting point for the alignments here was an alignment obtained with the algebraic method. Due to the mathematical difference of both methods, their solutions for the individual mirror misalignments differ. Figure 5.25 also shows for each fill the Cherenkov angle resolution obtained with the L2 regularisation method against the resolution obtained with the algebraic method. The corresponding resolutions are almost identical.

Conclusion

There is no visible difference in performance of the algebraic method and the L2 regularisation method for RICH1. This make intuitively sense, since the problem of identifying the individual mirror misalignments is not as complex for RICH1 as for RICH2.

For RICH2 both methods yield the same Cherenkov angle resolution but the L2 regularisation method reaches convergence in fewer iterations and is thus faster. Since

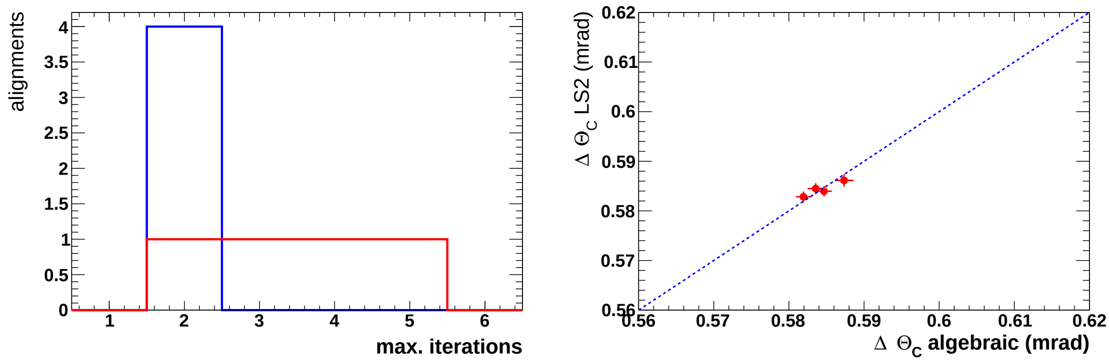


Figure 5.25: Right: *Distributions of the iterations needed to reach convergence for the alignments using the algebraic method (red) and the alignments using the L2 regularisation method (blue) for RICH2.* Left: *The Cherenkov angle resolution obtained with the L2 regularisation method against the resolution obtained with the algebraic method for each fill for RICH2.*

time is of crucial importance during the real-time alignment the L2 regularisation method is chosen for the automatic alignment procedure.

5.9 First Update of the LHCb Conditions Database in 2016

The first year of Run II (2015) was mainly used to implement the RICH mirror alignment into the online framework and to improve its performance. In the second year of Run II (2016) the RICH mirror alignment was ready to run automatically for each fill. To provide an adequate starting point, an alignment was performed at the beginning of the data-taking period which started from a *very misaligned database entry*. During this alignment procedure the magnification coefficients were recalculated for each iteration. The LHCb conditions database was updated with the thus determined RICH mirror orientations.

5.9.1 RICH1

The first alignment of the RICH1 mirrors performed on 2016 data took eight iterations to converge. The development of the Cherenkov angle resolution over the iterations is shown in Figure 5.26. The Cherenkov angle resolution improves with each iteration until it converges to about 1.72 mrad which is close to the theoretical optimum listed in Section 5.4. This indicates a successful alignment procedure.

Figure 5.27 shows the development of the corrections to the mirror orientations over the iterations for each primary and secondary mirror. The orientations for each mirror show acceptable convergence behaviour over the iterations, also indicating a successful alignment procedure.

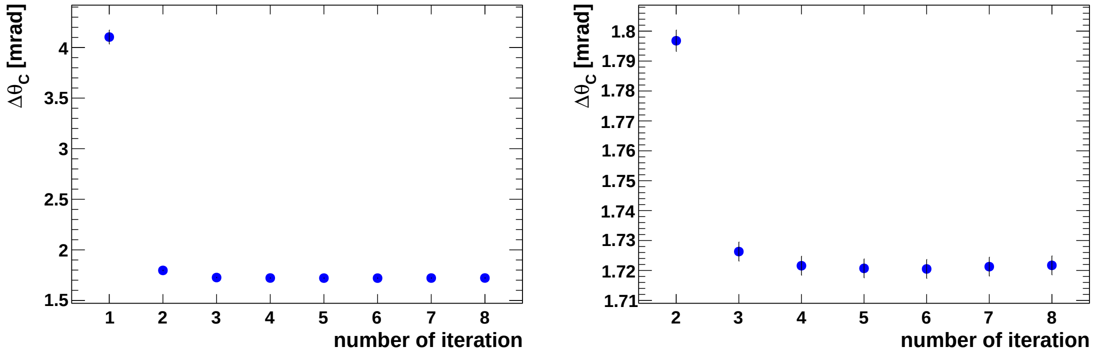


Figure 5.26: *Development of the Cherenkov angle resolution over the iterations for the first alignment of RICH1 with 2016 data. The alignment started from a database entry with a big misalignment which explains the Cherenkov angle resolution of 4* mrad *in the first iteration. The right figure is a zoom into the left figure without the first iteration.*

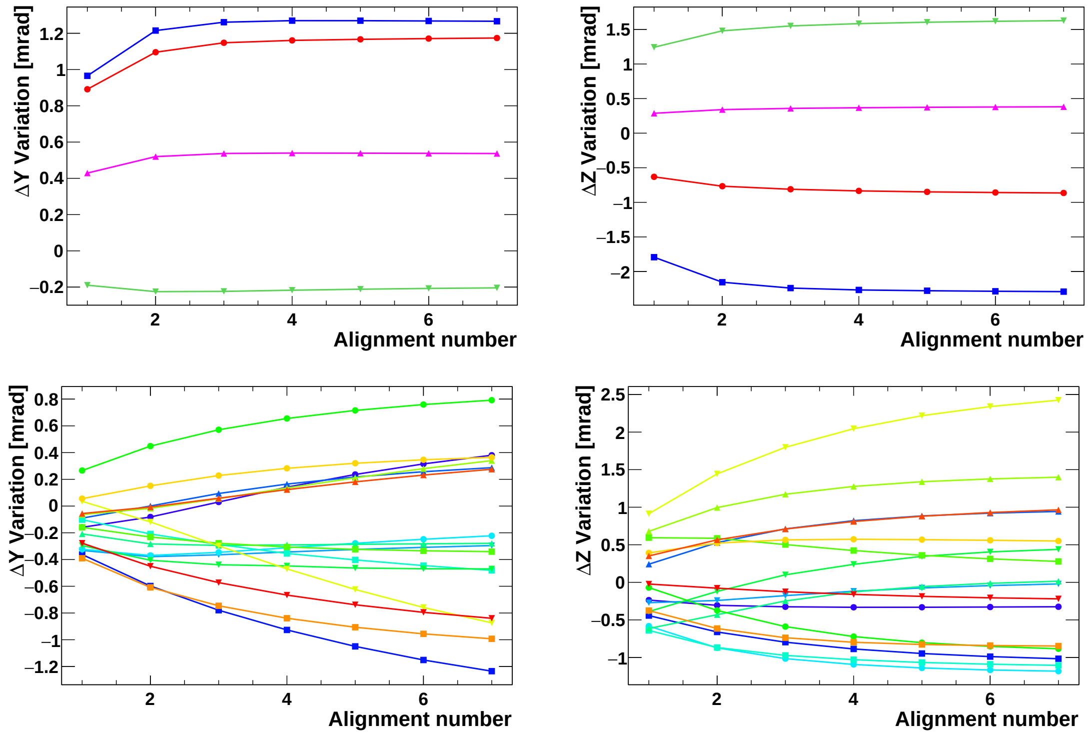


Figure 5.27: *Development of the corrections to the mirror orientations over the iterations for each primary (top) and secondary (bottom) mirror of RICH1 in y (left) and z (right) direction. Each colour represents a different mirror.*

5.9.2 RICH2

The first alignment of the RICH2 mirrors performed on 2016 data took three iterations to converge. The development of the Cherenkov angle resolution over the iterations is shown in Figure 5.28. The Cherenkov angle resolution improves with each iteration until it converges to about 0.66 mrad which is better than the theoretical optimum listed in Section 5.4. This indicates a successful alignment procedure.

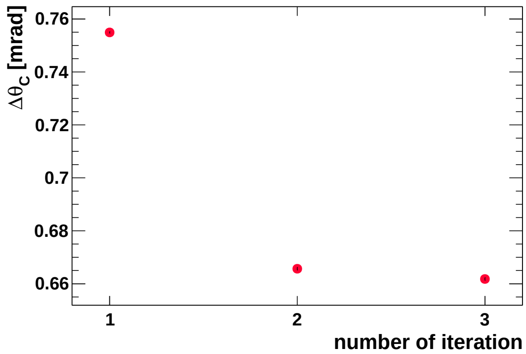


Figure 5.28: *Development of the Cherenkov angle resolution over the iterations for the first alignment of RICH2 with 2016 data. The alignment started from a database entry with a big misalignment which explains the Cherenkov angle resolution of 0.755* mrad *in the first iteration.*

Figure 5.29 shows the development of the corrections to the mirror orientations over the iterations for each primary and secondary mirror. The orientations for each mirror show acceptable convergence behaviour over the iterations, also indicating a successful alignment procedure.

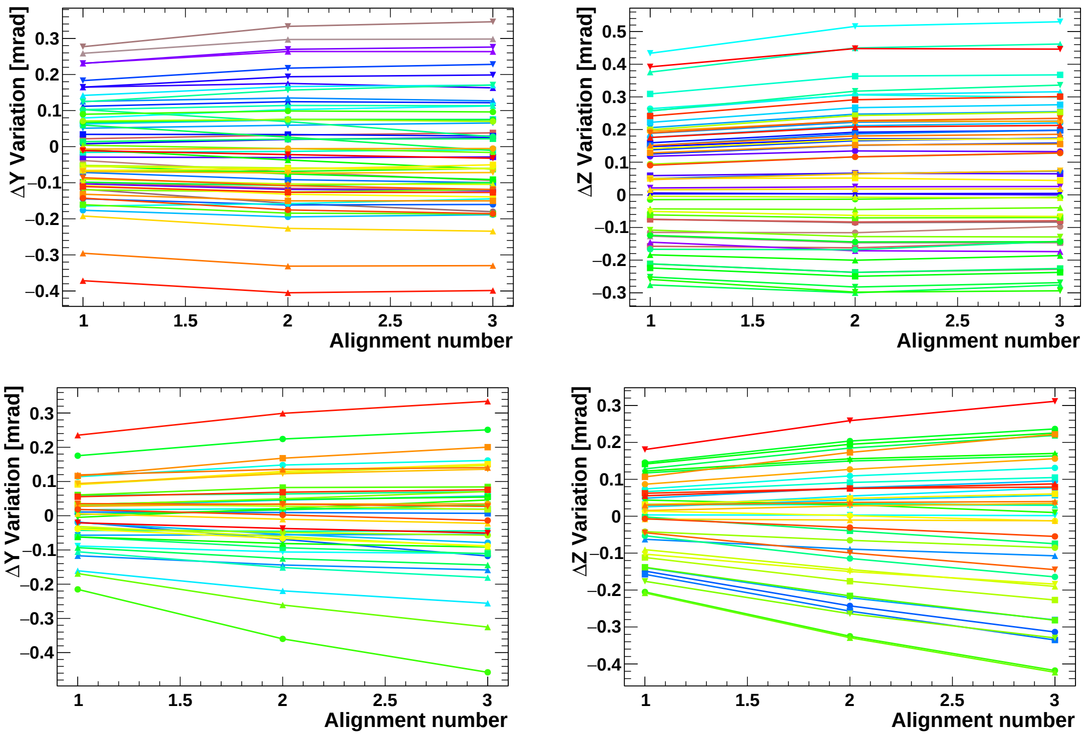


Figure 5.29: *Development of the corrections to the mirror orientations over the iterations for each primary (top) and secondary (bottom) mirror of RICH2 in y (left) and z (right) direction. Each colour represents a different mirror.*

5.9.3 Conclusion

The first alignment procedures performed on 2016 data were successful for both RICH1 and RICH2. The LHCb conditions database was updated with the thus determined RICH mirror orientations.

In the alignment procedures the magnification coefficients were calculated on the data for each iteration anew. The magnification coefficients for the last iteration were taken as predetermined set for the automatic alignment procedures performed in 2016.

5.10 Summary of the 2016 Data-taking Period

In the second year of Run II (2016) the alignment was started automatically for each fill[3] after the required number of events was collected by the HLT1 lines. The alignment procedure for RICH1 is started first. The alignment procedure for RICH2 is started after the RICH1 alignment procedure has finished.

While the framework for the automatic running of the alignment procedure for each fill was in place for 2016, the automatic update of the LHCb conditions database for the RICH mirror alignment was only put in place for 2017. The LHCb conditions database was only updated once in 2016 for the orientations of the RICH mirrors with the database entry produced in the procedure presented in Section 5.9. All alignments performed automatically in Run II started from this database entry. This allows the study of the mirror orientations throughout the year as well as the evaluation of the stability of the alignment procedure.

For the data-taking period of 2016, 44 alignment procedures are evaluated for RICH1 and RICH2, respectively. In the following, the time needed to perform an alignment is assessed as well as the development of the mirror orientations and the Cherenkov angle resolutions over the year.

Duration of Alignment Procedures

Figure 5.30 shows distribution of iterations needed to reach convergence of the alignment procedure for RICH1 and RICH2 as well as the distribution of time needed for the alignment procedures. All alignment procedures of RICH2 converged within one iteration and took 14 min on average. The alignment procedures of RICH1 needed between one and five iterations to converge and took 22 min on average.

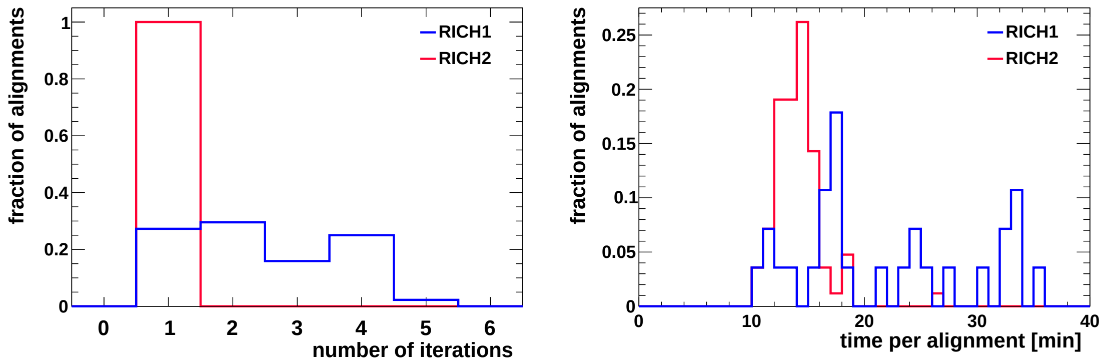


Figure 5.30: Left: *Distribution of iterations needed for the alignment procedure of RICH1 (blue) and RICH2 (red) to converge.* Right: *Distribution of time needed for the alignment procedure of RICH1 (blue) and RICH2 (red) to converge.*

[3] A fill takes on average about 12 hours.

Mirror Orientations

Figures 5.31 and 5.32 show the misalignment of the individual mirrors of RICH1 for each alignment in 2016. Figures 5.33 and 5.34 show the misalignment of the individual mirrors of RICH2 for the same period. Since the LHCb conditions database was only updated once at the beginning of the year, all the misalignments are computed with respect to the same conditions database. Thus these graphs can be used to evaluate the mirror orientations throughout the year.

The polarity of the LHCb magnet is changed regularly between *magnet up* and *magnet down* during the data taking. The figures show that the magnet polarity has a clear impact on the misalignments determined by the alignment procedure for RICH1. This effect occurs because the field of the LHCb magnet reaches into the RICH1 detector. Since the magnetic field does not range into the RICH2 detector, no difference can be seen in misalignments between the two magnet polarities.

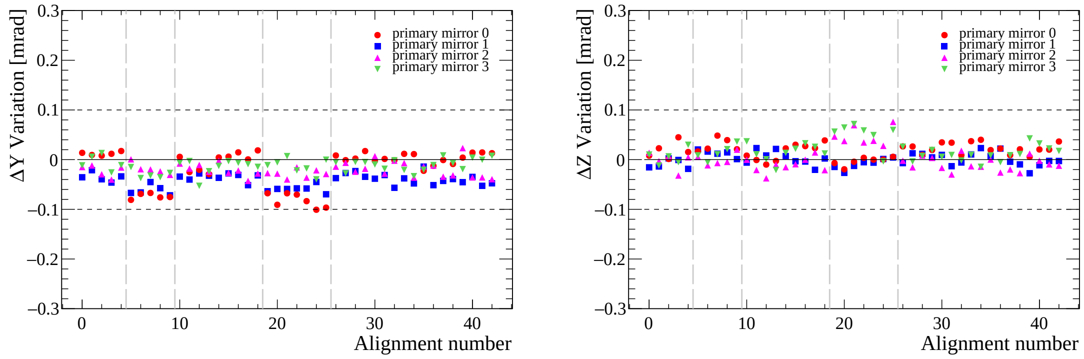


Figure 5.31: *Misalignments of the RICH1 primary mirrors in y (left) and z (right) for the alignments performed throughout 2016. The dashed lines indicate the convergence criteria. The changes of the magnet polarity at alignments 5, 10, 19 and 25 are distinguishable in the y misalignments.*

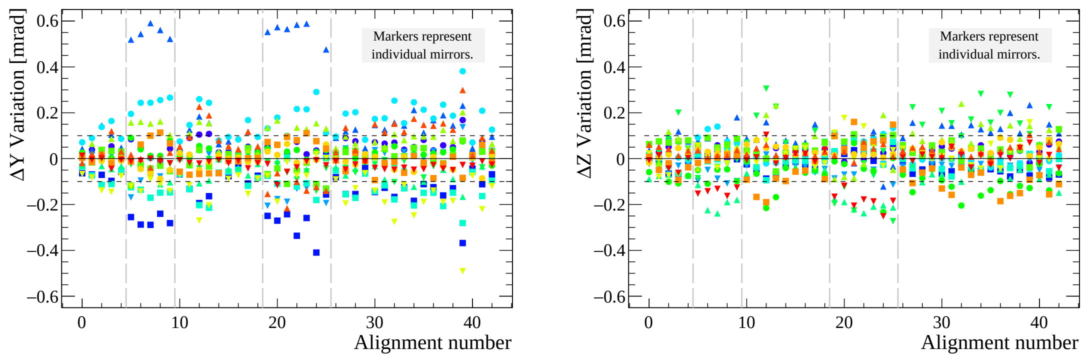


Figure 5.32: *Misalignments of the RICH1 secondary mirrors in y (left) and z (right) for the alignments performed throughout 2016. The dashed lines indicate the convergence criteria. The changes of the magnet polarity at alignments 5, 10, 19 and 25 are clearly distinguishable in the y misalignments.*

Figures 5.33 and 5.34 show the misalignment of the individual mirrors of RICH2 for each alignment in 2016. The procedure and the mirrors are extremely stable and show no dependency on the magnet polarity. The only visible effect on the RICH2 mirror orientations comes from the different data-taking conditions at the beginning of the year[4]. It can be seen in the figures that the first six alignments were performed on data taken under different conditions than the data for following alignments.

