A Model-Independent Approach to Mixing in Prompt $D^0 \to K^0_S\pi^+\pi^-$ Decays at LHCb

Nicholas Torr

St. Catherine's College, Oxford

Thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy at the University of Oxford

Trinity Term, 2014

Abstract

This thesis presents a measurement of the charm mixing parameters x_D and y_D in prompt $D^0 \to K^0_S\pi^+\pi^-$ decays using 1 fb^{-1} of data collected by the LHCb experiment during 2011.

Mixing in charm is predicted to be small within the Standard Model, but there are significant uncertainties associated with calculating the long range contributions to the decay. Recent measurements made by LHCb and others have confirmed that mixing in charm exists at a rate of less than 1 %. With LHCb due to collect more data and Belle II being commissioned, the reduction of systematic uncertainties will become increasingly important.

The $D^0 \to K^0_\mathrm{S}\pi^+\pi^-$ decay provides sensitivity to both the magnitude and relative sign between the mixing parameters. It is also one of the few channels that can measure x_D directly. It is therefore crucial to study this mode in detail as more data becomes available. The work presented in this thesis utilises a model-independent description of the $K^0_\mathrm{S}\pi^+\pi^-$ Dalitz plot decay for the first time in the context of charm mixing. Previous mixing measurements with this final state have used a Dalitz plot amplitude model, and the associated systematic uncertainty is not straight forward to estimate or control. In its place, this analysis uses external, statistically-limited measurements of the strong-phase difference between D^0 and $\overline{D}^0$ obtained by CLEO as input. In addition, a data-driven technique is used to correct for decay time biases induced by the selection removing any systematic effects due to extracting this from simulated data. As the amount of available data increases, both of these techniques will become vital to improving our understanding of mixing in charm.

In the CP convention used by BaBar and adopted for this thesis, the measured mixing parameters are

$$x_D = -\left(0.863 \pm 0.527\,(\mathrm{stat.}) \pm 0.171\,(\mathrm{syst.})\right)\%,$$

$$y_D = -\left(0.026 \pm 0.463\,(\mathrm{stat.}) \pm 0.134\,(\mathrm{syst.})\right)\%.$$

Both x_D and y_D are consistent with the current world averages.

For Annabelle.

Acknowledgements

First and foremost I would like to say a massive thank you to my supervisor Matthew Charles for his advice, endless patience and support throughout my D.Phil studies, even once his time at Oxford had come to an end. I would also like to thank Malcolm John, Guy Wilkinson and Neville Harnew for their advice on physics and moral support.

Thank you to Sue Geddes, Kim Proudfoot and the rest of the administrative staff at Oxford for helping make this work possible. I'd also like to thank the STFC for funding my studies and for providing me with the opportunity to live and work at CERN.

I would like to thank the entire LHCb collaboration for their excellent work in building and running the experiment as well as providing the infrastructure that has made this analysis work possible. I'd especially like to thank Antonis Papanestis, Carmelo D'Ambrosio and the RICH working group for the fun I had tinkering at the pit, the Charm WG for providing clear and useful feedback on my analysis work and the DIRAC team for all their help in implementing the swimming in central production.

A big thank you to Jordi Garra Tico and Tomas Pilar for their hard work and advice relating to the analysis, Vava Gligorov for his patient guidance with all things swimming related, and Marco Gersabeck, Michael Alexander, Mike Sokoloff, Brian Meadows and many others I do not have the space to mention for helping me with the plethora of technical and physics related problems I encountered.

I would like to thank everyone I met at CERN during my LTA for making it such a wonderful and enlightening experience with a special thanks to the original Ferney-Voltaire massive for the good company, plentiful trips to Paddy's and the late night poker sessions.

Warm thanks to all of the postgraduate students and post-docs at Oxford for making my time there a wonderful experience. Thanks to Chris Thomas and Denis Derkach for their eagerness in answering my questions no matter how dumb they were, and thanks to Alexandru Dafinca, Rhorry Gauld, Ed Greening, Dave Hall, Nazim Hussain, Lucy Kogan and Ed Smith for making my D.Phil experience all the more enjoyable.

Thanks to Benedict Allbrooke for motivating me to to pursue particle physics at the graduate level, for all his help revising for my interview and for the home made beer.

A huge thank you to my ever increasing family for all their tireless love and continuing support (financial and otherwise). I couldn't have achieved any of this without them.

Finally, I would like to give my endless thanks to Anni for her love and overwhelming support right from day one, for her positive influence, and for remaining the wonderful person that she alway has been. I will always be grateful to you for putting up with me.

Contents

List of Figures

List of Tables

Preface

The Standard Model of particle physics is one of the most successful physical theories ever developed. It describes the interactions of fundamental particles on the quantum scale and its predictions have withstood generations of rigorous experimental testing. However, the Standard Model is known to be an incomplete theory, not least because of its inability to account for the CP violating effects that have resulted in the dominance of matter over anti-matter in the visible universe.

The LHCb experiment at the LHC is designed to study the relationship between matter and anti-matter by performing measurements of the rate of mixing and CP-violation in heavy meson decays with unrivalled precision. LHCb has already amassed some of the largest samples of D and B mesons to date and is due to carry on collecting data at higher centre of mass energies and luminosities for many years to come. As the amount of available data increases, systematic uncertainties will begin to dominate these measurements. Therefore, the reduction of systematic uncertainties will become ever more critical in the future.

Presented in this thesis is a measurement of the rate of mixing in $D^0 \to K^0_\mathrm{S}\pi^+\pi^-$ decays through a time-dependent Dalitz plot analysis using 1 fb^{-1} of data collected by LHCb during 2011. A model-independent description of the Dalitz plot distribution developed and measured by the CLEO collaboration is used for the first time in the context of charm mixing. In addition, this is the first analysis to utilise a data-driven decay time acceptance correction technique in multi-body charm decays at LHCb. Both of these analysis techniques will become vital in reducing systematic uncertainties on the measured rate of mixing in $D^0 \to K^0_\mathrm{S}\pi^+\pi^-$ decays as LHCb continues collect data.

Chapter 1 provides an overview of the Standard Model of particle physics with a focus on the concepts relevant to the work presented in this thesis. Chapter 2 provides a more detailed description of mixing in $D^0 \to K^0_\mathrm{S}\pi^+\pi^-$ decays and introduces the concepts behind the model-independent Dalitz plot analysis technique. Chapter 3 describes the LHCb detector and its data acquisition and processing algorithms. Chapter 4 describes the dataset used in the analysis, the selection that is applied and a study of potential sources of background. Chapter 5 describes the fit procedure used to extract the rate of mixing in $D^0 \to K^0_\mathrm{S}\pi^+\pi^-$ decays and summarises the results. Chapter 6 describes the various systematic uncertainties associated with the analysis technique and how they are evaluated. And finally, a brief summary and discussion of the results is given in Chapter 7.

Chapter 1

The Standard Model of Particle Physics

The Standard Model of particle physics is an effective field theory that describes matter interactions at the quantum scale. Developed over the latter half of the last century it has proven to be one of the most successful physical theories ever created. However, it is known that the Standard Model is not a complete theory as there are several observed phenomena that it cannot account for. The following chapter provides a brief introduction to the Standard Model[1].

1.1 Introduction

Since the dawn of civilisation mankind has sought to understand the workings of the universe. Through the development of basic mathematics and metaphysics by the ancient polymaths to modern physical theories, mankind has made great leaps towards answering some of the deep questions relating to the nature of our existence, of "life, the universe and everything". The Standard Model (SM) of particle physics is one of the most successful physical theories ever developed. It describes the nature of the fundamental building blocks of matter and the interactions between them. The SM is built on the basis of atomic theory and the idea that all matter is built from elementary particles. The concept of the atom was first developed

---

[1] Many sources were used to compile this summary, but the formalism adopted follows closely what is documented in [1] and [2].

by the ancient Greeks several thousand years ago. However, the development of a physical theory to describe the behaviour of atoms did not begin in earnest until the late 1800s with the work of J. Dalton [3], E. Rutherford [4] and J. J. Thompson [5]. With the advent of Quantum Mechanics and Special Relativity during the early 20th century the mathematical underpinnings of the Standard Model were developed and the scientific community began to realise the powerful insights their new theories could offer. One example is the prediction of the existence of anti-matter by P. Dirac's [6] extension of the Schrödinger-Klein-Gordon interpretation of relativistic quantum mechanics via his equation[2]

$$i\frac{\partial\psi}{\partial t} = (-i\boldsymbol{\alpha}\cdot\boldsymbol{\nabla} + \beta m)\psi, \tag{1.1}$$

where $\boldsymbol{\alpha}$ and β are 2×2 matrices and ψ is a 4-component Dirac spinor (this is described in more detail in Section 1.4). The existence of anti-matter was later confirmed by C. Anderson's experimental discovery of the positron [7]. During the latter half of the 20^{th} century, the discovery of subatomic structure in the form of quarks [8, 9] paved the way for the formulation of the SM as it exists today. The SM has consistently re-asserted its predictive power by predicting the existence of heavy quarks [10, 11], the τ-neutrino [12, 13] and a plethora of other fundamental and composite particle types prior to their confirmation by experiment. It has also had a major impact in the field of cosmology and in the development of the big bang theory [14]. However, despite its many successes, it is known that the SM is not a complete theory of matter. We know there are four fundamental forces present in nature; the electromagnetic force, the strong and weak nuclear forces, and gravity. The SM describes particle interactions via the former three forces, but it does not provide a framework to model gravitational interactions. It also cannot account for newly observed phenomena such as the strong evidence for the existence of dark matter [15] and the discovery of neutrino mixing [16, 17] (although the latter can be added as an extension). However, one of the most fundamental shortcomings of the SM is its inability to account for the dominance of matter over anti-matter in the visible universe. Therefore, understanding the relationship between matter and anti-matter is crucial to the future development of the SM and is the focus of

---

[2] Natural units are assumed throughout this chapter such that $\hbar = c = 1$.

the work presented in this thesis.