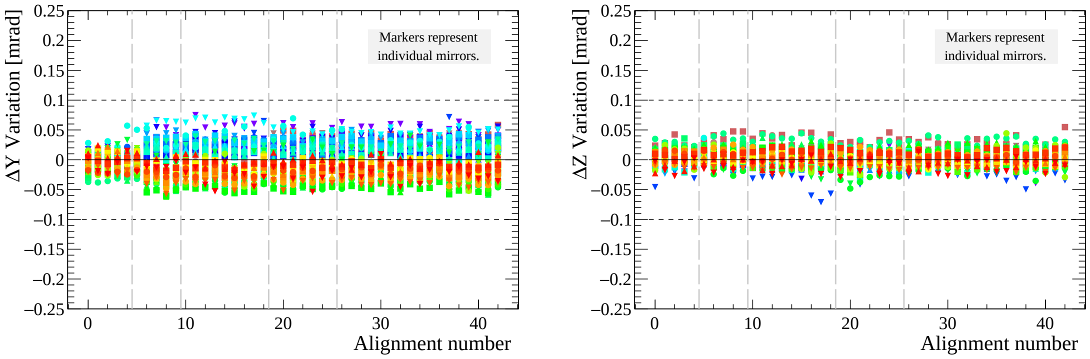


Figure 5.33: *Misalignments of the RICH2 primary mirrors in y (left) and z (right) for the alignments performed throughout 2016. The dashed lines indicate the convergence criteria. The data-taking conditions for the first six alignments differed from the conditions for the rest of the alignments.*

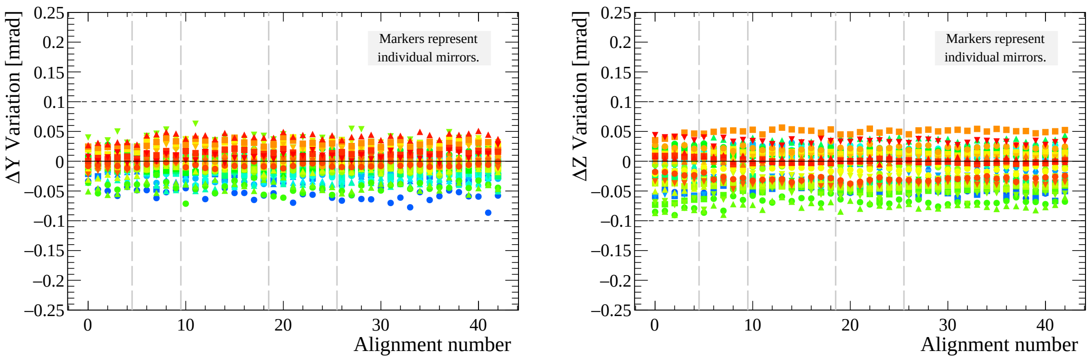


Figure 5.34: *Misalignments of the RICH2 secondary mirrors in y (left) and z (right) for the alignments performed throughout 2016. The dashed lines indicate the convergence criteria. The data-taking conditions for the first six alignments differed from the conditions for the rest of the alignments.*

5.10.1 Cherenkov Angle Resolution

The Cherenkov angle resolutions for RICH1 and RICH2 as computed for the monitoring plots after each alignment are shown in Figure 5.35. The average resolution is 1.71 mrad for RICH1 and 0.66 mrad for RICH2. As for the RICH1 mirror misalignments, the RICH1 Cherenkov angle resolution shows a clear dependence on the magnet polarity. This was already observed in Run I. No such effect can be discerned for RICH2.

[4]The data taking conditions at the beginning of the year include e.g. a lower track multiplicity in the individual events and therefore a lower occupancy in the RICH detectors.

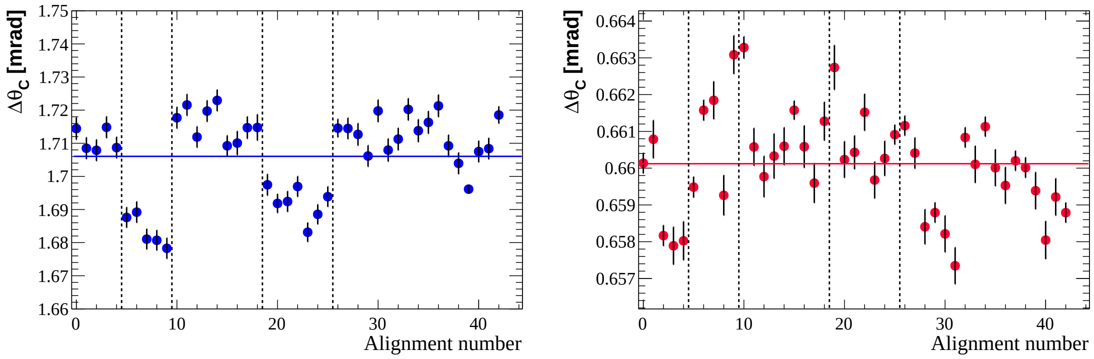


Figure 5.35: *Development of the Cherenkov angle resolution for RICH1 (left) and RICH2 (right) over the data-taking period of 2016. The solid horizontal line shows the average resolution while the dashed vertical lines indicate the magnet polarity switches. The RICH1 Cherenkov angle resolution shows a clear dependence on the magnet polarity while no such effect can be discerned for RICH2.*

5.10.2 Conclusion

The real-time alignment procedure of the RICH mirror systems was very successful during 2016. The results of 44 alignment procedures performed throughout the year were evaluated in order to gain a better understanding of the procedure itself as well as of the data-taking conditions. The alignment procedures took on average 22 min and 14 min for RICH1 and RICH2 respectively which is an acceptable amount of time during the real-time data-taking, and a dramatic improvement from several days in Run I. The mirror orientations determined by the alignment procedures for RICH2 are very stable while the RICH1 mirror orientations show a dependence on the polarity of the LHCb magnet. The average Cherenkov angle resolutions were 1.71 mrad for RICH1 and 0.66 mrad for RICH2. These resolutions are comparable with the minimal resolutions possible to obtain with a perfectly aligned mirror system.

This study led to the decision to update the alignment constants in the LHCb conditions database after each magnet polarity change for the remaining time of Run II.

5.11 Conclusion and Outlook

The alignment of the RICH mirror systems was successfully implemented into the sequence of real-time alignment procedures for LHCb. The alignment procedure has been improved at different points, leading to the procedure taking about 20 min to converge, compared to several days in Run I.

The information gathered by the frequent running of the alignment in 2016 is used to further the understanding of the alignment procedure as well as the understanding of the LHCb detector itself. The development of the mirror misalignments throughout the year (see in Section 5.10) has already been used to optimise the convergence criteria. This allows to obtain maximum precision on the Cherenkov angle while avoiding sensitivity to statistical fluctuations. The same information has also been used to tune thresholds for automatic updates of the LHCb conditions database in 2017.

The RICH mirror alignment will gain even more importance after the upgrade of the LHCb detector, since the overall Cherenkov angle resolution will be smaller. Two main changes of the RICH detectors will contribute to the reduced resolution. The first change is the transition from the hybrid photon detectors to multiple-anode photon multipliers (MaPMT) in both RICH detectors. The new photon detectors accept a smaller range of wavelength, thus greatly reducing the uncertainty from chromatic dispersion. The second change concerns the entire optical system of the RICH1 detector, which will be optimised for having only the C_4F_{10} gas as radiator (the aerogel that was removed before Run II will not be reintroduced). The optical system of RICH1 will spread out the image, which reduces the occupancy of the photon detectors as well as increases the size of the Cherenkov rings. The latter effect indirectly reduces the uncertainty from the finite pixel size. Additionally, the tilt of the primary mirrors, will be reduced which reduces the emission point uncertainty. With the reduced Cherenkov angle resolution the alignment procedure has to be more precise to achieve the best possible performance of the particle identification.

6. Study of the sensitivity to the CKM angle γ using $B^{\pm}\to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ decays at LHCb

This chapter is dedicated to the study of the sensitivity to the CKM angle γ that could be obtained with $B^{\pm}\to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ decays at LHCb. In this analysis the sensitivity to γ is obtained by dividing the $D\to 2\pi^{+}2\pi^{-}$ phase space into bins and evaluating the variation of $B^{-}\to D(\to 2\pi^{+}2\pi^{-})K^{-}$ yields and $B^{+}\to D(\to 2\pi^{+}2\pi^{-})K^{+}$ yields over the $2\pi^{+}2\pi^{-}$ bins.

In order to estimate the expected $B^{\pm}\to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ yields for the full Run II dataset of the LHCb experiment, the 2016 dataset is analysed. The resulting yield from the 2016 dataset is then extrapolated to the full, expected Run II luminosity.

The sensitivity study is the performed by simulating the distributions of $B^{-}\to D(\to 2\pi^{+}2\pi^{-})K^{-}$ events and $B^{+}\to D(\to 2\pi^{+}2\pi^{-})K^{+}$ events over the bins and extracting the values of r_B, δ_B and γ with a fit. Three sets of $B^{-}\to D(\to 2\pi^{+}2\pi^{-})K^{-}$ and $B^{+}\to D(\to 2\pi^{+}2\pi^{-})K^{+}$ event yields are simulated, namely the yield expected in 2016, the yield expected in the entire Run II data sample and the yield expected in the Run I and Run II data sample combined. Additionally, the propagated contribution from the uncertainty on the hadronic parameters of the $D^{0}\to 2\pi^{+}2\pi^{-}$ decay is determined.

The first section outlines the strategy used to estimate the sensitivity to γ using simulated samples of $B^{\pm}\to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ decays. The second section lists the data and simulation samples used in this analysis. The third section explains the topology of b hadron decays in proton-proton collisions. The fourth section summarises the LHCb reconstruction and preselection of the $B^{\pm}\to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ candidates. The fifth selection explains the selection criteria applied after the LHCb preselection, including a multivariate analysis technique. The sixth section presents the extraction of the $B^{-}\to D(\to 2\pi^{+}2\pi^{-})K^{-}$ and $B^{+}\to D(\to 2\pi^{+}2\pi^{-})K^{+}$ signal yields. The seventh section describes the methods used to estimate the sensitivity to γ and summarises the results. The last section offers a conclusion of the results.

6.1 Strategy

As outlined Chapter 1, the CKM angle γ can be determined using $B^{\pm}\to DK^{\pm}$ decays where the D meson decays to a final state accessible to both D^{0} and $\overline{D}^{0}$ mesons.

Sensitivity to γ is gained through the interference of the $B^{\pm} \to D^0K^{\pm}$ and the $B^{\pm} \to \overline{D}^0K^{\pm}$ transition. The strength of this interference varies over the D meson decay phase space. This leads to an increased sensitivity to γ when dividing the D meson decay phase space into different regions as opposed to integrating over the entire phase space.

The formalism introduced in Section 1.4.3 of Chapter 1 shows how γ can be measured from the variation of the yields of $B^{\pm} \to D(\to 2\pi^+2\pi^-)K^{\pm}$ decays over the $D \to 2\pi^+2\pi^-$ phase space. If the $D \to 2\pi^+2\pi^-$ phase space is divided into bins, the number of $B^- \to D(\to 2\pi^+2\pi^-)K^-$ events $N_i^{B^-\to DK^-}$ in bin i is given by

$$N_i^{B^-\to DK^-} = h_{B^-}\left(T_{-i}^f r_B^2 + T_i^f + 2\sqrt{T_i^f T_{-i}^f}(c_i^f x_- + s_i^f y_-)\right) \tag{6.1}$$

and the number of $B^+ \to D(\to 2\pi^+2\pi^-)K^+$ events $N_i^{B^+\to DK^+}$ in bin i is given by

$$N_i^{B^+\to DK^+} = h_{B^+}\left(T_i^f r_B^2 + T_{-i}^f + 2\sqrt{T_i^f T_{-i}^f}(c_i^f x_+ - s_i^f y_+)\right) \tag{6.2}$$

where h_{B^-} and h_{B^+} are independent normalisation factors. The T_i^f, c_i^f and s_i^f are the $D^0 \to 2\pi^+2\pi^-$ hadronic parameters that have been determined previously with data recorded by the CLEO-c experiment [46]. The hadronic parameters are used as input to this analysis.

The binning scheme used to divide the $2\pi^+2\pi^-$ phase space is chosen to be the $\Delta\delta$ binning described in Section 1.6 of Chapter 1. The phase space is divided into bins according to the strong-phase difference between $D^0 \to 2\pi^+2\pi^-$ and $D^0 \to 2\pi^+2\pi^-$ decays as predicted by a recently published amplitude model [24]. The binning scheme with $N = 5$ bins is chosen, where the index i in Equations 6.1 and 6.2 goes from $-N$ to N. This binning scheme was also used in the measurement of the hadronic parameters of the $D^0 \to 2\pi^+2\pi^-$ decay in Reference [46].

The distributions of $B^- \to D(\to 2\pi^+2\pi^-)K^-$ events and $B^+ \to D(\to 2\pi^+2\pi^-)K^+$ events over the $2\pi^+2\pi^-$ bins can be measured at the LHCb experiment. Then a fit can be performed to the distributions to extract γ together with r_B and δ_B. The normalisation factors h_{B^-} and h_{B^+} are also free parameters in the fit.

The first step towards measuring the CKM angle γ is the study of the sensitivity to γ that can be obtained with data recorded by the LHCb experiment. Therefore the expected $B^- \to D(\to 2\pi^+2\pi^-)K^-$ and $B^+ \to D(\to 2\pi^+2\pi^-)K^+$ yields have to be estimated. From these yields the distribution of $B^- \to D(\to 2\pi^+2\pi^-)K^-$ and $B^+ \to D(\to 2\pi^+2\pi^-)K^+$ events over the $2\pi^+2\pi^-$ bins can be simulated and γ can be extracted from a fit.

The yield of $B^{\pm} \to D(\to 2\pi^+2\pi^-)K^{\pm}$ events in the Run I data sample has previously been measured to be 1 500 $B^{\pm} \to D(\to 2\pi^+2\pi^-)K^{\pm}$ [6]. The yield of $B^{\pm} \to D(\to 2\pi^+2\pi^-)K^{\pm}$ events in Run II is estimated in the following sections by selecting $B^{\pm} \to D(\to 2\pi^+2\pi^-)K^{\pm}$ candidates from the 2016 data sample and extrapolating the result to the full expected luminosity of Run II. The $B^{\pm} \to D(\to 2\pi^+2\pi^-)K^{\pm}$ candidates are preselected by a dedicated LHCb stripping line, then a further selection is applied that includes a multivariate analysis. The number of $B^- \to D(\to$

$2\pi^+2\pi^-)K^-$ and $B^+ \to D(\to 2\pi^+2\pi^-)K^+$ events in the data sample is extracted by fitting a probability distribution (PDF) to the distribution of the reconstructed mass of the $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ candidates. The PDF contains three components, one signal component and two different background components.

The sensitivity to the CKM angle γ is determined by generating distributions of $B^- \to D(\to 2\pi^+2\pi^-)K^-$ and $B^+ \to D(\to 2\pi^+2\pi^-)K^+$ events over the $2\pi^+2\pi^-$ bins according to the yields of the three LHCb data-taking scenarios. These distributions are varied according to statistical fluctuations to extract the statistical uncertainty on γ. Additionally, there is an uncertainty on γ that comes from the measured uncertainties on the hadronic parameters of the $D \to 2\pi^+2\pi^-$ decay. This uncertainty is determined by randomly varying the hadronic parameters in the fitting procedure within their measured uncertainties and determining γ each time.

6.2 Data and simulation samples

In this section the data samples and the simulation samples used in the estimation of the $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ events yields are presented.

6.2.1 Recorded data sample

The dataset used in this analysis is the $\approx 1.6\,\mathrm{fb}^{-1}$ of proton-proton collision data collected by LHCb in 2016. The center of mass energy of the proton-proton collision is 13 TeV.

6.2.2 Simulated data samples

All samples of simulated data are produced within the LHCb framework. Different Monte Carlo samples are used for different purposes, such as the training of multivariate analysis tool, the estimation of the reconstruction and selection efficiencies and the identification of different sources of background.

At LHCb simulated events are generated using the Gauss software package [77]. Therein, proton-proton collisions are simulated with Pythia [50], the decay of B hadrons is made with EvtGen [37] and the detector simulation is implemented in Geant4 [38].

All samples are generated with a non-resonant amplitude model for the D decay. For the signal Monte Carlo to more accurately match the signal in data, the signal Monte Carlo samples for $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ are reweighted according to Equation 1.31 with the weights $\mathcal{W}(B^- \to DK^-)$ for $B^- \to DK^-$ events given by

$$\mathcal{W}(B^- \to DK^-) = {A^f_{\mathbf{p}}}^2 + r_B^2\, {\bar{A}^f_{\mathbf{p}}}^2 + 2\, A^f_{\mathbf{p}}\, \bar{A}^f_{\mathbf{p}} \left[x_- \cos(\Delta\delta^f_{\mathbf{p}}) + y_- \sin(\Delta\delta^f_{\mathbf{p}})\right] \ . \tag{6.3}$$

The weights $\mathcal{W}(B^+ \to DK^+)$ for $B^+ \to DK^+$ events are defined equivalently according to Equation 1.32 as

$$\mathcal{W}(B^+ \to DK^+) = r_B^2\, {A^f_{\mathbf{p}}}^2 + {\bar{A}^f_{\mathbf{p}}}^2 + 2\, A^f_{\mathbf{p}}\, \bar{A}^f_{\mathbf{p}} \left[x_+ \cos(\Delta\delta^f_{\mathbf{p}}) + y_+ \sin(\Delta\delta^f_{\mathbf{p}})\right] \ . \tag{6.4}$$

The values for the D^0 and $\overline{D}^0$ amplitudes $A^f_{\mathbf{p}}$ and $\bar{A}^f_{\mathbf{p}}$ at the point $\mathbf{p}$ in phase space are taken from the $D^0 \to 2\pi^+2\pi^-$ amplitude model [24]. The values for r_B and δ_B are taken from the UTfit collaboration as 0.1025 and 137.0°, respectively [14]. Three different sets of event weights are produced for three different values for the CKM angle γ to avoid introducing a bias, namely 60°, 70° and 80°. An example of the distribution of events in the plane spanned by the square of the invariant masses of two opposite sign pion pairs $m^2(\pi_1^+\pi_1^-)$ *vs.* $m^2(\pi_2^+\pi_2^-)$ before and after the reweighting procedure is shown in Figure 6.1. The resonant structures from the ρ^0 resonance decaying to two opposite sign pions is clearly visible after the reweighting procedure.

Apart from the signal decay, the decays $B^\pm \to D(\to K^0_S\pi^+\pi^-)K^\pm$ and $B^\pm \to D(\to 2\pi^+2\pi^-)\pi^\pm$ are generated to evaluate their possible contribution to the signal yield and develop veto selections.

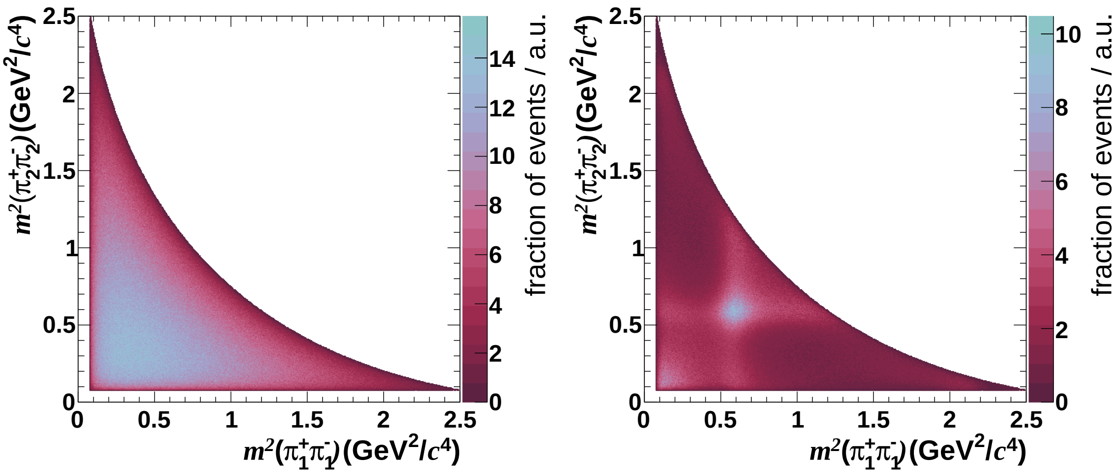


Figure 6.1: *Distribution of events over the plane spanned by $m^2(\pi^+\pi^-)$ vs. $m^2(\pi^+\pi^-)$ before (left) and after (right) the reweighting procedure. The resonant structures from the $\rho^0 \to \pi^+\pi^-$ decay is clearly visible after the reweighting procedure.*

6.3 Topology of the signal decay

A schematic illustration of the topology of $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ decays is shown in Figure 6.2. B hadrons are produced predominantly in the forward region (see Figure 4.1 in Chapter 4). Since they are produced with a large momentum and have a lifetime of 1.6 ps they fly several millimetres before decaying. This results in a detectable secondary vertex.

In $B^\pm \to DK^\pm$ decays a significant amount of the $B^\pm$ mass is transferred to the kinetic energy of its daughters. This means that the D meson with its lifetime of 410 fs also travels a short distance before decaying. Additionally, the $B^\pm$ mesons are created predominantly in the forward direction. The forward boost of the $B^\pm$ is transferred to its daughters meaning that the D decay vertex lies further downstream than the $B^{\pm}$ decay vertex. This can be used to distinguish the signal decay from events with the same or similar final states that do not contain a D meson.

The $B^{\pm}$ meson has a large mass compared to its daughters. This means that even though the $B^{\pm}$ and its daughters are usually boosted in the forward direction, the daughters have a significant transverse momentum component. This can be used to differentiate between the signal decay and combinatorial background. The relatively high transverse momentum of the final state particles and the flight distance of the $B^{\pm}$ and D mesons also leads to comparably large impact parameters with respect to the primary vertex for the D candidates and all final state tracks. Since the $B^{\pm}$ candidate originates from the primary vertex it has a comparably small impact parameter.

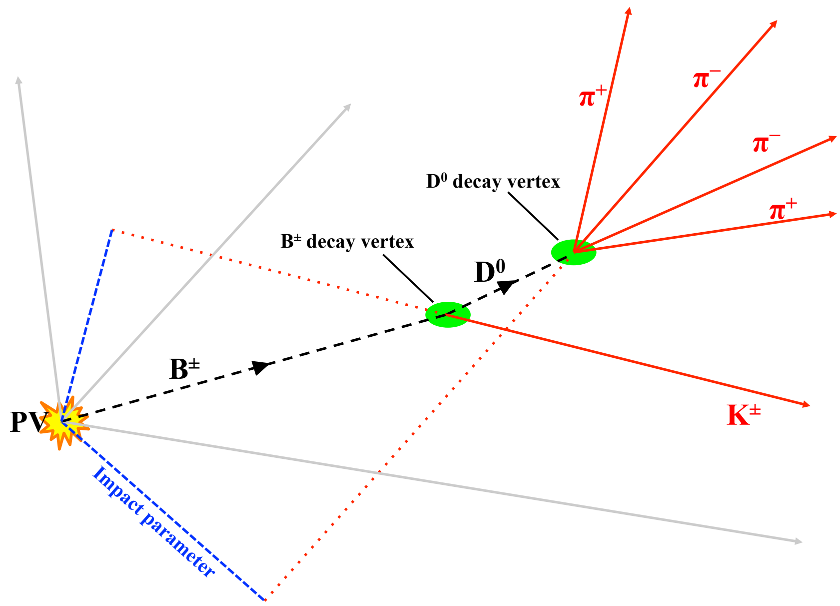


Figure 6.2: *Schematic illustration showing the topology of a $B^{\pm} \to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ decay at LHCb. The flight distances of the $B^{\pm}$ and D mesons are indicated by the black dashed lines. The impact parameters of the final state particles are shown for the $K^{\pm}$ and for one π^{-}. The grey arrows represent particles originating from the primary vertex which do not belong to the signal decay. These particles are also used to constrain the position of the primary vertex.*

6.4 The LHCb Reconstruction and Preselection

This section presents the reconstruction and selection of $B^{\pm} \to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ events within the LHCb framework. The first stage is the LHCb trigger selection and the second stage the LHCb stripping. Additionally, a kinematic fit is performed during the reconstruction of the events to better constrain the position of the signal events in the $2\pi^{+}2\pi^{-}$ phase space.

6.4.1 The LHCb Trigger Selection

The LHCb data acquisition has three trigger stages that were introduced in Section 4.2. For each trigger stage a selection is chosen.

First, the trigger selection requires that the signal candidates trigger the `L0 Hadron` line (see Section 4.2.1) or that the `L0` was triggered independently of the signal decay. The `L0 Hadron` line is activated when one cluster in the hadronic calorimeter passes a transverse energy threshold. The transverse energy of a cluster is calculated from the sum of the transverse energy in the hadronic calorimeter with the transverse energy the corresponding electromagnetic calorimeter cells.

Second, the signal candidates have to trigger the `HLT1 TrackMVA` line or the `HLT1 TwoTrackMVA` line [67]. The first line places selection criteria on individual tracks and selects events with at least one track that passes its selection criteria. First a rectangular selection is applied to the transverse momentum p_T, the quality of the impact parameter χ^2_{IP} and the track fit quality χ^2_{track}/n_{dof} of the track. Subsequently a multivariate analysis tool is used to distinguish signal candidates from background events in the plane spanned by the quality of the impact parameter and the transverse momentum. The `HLT1 TwoTrackMVA` searches for two tracks that form a vertex that is displaced with respect to the primary vertex. First a set of rectangular selection criteria, similar to the `HLT1 TrackMVA`, is applied on the individual tracks. Selection criteria are then placed on the combination of two tracks, such as the vertex fit quality χ^2 and the composite transverse momentum p_T. Subsequently a multivariate analysis tool is used.

Third, the signal candidates are required to trigger one of the four `HLT2 Topological` lines [67]. The `HLT2 Topological` lines were designed to trigger specifically on n-body B decays (n = 2, 3, 4) with at least two charged children and handle the possible omission of child particles. Since the HLT2 performs a full reconstruction of the events, the entire signal decay including all tracks and possible displaced vertices can be evaluated. As a first step the HLT2 topological lines perform a rectangular preselection on variables related to the charged particle tracks, such as transverse momentum p_T and impact parameter significance χ^2_{IP}, and on variables from displaced vertices, such as the quality of the vertex fit χ^2_{SV}. Then a multivariate analysis tool in form of a MatrixNet [78] is used to more efficiently separate signal candidates from background.

6.4.2 The LHCb Preselection: the Stripping Line

The LHCb stripping offers an efficient preselection of the entire dataset. The stripping line used in this analysis is the `B2D0KD2HHHH` line that selects $B^\pm \to DK^\pm$ decays with the D meson decaying to a hadronic four-body final state. The $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ candidates are build by first selecting tracks for charged pion candidates and combining two opposite-sign pion pairs to form a D meson candidate. The D meson candidate is then paired with a charged kaon candidate to give the $B^\pm$ meson candidate.

Selection criteria are applied to all particle momenta p, transverse momenta p_T and impact parameter significances χ^2_{IP}. Selection criteria are also placed on the track parameters of the final state particles, namely on the quality of the track fit χ^2_{track}/n_{dof} and the probability P_{ghost} that a track is not associated to a real particle [79]. Additionally, a loose selection is applied on the discriminant particle identification variable $DLL_{K-\pi}$ of the charged pion candidates (see Section 4.1.2.4).

Further selections are applied to variables of the composite particles such as the angle between the direction of the composite candidate momentum and the direction between the primary vertex and the particle's decay vertex θ_{flight}, masses of the $B^{\pm}$ and D candidates $m_B^{\pm}$ and m_D, the fit quality of the D decay vertex χ^2_{vtx}/N_{dof} and the decay-time $\tau_{B^{\pm}}$ of the $B^{\pm}$ meson. All selection criteria of the LHCb stripping for the $B^{\pm} \to D(\to 2\pi^+2\pi^-)K^{\pm}$ candidates are listed in Table 6.1.

Particle	Parameter	Selection criteria
$\boldsymbol{\pi^{\pm}}$	p	$> 1000\,\mathrm{MeV}$
	p_T	$> 100\,\mathrm{MeV}$
	$DLL_{K-\pi}$	< 20
	χ^2_{track}	< 4
	χ^2_{IP}	> 4.0
	P_{ghost}	< 0.4
$\boldsymbol{K^{\pm}}$	p	$> 5000\,\mathrm{MeV}$
	p_T	$> 500\,\mathrm{MeV}$
	χ^2_{track}	< 4
	χ^2_{IP}	> 4.0
	P_{ghost}	< 0.4
$\boldsymbol{D}$	p_T	$> 1800\,\mathrm{MeV}$
	m_D	$[1764.84, 1964.84]\,\mathrm{MeV}$
	χ^2_{vtx}/N_{dof}	< 10
	θ_{flight}	> 0
$\boldsymbol{B^{\pm}}$	p_T	$> 1000\,\mathrm{MeV}$
	$m_{B^{\pm}}$	$[4750, 7000]\,\mathrm{MeV}$
	$\tau_{B^{\pm}}$	$> 0.2\,\mathrm{ps}$
	χ^2_{IP}	< 25
	θ_{flight}	>0.999

Table 6.1: *Selection criteria for all particles for $B^{\pm} \to D(\to 2\pi^+2\pi^-)K^{\pm}$ candidates in the LHCb stripping framework.*

The combined reconstruction and selection efficiency of the trigger lines and the stripping line is determined on the signal Monte Carlo sample to be 3.5% for the $B^{\pm} \to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ decay.

6.4.3 Kinematic Fitting

The LHCb reconstruction software DaVinci provides the option of performing fits on kinematic objects like tracks. The algorithm for the kinematic fitting is provided by the DecayTreeFitter package [80]. The algorithm parameterises a complete decay chain in terms of vertex positions, decay lengths and momentum parameters. All free parameters are then fitted simultaneously, taking into account the relevant constraints, such as the measured parameters of the final state tracks and photons, and momentum and energy conservation at each vertex.

The kinematic fit is used to more accurately reconstruct the position of the $B^{\pm} \to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ events in the $2\pi^{+}2\pi^{-}$ phase space, or more specifically, in which $2\pi^{+}$ $2\pi^{-}$ bin the individual $B^{\pm} \to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ event lies. For this purpose, in addition to the constraints listed above, the four charged pions are constrained to the nominal D meson mass.

The effect of the kinematic fit on the reconstructed position of the $B^{\pm} \to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ events in the $2\pi^{+}2\pi^{-}$ phase space is quantified with signal Monte Carlo. Without the kinematic fit an average of 92% of the $B^{\pm} \to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ candidates are reconstructed in the correct bin while with the kinematic fit an average of 96% of the candidates are reconstructed in the correct bin. Figure 6.3 shows the purity of events per bin, i.e. the number of events that were produced in a given bin divided by the total number of events reconstructed in that bin.

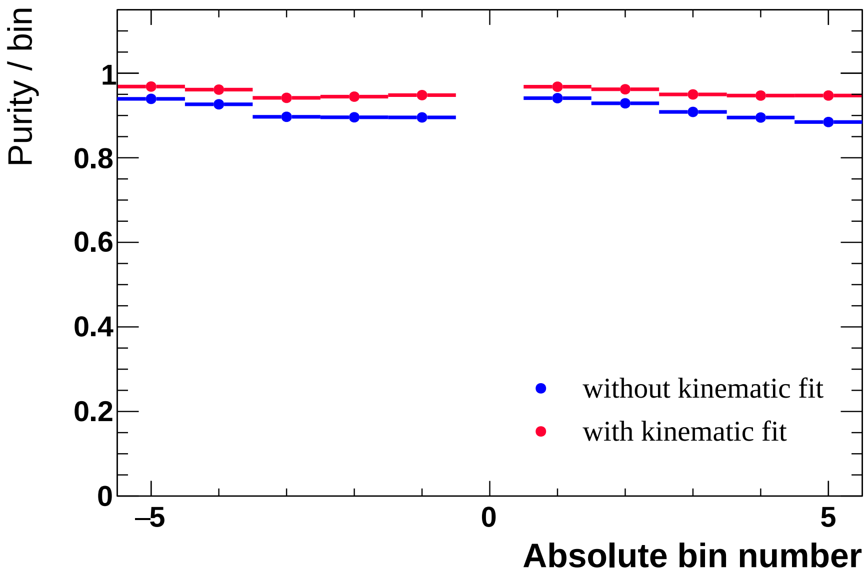


Figure 6.3: *Distribution of purity of events per bin as determined on Monte Carlo for $B^{\pm} \to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ candidates without the kinematic fitting (blue) and with the kinematic fitting (red).*

6.5 Selection of $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ Events for the Sensitivity Study

An additional selection has to be applied to the $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ candidates after the LHCb stripping since the signal to background ratio is so small as can be seen in Figure 6.4. The additional selection of the $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ candidates consists of a preselection, a veto selection against $B^\pm \to D(\to K^0_S\pi^+\pi^-)K^\pm$ decays, a multivariate analysis and a set of selection criteria on the particle identification.

6.5.1 Preselection

Before the Boosted Decision Tree is trained, as explained in Section 6.5.2, a set of rectangular selection criteria is applied. This selection is designed to keep the signal efficiency very high while removing any candidates that are almost certainly background.

Selections are placed on the particle identification discriminant variables ProbNN (see Section 4.1.2.4) of all final state particles as well as on the impact parameter significances χ^2_{IP} of the kaon candidates and the D and $B^\pm$ candidates, the transverse momentum of the D candidate p_T, the flight distance significance χ^2_{FD} of the D and $B^\pm$ candidates and the quality of the kinematic fit χ^2_{DTF}. Additionally, a relatively strict selection is applied to the D mass to further remove $B^\pm \to D(\to K^\pm\pi^\mp\pi^+\pi^-)K^\pm$ decays and events with a random combination of four pions that do not originate from a D meson.

Another discriminating variable is the fight distance significance of the D meson in the z direction, defined as

$$\chi^2_{FD\,z} = \frac{z^D_{decay\ vtx} - z^D_{origin\ vtx}}{\sqrt{\sigma^2(z^D_{decay\ vtx}) - \sigma^2(z^D_{origin\ vtx})}} \tag{6.5}$$

where $z^D_{origin\ vtx}$ is the z coordinate of the point of origin of the D meson candidate and $z^D_{decay\ vtx}$ is the z coordinate of the decay vertex of the D meson candidate. As mentioned in Section 6.3, this quantity can be used to reject events where the final state particles come directly from the $B^\pm$ meson decay, so-called *charmless* decays.

All preselection criteria are listed in Table 6.2. The preselection is evaluated on the signal Monte Carlo sample and the *low-mass* sideband of the data sample defined as $m_{B^-} < 5000\,\text{MeV}$ and the *high-mass* sideband of the data sample defined as $m_{B^-} > 5500\,\text{MeV}$. The high-mass sideband consists mainly of combinatorial background events while the low-mass sideband has a significant contributions from partially reconstructed decays of the type $B \to D^{(*)}X$ containing a real B hadron. The signal efficiency obtained from the Monte Carlo sample is 93.4% while 91.0% of events in the low-mass sideband and 93.2% of events in the high-mass sideband are removed. The greatest loss of signal efficiency comes from the selection on the mass of the D candidates which removes 5.5% of signal decays. The distribution of the $B^\pm$ mass before and after the selection is shown in Figure 6.4.

Particle	Parameter	Selection criteria
$\pi^{\pm}$	$ProbNN_{\pi}$	> 0.02
	$ProbNN_{K}$	< 0.96
$K^{\pm}$	$ProbNN_{K}$	> 0.02
	$ProbNN_{\pi}$	< 0.96
	χ^2_{IP}	> 5
D	p_T	$> 800\,\text{MeV}$
	m_D	$[1839.84, 1889.84]\,\text{MeV}$
	χ^2_{IP}	> 1.5
	χ^2_{FD}	> 50
	$\chi^2_{FD\,z}$	> -4
$B^{\pm}$	χ^2_{IP}	< 20
	χ^2_{FD}	> 30
	θ_{flight}	> 0.9995
	χ^2_{DTF}	< 60

Table 6.2: *Preselection criteria for all particles for $B^{\pm} \to D(\to 2\pi^+2\pi^-)K^{\pm}$ candidates in the LHCb stripping framework. This selection aims at retaining as much signal as possible.*

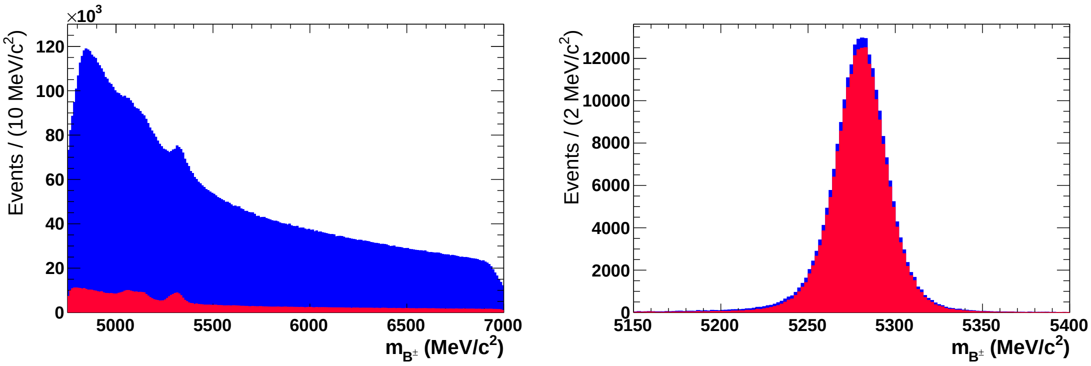


Figure 6.4: *Distribution of events over the $B^{\pm}$ mass range for the 2016 LHCb dataset (left) and signal Monte Carlo (right). The blue distributions are the candidates after the trigger and the LHCb stripping while the red distributions are the candidates after the preselection.*

6.5.2 Boosted Decision Tree

In order to efficiently select $B^{\pm} \to D(\to 2\pi^+2\pi^-)K^{\pm}$ candidates a multivariate analysis (MVA) is used. The MVA is especially powerful in rejecting the combinatorial background, that is background that results from the false combination of tracks in the detector. The MVA chosen for this analysis is the Boosted Decision Tree (BDT) with gradient boost which is implemented in the TMVA software package [81].

Three different BDTs are trained, one for each reweighted $B^{\pm} \to D(\to 2\pi^+2\pi^-)K^{\pm}$ Monte Carlo sample defined in Section 6.2.2. The responses of the three BDTs can be averaged in the selection procedure to avoid biasing the selection toward a given value of γ.