1.2 Particle Content

The particles that make up the Standard Model can be split into two categories; fermions and bosons. Each particle has a set of quantum numbers that determines the rules under which it can interact within the SM. One such quantum number is spin. Fermions are spin-1/2 particles and are the fundamental building blocks of matter. Fermions obey the Pauli Exclusion Principle [18] which states that no two fermions in a closed system can occupy the same quantum state. Any wave function describing a fermion must be anti-symmetric under particle exchange. Fermions are divided into two groups; quarks and leptons. Each quark and lepton has an anti-matter cousin of the same mass, but whose quantum numbers (apart from spin) are of opposite sign. In addition, quarks and leptons are divided into three generations with similar properties, but occupy different levels in mass. Interactions between quarks and leptons are mediated via spin-1 particles known as gauge bosons. Bosons do not obey the Pauli exclusion principle, but obey the rules of Bose-Einstein statistics [19, 20] such that any number of non-interacting particles in a closed system can be in the same quantum state. The wave function for bosons must be symmetric under particle exchange. The rules under which fermions interact and the types of boson involved differ for quarks and leptons and are also dependent on the individual properties of a given fermion, i.e. its quantum numbers and its mass.

The SM describes matter interactions via three of the four fundamental forces. The electromagnetic force is mediated by the massless photon, γ, which only interacts with particles that have electric charge. The weak nuclear force is mediated by the massive Z^0 and $W^\pm$ bosons and only interacts with particles that carry weak isospin and weak hypercharge. Finally, the strong nuclear force is mediated by the massless gluon, g, which only interacts with particles that have colour charge. Quarks carry all types of charge and can interact via any of the three forces. Leptons do not carry colour charge and therefore do not interact via the strong force. However, leptons that carry electric charge can interact via the electromagnetic and the weak nuclear force whereas neutral leptons (neutrinos) only interact via the weak nuclear force. In addition to the spin-1 bosons, a spin-0 boson known as the Higgs boson was predicted. The Higgs is a product of electroweak symmetry breaking which gives mass to the weak gauge bosons and fermions via interaction with the Higgs field. The discovery of a boson with mass $\sim$ 125 GeV/c^2 that is consistent with the SM Higgs was made independently by both the ATLAS and CMS collaborations in July 2012 [21, 22] although further studies of its properties are required to confirm it is indeed the Higgs predicted by the SM. Tables 1.1, 1.2 and 1.3 list the properties of the SM quarks, leptons and gauge bosons, respectively. The following sections describe the mathematical structure of the SM and the rules that govern the interactions of each of the three fundamental forces accommodated within the SM.

Generation	Name	Mass (MeV/c^2)	Electric Charge	Spin	Colour Charge	I_3
First	Up, u	$2.3^{+0.7}_{-0.5}$	2/3	1/2	r, g, b	+1/2
	Down, d	$4.8^{+0.5}_{-0.3}$	−1/3	1/2	r, g, b	−1/2
Second	Charm, c	1275 ± 25	2/3	1/2	r, g, b	+1/2
	Strange, s	95 ± 5	−1/3	1/2	r, g, b	−1/2
Third	Top, t	173070 ± 880^a	2/3	1/2	r, g, b	+1/2
	Bottom, b	4180 ± 30	−1/3	1/2	r, g, b	−1/2

---

[a] The t quark mass is the only direct measurement with all others calculated from the $\overline{\text{MS}}$ scheme.

Table 1.1: Properties of each quark in the SM [23]. I_3 is the third component of weak isospin for left handed fields.

Generation	Name	Mass (MeV/c^2)	Electric Charge	Spin	I_3	Colour Charge	Lepton # e	μ	τ
Electron	e^-	0.51	−1	1/2	−1/2	0	1	0	0
	ν_e	$< 2.2 \times 10^{-6}$	0	1/2	+1/2	0	1	0	0
Muon	μ^-	105.7	−1	1/2	−1/2	0	0	1	0
	ν_μ	$< 170 \times 10^{-3}$	0	1/2	+1/2	0	0	1	0
Tau	τ^-	1776.8	−1	1/2	−1/2	0	0	0	1
	ν_τ	< 15.5	0	1/2	+1/2	0	0	0	1

Table 1.2: Properties of each lepton in the SM [23]. I_3 is the third component of weak isospin for left handed fields.

Force	Boson	Mass (GeV/c^2)	Electric Charge	I_3	Colour Charge
EM	Photon, γ	0	0	0	0
Weak	Z^0	91.188 ± 0.002	0	0	0
	$W^\pm$	80.385 ± 0.150	±1	±1	0
Strong	Gluon, g	0	0	0	8 unique combinations

Table 1.3: Properties of each gauge boson in the SM [23]. I_3 is the third component of weak isospin for left handed fields.

1.3 Quantum Field Theory

The SM is formulated as an effective quantum field theory (QFT). The QFT description of particle interactions arises from the need to unify Special Relativity and Quantum Mechanics, and there is no evidence to suggest that the QFT description of nature is inadequate. Indeed, many extensions to the SM that are designed to deal with its shortcomings, for example Supersymmetry [24], are also QFTs. It is denoted an “effective” QFT as it describes, effectively, particle interactions at energy scales up to a few TeV. It is unclear whether this is sufficient at much higher energies when a quantum description of gravity becomes important[3].

The behaviour of any dynamic system modelled using a field theory (quantum or classical) is described via a Lagrangian density, $\mathcal{L}$, where

$$\mathcal{L} = T - V, \tag{1.2}$$

and T and V are the kinetic and potential energy of the system, respectively. Field theories can be thought of as continuous fields that associate some mathematical object to every point in space-time. This formalism is especially useful when considering point particle interactions as conservation laws can be determined from invariances of the Lagrangian density (more on this in Section 1.9). Additionally, the equations of motion for any system

---

[3] For example, it is known that the current formulation of the SM would be insufficient to describe matter interactions during the very early universe because of the strength of gravity at that time.

can be defined using the *Principle of Least Action.* This states that the path taken by a system between two points in time is the one for which the action is a minimum or, mathematically, $\delta \int_{t_1}^{t_2} \mathcal{L} \mathrm{d}t = 0$. In the case of a relativistic QFT such as the SM, one can apply this principle via perturbation theory to calculate scattering amplitudes for particle interactions. Applying this to a non-interacting scalar field ϕ results in the Euler-Lagrange field equation

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) = \frac{\partial \mathcal{L}}{\partial \phi} \tag{1.3}$$

where ∂_μ is the covariant derivative across all four space-time coordinates. The SM Lagrangian[4] can be broken down into components that describe the interactions between fermions and bosons via each of the three forces and the origin of their mass via interaction with the Higgs field such that,

$$\mathcal{L}_{\mathrm{SM}} = \mathcal{L}_{\mathrm{EW}} + \mathcal{L}_{\mathrm{QCD}} + \mathcal{L}_{\mathrm{Higgs}} + \mathcal{L}_{\mathrm{Yukawa}} \tag{1.4}$$

where interactions via the electromagnetic and weak nuclear force are described in the first term, interactions via the strong force in the second term and interaction with the Higgs field in the final two terms. The formulations of the terms are described in the following sections. As mentioned, any Lagrangian that is invariant under local and/or global gauge transformations will have resulting conservation laws associated with the symmetry. The SM Lagrangian is invariant under local gauge transformations of the symmetry group

$$\mathrm{SU}(3)_C \otimes \mathrm{SU}(2)_I \otimes \mathrm{U}(1)_Y \tag{1.5}$$

where $\mathrm{SU}(3)_C$ and $\mathrm{SU}(2)_I \otimes \mathrm{U}(1)_Y$ are the symmetry groups of the strong and the electroweak interactions, respectively. These symmetries correspond to the conservation of colour charge (C), weak isospin (I) and weak hypercharge (Y). The latter two conservation laws also imply conservation of electric charge due to the relation $Q = Y + I_3$ where I_3 is the third component of weak isospin (see Section 1.6).