In this section the mathematical formalism of the BDT is explained. Furthermore the input variables, and the data and simulation samples used to train the BDTs are presented. Then the results of the trained BDTs are summarised.

6.5.2.1 Boosted Decision Tree: a Multivariate Method

To separate the signal events from the surplus of background events a set of n discriminative variables is used. Classically each of these variables is examined independently of the others and the total selection of n rectangular cuts is applied.

The multivariate analysis methods combine the information on the different discriminate variables into one single classifier. The selection can have either a linear or non-linear shape in the n-dimensional space of the variables. This principle is illustrated in Figure 6.5 [82].

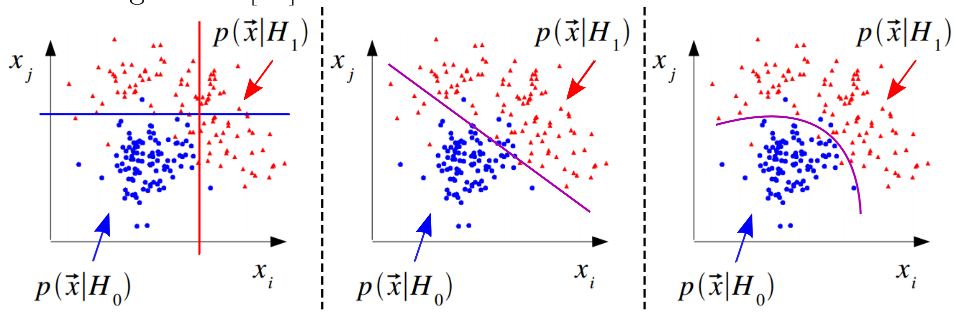


Figure 6.5: *Illustration of the multivariate analysis for two variables x_i and x_j and two data types H_0 and H_1 (for example signal and background). The left plot shows a set of rectangular cuts as used in a classic selection. The middle and the right plot represent a selection from a multivariate analysis where the combination of variables is used to find the optimal selection. The middle plot shows a linear discriminant (Fischer discriminant) and the right plot shows a non-linear discriminant (Boosted Decision Tree, Neural Networks, etc.).* [82]

The multivariate method used in this analysis is the Boosted Decision Tree (BDT). The BDT is a weighted sum of m simple decision trees. Each decision tree classifies a given event with n variables $\mathbf{x}$ as either background or signal with the output of -1 and 1, respectively. The response $F(\mathbf{x})$ of the full BDT is the weighted sum of the responses of the individual decision trees. The weight for the individual tree is determined by how well this tree separates the signal and the background in the training sample. Small values of $F(\mathbf{x})$ indicate that a certain event is more background-like while large values indicate a more signal-like characteristics.

The BDT is trained on a sample of events which are already classified as signal or background. During the training process, each decision tree aims at minimising the weighted misclassification rate. When training the first tree all events are given the same weight. The subsequent tree is trained on a modified event sample where the previously misclassified events are given a bigger weight.

Different methods exists to determine the values of the weights assigned to misclassified events. The BDT used in this analysis implements the *GradientBoost* method. The GradientBoost does not over-penalise misclassified events and therefore performs very well in a noisy environment, that is an environment where some background events tend to look like signal. This makes the GradientBoost BDT very robust with respect to overtraining. A more detailed description of BDTs and the GradientBoost method is provided in Appendix C.

6.5.2.2 The Input Variables

The variables for the BDT must have discriminative power between the combinatorial background and the $B^{\pm} \to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ signal.

In accordance with the description of the topology of the signal decay given in Section 6.3, the variables chosen for the BDT are the transverse momenta p_T of all particles, the logarithms of the impact parameter significances χ^2_{IP} of all particles, the logarithms of the flight distance significances χ^2_{FD} of the $B^{\pm}$ and the D meson candidates, the logarithms of the quality of the D and $B^{\pm}$ decay vertex, the θ_{flight}[1] of the $B^{\pm}$ and the D meson candidates and the quality χ^2_{DTF} of the kinematic fit to the whole decay. An additional variable that has been found to carry discriminative power [6] is the imbalance of transverse momentum I_{p_T} around the $B^{\pm}$ candidate, defined as

$$I_{p_T} = \frac{p_T(B^{\pm}) - \sum p_T}{p_T(B^{\pm}) + \sum p_T} \tag{6.6}$$

where the sum is taken over tracks lying within a cone around the $B^{\pm}$ candidate, excluding the tracks related to the signal. The cone is defined by a circle with a radius of 1.5 units in the plane of pseudorapidity and azimuthal angle (expressed in radians). This variable can be used to select $B^{\pm}$ candidates that are either isolated from the rest of the event, or consistent with a recoil against another b hadron.

All variables used as input to the BDT are listed in Table 6.3 and a selection is displayed in Figure 6.6 for background and signal respectively.

The particle identification variables ProbNN are not used in the training of the BDTs but applied later independently. That is because they are used to reject specific background such as $B^{\pm} \to D(\to 2\pi^{+}2\pi^{-})\pi^{\pm}$ events.

[1] As mentioned in the previous section, θ_{flight} is the angle between the direction of the composite candidate (here $B^{\pm}$ or D meson candidate) momentum and the direction between the primary vertex and the composite particle's decay vertex.

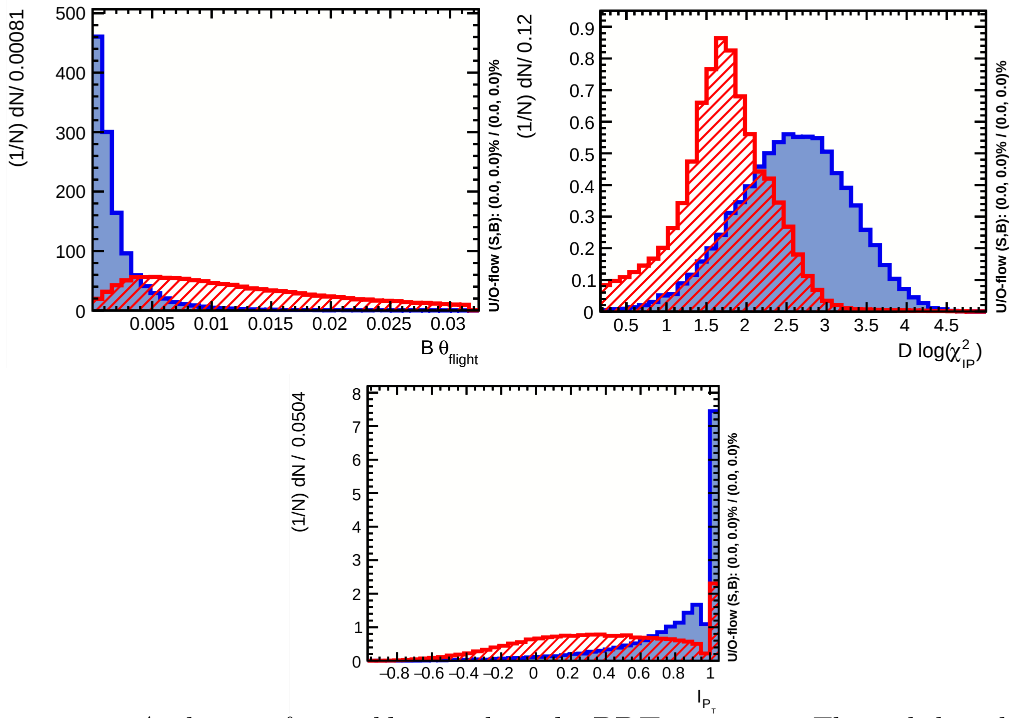


Figure 6.6: *A choice of variables used in the* BDT *training. The red distribution represents the combinatorial background while the blue distribution comes from the signal.* Top left: θ_{flight} *of the* $B^{\pm}$ *meson candidate. This variables carries the most discriminative power.* Top right: *Logarithm of the impact parameter significance of the D meson candidate.* Bottom: *Transverse momentum imbalance around the* $B^{\pm}$ *meson candidate.*

Particle	Variable					
B^0	p_T,	$log(\chi^2_{IP})$,	$log(\chi^2_{FD})$,	$log(\chi^2_{Vertex})$,	θ_{flight}	
D	p_T,	$log(\chi^2_{IP})$,	$log(\chi^2_{FD})$,	$log(\chi^2_{Vertex})$,	θ_{flight},	$\chi^2_{FD\,z}$
K	p_T,	$log(\chi^2_{IP})$				
π	p_T,	$log(\chi^2_{IP})$				
	I_{p_T},	χ^2_{DTF}				

Table 6.3: *Variables used in the Boosted Decision Tree.*

6.5.2.3 The Signal and Background Samples for the BDT Trainings

The signal sample for the BDT is taken from the $B^{\pm} \to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ Monte Carlo. Three different BDTs are trained, one for each of the reweighted Monte Carlo samples presented in Section 6.2.2.

The background sample is taken from the data. The BDT is very effective against combinatorial background and should hence be trained on combinatorial background only. Therefore the data for the background is taken from the upper sideband of the

2016 data sample. Since the events used in the training of the BDT cannot be used in the later analysis, the sideband is chosen far enough away from the signal region to not be needed in the fit to extract the signal yields (see Section 6.6). Therefore the upper sideband defined as any candidate with a reconstructed $B^{\pm}$ mass greater than $5900\,\mathrm{GeV}/c^2$.

Both the signal sample and the background sample are halved. One half is used for the training of the BDTs while the other half is used to test the performance of the BDTs, particularly to test if a BDT was overtrained.

6.5.2.4 Results of the BDT Trainings

Figure 6.7 shows the response of the trained BDTs on the signal and the background. It can be seen from the agreement of the distributions of the testing and the training samples that the BDTs were not overtrained.

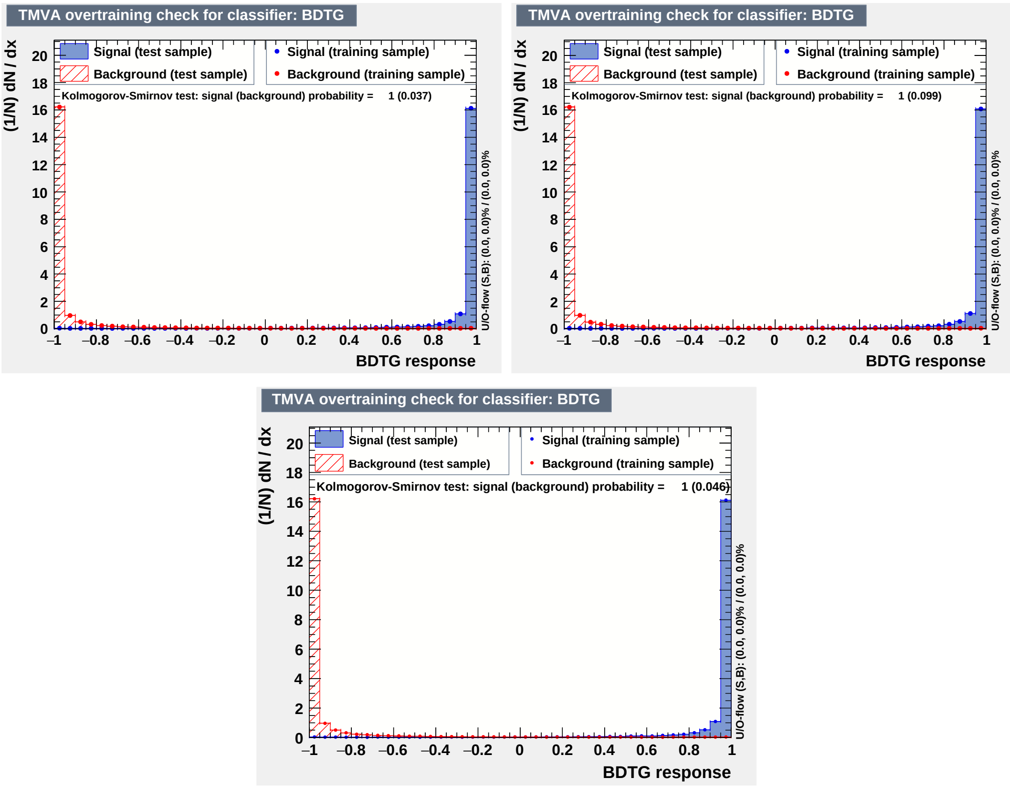


Figure 6.7: *Response on the signal and the background of the* BDTs *trained on the signal Monte Carlo weighted with $\gamma = 60°$ (top left), $\gamma = 70°$ (top right) and $\gamma = 80°$ (bottom). The graphs show the response on the background training sample (red dots), the response on the background testing sample (red area), the response on the signal training sample (blue dots) and the response on the signal testing sample (blue area). The response for the testing and the training samples coincide for signal and background respectively which implies that the* BDTs *were not overtrained.*

Figure 6.8 shows the background rejection efficiency as a function of the signal selection efficiency obtained from applying the trained BDTs on the testing samples.

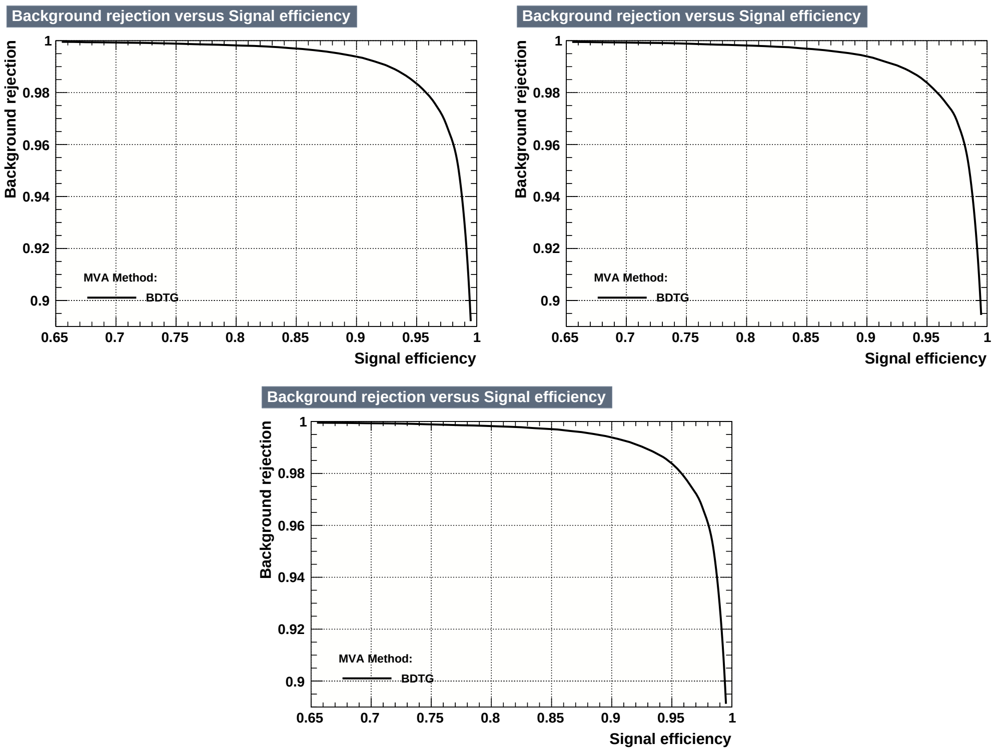


Figure 6.8: *Background rejection efficiency as a function of the signal selection efficiency obtained from the testing samples for the* BDTs *trained on the signal Monte Carlo weighted with* $\gamma = 60°$ *(top left),* $\gamma = 70°$ *(top right) and* $\gamma = 80°$ *(bottom).*

6.5.3 Selection of $B^{\pm} \to D(\to 2\pi^+ 2\pi^-)K^{\pm}$ Events for the Sensitivity Study

In order to estimate the number of $B^{\pm} \to D(\to 2\pi^+ 2\pi^-)K^{\pm}$ decays in the 2016 LHCb data sample further selection criteria are placed on the particle identification variable ProbNN of the final state particles, on the value of the BDT responses and on the variable $\chi^2_{FD\,z}$ defined in Equation 6.5. The particle identification discriminant variables from the neural network for the pion hypothesis ProbNN_π and for the kaon hypothesis $\mathrm{ProbNN}_\mathrm{K}$ are combined for the charged pion candidates in the final state according to

$$\mathrm{ProbNN}_\pi \cdot (1 - \mathrm{ProbNN}_K)\ . \tag{6.7}$$

The selection on the particle identification information of the charged kaon is discussed in the context of the $B^{\pm} \to D(\to 2\pi^+ 2\pi^-)\pi^{\pm}$ veto selection in Section 6.5.5.

The values of the three BDTs– each corresponding to a different value of γ used in the reweighting procedure for the signal Monte Carlo – are averaged according to

$$BDT^{av} = \frac{BDT_{\mathcal{W}_1} + BDT_{\mathcal{W}_3} + BDT_{\mathcal{W}_3}}{3}\ . \tag{6.8}$$

Figure 6.9 shows the distributions of the combined ProbNN variable for one of the charged pions, the averaged response of the BDTs and the variable $\chi^2_{FD\,z}$ for the signal Monte Carlo sample, the low $B^{\pm}$ mass sideband and the high $B^{\pm}$ mass sideband in data. The vertical black lines indicate the chosen selection criteria. The combination of these three selection criteria removed 84% of candidates in the low $B^{\pm}$ mass sideband and 98% of candidates in the high $B^{\pm}$ mass background, while accepting 70% of signal decays. The values for the selection criteria are also summarised in Table 6.4.