---

[4] The term “Lagrangian” hereby replaces Lagrangian density for brevity.

1.4 Quantum Electrodynamics

Quantum Electrodynamics (QED) describes the fundamental interactions of the electromagnetic force as a QFT. It was chiefly developed by the great Richard Feynman[5] during the 1960s [25] and has proven to be by far the most robust theory against experiment ever developed [26]. Starting with the Dirac equation given in Equation 1.1 one can construct a Lagrangian for non-interacting spin-1/2 particles

$$\mathcal{L} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi \tag{1.6}$$

where γ_μ are the gamma matrices, ψ is the Dirac spinor for free fermions ($\bar{\psi} = \psi^{\dagger}\gamma^{0}$ is the adjoint spinor) and m is the fermion mass. Equation 1.6 describes the kinematics of the system and is invariant under global gauge transformations. However, it is not invariant under local gauge transformations. Invariance can be restored by introducing a gauge field, A_μ which transforms under local gauge transformations as

$$\mathcal{A}_{\mu} \to \mathcal{A}_{\mu} + \frac{1}{q}\partial_{\mu}\psi \tag{1.7}$$

such that

$$\mathcal{L} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi + q\bar{\psi}\gamma^{\mu}\psi\mathcal{A}_{\mu}. \tag{1.8}$$

The final term in Equation 1.8 describes the interaction between fermions and the gauge field, A_μ with coupling q. Adding a self interaction or mass term for the gauge field would break the local gauge invariance which implies that QED is an abelian gauge theory and that the corresponding boson of the gauge field, the photon, is massless. One can add a term which describes the kinematics of the free gauge field (i.e photons propagating through space) which completes the Lagrangian for electromagnetic interactions of spin-1/2 particles,

$$\mathcal{L}_{\mathrm{EM}} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi + q\bar{\psi}\gamma^{\mu}\psi\mathcal{A}_{\mu} - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \tag{1.9}$$

---

[5] As well as being a ground-breaking physicist he was renowned for being a womaniser and a talented percussionist. Truly a man after my own heart.

where $F^{\mu\nu} = \partial^\mu \mathcal{A}^\nu - \partial^\nu \mathcal{A}^\mu$ is the EM field tensor. Note that $\mathcal{L}_{\mathrm{EM}}$ does not appear in Equation 1.4 as the electromagnetic interaction is incorporated into the electroweak term.

The development of QED laid the fundamentals of the Feynman rules for calculating scattering amplitudes and cross-sections. These rules provide a toolkit for associating propagators and vertex factors to Feynman diagrams for a given process. By assembling the relevant components from a given diagram one can calculate an amplitude for the process and take the modulus squared to calculate the rate. A complete description of the Feynman rules can be found in [27]. It is worth noting that these rules are not only applicable in QED, but also in QCD and Electroweak theory which are described in the following sections.

1.5 Quantum Chromodynamics

Quantum Chromodynamics describes particle interactions via the strong force. Gluons are the force carrier of the strong force and they only interact with particles that carry colour charge. The need for these additional quantum numbers in the quark model stems from different behaviour observed in two forms of observed quark matter; baryons and mesons. Baryons consist of three quarks (or three anti-quarks) bound together and act like fermions. Mesons consist of a quark anti-quark pair and act like bosons. The observation of several baryon states such as the Δ^{++} that are constructed of like quarks (in this case uuu) with parallel spins were initially thought to violate the Pauli exclusion principle. However, M. Han, Y. Nambu and O. Greenberg independently resolved the problem by suggesting that quarks carry an additional $\mathrm{SU}(3)_C$ gauge degree of freedom i.e. colour charge[6] [29, 30]. There are three colours in QCD, red, green and blue, such that the quark spinor is given by ψ_i where i runs over r, g, b. Quarks carry colour and anti-quarks carry anti-colour such that all mesons are colour neutral objects. Similarly, a combination of three quarks that have different colours results in zero total colour rendering baryons colour neutral as well. The

---

[6] Although this was the first time that this extra freedom was given a formal mathematical definition, it was first suggested that quarks may have additional quantum numbers by B. V. Struminsky in his PhD thesis examining the Ω^- baryon [28].

fact that isolated states are colour singlets means that quarks cannot exist as free objects in nature, a phenomenon known as *Confinement*. As gluons also carry colour charge they also obey the confinement principle. There exist 8 individual gluons, $\mathcal{A}^a_\mu$, that have different couplings to different colours with the couplings described via matrices T^a. The couplings are non-commuting and obey the relation,

$$\left[T^a, T^b\right] = if^{abc}T^c \tag{1.10}$$

where f^{abc} are structure constants which have anti-symmetric behaviour. The non-commuting property of the QCD charges and existence of multiple gauge bosons render QCD a non-abelian gauge theory. The QCD Lagrangian is given by

$$\mathcal{L}_{\text{QCD}} = \bar{\psi}_i(i(\gamma^\mu D_\mu)_{ij} - m\delta_{ij})\psi_j - \frac{1}{4}F^a_{\mu\nu}F^{\mu\nu}_a \tag{1.11}$$

where $F^a_{\mu\nu}$ is the gluon field strength tensor which is given by

$$F^a_{\mu\nu} = \partial_\mu \mathcal{A}^a_\nu - \partial_\nu \mathcal{A}^a_\mu + g_s f^{abc}\mathcal{A}^b_\mu \mathcal{A}^c_\nu \tag{1.12}$$

where g_s is the strong coupling constant and D_μ is a covariant derivative which obeys the SU(3) symmetry under local gauge transformations, and is given by

$$D_\mu = \left(\partial_\mu - ig_s\mathcal{A}^a_\mu \frac{1}{2}\lambda^a\right). \tag{1.13}$$

In Equation 1.13, λ^a are 3×3 Gell-Mann matrices which are related to the colour matrices via $T^a = \lambda^a/2$. One should note that adding a mass matrix for the gauge fields would break local gauge invariance, implying that the gluon is massless. Because gluons interact with all particles that carry colour they can form self- interaction vertices (another common property of non-abelian gauge theories). This self-interaction results in an anti-screening effect that is not present in abelian gauge theories such as QED. Photons have the effect of screening the electric charge of quarks which results in the strength of the electromagnetic interaction weakening over long distances (low energies). Gluons have the opposite effect and strengthen the strong force over long distances. This is known as *Asymptotic Freedom* and results in the coupling constant of QCD, α_s, having an energy dependence. This causes normalisation problems at low energies where α_s is strongest and the perturbative approach to calculating scattering amplitudes using the Feynman rules breaks down. Asymptotic freedom also provides an explanation for the confinement effect mentioned earlier in this section. A process known as *Hadronisation* (also known as *Fragmentation*) occurs when the distance between two quarks increases to the point where α_s is so strong that it is energetically favourable for a new quark-antiquark pair to be created. Experimental evidence of confinement comes from failed searches for isolated quarks [31], but also from the hadronisation of gluons produced in particle interactions at colliders into showers of particles commonly known as jets.

1.6 Electroweak Interactions

The GWS (Glashow-Weinberg-Salam) theory of electroweak (EW) interactions combines the electromagnetic and the weak nuclear forces into a single theory. The electroweak interaction obeys the $\mathrm{SU}(2)_I \otimes \mathrm{U}(1)_Y$ symmetry group and has four resulting gauge fields; three weak isospin bosons, W^i_μ, and one weak hypercharge boson, B_μ.