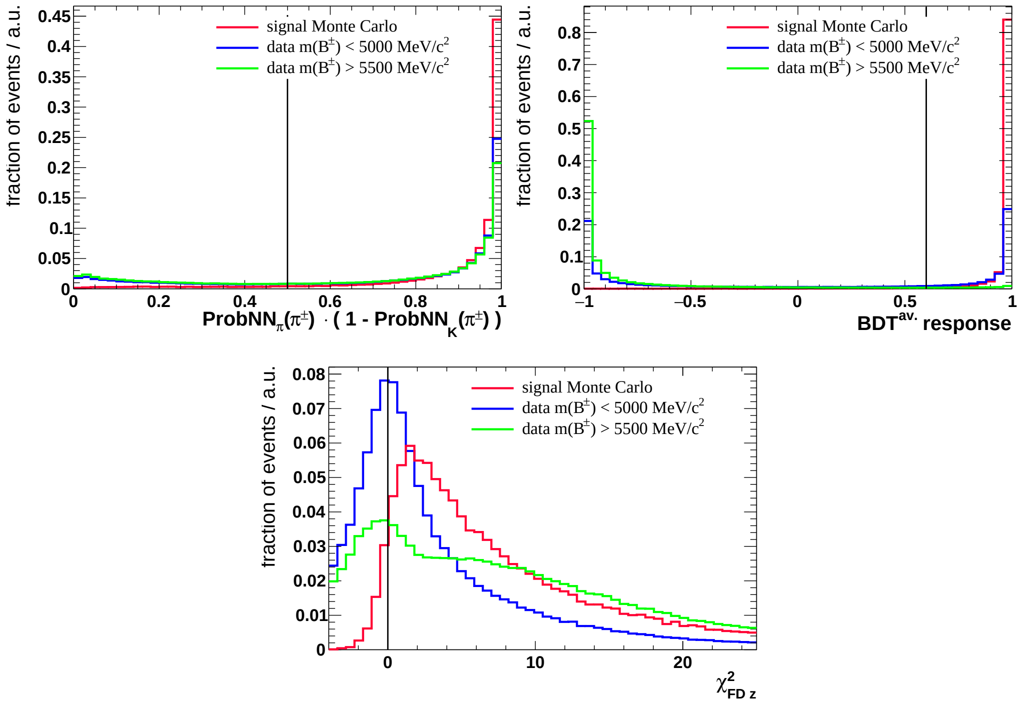


Figure 6.9: *Distributions of the discriminant variables used in the final selection of the $B^{\pm} \to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ candidates for the signal Monte Carlo sample (red), the low $B^{\pm}$ mass background (blue) and the high $B^{\pm}$ mass background (green) in data. The vertical black line indicates the chosen selection criterion.* Top left: *Distribution of the combined* ProbNN *variable for one of the charged pions.* Top right: *Distribution of the averaged* BDT *responses.* Bottom: *Distribution of $\chi^2_{FD\,z}$.*

Particle	Parameter	Selection Criterion
$\pi^{\pm}$	$\mathrm{ProbNN}_{\pi} \cdot (1 - \mathrm{ProbNN}_{K})$	> 0.55
D	$\chi^2_{FD\,z}$	> 0
	$\frac{BDT_{\mathcal{W}_1}+BDT_{\mathcal{W}_2}+BDT_{\mathcal{W}_3}}{3}$	> 0.6

Table 6.4: *Selection criteria for the $B^{\pm} \to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ events.*

6.5.4 The $B^{\pm} \to D(\to K^0_S\pi^+\pi^-)K^{\pm}$ Veto Selection

A selection is applied to remove candidates where the D meson decayed into $K^0_S\pi^+\pi^-$ instead of $2\pi^+2\pi^-$. Since the branching ratio of $D^0 \to K^0_S\pi^+\pi^-$ is four times bigger than the branching ratio of $D^0\to 2\pi^+2\pi^-$, $D\to K^0_S\pi^+\pi^-$ is an important background. The LHCb `B2D0KD2HHHH` stripping line and the preselection from Section 6.5.1 already removes 99.90% of $B^{\pm}\to D(\to K^0_S\pi^+\pi^-)K^{\pm}$ events. The K^0_S mesons in the remaining sample have a short flight distance, making the $D\to K^0_S\pi^+\pi^-$ decay look like a $D\to 2\pi^+2\pi^-$ decay.

In accordance with the selection used in the measurement of the hadronic parameters of $D^0\to 2\pi^+2\pi^-$ from Reference [22], a rectangular selection is applied to the invariant mass $m(\pi^+\pi^-)$ of each pair of opposite sign pions. The region around the K^0_S mass is excluded with $480\,\mathrm{MeV}/c < m(\pi^+\pi^-) < 505\,\mathrm{MeV}/c$. The distribution of events over the invariant mass range of one example opposite-sign pion pair for $B^{\pm}\to D(\to 2\pi^+2\pi^-)K^{\pm}$ decays and $B^{\pm}\to D(\to K^0_S\pi^+\pi^-)K^{\pm}$ decays reconstructed as $B^{\pm}\to D(\to 2\pi^+2\pi^-)K^{\pm}$ is shown in Figure 6.10.

The $D\to K^0_S\pi^+\pi^-$ veto selection reduces the $B^{\pm}\to D(\to K^0_S\pi^+\pi^-)K^{\pm}$ selection efficiency to 0.002%, while removing 9.5% of the signal decays. This means that the expected contribution from $B^{\pm}\to D(\to K^0_S\pi^+\pi^-)K^{\pm}$ decays is less than 0.5% of the signal yield.

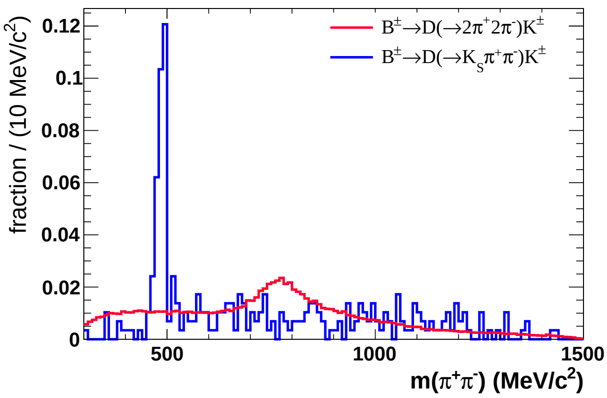


Figure 6.10: *Distribution of events over the $B^{\pm}$ mass range for the $B^{\pm}\to D(\to 2\pi^+2\pi^-)K^{\pm}$ Monte Carlo sample (red) and the $B^{\pm}\to D(\to K^0_S\pi^+\pi^-)K^{\pm}$ Monte Carlo reconstructed as $B^{\pm}\to D(\to 2\pi^+2\pi^-)K^{\pm}$ (blue). The peak in the distribution of the $B^{\pm}\to D(\to 2\pi^+2\pi^-)K^{\pm}$ Monte Carlo at about 770 MeV/c^2 is the ρ^0 resonance.*

6.5.5 The $B^{\pm}\to D(\to 2\pi^+2\pi^-)\pi^{\pm}$ Veto Selection

The $B^{\pm}\to D\pi^{\pm}$ decay is an important source of background since its branching ratio is over ten times larger than the $B^{\pm}\to DK^{\pm}$ branching ratio [1]. The $B^{\pm}\to D\pi^{\pm}$ decay has almost the same topology as the $B^{\pm}\to DK^{\pm}$ decay but a very different *CP* asymmetry. It is therefore important to remove the $B^{\pm}\to D\pi^{\pm}$ events from the $B^{\pm}\to DK^{\pm}$ sample.

Since no particle identification variables are used in the training of the BDT, the BDT cannot distinguish between $B^{\pm}\to D(\to 2\pi^+2\pi^-)K^{\pm}$ events and $B^{\pm}\to D(\to 2\pi^+2\pi^-)\pi^{\pm}$ events. In order to remove $B^{\pm}\to D\pi^{\pm}$ events a selection is applied on the particle identification variables from the neural network ProbNN (see Section 4.1.2.4) of the charged kaon candidate. The variables representing a kaon hypothesis ProbNN_K and the pion hypothesis ProbNN_π are combined according to give

$$\text{ProbNN}_K \cdot (1 - \text{ProbNN}_\pi)\ . \tag{6.9}$$

The selection is studied on the $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ Monte Carlo sample and the $B^\pm \to D(\to 2\pi^+2\pi^-)\pi^\pm$ Monte Carlo sample that is reconstructed as $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ events. The distribution of the combined ProbNN variable of the charged kaon candidate for the $B^\pm \to DK^\pm$ events and the $B^\pm \to D\pi^\pm$ events is shown in Figure 6.11. The selection criterion is chosen to be greater than 0.8 such that it reduces the expected $B^\pm \to D\pi^\pm$ contribution in the data to less than 1% of the $B^\pm \to DK^\pm$ contribution. This selection removes an additional 50% of the signal yields. The effect of the selection is illustrated in Figure 6.12 on the Monte Carlo samples.

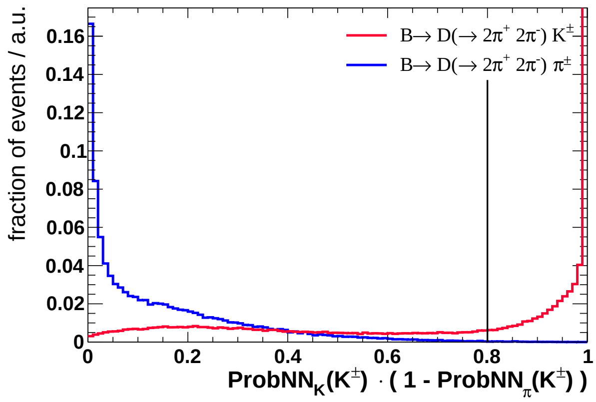


Figure 6.11: *Distribution of the combined* ProbNN *variable for the charged kaon candidate for the signal Monte Carlo sample (red) and the $B^\pm \to D(\to 2\pi^+2\pi^-)\pi^\pm$ Monte Carlo reconstructed as $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ events. The vertical black line indicates the chosen selection criterion.*

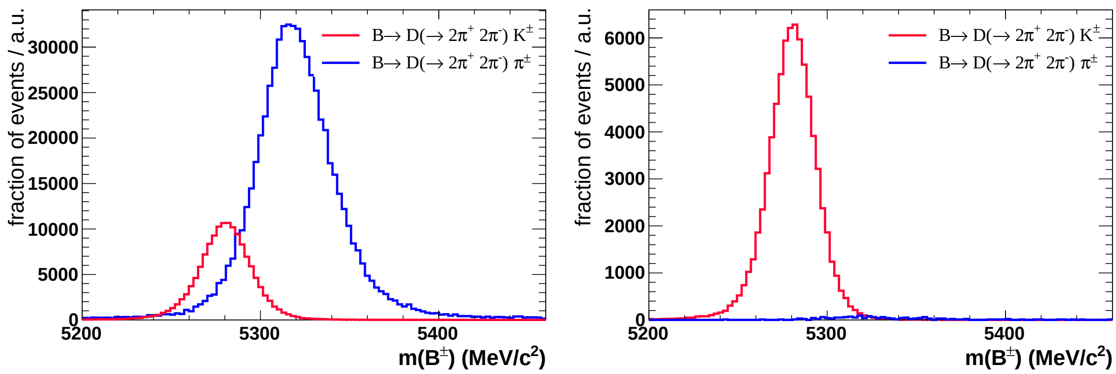


Figure 6.12: *Distribution of the invariant mass of the $B^\pm$ candidate for the signal Monte Carlo sample (red) and the $B^\pm \to D(\to 2\pi^+2\pi^-)\pi^\pm$ Monte Carlo reconstructed as $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ events (blue). The Monte Carlo distributions represent the expected ratio of $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ events and $B^\pm \to D(\to 2\pi^+2\pi^-)\pi^\pm$ events in data before the selection on the combined* ProbNN *variable for the charged kaon candidate (left) and after the selection (right). The expected contribution of $B^\pm \to D(\to 2\pi^+2\pi^-)\pi^\pm$ events is reduced to less than 1% of the $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ yields.*

6.6 Estimation of the $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ Event Yield

In order to study the sensitivity to the CKM angle γ the number of reconstructed and selected $B^- \to D(\to 2\pi^+2\pi^-)K^-$ events and $B^+ \to D(\to 2\pi^+2\pi^-)K^+$ events in the $1.6\,\mathrm{fb}^{-1}$ of data collected by LHCb in 2016 are estimated. To do this, a probability density function (PDF) is fitted to the reconstructed B^+ and B^- mass distributions. The PDFs are composed of three different components, namely the signal component PDF^{signal}, the combinatorial background component PDF^{comb} and the partially reconstructed background component $PDF^{partially}$. The PDFs are summed to give

$$PDF = N_{signal} \cdot PDF^{signal} + N_{comb} \cdot PDF^{comb} + N_{partially} \cdot PDF^{partially} \tag{6.10}$$

where N_{signal}, N_{comb} and $N_{partially}$ are the events yields of the signal, the combinatorial background and the partially reconstructed background in the data sample, respectively. The contributions from $B^\pm \to D(\to 2\pi^+2\pi^-)\pi^\pm$ and $B^\pm \to D(\to K^0_S\pi^+\pi^-)K^\pm$ events is assumed to be negligible and is not accounted for in the fit. The individual PDFs and the total PDF are built and fitted with the RooFit package [83] within Root. The individual PDFs are explained in the following.

6.6.1 The Signal PDF

The PDF of the signal component $PDF^{sig.}$ is represented by the sum of two Gaussian distributions which share the same mean as

$$PDF^{sig.} = a_{frac}G(m_B|\mu, \sigma_1) + (1 - a_{frac})G(m_B|\mu, \sigma_2) \tag{6.11}$$

where the fraction a_{frac} between the two Gaussian distributions is determined from a fit to the signal Monte Carlo. The widths of the Gaussian distributions are related via

$$\sigma_2 = a_\sigma \cdot \sigma_1 \tag{6.12}$$

where the fraction a_σ between the widths of the two Gaussian distributions is also determined from a fit to the signal Monte Carlo. Since there are no distinguishable differences in the B^+ and B^- mass distributions in Monte Carlo, the same shape parameters are used to define their PDFs. The mean of the Gaussians are left free in the fit to the data.

The fit to the signal Monte Carlo is shown in Figure 6.13 and the resulting values for the parameters of the PDF^{sig} are listed in Table 6.5. The extracted values for a_{frac} and a_σ are 0.90 and 2.25, respectively.

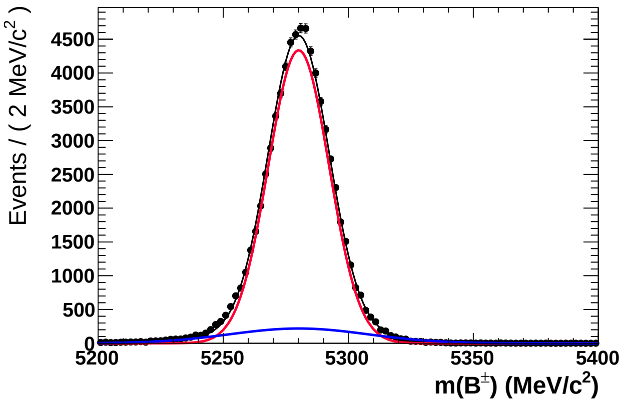


Figure 6.13: *Distribution of the $B^{\pm}$ mass in Monte Carlo (black dots) and the double Gaussian function (black line) fitted to the Monte Carlo. The two individual Gaussian functions are shown as the blue and the red line.*

Parameter	Fit Result
μ	(5280.19 ± 0.05) MeV/c^2
σ_1	(12.12 ± 0.07) MeV/c^2
a_σ	2.25 ± 0.01
a_{frac}	0.90 ± 0.01

Table 6.5: *Result of the fit of the signal component $PDF^{sig.}$ to the $B^{\pm} \to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ signal Monte Carlo sample.*

6.6.2 The Combinatorial Background PDF

The PDF of the combinatorial background is described by an exponential function as

$$PDF^{comb.} = e^{-b\, m_B} \tag{6.13}$$

where the slope b of the exponential function is a free parameter in the fit.

6.6.3 The Partially Reconstructed Background PDF

The partially reconstructed background consists of decays of the type $B \to D^{(*)}X$ where one or more particles are not reconstructed or misidentified. For the purpose of estimating the $B^{\pm} \to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ event yields in 2016 LHCb data sample, the shape of the partially reconstructed background is approximated by a Gaussian distribution

$$PDF^{part.} = G(m_B|\mu, \sigma) \tag{6.14}$$

where all parameters are free in the fit.

6.6.4 Results

Figure 6.14 shows the B^- and B^+ mass distributions in the reconstructed and selected 2016 data sample and Table 6.6 lists the resulting values for the different parameters of the PDF. From the fit, the number of $B^- \to D(\to 2\pi^+2\pi^-)K^-$ decays in the 2016 LHCb data sample is estimated to be 1193 ± 41 and the number of $B^+ \to D(\to 2\pi^+2\pi^-)K^+$ decays is estimated to be 1149 ± 41.

The $B^{\pm} \to D(\to 2\pi^+2\pi^-)K^{\pm}$ yield of $\approx 2\,200$ events in the $1.6\,\mathrm{fb}^{-1}$ of data recorded by LHCb in 2016 surpasses the yield of 1 500 events in the $3\,\mathrm{fb}^{-1}$ of data recorded in Run I. This is due to the increased $b\bar{b}$ production cross section at the center-of-mass energy of 13 TeV in Run II with respect to the 7 and 8 TeV in Run I. Additionally, the new trigger scheme in Run II allows a full offline-like reconstruction in the high level trigger scheme. In combination with the full alignment and calibration of the LHCb detector, this leads to increased selection efficiencies.

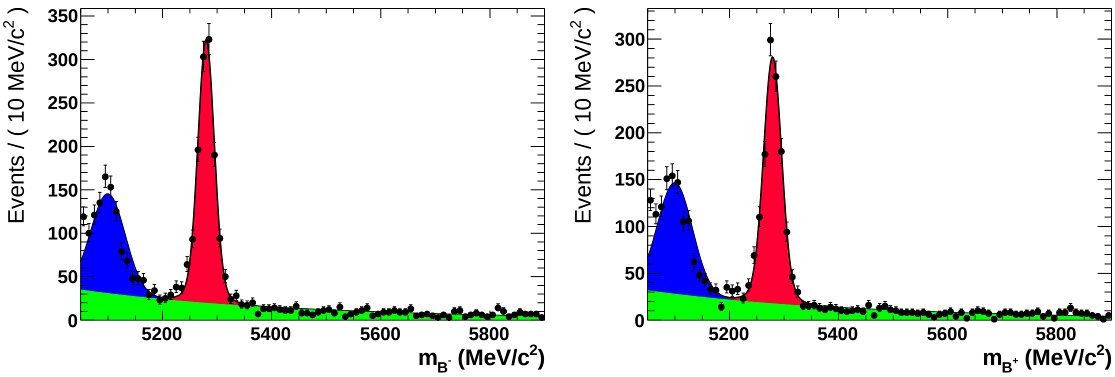


Figure 6.14: *Distribution of the B^- (left) and B^+ (right) invariant mass of the $B^{\pm} \to D(\to 2\pi^+2\pi^-)K^{\pm}$ candidates that were reconstructed and selected from the 2016 LHCb data sample. The black dots represent the data while the black line shows the fitted PDF. The individual PDFs for the the signal (red area), the combinatorial background (green area) and the partially reconstructed background (blue area) are also shown.*

Parameter	B^- Fit Result	B^+ Fit Result
$N^{sig.}$	1193.51 ± 41.26	1149.57 ± 40.55
$\mu^{sig.}$	(5279.85 ± 0.55) MeV/c^2	(5278.87 ± 0.62) MeV/c^2
$\sigma_1^{sig.}$	(14.85 ± 0.55) MeV/c^2	(16.49 ± 0.60) MeV/c^2
$N^{comb.}$	1238.66 ± 73.05	1077.38 ± 68.28
$b^{comb.}$	$(-2.36 \pm 0.23) \cdot 10^{-3}$	$(-2.26 \pm 0.25) \cdot 10^{-3}$
$N^{part.}$	857.82 ± 55.50	922.98 ± 52.59
$\mu^{part.}$	(5090.44 ± 2.15) MeV/c^2	(5086.75 ± 2.55) MeV/c^2
$\sigma^{part.}$	(33.25 ± 2.79) MeV/c^2	(35.65 ± 2.82) MeV/c^2

Table 6.6: *Result of the fit to B^- (left) and B^+ (right) invariant mass distributions of the $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ candidates that were reconstructed and selected from the 2016 LHCb data sample.*

6.7 Extraction of the Sensitivity on the CKM Angle γ

In order to determine the sensitivity of the $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ events collected by LHCb to γ a series of pseudo-experiments is used. First a distribution of $B^- \to D(\to 2\pi^+2\pi^-)K^-$ yields $N_i^{B^-\to DK^-}$ and $B^+ \to D(\to 2\pi^+2\pi^-)K^+$ yields $N_i^{B^+\to DK^+}$ in each bin of the $2\pi^+2\pi^-$ phase space is generated according to

$$N_i^{B^-\to DK^-} = h_{B^-}\left(T^f_{-i} r_B^2 + T^f_i + 2\sqrt{T^f_i T^f_{-i}}(c^f_i x_- + s^f_i y_-)\right) \tag{6.15}$$

and

$$N_i^{B^+\to DK^+} = h_{B^+}\left(T^f_i r_B^2 + T^f_{-i} + 2\sqrt{T^f_i T^f_{-i}}(c^f_i x_+ - s^f_i y_+)\right) \ , \tag{6.16}$$

where h_{B^-} and h_{B^+} are independent normalisation factors and the T^f_i, c^f_i and s^f_i are the $D^0 \to 2\pi^+2\pi^-$ hadronic parameters that have been determined previously with data recorded by the CLEO-c experiment [46]. The input values for the generation of the $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ distributions over the $2\pi^+2\pi^-$ bins are chosen to be $r_B = 0.1$, $\delta_B = 140°$ and $\gamma = 70°$. The normalisation factors h_{B^-} and h_{B^+} are calculated such that the total number of $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ decays corresponds to the number expected in one of the three LHCb data-taking periods considered.

To generate toy datasets, the expected number of decays in each bin (found from Equation 6.15 and Equation 6.16) is varied 20 000 times according to a Poisson distribution. A fit is performed to each toy dataset to extract the normalisation factors, r_B, δ_B and γ.

The fit is performed by minimising the total χ^2, defined by

$$\chi^2 = \sum_i^{bins} \frac{\left(N_i^{B^-\to DK^-} - N_i^{B^-\to DK^-\,fit}\right)^2}{\sigma^2(N_i^{B^-\to DK^-})} \tag{6.17}$$

$$+ \sum_i^{bins} \frac{\left(N_i^{B^+\to DK^+} - N_i^{B^+\to DK^+\,fit}\right)^2}{\sigma^2(N_i^{B^+\to DK^+})} \,. \tag{6.18}$$

Constraining the hadronic parameters of the D decay in the fit introduces a significant bias [16]. Therefore the hadronic parameters are fixed in the fit and the effect of their uncertainties on γ is determined using an alternative method. In this method the fit for toy dataset is repeated 200 times, where for each fit the hadronic parameters are randomly sampled from their associated covariance matrix.

The sensitivity study is performed for one set of r_B, δ_B and γ values and for three different $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ events yields.

6.7.1 Validation of the Fitting Procedure

The fitting procedure is validated by generating toy datasets with larger numbers of signal decays, specially, 100 000 $B^- \to D(\to 2\pi^+2\pi^-)K^-$ and 100 000 $B^+ \to D(\to 2\pi^+2\pi^-)K^+$.

The distributions of the pulls of the pseudo-experiments — i.e. the difference of the values of r_B, δ_B and γ used to generate the *true* distribution of yields per $2\pi^+2\pi^-$ bin and the fitted values divided by the uncertainty provided by the fit — are shown in Figure 6.15. The means and widths of the pull distributions are listed in Table 6.7. The means of all distributions are close to 0 and their widths are close to 1. This indicates that there is no significant bias in the fitting procedure and that the statistical uncertainty on r_B, δ_B and γ is calculated correctly by the fitter.

Variable	Mean	Standard Deviation
r_B	0.02	0.99
δ_B	0.00	1.00
γ	0.03	1.00

Table 6.7: *Means and widths of the pull distributions for r_B, δ_B and γ of the pseudo-experiments used to validate the fitting procedure.*

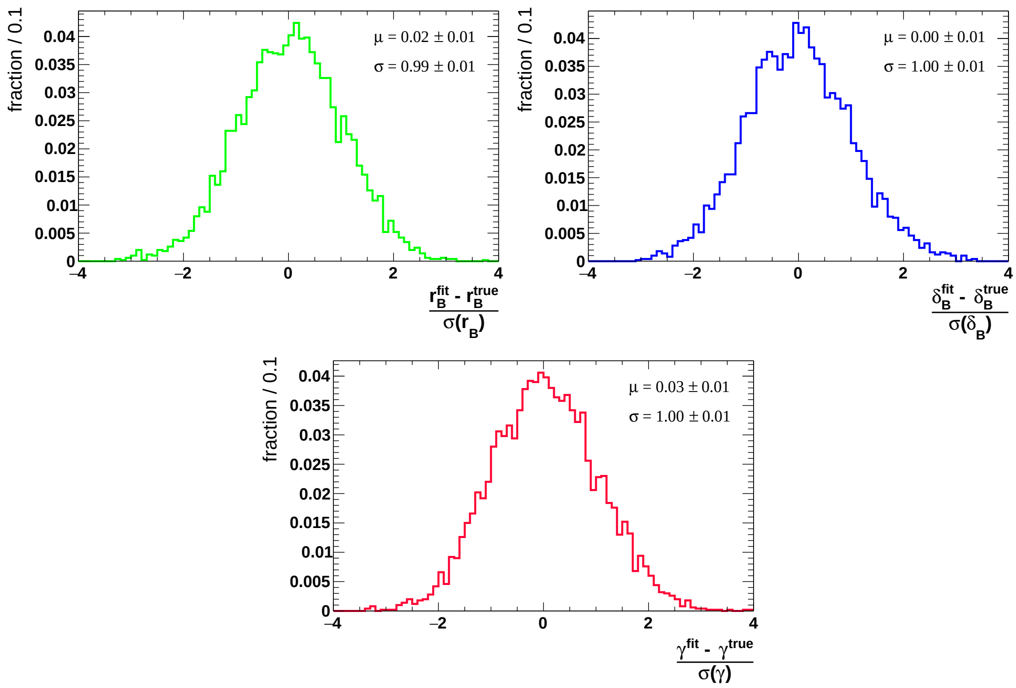


Figure 6.15: *Distributions of the pulls of the pseudo-experiments from the pseudo-experiments used to evaluate the fitting procure for r_B (top left), δ_B (top right) and γ (bottom).*

6.7.2 The Statistical Sensitivity

The sensitivity to the CKM angle γ is estimated by generating toy datasets with the number of signal decays expected in the three LHCb data-taking scenarios, namely the yields expected in 2016 LHCb data sample, in the entire Run II LHCb data sample and in the Run I and Run II LHCb data samples combined.

As shown in Section 6.6, the expected yield of $B^- \to D(\to 2\pi^+2\pi^-)K^-$ decays and $B^+ \to D(\to 2\pi^+2\pi^-)K^+$ decays in the 1.6 fb^{-1} of data collected in 2016 is 1 100 events each. This can be extrapolated to the expected yield for the full Run II luminosity of $\approx$ 5 fb^{-1} to 3 400 events for each b flavour. The $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ yields for Run I were already determined in a previous analysis to be 750 for each b flavour [6]. The event yields used for the three different scenarios are summarised in Table 6.8.

Figure 6.16 shows the difference between the input value and the fitted value for r_B, δ_B and γ for all three scenarios. Table 6.9 lists the means and the widths of these distributions. The statistical uncertainty on γ that can be obtained with the 2016 data is 20° which is reduced to 10° with the expected full Run II dataset and even further reduced to 9° when adding the Run I dataset. The means of the distributions vary from zero which shows that there is a small bias on the fitted values. This bias is 4° for γ for the 2016 data which corresponds to 20% of the standard deviation. For the combined Run I and Run II this bias is reduced to 1° which corresponds to 11% of the standard deviation.

Data Sample	Expected $B^- \to D(\to 2\pi^+2\pi^-)K^-$ / $B^+ \to D(\to 2\pi^+2\pi^-)K^+$ Yields
2016 data	1 100
Run II data	3 400
Run I+ Run II data	4 200

Table 6.8: *$B^- \to D(\to 2\pi^+2\pi^-)K^-$ and $B^+ \to D(\to 2\pi^+2\pi^-)K^+$ events yields used in the three scenarios to estimate the statistical sensitivity to the CKM angle γ with LHCb data.*

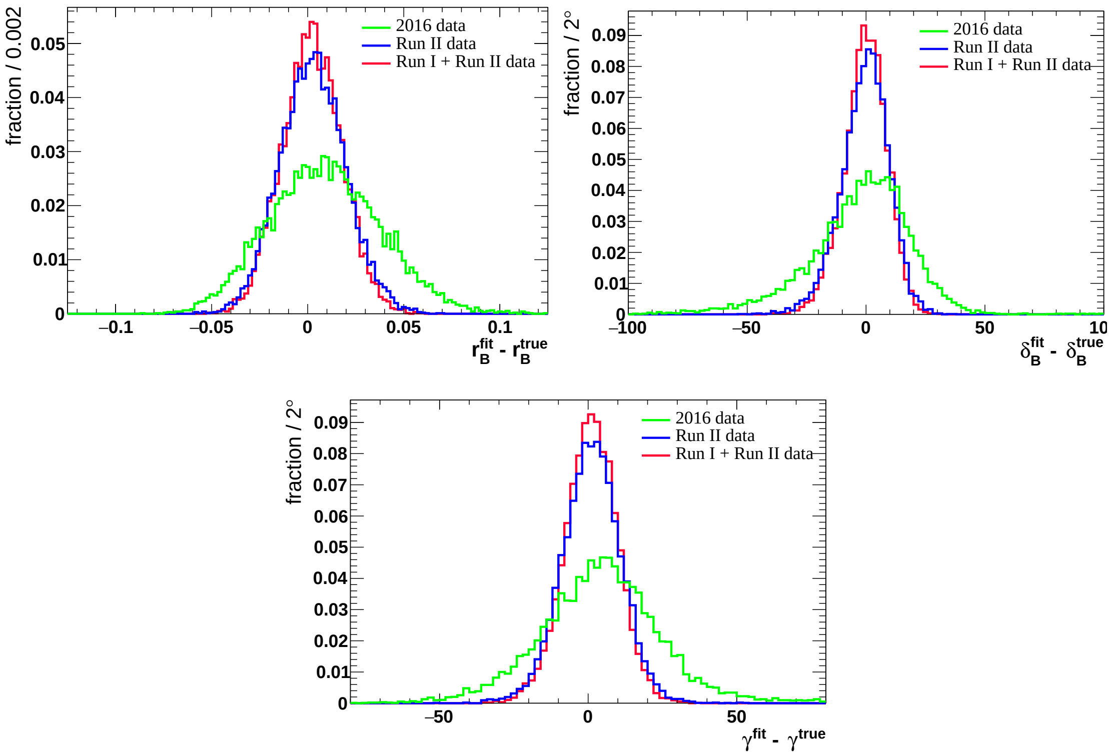


Figure 6.16: *Distributions difference of the input value and the value determined by the fitter of the pseudo-experiments for r_B (top left), δ_B (top right) and γ (bottom). Three different scenarios are shown, namely the results on the 2016 data sample (green line), the expected Run II data sample (blue line) and the combined Run I and Run II data sample (red line).*

Dataset	Variable	μ	σ
2016 data	r_B	0.01 ± 0.00	0.03 ± 0.00
	δ_B	-2.61° ± 0.16°	21.31° ± 0.11°
	γ	4.04° ± 0.15°	20.30° ± 0.10°
Run II data	r_B	0.00 ± 0.00	0.02 ± 0.00
	δ_B	-0.54° ± 0.07°	10.21° ± 0.05°
	γ	1.07° ± 0.07°	9.98° ± 0.051°
Run I + Run II data	r_B	0.00 ± 0.008	0.02 ± 0.00
	δ_B	-0.36° ± 0.07°	9.07° ± 0.05°
	γ	1.00° ± 0.06°	8.92° ± 0.05°

Table 6.9: *Means and widths of the distributions difference of the input value and the value determined by the fitter of the pseudo-experiments.*

Figure 6.17 illustrates the distribution of the fit results in the plane spanned by γ and δ_B for the three different datasets.

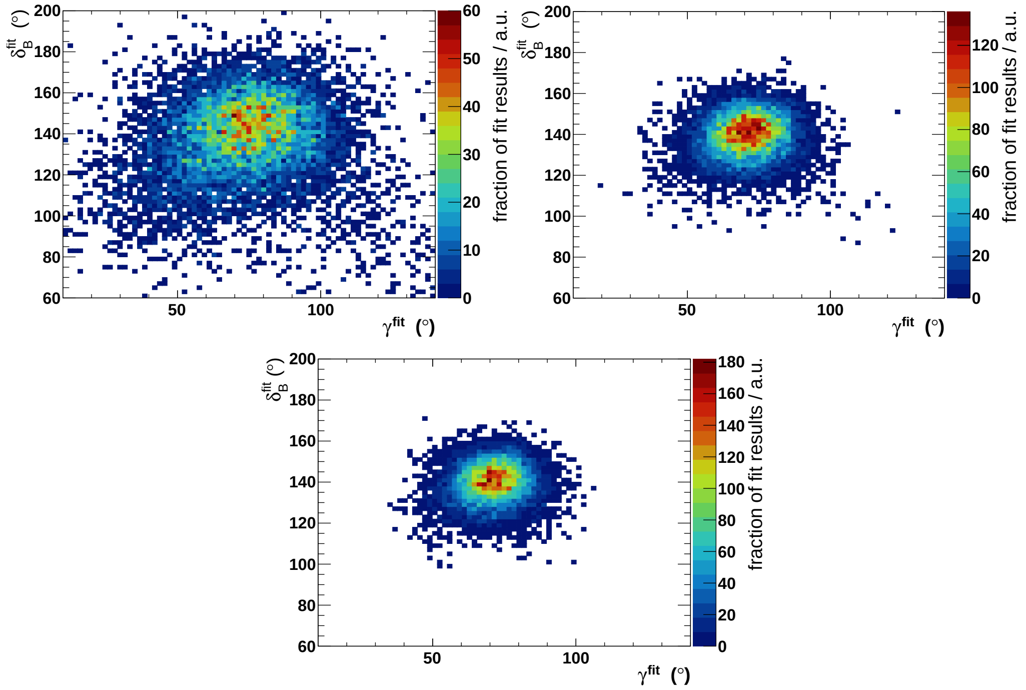


Figure 6.17: *Distribution of the results of the fitting procedure in the plane spanned by γ and δ_B for the 2016 dataset (top left), the expected full Run II dataset (top right) and the combined Run I and Run II dataset (bottom).*

6.7.3 The Systematic Uncertainty from Hadronic Parameters

The systematic uncertainty on γ from the measured uncertainty on the $D \to 2\pi^+2\pi^-$ hadronic parameters is studied by generating toy datasets with the number of signal decays in the 2016 LHCb data sample.

Each generated distribution is then fitted 200 times to extract r_B, δ_B and γ where for each fit the hadronic parameters are varied according to their covariance matrix. The results of r_B, δ_B and γ for each fit are filled into a histogram. The mean of the histogram is chosen as the fitted value for the given dataset and the width of the histogram is taken as the systematic uncertainty on the respective variable r_B, δ_B or γ.

This procedure yields an uncertainty of 0.02 on r_B, 15° on δ_B and 21° on γ. This uncertainty is comparable with the statistical uncertainty on γ for the 2016 data sample but would dominate the uncertainty for any bigger dataset.

6.8 Next Steps Towards the Measurement of the CKM Angle γ

The study of the sensitivity to the CKM angle γ is the first step towards the measurement of γ using the GGSZ approach with $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ decays at LHCb. In order to obtain a final measurement on γ four more steps have to be executed.

1. The $B^+ \to D(\to 2\pi^+2\pi^-)K^+$ events and $B^- \to D(\to 2\pi^+2\pi^-)K^-$ events have to be divided into 5 pairs of bins according to their position in the $D \to 2\pi^+2\pi^-$ phase space.

2. The distribution of the reconstruction and selection efficiency of the $B^+ \to D(\to 2\pi^+2\pi^-)K^+$ and $B^- \to D(\to 2\pi^+2\pi^-)K^-$ decays over the bins has to be determined. This can be done using a mainly data-driven method with an additional correction applied from Monte Carlo. The efficiency is first determined by comparing the reconstructed and selected $B^\pm \to D(\to 2\pi^+2\pi^-)\pi^\pm$ yields with the yields that are expected under the assumption that the *CP* violation is negligible in this channel. Any possible difference due to the pion in the $B^\pm \to D\pi^\pm$ final state can be corrected by comparing the efficiency of $B^\pm \to DK^\pm$ decays and $B^\pm \to D\pi^\pm$ decays in Monte Carlo. Thus, this method reduces the dependency of the calculated efficiency on the Monte Carlo samples.

3. A simultaneous fit to the invariant mass distributions of the B^+ and B^- meson candidates is performed for each bin of the $D \to 2\pi^+2\pi^-$ phase space to extract r_B, δ_B and γ. To this purpose, a fit such as described in Section 6.6 is performed to the B^+ and B^- invariant mass distribution, with the difference that the number of signal events in each bin is described by the product of the reconstruction and selection efficiency with $N_i^{B^-\to DK^-}$ from Equation 6.1 and $N_i^{B^+\to DK^+}$ from Equation 6.2, respectively. The shape of the signal component

can be further constrained using the $B^\pm \to D(\to 2\pi^+2\pi^-)\pi^\pm$ decay channel which has a much higher statistics than the $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ channel. Possible differences between the two channels due to the different hadron in the final state can be compensated with the use of the Monte Carlo samples. Additionally, the shape of the shape of the partially reconstructed background can be further constraint with the use of a Monte Carlo sample that consists of a mixture of all decays that can contribute to this background.

4. As a last step the systematic uncertainties related to this analysis procedure have to be evaluated. These include uncertainties related to the correctness of the Monte Carlos samples used in the analysis ad well as the models that are used to fit the signal and different background components. The effects of possible – albeit very small – *CP* violation in $B^\pm \to D\pi^\pm$ decays have also to be studied and accounted for.

6.9 Conclusion and Outlook

In this chapter a study is performed on the sensitivity to the CKM angle γ that can be obtained using the GGSZ approach with $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ decays at LHCb. In the GGSZ approach the phase space of the D meson decay is divided into bins and the sensitivity to the CKM angle γ is obtained by analysing the variation of the yields of $B^- \to D(\to 2\pi^+2\pi^-)K^-$ decays and $B^+ \to D(\to 2\pi^+2\pi^-)K^+$ decays over the bins. For this, the D decay is parameterised in terms of the hadronic parameters.

The study in this chapter is performed for the expected $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ decay yields for the 2016 LHCb dataset, the expected full Run II LHCb dataset and the combined Run I and Run II LHCb dataset. To this purpose, the yield of $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ decays is extracted from the 2016 LHCb data sample and extrapolated to the full luminosity expected for Run II. The resulting statistical uncertainties on γ are found to be 20°, 10° and 9° for the three datasets, respectively. There is an additional uncertainty on γ due to the measured uncertainty on the $D \to 2\pi^+2\pi^-$ hadronic parameters. This systematic uncertainty is estimated to be 20° and would dominate the uncertainty on γ for any dataset that is bigger than the 2016 dataset.

The current best measurements of γ are from LHCb [16], BaBar [84] and Belle [85]. These analyses use $B^\pm \to D^{(*)}K^{(*)\pm}$ decays where the D meson is reconstructed as $K^0_S\pi^+\pi^-$ and $K^0_S K^+K^-$. The result of the LHCb analysis is $\gamma = (62^{+15}_{-14})^\circ$ [2]. The results of the BaBar analysis is $\gamma = 68^\circ \pm 14^\circ\,(\text{stat}) \pm 4^\circ\,(\text{syst}) \pm 3^\circ\,(\text{model})$ and of the Belle analysis is $\gamma = 78^\circ{}^{+11^\circ}_{-12^\circ}\,(\text{stat}) \pm 4^\circ\,(\text{syst}) \pm 9^\circ\,(\text{model})$, where the last uncertainty is due to the modelling of the D decay. The measurement of γ using the GGSZ approach with $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ decays collected by LHCb in Run II would thus yield a comparable statistical uncertainty but suffer from a much bigger uncertainty from the modelling of the D decay.