The weak nuclear interaction is the only process that discriminates between the "handedness" of fermions and interacts solely with left-handed fermions which form doublets

$$\boldsymbol{\chi}_L = \begin{pmatrix} \nu_L \\ l_L \end{pmatrix}, \begin{pmatrix} u_L \\ d'_L \end{pmatrix} \quad \text{with adjoint} \quad \bar{\boldsymbol{\chi}}_L = \begin{pmatrix} \bar{\nu}_L \\ \bar{l}_L \end{pmatrix}, \begin{pmatrix} \bar{u}_L \\ \bar{d}'_L \end{pmatrix}, \tag{1.14}$$

where the prime on the down-type quark component denotes it as a flavour eigenstate as opposed to the mass eigenstate of the up-type quark component (see Section 1.8). The upper and lower members of the doublet have weak isospin $I_3 = \pm 1/2$, respectively. The weak currents are defined as,

$$J^i_\mu = \bar{\boldsymbol{\chi}}_L \gamma^\mu \frac{1}{2} \tau^i \boldsymbol{\chi}_L \tag{1.15}$$

where γ^μ are the gamma matrices, τ^i are the weak isospin generators which take on the form of the 2 × 2 Pauli spin matrices and the index $i = 1, 2, 3$. The isospin generators obey the commutation relation,

$$[\frac{1}{2}\tau^i, \frac{1}{2}\tau^j] = i\epsilon_{ijk}\frac{1}{2}\tau^k, \tag{1.16}$$

with isospin raising and lowering operators $\tau^\pm = \frac{1}{2}(\tau^1 \pm i\tau^2)$. The interactions that raise and lower the total charge can be written in terms of the first and second components of the weak current giving

$$J_\mu^\pm = \bar{\boldsymbol{\chi}}_L\gamma^\mu\tau^\pm\boldsymbol{\chi}_L = J_\mu^1 \pm iJ_\mu^2. \tag{1.17}$$

This is the charged interaction current. The neutral interaction current for the weak force is then given by the third component of the weak current. However, the electromagnetic interaction is also a neutral current which interacts with both left-handed and right-handed fermions, the latter of which form singlets $\chi_R = \{l_R, q_R\}$ where l_R are the charged lepton mass eigenstates and q_R the quark mass eigenstates. To combine this with the weak neutral current that only interacts with left-handed fermions, an extra current known as weak hypercharge is added to J_μ^3. Defining the electromagnetic current as

$$J_\mu^{EM} = Q(\bar{\chi}_L\gamma^\mu\chi_L + \bar{\chi}_R\gamma^\mu\chi_R) = J_\mu^3 + \frac{1}{2}J_\mu^Y, \tag{1.18}$$

where χ_L is the lower element of $\boldsymbol{\chi}_L$, one can derive an expression for the hypercharge current J_μ^Y and form the familiar relation

$$Q = I^3 + \frac{Y}{2}, \tag{1.19}$$

where Q is the electric charge of the particle in units of e, I^3 is its third component of weak isospin and Y is its weak hypercharge. Introducing four vector gauge bosons, W_μ^i (i = 1,2,3) with coupling g for the SU(2)$_I$ field and B_μ with coupling g' for the U(1)$_Y$ field, one can write down the fermion-boson term in the EW Lagrangian

$$\mathcal{L}_{\text{fermion}} = \bar{\boldsymbol{\chi}}_L^f\gamma^\mu[i\partial_\mu - \frac{g}{2}(\tau^i \cdot W_\mu^i) + \frac{g'}{2}B_\mu]\boldsymbol{\chi}_L^f + \bar{\chi}_R^f\gamma^\mu[i\partial_\mu + g'B_\mu]\chi_R^f, \tag{1.20}$$

where $\tau^i = (\tau^1, \tau^2, \tau^3)$ and the sum over all fermion generations f is assumed. By substituting Equation 1.17 into Equation 1.20 one can split the fermion interaction terms into charged and neutral currents. One can then define the charged vector bosons in terms of the first two SU(2)$_I$ bosons

$$W^{\pm}_{\mu} = \frac{1}{\sqrt{2}}(W^1_\mu \pm W^2_\mu) \tag{1.21}$$

and the neutral bosons in terms of the third SU(2)$_I$ boson W^3_μ and the U(1)$_Y$ boson B_μ giving

$$\begin{aligned} Z^0_\mu &= \cos\theta_W W^3_\mu - \sin\theta_W B_\mu \\ A^0_\mu &= \sin\theta_W W^3_\mu + \cos\theta_W B_\mu, \end{aligned} \tag{1.22}$$

where θ_W is the weak mixing angle. These are the bosons of the physical fields $W^{\pm}$, Z^0 and A^0_μ is the photon.

The final part of the EW Lagrangian contains the kinetic energy terms for the four gauge fields and the self-interaction terms for the SU(2)$_I$ gauge fields:

$$\mathcal{L}_{\text{EW}} = \bar{\boldsymbol{\chi}}^f_L \gamma^\mu [i\partial_\mu - \frac{g}{2}(\tau^i \cdot W^i_\mu) + \frac{g'}{2} B_\mu] \boldsymbol{\chi}^f_L + \bar{\chi}^f_R \gamma^\mu [i\partial_\mu + g' B_\mu] \chi^f_R - \frac{1}{4} W^i_{\mu\nu} W^{i\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \tag{1.23}$$

where

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{1.24}$$

is the EM field tensor and

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + g\epsilon_{ijk} W^j_\mu W^k_\nu \tag{1.25}$$

is the SU(2)$_I$ field tensor with the final term describing the self-interaction vertices of the weak vector bosons. Note that this formulation of the EW Lagrangian assumes that the gauge bosons are massless as adding mass terms would break the gauge invariance of the Lagrangian in Equation 1.23. However, the physical weak gauge bosons have measured masses of approximately $M_{W^{\pm}} = 80.4$ GeV/c^2 and $M_{Z^0} = 91.2$ GeV/c^2, respectively. This issue can be solved by spontaneously breaking the $\mathrm{SU}(2)_I$ symmetry using the so-called Higgs mechanism.

1.7 The Higgs Mechanism

The Higgs mechanism [32, 33] is a solution to the issue of adding boson mass terms to the electroweak Lagrangian. The technique involves introducing a weak isospin doublet of complex scalar Higgs fields

$$\mathbf{\Phi} = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad \text{with adjoint} \quad \mathbf{\Phi}^\dagger = \begin{pmatrix} \phi^- \\ \bar{\phi}^0 \end{pmatrix}. \tag{1.26}$$

The masses of the EW gauge bosons are generated from their interaction with the Higgs fields. This interaction can be described by the Lagrangian

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \mathbf{\Phi})^\dagger D^\mu \mathbf{\Phi} - V(\mathbf{\Phi}) \tag{1.27}$$

where the simplest $\mathrm{SU}(2)_I \otimes \mathrm{U}(1)_Y$ invariant scalar potential is

$$V(\mathbf{\Phi}) = -\mu^2(\mathbf{\Phi}^\dagger\mathbf{\Phi}) + \lambda(\mathbf{\Phi}^\dagger\mathbf{\Phi})^2. \tag{1.28}$$

Minimising the potential for the case where $\mu^2 > 0$ gives

$$\mathbf{\Phi}^\dagger\mathbf{\Phi} = \frac{\mu^2}{2\lambda}. \tag{1.29}$$

Now we select a vacuum expectation value (vev) for the real scalar fields such that the potential is minimised, the $\mathrm{SU}(2)_I \otimes \mathrm{U}(1)_Y$ symmetry is broken, but the $\mathrm{U}(1)_Y$ symmetry is preserved. This is done by setting the vevs for ϕ_1, ϕ_2 and ϕ_4 to zero and assigning a non-zero vev, v, to the neutral field ϕ_3 such that

$$\mathbf{\Phi} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ H + v \end{pmatrix}, \tag{1.30}$$

where H is the neutral scalar Higgs field and the vev v is given by

$$v = \frac{\mu}{\sqrt{\lambda}}. \tag{1.31}$$

Note that this result is achieved by using a special gauge known as the *Unitary Gauge* [34] which allows the remaining fields of the Higgs doublet to be ignored. The resulting Higgs Lagrangian can be written in terms of the physical gauge boson fields giving

$$\begin{aligned} \mathcal{L}_{\text{Higgs}} =& \frac{1}{2}\partial_\mu H \partial^\mu H + \frac{1}{4} g^2 (H^2 + 2vH + v^2) W_\mu^+ W^{-\mu} \\ &+ \frac{1}{8}(g^2 + g'^2)(H^2 + 2vH + v^2) Z_\mu^0 Z^{0\mu} \\ &- \mu^2 H^2 - \frac{\lambda}{4}(H^4 + 4vH^3). \end{aligned} \tag{1.32}$$

Note there are no terms involving the gauge field A_μ^0 keeping the photon massless as prescribed. The masses of the gauge bosons can then be read from Equation 1.32 by selecting the coefficients of the terms $W_\mu^+ W^{-\mu}$ and $Z_\mu^0 Z^{0\mu}$ which give

$$M_{W^\pm} = \frac{1}{2} g v \qquad \text{and} \qquad M_{Z^0} = \frac{1}{2} \frac{gv}{\cos\theta_W}. \tag{1.33}$$

Similarly, the mass of the Higgs[7] can be read as the coefficient of the H^2 term giving $M_H = \sqrt{2}\mu$. Note that there are terms in the Higgs Lagrangian that allow for self-interaction vertices of the Higgs (H^3 and H^4) and vertices involving the Higgs and pairs of vector bosons ($V^2 H$ and $V^2 H^2$). The following section will show how the Higgs mechanism can be extended to give mass to fermions.

---

[7] It should be noted that, although a term for the Higgs mass can be written down, it still remains a free parameter which can only be fully determined by experiment.

1.8 Fermion Masses & The CKM Matrix

As was the case for the weak gauge bosons, fermion fields also get their mass by interacting with the Higgs field. The couplings of fermions to the Higgs are included in the SM Lagrangian by adding so-called Yukuwa terms for each generation of fermions. The Yukuwa term for a single fermion takes on the form,

$$\mathcal{L}_{\text{Yukuwa}} = -y^f \bar{\boldsymbol{\chi}}^f_L \boldsymbol{\Phi} \chi^f_R + y \boldsymbol{\chi}^f_L \boldsymbol{\Phi} \bar{\chi}^f_R = \frac{y^f}{\sqrt{2}} \bar{\chi}\chi, \tag{1.34}$$

where y^f is the Yukuwa coupling of the fermion f. The mass of the fermion is then read off as the coefficient of the $\bar{\chi}\chi$ term giving $m_f = \frac{y^f}{\sqrt{2}}$. Note that this approach only works for down type quarks and charged leptons due to the upper component of $\boldsymbol{\Phi}$ being zero for up type quarks. However, by exploiting $\mathrm{SU}(2)_I$ symmetry one can replace the regular Higgs doublet with

$$\boldsymbol{\Phi}^c = \epsilon_{ij} \boldsymbol{\Phi}^{j*} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ v + H^* \end{pmatrix} = \begin{pmatrix} v + H^* \\ 0 \end{pmatrix} \tag{1.35}$$

The description given thus far is perfectly suitable when considering one generation of fermion interacting via the SM gauge bosons. However, this formalism must be extended to include the additional generations discussed in Section 1.2. In addition, the formalism must account for the fact that the fermion doublets of higher generations must couple to the weak gauge bosons in an identical fashion as for the first generation, but must couple to the Higgs field differently to successfully reproduce their different masses. To do this, the Yukuwa couplings are expressed as matrices in flavour space. These matrices exist for quarks in the form of the CKM matrix [35] and leptons in the form of the PMNS matrix[8] [36], although the latter is beyond the scope of this thesis. Introducing these matrices for up and down type quarks into Equation 1.34 gives

$$\mathcal{L}_{\text{Yukuwa}} = \sum_{1,2,3} - \left[\bar{\boldsymbol{\chi}}^f_L Y^D_{ff'} \boldsymbol{\Phi} d^f_R + \bar{\boldsymbol{\chi}}^f_L Y^U_{ff'} \boldsymbol{\Phi}^c u^f_R + \text{h.c.} \right], \tag{1.36}$$

---

[8] The PMNS lepton mixing matrix is the required extension to the SM to account for neutrino masses and mixing.

where u and d are the up-type and down-type flavour singlets, $Y^U_{ff'}$ and $Y^D_{ff'}$ are the matrices of Yukuwa couplings for up-type and down-type quarks, respectively, and the sum runs over all three generations. Non-zero diagonal elements in these matrices result in mixing between the quark generations, allowing for flavour changing processes in weak charged current decays. The mass eigenstates (d', s', b') can be expressed in terms of the flavour eigenstates (d, s, b) by a rotation in flavour space expressed as

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.37}$$

where the CKM matrix elements have been measured as [23]

$$V_{CKM} = \begin{pmatrix} |V_{ud}| = 0.97472 \pm 0.00015 & |V_{us}| = 0.22534 \pm 0.00065 & |V_{ub}| = 0.00351^{+0.0015}_{-0.00014} \\ |V_{cd}| = 0.22520 \pm 0.00065 & |V_{cs}| = 0.97344 \pm 0.00016 & |V_{cb}| = 0.0412^{+0.0011}_{-0.0005} \\ |V_{td}| = 0.00867^{+0.00029}_{-0.00031} & |V_{ts}| = 0.0404^{+0.0011}_{-0.0005} & |V_{tb}| = 0.999146^{+0.000021}_{-0.000046} \end{pmatrix}.$$

Taking the squared modulus of a particular element of the CKM matrix $|V_{ij}|^2$ gives the probability of a transition from quark flavour $i \to j$. There are 9 elements in the CKM matrix and each of these can be complex giving 18 free parameters in total. However, as the CKM matrix is unitary (i.e. $VV^\dagger = 1$) this can be reduced by a factor of 2. It should also be noted that the unitarity of the CKM matrix is what prevents tree-level flavour changing neutral currents and implies that the sum of all couplings of the up-type quarks to the down-type quarks is constant for all generations. This latter property is known as *Weak Universality.* By exploiting $\mathrm{SU}(2)_I$ symmetry and rotating the quark fields, a further 5 degrees of freedom can be removed leaving just 4 free parameters. The CKM matrix can then be re-written as

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\phi} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\phi} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\phi} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\phi} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\phi} & c_{23}c_{13} \end{pmatrix} \tag{1.38}$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$ are the cosine and sine, respectively, of the mixing angles between generations i and j, of which there are three, and ϕ is a complex phase that allows for CP-violation (CPV) to occur (see Section 1.9). There are alternative parameterisations that can be adopted [37]; the standard parameterisation in Equation 1.38 is taken from [23]. The unitarity of the CKM matrix implies that the off-diagonal elements that are responsible for the quark mixing must obey the rule

$$\sum_k V_{ki} V_{kj}^* = 0 \tag{1.39}$$

which gives rise to six different expressions that describe triangles in the complex plane. Each triangle is unique, but they all have the same area which is equal to the magnitude of the CP-violating phase ϕ. Therefore, measuring the properties of these triangles is crucial to understanding the amount of CP-violation within the SM. In practice, only one triangle needs to be measured very precisely to measure the rate of CPV. The triangle that is most commonly used is known as the *unitarity triangle* which is expressed by the relation

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \tag{1.40}$$

The parameters that describe the amount of CPV are the angles of this triangle which are given by

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right) \tag{1.41}$$

$$\beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right) \tag{1.42}$$

$$\gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \tag{1.43}$$

The sum of these angles should be 180°. Any deviation from this could be a sign of CP-violation from another source not accounted for in the CKM matrix. Measuring the properties of the unitarity triangle, in particular γ, is one of the key physics goals of LHCb. Section 2.2.1 describes one approach to measuring γ using $B \to DK$ decays which has been

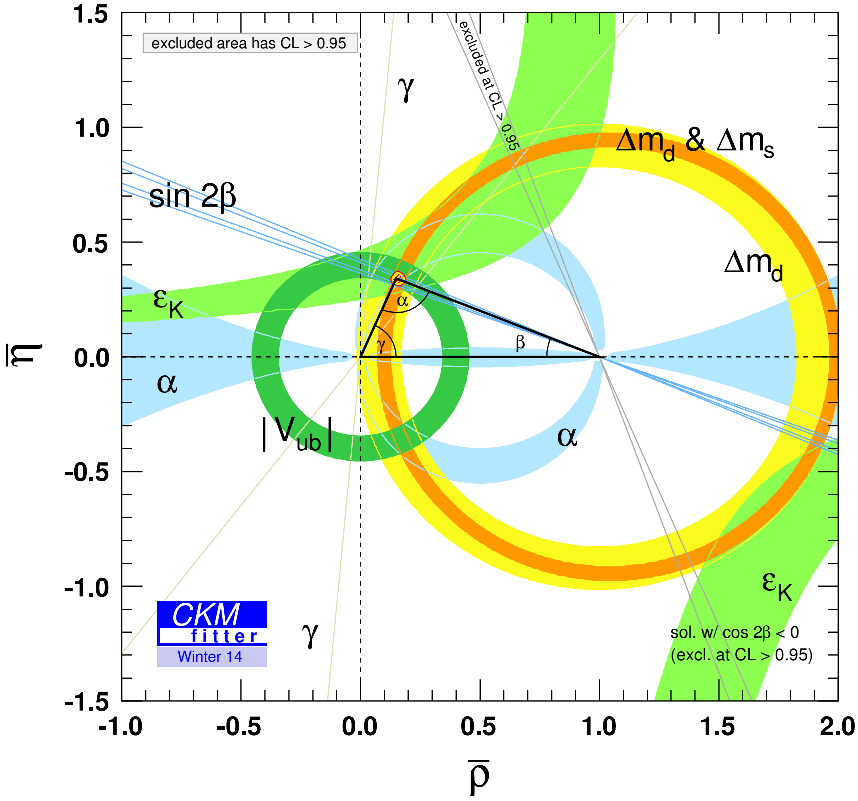


Figure 1.1: Latest results of a global fit to all available measurements of the unitarity triangle [40]. The parameters $\bar{\rho}$ and $\bar{\eta}$ are related to the CKM matrix elements by $\bar{\rho} + i\bar{\eta} = -\frac{V_{ud}V^*_{ub}}{V_{cd}V^*_{cb}}$.

used to constrain γ [38]. All the available measurements of the angles of the unitarity triangle are combined by groups such as UTFit [39] and CKMFitter [40]. The latest results at the time of writing from CKMFitter are shown in Figure 1.1.