The uncertainties on the hadronic parameters could be reduced through additional constraints from BESIII and LHCb. The current BESIII dataset recorded in e^+e^-

[2] The uncertainty quoted in Reference [16] is not divided into individual uncertainties.

collisions at the $\psi(3770)$ resonance is $2.9\,\mathrm{fb}^{-1}$ and a further $7\,\mathrm{fb}^{-1}$ is planned to be recorded in the future. These datasets correspond to 3.5 and 12 times the amount collected by CLEO-c, respectively, and could significantly reduce the uncertainty on the hadronic parameters of the $D \to 2\pi^+2\pi^-$ decay. Additionally, the flavour tagged fractions T_i could be measured at LHCb using $D^{*+} \to D^0\pi^+$ decays and its conjugate where the charge of the pion defines the flavour of the D meson, and with semi-leptonic B decays such as $B^+ \to \overline{D}^0\mu^+\nu_\mu$ decays and its complex conjugate where the charge of the lepton defines the flavour of the D meson. Since the uncertainties on the hadronic parameters are still dominated by the statistical uncertainty, the addition of the LHCb dataset could significantly improve the measurements and make the measurement of γ using the GGSZ method on $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ decays a valuable contribution to the overall constraint on γ.

Conclusion

Three subjects were covered in this thesis, namely the measurement of the *CP*-even fraction of the $D^0 \to 2\pi^+2\pi^-$ decay, the real-time alignment of the LHCb RICH mirror systems and the estimation of the sensitivity to the CKM angle γ that can be obtained using $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ decays at LHCb.

The measurement of parameters associated with *CP* violation in the Standard Model is currently one of the principal objectives in flavour physics. Many of these parameters, such as the CKM angle γ, appear as complex phases in the theory and are difficult to determine. It is therefore important to measure these parameters in many different and independent analyses and to combine and compare the results. The work in this thesis is a step towards the measurement of the CKM angle γ using $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ decays at LHCb. The formalism developed in Chapter 1 shows that the CKM angle γ appears in combination with quantities related to the D meson decay. In order to avoid the difficulty of modelling the D meson decay amplitude over the five-dimensional phase space of the $D^0 \to 2\pi^+2\pi^-$ decay, the model-independent approach of the GGSZ method is chosen in this thesis. In this approach the phase space of the D meson decay is divided into bins and the $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ decay width is integrated over each bin. The resulting D meson decay related quantities are called hadronic parameters and can be measured directly.

The first analysis presented in this thesis is the measurement of the *CP*-even fraction $F_{4\pi}^+$ of the $D^0 \to 2\pi^+2\pi^-$ decay. The *CP*-even fraction $F_{4\pi}^+$ was measured using quantum correlated $D\overline{D}$ decays collected by the CLEO-c experiment at Cornell University. In this analysis one of the correlated D mesons was reconstructed as $D \to 2\pi^+2\pi^-$ while the other D meson was reconstructed as a *CP*-mixed final state, namely $D \to K_\mathrm{S}^0\pi^+\pi^-$ or $D \to K_\mathrm{L}^0\pi^+\pi^-$. The phase space of the $D \to K_\mathrm{S,L}^0\pi^+\pi^-$ decays was divided into bins and the *CP*-even fraction of $D \to 2\pi^+2\pi^-$ was measured by analysing the variation of the signal events over the $K_\mathrm{S,L}^0\pi^+\pi^-$ bins. This analysis was the first measurement of the *CP*-even fraction of $D^0 \to 2\pi^+2\pi^-$ using *CP*-mixed tags. The result of the analysis was $F_{4\pi}^+ = 0.755 \pm 0.050\,(\mathrm{stat}) \pm 0.029\,(\mathrm{syst})$ and was been published in Reference [45]. The value of $F_{4\pi}^+$ has already been used in a measurement of the *CP* observables in $B^\pm \to DK^\pm$ and $B^\pm \to D\pi^\pm$ decays with two-and four-body D decays at LHCb [6]. The measurement of $F_{4\pi}^+$ was also the basis for the measurement of the hadronic parameters of the $D \to 2\pi^+2\pi^-$ decay presented in Reference [46].

The second analysis in this thesis is the study of the sensitivity to the CKM angle γ that can be achieved with the GGSZ method applied to $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ decays recorded by LHCb. In this study the distribution of $B^\pm \to D(\to 2\pi^+2\pi^-)K^\pm$ decays over the $2\pi^+\,2\pi^-$ bins is simulated using the expected yield for three LHCb data-taking scenarios, namely 2016 LHCb dataset, the expected Run II LHCb dataset and the combined Run I and Run II LHCb dataset. To this purpose, the yield of $B^{\pm} \to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ decays is extracted from the 2016 LHCb data sample and extrapolated to the full luminosity expected for Run II. The resulting statistical uncertainties on the CKM angle γ are found to be 20°, 10° and 9° for the three datasets, respectively. There is an additional uncertainty on γ due to the measured uncertainty on the $D \to 2\pi^{+}2\pi^{-}$ hadronic parameters. This systematic uncertainty is estimated to be 21° and would dominate the uncertainty on the CKM angle γ for any dataset bigger than the 2016 dataset. The measurement of the CKM angle γ using the GGSZ method on $B^{\pm} \to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ decays is a valuable contribution to the overall constraint on γ.

Particle identification, especially separation of charged pions and kaons, is of crucial importance for the measurement of the CKM angle γ using $B^{\pm} \to D(\to 2\pi^{+}2\pi^{-})K^{\pm}$ decays. In LHCb, the main source of particle identification information for charged kaons and pions are the two ring imaging Cherenkov (RICH) detectors. Both detectors possess intricate optical systems for detecting Cherenkov photons. In order for the particle identification to function optimally, the position of all optical components has to be known to the best precision. This is achieved with the data-driven RICH mirror alignment procedure. As part of this thesis, the LHCb RICH mirror alignment was implemented into the LHCb online computing framework, enabling the mirror alignment to be performed in real-time during the data-taking period. Additionally, the alignment procedure was improved at different points leading to the procedure taking about 20 min to converge, compared to several days in Run I. The information gained in the frequent running of the alignment procedure during the 2016 data-taking period was analysed and has already been used to further optimise the RICH mirror alignment procedure.

A. Appendix

The weak eigenstates for the K^0_S and K^0_L meson can be expressed in terms of the mass eigenstates as

$$|K^0_S\rangle = \frac{1}{\sqrt{2}}\left(|\overline{K}^0\rangle + |K^0\rangle\right)$$
$$|K^0_L\rangle = \frac{1}{\sqrt{2}}\left(|\overline{K}^0\rangle - |K^0\rangle\right) \ . \tag{A.1}$$

This means that the D^0 meson decay amplitude to $K^0_S\pi^+\pi^-$ and $K^0_L\pi^+\pi^-$ can be written as

$$\mathcal{A}(D^0 \to K^0_S\pi^+\pi^-) = \frac{1}{\sqrt{2}}\left(\mathcal{A}(D^0 \to \overline{K}^0\pi^+\pi^-) + \mathcal{A}(D^0 \to K^0\pi^+\pi^-)\right)$$
$$\mathcal{A}(D^0 \to K^0_L\pi^+\pi^-) = \frac{1}{\sqrt{2}}\left(\mathcal{A}(D^0 \to \overline{K}^0\pi^+\pi^-) - \mathcal{A}(D^0 \to K^0\pi^+\pi^-)\right) \tag{A.2}$$

and the decay amplitude of $\overline{D}^0$ mesons to $K^0_S\pi^+\pi^-$ and $K^0_L\pi^+\pi^-$ is given by

$$\mathcal{A}(\overline{D}^0 \to K^0_S\pi^+\pi^-) = \frac{1}{\sqrt{2}}\left(\mathcal{A}(\overline{D}^0 \to \overline{K}^0\pi^+\pi^-) + \mathcal{A}(\overline{D}^0 \to K^0\pi^+\pi^-)\right)$$
$$\mathcal{A}(\overline{D}^0 \to K^0_L\pi^+\pi^-) = \frac{1}{\sqrt{2}}\left(\mathcal{A}(\overline{D}^0 \to \overline{K}^0\pi^+\pi^-) - \mathcal{A}(\overline{D}^0 \to K^0\pi^+\pi^-)\right) \ . \tag{A.3}$$

The amplitudes $\mathcal{A}(D^0 \to K^0\pi^+\pi^-)$ and $\mathcal{A}(\overline{D}^0 \to \overline{K}^0\pi^+\pi^-)$ are doubly Cabibbo suppressed while the amplitudes $\mathcal{A}(D^0 \to \overline{K}^0\pi^+\pi^-)$ and $\mathcal{A}(\overline{D}^0 \to K^0\pi^+\pi^-)$ are Cabibbo favoured. By neglecting the Cabibbo suppressed amplitudes the D^0 decay amplitudes to $K^0_S\pi^+\pi^-$ and $K^0_L\pi^+\pi^-$ become

$$\mathcal{A}(D^0 \to K^0_S\pi^+\pi^-) = \frac{1}{\sqrt{2}}\mathcal{A}(D^0 \to \overline{K}^0\pi^+\pi^-)$$
$$\mathcal{A}(D^0 \to K^0_L\pi^+\pi^-) = \frac{1}{\sqrt{2}}\mathcal{A}(D^0 \to \overline{K}^0\pi^+\pi^-) \tag{A.4}$$

and the $\overline{D}^0$ decay amplitudes to $K^0_S\pi^+\pi^-$ and $K^0_L\pi^+\pi^-$ become

$$\mathcal{A}(\overline{D}^0 \to K^0_S\pi^+\pi^-) = \frac{1}{\sqrt{2}}\mathcal{A}(\overline{D}^0 \to K^0\pi^+\pi^-)$$
$$\mathcal{A}(\overline{D}^0 \to K^0_L\pi^+\pi^-) = -\frac{1}{\sqrt{2}}\mathcal{A}(\overline{D}^0 \to K^0\pi^+\pi^-) \ . \tag{A.5}$$

Under this approximation is follows that

$$\mathcal{A}(D^0 \to K^0_S\pi^+\pi^-) = \mathcal{A}(D^0 \to K^0_L\pi^+\pi^-)$$
$$\mathcal{A}(\overline{D}^0 \to K^0_S\pi^+\pi^-) = -\mathcal{A}(\overline{D}^0 \to K^0_L\pi^+\pi^-) \ . \tag{A.6}$$

B. Appendix

Degeneracy of system of equations for the RICH2 mirror system

For rotations α_p^y, α_p^z of the primary mirror p around y, z respectively, and rotations β_s^y, β_s^z of the secondary mirror s around y, z respectively, the misalignment on the detector plane is expressed as

$$\Theta_{ps}^y = A_{ps}^y \alpha_p^y + B_{ps}^y \beta_s^y + a_{ps}^z \alpha_p^z + b_{ps}^z \beta_s^z$$
$$\Theta_{ps}^z = A_{ps}^z \alpha_p^z + B_{ps}^z \beta_s^z + a_{ps}^z \alpha_p^y + b_{ps}^y \beta_s^y$$

where A_{ps}^y, B_{ps}^y, A_{ps}^z, B_{ps}^z, a_{ps}^y, b_{ps}^y a_{ps}^z and b_{ps}^z are the magnification coefficients. The RICH2 detector has 56 primary and 40 secondary mirrors. This results in a total of 192 unknowns in total (rotations in y and z for each mirror).

The set of equations for Θ^y_{ps} can be written in matrix formalism as

$$\begin{pmatrix} \Theta^y_{0,0} \\ \Theta^y_{1,0} \\ \vdots \\ \Theta^y_{56,0} \\ \Theta^y_{0,1} \\ \Theta^y_{1,1} \\ \vdots \\ \Theta^y_{56,1} \\ \vdots \\ \vdots \\ \Theta^y_{56,40} \end{pmatrix} = \mathbf{Y} \begin{pmatrix} \alpha^y_0 \\ \alpha^y_1 \\ \vdots \\ \alpha^y_{56} \\ \beta^y_0 \\ \beta^y_1 \\ \vdots \\ \beta^y_{40} \\ \alpha^z_0 \\ \alpha^z_1 \\ \vdots \\ \alpha^z_{56} \\ \beta^z_0 \\ \beta^z_1 \\ \vdots \\ \beta^z_{40} \end{pmatrix}$$

and similarly for Θ^z_{ps}.

If the magnification coefficients are the same for every mirror combination, i.e. $A^y_{ps} = A^y$, it can now be shown that the order of the matrices $\mathbf{Y}$ and $\mathbf{Z}$ is 95 each. The matrix can be reduced to 95 independent rows

$$\begin{pmatrix}
A^y & 0 & \cdots & 0 & B^y & 0 & \cdots & 0 & a^z & 0 & \cdots & 0 & b^z & 0 & \cdots & 0 \\
0 & A^y & \cdots & 0 & B^y & 0 & \cdots & 0 & 0 & a^z & \cdots & 0 & b^z & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & A^y & B^y & 0 & \cdots & 0 & 0 & \cdots & a^z & b^z & 0 & \cdots & 0 & \\
A^y & 0 & \cdots & 0 & 0 & B^y & \cdots & 0 & a^z & 0 & \cdots & 0 & 0 & b^z & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
A^y & 0 & \cdots & 0 & 0 & 0 & \cdots & B^y & a^z & 0 & \cdots & 0 & 0 & 0 & \cdots & b^z
\end{pmatrix}$$

Any other row can be expressed as a linear combination of the rows in this matrix.

To show this, consider the row for the combination of primary mirror $i \neq 0$ and secondary mirror $j \neq 0$,

$$\begin{array}{cccccccccccccccccccc}
 & & \mathbf{i} & & & & & \mathbf{56+j} & & & & & \mathbf{56+40+i} & & & & & \mathbf{2 \cdot 56+40+j} & & \\
\big(0 & \cdots & A^y & \cdots & 0 & 0 & \cdots & B^y & \cdots & 0 & 0 & \cdots & a^z & \cdots & 0 & 0 & \cdots & b^z & \cdots & 0\big)
\end{array}.$$

This row can be expressed as the linear combination of three rows from matrix $\mathbf{Y}$ by first adding the row for the combination of primary mirror 0 and secondary mirror j and the row for the combination of primary mirror i and secondary mirror 0

$$\begin{array}{cccccccccccccccccccc}
\mathbf{0} & & \mathbf{i} & & & & & \mathbf{56+j} & & & \mathbf{56+40} & & \mathbf{56+40+i} & & & & & \mathbf{2 \cdot 56+40+j} & & \\
\big(A^y & \cdots & 0 & \cdots & 0 & 0 & \cdots & B^y & \cdots & 0 & a^z & \cdots & 0 & \cdots & 0 & 0 & \cdots & b^z & \cdots & 0\big) \\
 & & & & & & & & & & + & & & & & & & & & \\
\big(0 & \cdots & A^y & \cdots & 0 & B^y & \cdots & 0 & \cdots & 0 & 0 & \cdots & a^z & \cdots & 0 & b^z & \cdots & 0 & \cdots & 0\big) \\
 & & & & & & & & & & = & & & & & & & & & \\
\big(A^y & \cdots & A^y & \cdots & 0 & B^y & \cdots & B^y & \cdots & 0 & a^z & \cdots & a^z & \cdots & 0 & b^z & \cdots & b^z & \cdots & 0\big)
\end{array}$$

The the row for the combination of primary mirror 0 and secondary mirror 0 is subtracted

$$
\begin{array}{cccccccccccccccccccccc}
 & \mathbf{0} & & \mathbf{i} & & & & & \mathbf{56+j} & & & \mathbf{56+40} & & \mathbf{56+40+i} & & & & & \mathbf{2\cdot 56+40+j} & & & \\
\Big(& A^y & \cdots & A^y & \cdots & 0 & B^y & \cdots & B^y & \cdots & 0 & a^z & \cdots & a^z & \cdots & 0 & b^z & \cdots & b^z & \cdots & 0 & \Big) \\
 & & & & & & & & & & & - & & & & & & & & & & \\
\Big(& A^y & \cdots & 0 & \cdots & 0 & B^y & \cdots & 0 & \cdots & 0 & a^z & \cdots & 0 & \cdots & 0 & b^z & \cdots & 0 & \cdots & 0 & \Big) \\
 & & & & & & & & & & & = & & & & & & & & & & \\
\Big(& 0 & \cdots & A^y & \cdots & 0 & 0 & \cdots & B^y & \cdots & 0 & 0 & \cdots & a^z & \cdots & 0 & 0 & \cdots & b^z & \cdots & 0 & \Big)
\end{array}
$$

which yields the desired result. The same counts for the matrix $\mathbf{Z}$.

Thus, having 95 + 95 independent equations for 190 unknowns means that the system of equation degenerate. Now in reality the magnification coefficients for different mirror combinations differ from each other and the system of equations is not truly degenerate. But since the magnification coefficients don't differ by much the solution to the system of equations is unstable.

C. Appendix

C.1 Boosted Decision Tree: a Multivariate Method

To separate the signal events from the surplus of background events a set of n discriminative variables is used. Classically each of these variables is examined independently of the others and a cut for each variable is found that rejects most of the background while keeping as much signal as possible. The total selection is then a rectangular set of n cuts.

The multivariate analysis methods combine the information on the different discriminate variables into one single classifier. The selection has then a linear or non-linear shape in the n-dimensional space of the variables. This principle is illustrated in Figure C.1 [82].

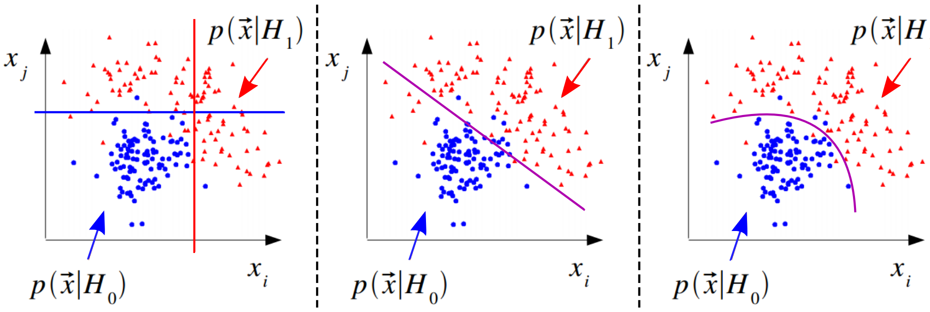


Figure C.1: *Illustration of the multivariate analysis for two variables x_i and x_j and two data types H_0 and H_1 (for example signal and background). The left plot shows a set of rectangular cuts as used in a classic selection. The middle and the right plot represent a selection from a multivariate analysis where the combination of variables is used to find the optimal selection. The middle plot shows a linear discriminant (Fischer discriminant) and the right plot shows a non-linear discriminant (Boosted Decision Tree, Neural Networks, etc.).* [82]

There are different multivariate methods of combining the variables to the final classifier. The multivariate method used in this analysis is the Boosted Decision Tree (BDT). The BDT is a weighted sum of m simple decision trees, called 'basic classifiers' or 'weak learners'. Each decision tree classifies a given event with n variables $\mathbf{x}$ as either background or signal with the output of $f(\mathbf{x}) = -1$ and $f(\mathbf{x}) = 1$ respectively. The response $F(\mathbf{x})$ of the final BDT– called the *boosted event classification* – is then

$$F(\mathbf{x}) = \frac{1}{m} \sum_{i=0}^{m} \ln(\alpha_i) f_i(\mathbf{x}) \tag{C.1}$$

where the α_i represent the weights associated to the i-th tree. Small values of $F(\mathbf{x})$ indicate that a certain events is more background-like while great values indicate a more signal-like structure.