1.9 CP-Violation

As mentioned in Section 1.3, the invariance of a Lagrangian under some transformation corresponds to a conservation law. These conservation laws are discussed in terms of symmetries and are defined via Noether's theorem [41]. There are three types of symmetry that are crucial to the SM. The first of consists of continuous symmetries such as spatial and temporal translations, which correspond to the conservation of linear momentum and energy, respectively. The second consists of internal symmetries such as the gauge transformations of $\mathrm{SU}(3)_\mathrm{C} \otimes \mathrm{SU}(2)_I \otimes \mathrm{U}(1)_Y$ under which the SM Lagrangian is invariant and which correspond to the conservation of colour charge, weak isospin and weak hypercharge, respectively, all of which are independent of space-time coordinates. All of these have been discussed in previous sections. However, there is an additional group of symmetries known as discrete symmetries which involve interchanges under the transformation of a discrete group. In the case of the SM there are three discrete symmetries that are of interest: Parity (P), Charge Conjugation (C) and Time Reversal (T). Charge conjugation is the exchange of particles for anti-particles, parity is the reversal of spatial coordinates and time reversal is quite simply the reversal of time. The SM Lagrangian is invariant under the combined CPT symmetry (i.e. applying each of C, P and T in succession). However, it is known that the individual symmetries can be violated. In particular, the CP symmetry is known to be violated by the weak interaction as was discovered experimentally in 1964 [42]. CP-violation is accommodated within the Standard Model by the complex phase ϕ discussed in the previous section. Experimentally, there are three manifestations of CP-violation:

1. Direct CP-violation (also known as CP-violation in the decay). This is where the decay amplitudes for CP conjugate states are not equal such that

   $$\mathcal{A}_f \neq \bar{\mathcal{A}}_{\bar{f}}. \tag{1.44}$$

   This can occur in charged and neutral hadrons and is independent of the particle's decay time.

2. CP-violation in mixing. This occurs when the mass eigenstates are not CP eigenstates. One can write the mass eigenstates for a neutral meson M in terms of the flavour eigenstates such that

   $$|M_1\rangle = p|M^0\rangle + q|\bar{M}^0\rangle \tag{1.45}$$

   $$|M_2\rangle = p|M^0\rangle - q|\bar{M}^0\rangle, \tag{1.46}$$

where p and q are complex numbers that satisfy $|p^2| + |q^2| = 1$. The rate of CPV can be parameterised by the ratio $|q/p|$ where any deviation from 1 would be a signal of indirect CPV.

3. CP-violation in the interference between mixing and decay. This can be parameterised as

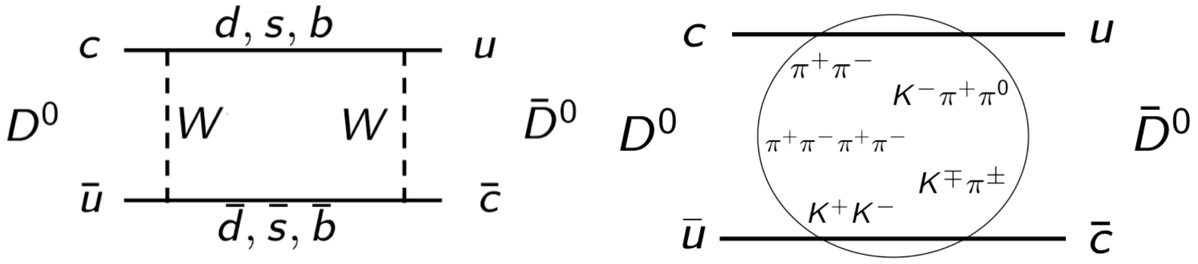
$$\frac{\mathcal{A}_f}{\bar{\mathcal{A}}_{\bar{f}}} \frac{p}{q} \tag{1.47}$$

where any deviation from 1 is a signal of CP-violation.

Types 2 and 3 are both referred to as indirect CP-violation.

1.10 Neutral Meson Mixing

Mixing is a process in which a neutral, heavy particle produced as a meson of one flavour (e.g. D^0) contains a non-zero component of the other flavour (e.g. $\overline{D}^0$) at some later time. This can occur via two types of interaction; short range interactions in which two virtual W bosons are exchanged and long-range hadronic interactions via an intermediate state which couples to both the initial meson and its charge conjugate. Figure 1.2 (a) depicts the Feynman box diagram of the short range interaction for a D^0 while 1.2 (b) presents a schematic diagram for long-range mixing with a set of example intermediate states through which mixing can occur. Mixing is known to occur in heavy neutral mesons that have net flavour (i.e. K^0, B^0, B_s^0 and D^0) and has been confirmed experimentally for each [43, 44, 45, 46]. As mixing in the charm sector is the primary focus of the work presented in this thesis, the following chapter provides a more in-depth description of D^0 mixing within the Standard Model paying particular attention to mixing in $D^0 \to K_S^0\pi^+\pi^-$ decays.



(a) (b)

Figure 1.2: Examples of interactions in which D^0 mixing can occur. (a) Feynman diagram involving the exchange of two W bosons. (b) Schematic of long-range interactions with a subset of possible intermediate states.

Chapter 2

A Model-Independent Approach to Mixing and CP-Violation

The following chapter describes a model-independent parameterisation of mixing in $D^0 \to K^0_S\pi^+\pi^-$ decays. An introduction to the Dalitz plot formalism is given as well as a description of the external inputs to the measurement presented in this thesis and how they were obtained. Finally, the current status of mixing and CP-violation measurements in charm is summarised with prospects for future measurements with the model-independent technique.

2.1 Parameterisation of Mixing and CP-Violation in $D^0 \to K^0_S\pi^+\pi^-$ Decays

A detailed description of the mixing formalism used in charm and particularly in $D^0 \to K^0_S\pi^+\pi^-$ decays is given in the following sections. To get a handle on the rate of mixing in multi-body charm decays, two essential ingredients are required; an understanding of how the decay rate behaves as a function of time and how the dynamics of the decay behave in the available momentum space. To understand the decay rate as a function of time one must begin with the time evolution of the mass eigenstates. The mass eigenstates can be written as a linear superposition of the flavour eigenstates:

$$|D_{1,2}\rangle = p\sqrt{1 \pm \epsilon}|D^0\rangle \pm q\sqrt{1 \mp \epsilon}|\overline{D}^0\rangle, \tag{2.1}$$

where ϵ is assumed to be zero due to CPT invariance and such that the normalisation $|p^2| + |q^2| = 1$ holds. Unlike in the B and K systems the rate of charm mixing is much slower than the decay rate. As the D^0 mesons produced at LHCb are flavour eigenstates it is necessary to work in the flavour basis

$$|D^0\rangle = \frac{|D_1\rangle + |D_2\rangle}{2p}, \tag{2.2}$$

$$|\overline{D}^0\rangle = \frac{|D_1\rangle - |D_2\rangle}{2q}, \tag{2.3}$$

where the flavour eigenstates are orthonormal such that

$$\langle D^0|D^0\rangle = 1,\ \langle\overline{D}^0|\overline{D}^0\rangle = 1 \text{ and } \langle D^0|\overline{D}^0\rangle = 0,\ \langle\overline{D}^0|D^0\rangle = 0. \tag{2.4}$$

The time evolution of the flavour eigenstates is governed by the time-dependent Schrödinger equation and can be written as

$$|D^0(t)\rangle = g_+(t)|D^0\rangle + \frac{q}{p}g_-(t)|\overline{D}^0\rangle, \tag{2.5}$$

$$|\overline{D}^0(t)\rangle = \frac{p}{q}g_-(t)|D^0\rangle + g_+(t)|\overline{D}^0\rangle, \tag{2.6}$$

where

$$g_\pm(t) = \frac{1}{2}\left[e^{-i(M_1 - i\Gamma_1/2)t} \pm e^{-i(M_2 - i\Gamma_2/2)t}\right] \tag{2.7}$$

and $M_{1,2}$ and $\Gamma_{1,2}$ are the masses and widths of the mass eigenstates $|D_{1,2}\rangle$. The rate of mixing in charm is parameterised by the dimensionless mixing parameters

$$x_D = \frac{M_2 - M_1}{\Gamma} \tag{2.8}$$

and

$$y_D = \frac{\Gamma_2 - \Gamma_1}{2\Gamma}, \tag{2.9}$$

where $\Gamma = (\Gamma_1 + \Gamma_2)/2$ is the average width of the mass eigenstates. Note that the CP formalism used in this analysis is the same as used in the current published mixing measurements from $K_S^0 h^+ h^-$ decays, i.e. from BaBar [47], but the mixing parameter definitions used in those analyses are

$$x = \frac{M_1 - M_2}{\Gamma} = -x_D \quad \text{and} \quad y = \frac{\Gamma_1 - \Gamma_2}{2\Gamma} = -y_D. \tag{2.10}$$

The Heavy Flavour Averaging Group (HFAG) [48] use the same mixing parameter definitions as used in this analysis, but a different CP convention. The consequence of these differences is that the mixing parameters defined using the formalism presented here will carry an extra factor of -1 compared to the current published results from analyses of $K_S^0 h^+ h^-$ decays and the world averages. The parameter x_D has contributions from short range box diagrams where two W bosons are exchanged and long range hadronic interactions (see Figures 1.2 (a) and (b)) whereas y_D is dominated by contributions from the latter. The decay amplitude of an initial D^0 state decaying to some final state f at a given time t can then be written as

$$\begin{aligned} A_{D^0 \to f}(t) &= \langle f | D^0 \rangle \langle D^0 | D^0(t) \rangle + \langle f | \overline{D}^0 \rangle \langle \overline{D}^0 | D^0(t) \rangle \\ &= \mathcal{A}_{D^0 \to f} \langle D^0 | D^0(t) \rangle + \mathcal{A}_{\overline{D}^0 \to f} \langle \overline{D}^0 | D^0(t) \rangle \\ &= \mathcal{A}_{D^0 \to f} \;\; g_+(t) + \mathcal{A}_{\overline{D}^0 \to f} \;\; \frac{p}{q} g_-(t) \end{aligned} \tag{2.11}$$

and equivalently for $\overline{D}^0$

$$A_{\overline{D}^0 \to f}(t) = \mathcal{A}_{D^0 \to f} \;\; \frac{q}{p} g_-(t) + \mathcal{A}_{\overline{D}^0 \to f} \;\; g_+(t). \tag{2.12}$$

In 2-body D^0 decays the time dependent amplitude can be expressed as the time-integrated parts, $\mathcal{A}_{D^0 \to f}$ and $\mathcal{A}_{\overline{D}^0 \to f}$, are constant across all of the available phase-space. However, for multi-body decays things are not so simple. There are extra degrees of freedom in momentum space where the decay amplitude can vary due to intermediate two body resonances.

Therefore, to construct an expression for the time-dependent decay rate of $D^0 \to K_S^0\pi^+\pi^-$ decays, one must understand how the decay amplitude varies across momentum space.

2.1.1 The Dalitz Plot

The Dalitz plot is an extremely powerful tool for analysing the amplitude structure of multi-body decays that proceed via intermediate resonant states. The technique was first developed by R. H. Dalitz when studying $K^+ \to \pi^+\pi^-\pi^+$ decays [49]. The idea was to measure the angular properties of the decay by looking at the various configurations of the three pions. The frequency with which various configurations occur is dependent on the spin of the K^+, so by measuring this, one can obtain information about the total angular momentum for the K^+. The fundamental idea behind the technique that Dalitz chose was to be able to describe the full dynamics of the decay, i.e. the frequency of various decay configurations, using as few kinematic variables as possible. In the case of 3-body decays involving spin-0 particles there are 12 degrees of freedom in the final state (3 four-vectors), but these can can be reduced to just 2. In the parent rest frame all the momenta lie in the same plane (rotational symmetry) removing 1 degree of freedom. Applying conservation of 4-momentum removes an additional 4. Finally, if the masses of the parent and daughter particles are known, then this removes a further 4 leaving just 2. For higher multiplicity decays the number of remaining degrees of freedom starts to increase rapidly. For instance, 4-body decays can only be reduced to 5 degrees of freedom and 5-body decays to just 8. The choice of variables to be used depends on what is being measured and the dynamics of the decay in question, but for 3-body hadronic decays these are generally taken as the invariant mass squared of pairs of the final state particles. An example Dalitz plot is shown in Figure 2.1. A more detailed discussion about Dalitz plots can be found in [23].

Labelling the final state particles 1 (K_S^0), 2 (π^+) and 3 (π^-), the decay rate as a function of position in the Dalitz plot can then be written as

$$d\Gamma = \frac{1}{(2\pi)^3 32 s^{3/2}} |\mathcal{M}|^2 dm_{12}^2 dm_{13}^2, \tag{2.13}$$

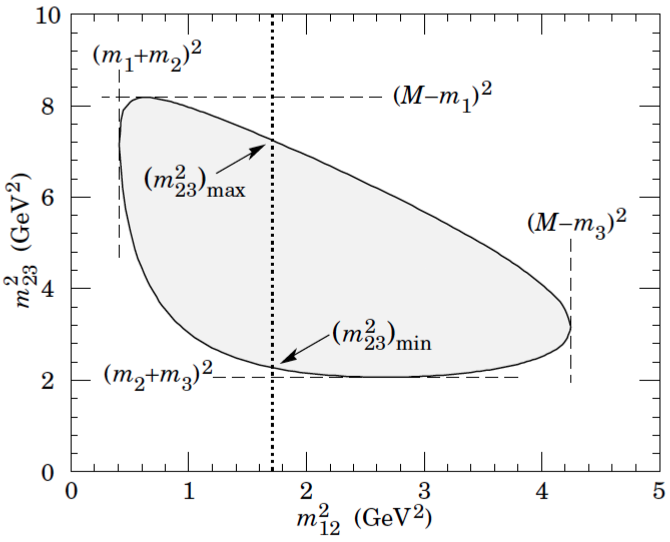


Figure 2.1: Example of a Dalitz plot for a three-body final state. The daughter particles are labelled 1,2,3 and the kinematic boundaries are indicated. The shaded region represents the kinematically allowed region of events. Figure taken from [23].

where m_{ij}^2 is the invariant mass squared of particles i and j and $\mathcal{M}$ is the matrix element describing the dynamics of the decay. The pre-factor containing $s^{3/2}$ is due to the decay kinematics. It can be seen that the density of events given a phase-space decay will be constant across the m_{12}^2, m_{13}^2 plane. Any structure in the Dalitz plot will be as a result of a change in $|\mathcal{M}|^2$ and therefore the Dalitz Plot distribution is only sensitive to the dynamics of the decay. Decays that do not pass through intermediate resonances are known as non-resonant (NR) or phase-space decays and are uniformly distributed across the Dalitz plot. The matrix element for these events is a constant. For decays that follow the chain $D \to rc, r \to ab$ where r is an intermediate resonance, the matrix element can be evaluated as

$$\mathcal{M}_r = Z(L, \boldsymbol{p}_c, \boldsymbol{p}_a) B_r^L(|\boldsymbol{p}_c|) B_r^L(|\boldsymbol{p}_a|) \mathcal{T}_r(m_{ab}), \tag{2.14}$$

where L is the orbital angular momentum between r and c, $\boldsymbol{p}_i$ is the momentum of daughter particle i in the r rest frame and m_{ab} is the invariant mass of the resonant pair. The Z term describes the angular distribution of a and c for the case where all the final state particles are spin-0 (as is the case for $D^0 \to K_S^0\pi^+\pi^-$). $B^L(p_i)_r$ are known as the Blatt-Weisskopf Barrier Factors [50, 51] or Form Factors. The maximum angular momentum achievable is limited in interactions with slow moving particles in which case the production of resonances with spin will be suppressed. The barrier factors weight the interaction rate to account for this effect. $\mathcal{T}_r$ describes the line shape of the resonant pair as a function of their invariant mass. If there are multiple resonant states contributing to the decay amplitude then one can simply take the sum of the matrix elements of each resonance rotated by a phase that describes the interference between the resonances i.e.