The BDT is trained on a sample of events that each have a label $y(\mathbf{x})$, meaning that they are already classified as signal or background. During the training process, each decision tree aims at minimising the weighted misclassification rate err. When training the first tree all events are given the same weight $\alpha = 1$. The subsequent tree i is trained on a modified event sample where the previously misclassified events are given a weight derived from the previous misidentification rate $err_{(i-1)}$

$$\alpha_i = \frac{1 - err_{(i-1)}}{err_{(i-1)}} \tag{C.2}$$

The entire sample is then renormalised such that the sum of the weights over all events remains normalised to 1.

C.1.1 Gradient Boost

The BDT used in this analysis implements the *GradientBoost* method. The *GradientBoost* performs very well in s noisy environment, that is and environment where some background events tend to look like signal. This is due to the fact that the loss-function (Equation C.4) varies smoothly and does not over-penalise misclassified events. As a consequence the *GradientBoost* is very robust with respect to overtraining.

BDTs implement a *loss-function* $L(F, y)$ which represents the deviation between the BDT response $F(\mathbf{x})$ and the true label $y(\mathbf{x})$ of a certain event. Furthermore, all events are given an individual weight corresponding to their loss-function.
The *GradientBoost* implements a binomial log-likelihood loss

$$L(F, y) = \ln(1 + e^{-2F(\mathbf{x})y}) \tag{C.3}$$

From here the average loss over the whole training sample with k events can be calculated as

$$\langle L \rangle^i = \sum_{l=0}^{k} \omega_l^i L_l^i \tag{C.4}$$

which is the analogous to the misclassification rate err. From $\langle L \rangle^i$ the boosting coefficient α_i for the i-th tree can be computed

$$\alpha_i = \frac{\langle L \rangle^i}{1 - \langle L \rangle^i} \tag{C.5}$$

This boosting factor is used to extract the weight ω_k^i for each event k in the i-th step

$$\omega_k^i = \omega_k^{(i-1)} \cdot \alpha_i^{1 - L_k^{(i-1)}} \tag{C.6}$$

By minimising the loss-function $L^i(F(\mathbf{x}), y)$ in each step, that is for each decision tree, the total BDT is optimised.

Bibliography

[1] C. Patrignani et al. Review of Particle Physics. *Chin. Phys.*, C40(10):100001, 2016.

[2] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012.

[3] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett.*, B716:1–29, 2012.

[4] Stephen P. Martin. A Supersymmetry primer. 1997. [Adv. Ser. Direct. High Energy Phys.18,1(1998)].

[5] R. N. Mohapatra and J. C. Pati. "Natural" left-right Symmetry. *Phys. Rev. D*, 11:2558–2561, May 1975.

[6] R. Aaij and etal. Measurement of CP observables in $B \to DK^{\pm}$ and $B \to D\pi^{\pm}$ with two- and four-body D decays. *Physics Letters B*, 760:117 – 131, 2016.

[7] F. Halzen and A.D. Martin. *Quarks and leptons: an introductory course in modern particle physics.* Wiley, 1984.

[8] Michael E Peskin and Daniel V Schroeder. *An introduction to quantum field theory; 1995 ed.* Westview, Boulder, CO, 1995. Includes exercises.

[9] Paul Langacker. Introduction to the Standard Model and Electroweak Physics. In *Proceedings of Theoretical Advanced Study Institute in Elementary Particle Physics on The dawn of the LHC era (TASI 2008): Boulder, USA, June 2-27, 2008*, pages 3–48, 2010.

[10] Emmy Noether. Invariant variation problems. *Transport Theory and Statistical Physics*, 1(3):186–207, 1971.

[11] Ling-Lie Chau and Wai-Yee Keung. Comments on the Parametrization of the Kobayashi-Maskawa Matrix. *Phys. Rev. Lett.*, 53:1802–1805, Nov 1984.

[12] Lincoln Wolfenstein. Parametrization of the Kobayashi-Maskawa Matrix. *Phys. Rev. Lett.*, 51:1945–1947, Nov 1983.

[13] J. Charles et al. CP violation and the CKM matrix: assessing the impact of the asymmetric B factories. *The European Physical Journal C - Particles and Fields*, 41(1):1–131, May 2005. updated results and plots available at: http://ckmfitter.in2p3.fr.

[14] *Unitarity Triangle analysis in the Standard Model from the UTfit Collaboration* , Proceedings of Science, 2016. updated results and plots available at: http://www.utfit.org/UTfit/WebHome.

[15] Update of the LHCb combination of the CKM angle γ using $B \to DK$ decays. Technical Report LHCb-CONF-2017-004. CERN-LHCb-CONF-2017-004, CERN, Geneva, Jul 2017.

[16] Roel Aaij et al. Measurement of the CKM angle γ using $B^{\pm} \to DK^{\pm}$ with $D \to K^0_S\pi^+\pi^-, K^0_S K^+K^-$ decays. *JHEP*, 10:097, 2014.

[17] *Charm CPV and rare decays at LHCb*, 2016.

[18] Michael Gronau and David London. How to determine all the angles of the unitarity triangle from $B^0 \to DK^0_S$ and $B^0_s \to D\phi$. *Physics Letters B*, 253(3):483 – 488, 1991.

[19] David Atwood, Isard Dunietz, and Amarjit Soni. Enhanced CP Violation with $B \to \mathrm{KD}^0(\overline{D}^0)$ Modes and Extraction of the Cabibbo-Kobayashi-Maskawa Angle γ. *Phys. Rev. Lett.*, 78:3257–3260, Apr 1997.

[20] Anjan Giri, Yuval Grossman, Abner Soffer, and Jure Zupan. Determining γ using $B^{\pm} \to \mathrm{DK}^{\pm}$ with multibody D decays. *Phys. Rev. D*, 68:054018, Sep 2003.

[21] A. Bondar and A. Poluektov. The use of quantum-correlated D^0 decays for ϕ_3 measurement. *The European Physical Journal C*, 55(1):51–56, May 2008.

[22] J. Libby et al. Model-independent determination of the strong-phase difference between D^0 and $\bar{D}^0 \to K^0_{S,L}h^+h^-$ $(h = \pi, K)$ and its impact on the measurement of the CKM angle γ/ϕ_3. *Phys. Rev.*, D82:112006, 2010.

[23] 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. D*, 78:034023, Aug 2008.

[24] P. d'Argent, N. Skidmore, J. Benton, J. Dalseno, E. Gersabeck, S. T. Harnew, P. Naik, C. Prouve, and J. Rademacker. Amplitude analyses of $D^0 \to \pi^+\pi^-\pi^+\pi^-$ and $D^0 \to K^+K^-\pi^+\pi^-$ decays. *Journal of High Energy Physics*, 2017(5):1–57, May 2017.

[25] M. S. Alam and etal. First measurement of the rate for the inclusive radiative penguin decay $b \to s\gamma$. *Phys. Rev. Lett.*, 74:2885–2889, Apr 1995.

[26] R. A. Briere et al. Cleo-c and cesr-c: A new frontier of weak and strong interactions. *CLNS 01/1742, 2001 (unpublished)*, 2001.

[27] D Peterson, K Berkelman, R Briere, G Chen, D Cronin-Hennessy, S Csorna, M Dickson, S von Dombrowski, K.M Ecklund, A Lyon, Sz Marka, T.O Meyer, J.R Patterson, A Sadoff, P Thies, E.H Thorndike, and D Urner. The {CLEO} {III} drift chamber. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*,

478(1â2):142 – 146, 2002. Proceedings of the ninth Int.Conf. on Instrumentation.

[28] M. Artuso, F. Azfar, A. Efimov, S. Kopp, R. Mountain, S. Schuh, T. Skwarnicki, S. Stone, G. Viehhauser, S. Anderson, A. Smith, Y. Kubota, E. Lipeles, T. Coan, J. Staeck, V. Fadeyev, I. Volobouev, and J. Ye. Test beam results from the CLEO III LiF-TEA RICH. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 419(2â3):577 – 583, 1998.

[29] W.R. Leo. *Techniques for Nuclear and Particle Physics Experiements.* Springer-Verlag Berlin Heidelberg New York, 1987.

[30] Martin J. Berger and Stephen M. Seltzer. Tables of energy losses and ranges of electrons and positrons. *NASA SP-3012*, 1964.

[31] Christopher M. Thomas. *Measurement of the strong-phase difference between D^0 and $\overline{D}^0$ decays to K^0_S K^+ K^- at CLEO-c and a determination of observables related to CP violation in $B^\pm \to DK^\pm$ decays at LHCb.* PhD thesis, St Johnâs College, University of Oxford and STFC Rutherford Appleton Laboratory, December 2011.

[32] A. H. Mahmood et al. Measurement of the b-meson inclusive semileptonic branching fraction and electron-energy moments. *Phys. Rev. D*, 70:032003, Aug 2004.

[33] D. Peterson. Charged particle tracking at cornell:gas detectors and reconstruction software, 2005.

[34] S. Dobbs et al. Measurement of absolute hadronic branching fractions of D mesons and $e^+e^- \to D\overline{D}$ cross sections at the $\psi(3770)$. *Phys. Rev. D*, 76:112001, Dec 2007.

[35] D. Rice, G. Codner, G. Dugan, R. Ehrlich, Z. Greenwald, Y. He, S. Henderson, Y. Li, V. Mejidzade, A. Mikhailichenko, N. Mistry, E. Nordberg, J. Rogers, D. Rubin, D. Sagan, E. Smith, A. Temnykh, and M. Tigner. Parameters for low energy operation of cesr. In *PACS2001. Proceedings of the 2001 Particle Accelerator Conference (Cat. No.01CH37268)*, volume 1, pages 374–376 vol.1, 2001.

[36] Sean Brisbane. *CLEO-c $D \to K^0_{S/L}\pi^+\pi^-$ Binned Dalitz-Plot Analyses Optimised for the CKM Angle γ Measurement and the Commissioning of the LHCb RICH Front-End Electronics.* PhD thesis, Oxford U., 2010.

[37] Anders Ryd, David Lange, et al. EvtGen, A Monte Carlo Generator for B -Physics. EvtGen V00-11-06, February 2004.

[38] V.N Ivanchenko. Geant4 toolkit for simulation of HEP experiments. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 502(2-3), April 2003.

[39] B. Aubert et al. Measurement of the Cabibbo-Kobayashi-Maskawa Angle γ in $B^{\mp} \to D^{(*)}K^{\mp}$ Decays with a Dalitz Analysis of $D \to K^0_S\pi^-\pi^+$. *Phys. Rev. Lett.*, 95:121802, Sep 2005.

[40] S. Eidelman et al. Review of Particle Physics. *Physics Letters B*, 592:1+, 2004.

[41] T. Gershon, J. Libby, and G. Wilkinson. Contributions to the width difference in the neutral D system from hadronic decays. *Phys. Lett.*, B750:338–343, 2015.

[42] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand. Parton fragmentation and string dynamics. *Physics Reports*, 97(2):31 – 145, 1983.

[43] Brian K. Heltsley. Fitevt: A user-friendly c++ interface for kinematic fitting. CBX 06-28, 2006.

[44] R.A. Briere et al. First model-independent determination of the relative strong phase between D^0 and $\overline{D}^0 \to K^0_S\pi^+\pi^-$ and its impact on the CKM angle γ/ϕ_3 measurement. *Phys. Rev. D*, 80:032002, Aug 2009.

[45] S. Malde, C. Thomas, G. Wilkinson, P. Naik, C. Prouve, J. Rademacker, J. Libby, M. Nayak, T. Gershon, and R.A. Briere. First determination of the CP content of $D^0 \to \pi^+\pi^-\pi^+\pi^-$ and updated determination of the CP contents of $D^0 \to \pi^+\pi^-\pi^0$ and $D^0 \to K^+K^-\pi^0$. *Physics Letters B*, 747:9 – 17, 2015.

[46] Samuel Harnew, Paras Naik, Claire Prouve, Jonas Rademacker, and David Asner. Model-independent determination of the strong phase difference between D^0 and $\bar{D}^0 \to \pi^+\pi^-\pi^+\pi^-$ amplitudes. 2017.

[47] R. Aaij and etal. Measurement of forward J/ψ production cross-sections in pp collisions at $\sqrt{s} = 13$ Tev. *Journal of High Energy Physics*, 2015(10):172, 2015.

[48] R. Aaij et al. Search for the rare decays $B^0_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$. CERN-PH-EP-2011-029, April 2011.

[49] The LHCb Collaboration. LHCb letter of intent. CERN/LHCC-95-5, August 1995.

[50] Torbjorn Sjostrand, Stephen Mrenna, and Peter Skands. PYTHIA 6.4 Physics and Manual. *JHEP 0605:026,2006*, May 2006.

[51] The LHCb Collaboration. The LHCb detector at the LHC. JINST 3 (2008) S08005 , 2008.

[52] The LHCb Collaboration. LHCb Vertex Locator Technical Design Report. CERN-LHCC-2001-011, May 2001.

[53] The LHCb Collaboration. LHCb Magnet Technical Design Report. CERN-LHCC-2000-007, January 2000.

[54] The LHCb Collaboration. LHCb Inner Tracker Technical Design Report. CERN-LHCC-2002-029, November 2002.

[55] The LHCb Collaboration. LHCb Outer Tracker Technical Design Report. CERN-LHCC-2001-024, September 2001.

[56] The LHCb Collaboration. LHCb RICH Technical Design Report. CERN-LHCC-2000-037, September 2000.

[57] The LHCb Collaboration. LHCb Calorimeters Technical Design Report. CERN-LHCC-2000-036, September 2000.

[58] Herve Terrier and Ivan Belyaev. Particle Identification with LHCb Calorimeters. LHCb 2003-092, August 2003.

[59] The LHCb Collaboration. LHCb Muon SystemTechnical Design Report. CERN-LHCC-2001-010, May 2001.

[60] G ; Felici G ; Murtas F ; Valente P ; Bonivento W (INFN Cagliari) ; Cardini A (INFN Cagliari) ; Lai A (INFN Cagliari) ; Pinci D (INFN Cagliari ; Cagliari U.) ; Saitta B (INFN Cagliari ; Cagliari U.) ; Bosio C (INFN Rome) Bencivenni. A triple-GEM detector with pad readout for the inner region of the first LHCb muon station. LHCb-MUON 2001-05, June 2001.

[61] Jiahang Zhong, Run-Sheng Huang, and Shih-Chang Lee. A program for the bayesian neural network in the {ROOT} framework. *Computer Physics Communications*, 182(12):2655 – 2660, 2011.

[62] Monica Pepe Altarelli. LHCb: detector performance and first physics results, 2011.

[63] Pablo Rodriguez Perez, for the LHCb VELO Group, and for the VELO Upgrade group. The LHCb Vertex Locator performance and Vertex Locator upgrade, October 2012.

[64] R. etal Aaij. Lhcb detector performance. *International Journal of Modern Physics A*, 30(07):1530022, 2015.

[65] The LHCb Collaboration. LHCb Trigger SystemTechnical Design Report. CERN-LHCC-2003-066, September 2003.

[66] Massimiliano Ferro-Luzzi and The LHCb Collaboration. The LHCb Trigger Strategy and Performance. ECONF C0304052:FO003,2003, June 2003.

[67] Tatiana Likhomanenko, Philip Ilten, Egor Khairullin, Alex Rogozhnikov, Andrey Ustyuzhanin, and Michael Williams. LHCb Topological Trigger Reoptimization. *Journal of Physics: Conference Series*, 664(8):082025, 2015.

[68] R. Aaij, S. Amato, L. Anderlini, S. Benson, M. Cattaneo, M. Clemencic, B. Couturier, M. Frank, V.V. Gligorov, T. Head, C. Jones, I. Komarov, O. Lupton, R. Matev, G. Raven, B. Sciascia, T. Skwarnicki, P. Spradlin, S. Stahl, B. Storaci, and M. Vesterinen. Tesla: An application for real-time data analysis in high energy physics. *Computer Physics Communications*, 208:35 – 42, 2016.

[69] G. Barrand, I. Belyaev, et al. GAUDI- A Software Architecture and Framework for building HEP Data Processing Applications. *Computer Physics Communications*, 140(1-2), October 2001.

[70] The LHCb Collaboration. The DAVINCI Project. LHCb-DaVinci.

[71] R W Forty and O Schneider. RICH pattern recognition. Technical Report LHCb-98-040, CERN, Geneva, Apr 1998.

[72] M. Adinolfi et al. Performance of the LHCb RICH detector at the LHC. *Eur. Phys. J.*, C73:2431, 2013.

[73] R. Aaij and etal. Measurement of b-hadron branching fractions for two-body decays into charmless charged hadrons. *Journal of High Energy Physics*, 2012(10):37, Oct 2012.

[74] LHCb technical design report: reoptimized detector design and performance, 2003.

[75] Nicholas Brook, Matthew Coombes, Thomas Hampson, and Anatoly Solomin. The software framework and method for the alignment of the LHCb RICH optical system using proton-proton collisions. Technical Report LHCb-INT-2013-007. CERN-LHCb-INT-2013-007, CERN, Geneva, Jan 2013.

[76] F Stagni, P Charpentier, and the LHCb Collaboration. The lhcb dirac-based production and data management operations systems. *Journal of Physics: Conference Series*, 368(1):012010, 2012.

[77] The Gauss project. http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/gauss/.

[78] P. Geurts and G. Louppe. Learning to rank with extremely randomized trees. In Olivier Chapelle, Yi Chang, and Tie-Yan Liu, editors, *Proceedings of the Learning to Rank Challenge*, volume 14 of *Proceedings of Machine Learning Research*, pages 49–61, Haifa, Israel, 25 Jun 2011. PMLR.

[79] Michel De Cian, Stephen Farry, Paul Seyfert, and Sascha Stahl. Fast neural-net based fake track rejection. Technical Report LHCb-PUB-2017-011. CERN-LHCb-PUB-2017-011, CERN, Geneva, Mar 2017.

[80] Wouter D. Hulsbergen. Decay chain fitting with a kalman filter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 552(3):566 – 575, 2005.

[81] Andreas Hoecker, Peter Speckmayer, Joerg Stelzer, Jan Therhaag, Eckhard von Toerne, and Helge Voss. TMVA: Toolkit for Multivariate Data Analysis. *PoS*, ACAT:040, 2007.

[82] Alexandra Martin Sanchez. *CP violation studies on the $B \to DK^{*0}$ decays and hadronic trigger performance with the LHCb detecor*. PhD thesis, Universite Paris-Sud 11, June 2013.

[83] W. Verkerke and D. Kirkby. The RooFit toolkit for data modeling. *ArXiv Physics e-prints*, June 2003.

[84] P. del Amo Sanchez et al. Evidence for direct CP violation in the measurement of the Cabibbo-Kobayashi-Maskawa angle gamma with $B^{\pm} \to D^{(*)}K^{(*)\pm}$ decays. *Phys. Rev. Lett.*, 105:121801, 2010.

[85] A. Poluektov et al. Evidence for direct CP violation in the decay $B \to D^{(*)}K$, $D \to K_{\mathrm{S}}^{0}\pi^{+}\pi^{-}$ and measurement of the CKM phase ϕ_3. *Phys. Rev.*, D81:112002, 2010.