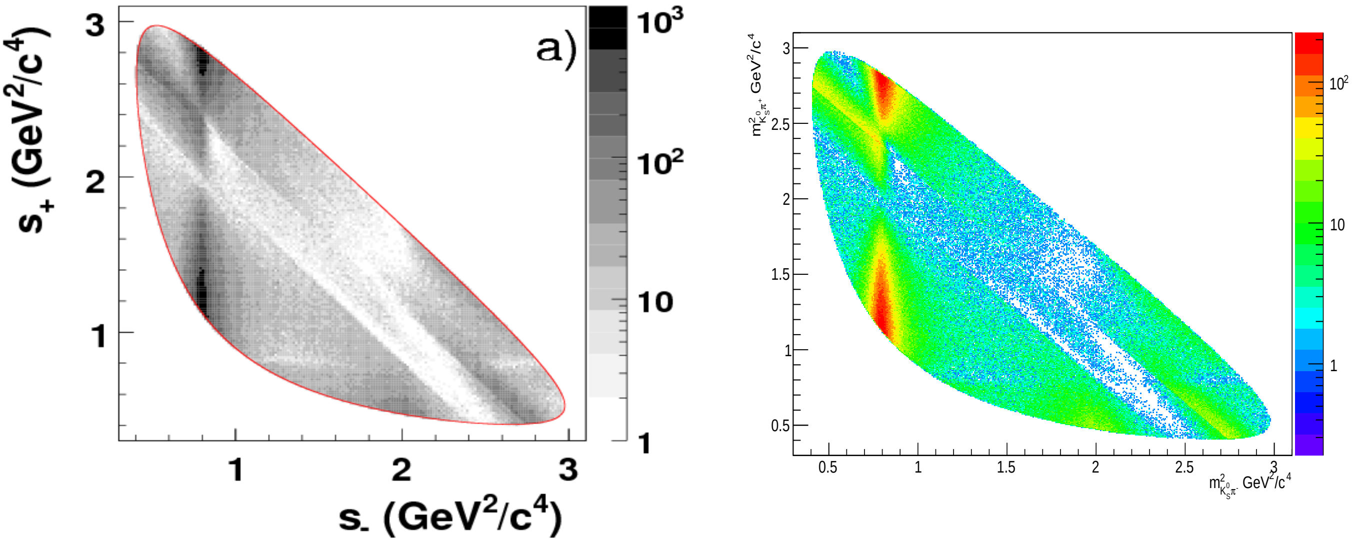
$$\mathcal{M} = \mathcal{M}_0 e^{i\delta_0} + \sum_r \mathcal{M}_r e^{i\delta_r}, \tag{2.15}$$

where the sum runs over all resonant states r, and the first term describes the non-resonant contribution to the decay. This is the so-called isobar approach to modelling decay amplitudes. Often the individual mass line shapes are modelled as Briet-Wigner (BW) distributions [52], but this approach can cause problems. If there is significant interference between the decay amplitudes then the coherent sum of individual BWs can violate unitarity. This problem is exacerbated in regions that contain broad overlapping resonances as even a small phase difference can lead to a large interference term (constructive or destructive). However, there are more advanced line shapes that can better model interference effects in these regions. For example, the K-matrix parameterisation can describe multiple overlapping resonances empirically and does not violate unitarity [53]. More realistic models have also been developed by studying experimental data. One example is the LASS [54] parameterisation which was originally introduced to model $K^-\pi^+$ resonances in $K^-p \to K^-\pi^+n$ interactions at the LASS experiment at SLAC, but can be generalised to describe the S-wave (spin-0) component in $K\pi$ resonances in any multi-body decay with a K and π in the final state.

Constructing a phase-space-dependent amplitude model for a multi-body decay involves selecting a set of resonances and associated line shapes that are representative of the data. The fit fraction for each resonance is determined by integrating the squared amplitude for that component across the kinematically allowed region and normalising such that the sum over the integrals for all resonant contributions is 1. Each contribution should be roughly equivalent to the relative branching fraction for that intermediate decay. By cross referencing with the measured branching fractions quoted by the PDG [23] one can estimate which modes are likely to be present in the data. Starting with a simple model containing the dominant contributions one can add and remove resonances that are physically allowed, but that the data may be less sensitive to, and fit each model to the data until they are in good agreement (usually by finding the model that gives the best χ^2/ndf). Fitting an amplitude model to the data involves floating the relative magnitudes and phases of each resonance with one fixed as a reference plus any shape parameters for the individual components i.e the mass and width for a resonance modelled with a Briet-Wigner. This trial and error approach is laborious and the only figure of merit for the model is the fit quality. In addition to this, any model will not only make assumptions about the type of resonances that contribute to the decay, but also about their respective line shapes, both of which will contribute to an additional systematic uncertainty on any physics parameters that are extracted with this technique. Despite this, model building of this type has been used extensively in mixing and CPV analyses with great success including several analyses on $D^0 \to K^0_S\pi^+\pi^-$ data performed by the BaBar, Belle and CLEO experiments [47, 55, 56]. The model developed in the most recent BaBar analysis [47] is cited frequently throughout this thesis, so a brief description is given here. This model consists of a series of resonances described using Briet-Wigners with the K-matrix formalism to describe the $\pi\pi$ S-wave component and a generalised LASS parameterisation for the $K\pi$ S-wave component in the $K^*_0(1430)^\pm$ decays. A list of each component and their fit fractions is given in Table 2.1. A χ^2 test of the model gave $\chi^2/ndf = 1.215$ for approximately 8,500 degrees of freedom. A visual comparison of the Dalitz distribution between the data collected by BaBar and toy Monte Carlo (see Section 3.2.9), that has been generated using the model they fit to it, is shown in Figure 2.2. It is clear that the model does a good job of describing the data. However, as available datasets get larger, reducing the systematic uncertainty associated with model building will become ever more important.

Component		Fit Fraction (%)
$K^*(892)^-$	(BW)	57.0
$\rho(770)^0$	(BW)	21.1
$K_0^*(1430)^-$	(LASS)	6.1
$K_2^*(1430)^-$	(BW)	1.9
$\omega(782)$	(BW)	0.6
$K^*(892)^+$	(BW)	0.6
$K^*(1680)^-$	(BW)	0.3
$f_2(1270)$	(BW)	0.3
$K_0^*(1430)^+$	(LASS)	< 0.1
$K_2^*(1430)^+$	(BW)	< 0.1
$\pi\pi$ S-wave	(K-matrix)	15.4

Table 2.1: List of resonant contributions and their fit fractions from the $K_S^0\pi^+\pi^-$ BaBar 2010 model [47].



(a) (b)

Figure 2.2: (a) Dalitz plot of $D^0 \to K_S^0\pi^+\pi^-$ data recorded by the BaBar collaboration [47]. (b) Dalitz plot of toy Monte Carlo generated using the model that best fits the data in (a). Note both are plotted with a log scale on the z-axis and that $s_+ = m^2_{K_S^0\pi^+}$ and $s_- = m^2_{K_S^0\pi^-}$.

2.1.2 Time-dependent Amplitude Analysis

Given that the full dynamics of a 3-body hadronic decay can be described using just two Dalitz plot variables, then the total time dependent amplitude of a 3-body decay can be described as follows. Returning to Equations 2.11 and 2.12 and replacing $\mathcal{A}$ with $\mathcal{A}(m_{12}^2, m_{13}^2)$ one has the following expressions for the time dependent decay amplitude

$$A_{D^0}(t, m_{12}^2, m_{13}^2) = \mathcal{A}_{D^0}(m_{12}^2, m_{13}^2) g_+(t) + \mathcal{A}_{\bar{D}^0}(m_{12}^2, m_{13}^2) \frac{p}{q} g_-(t) \tag{2.16}$$

$$A_{\bar{D}^0}(t, m_{12}^2, m_{13}^2) = \mathcal{A}_{D^0}(m_{12}^2, m_{13}^2) \frac{q}{p} g_-(t) + \mathcal{A}_{\bar{D}^0}(m_{12}^2, m_{13}^2) g_+(t). \tag{2.17}$$

The value of $\mathcal{A}_{D^0}(m_{12}^2, m_{13}^2)$ at a given value of m_{12}^2, m_{13}^2 can be expressed as

$$\mathcal{A}_{D^0}(m_{12}^2, m_{13}^2) = a_{12,13} e^{i\delta_{12,13}}, \tag{2.18}$$

where $a_{12,13}$ is the amplitude and $\delta_{12,13}$ is the relative phase at a given point in the Dalitz plot. Ignoring direct CP-violation the decay amplitudes for D^0 and $\bar{D}^0$ are related by,

$$\mathcal{A}_{\bar{D}^0}(m_{12}^2, m_{13}^2) = \mathcal{A}_{D^0}(m_{13}^2, m_{12}^2) = a_{13,12} e^{i\delta_{13,12}} \tag{2.19}$$

such that the Dalitz plot amplitude at $t = 0$ is invariant under the CP transformation. Here it is convenient to define the following expression,

$$\frac{q}{p} = r_{CP} e^{i\alpha_{CP}}, \tag{2.20}$$

where r_{CP} and α_{CP} parameterise the amount of indirect CPV. Substituting these into Equation 2.16 results in a total decay amplitude for $D^0 \to K_S^0 \pi^+ \pi^-$ of

$$A_{D^0 \to K_S^0 \pi^+ \pi^-}(t, m_{12}^2, m_{13}^2) = a_{12,13} e^{i\delta_{12,13}} g_+(t) + r_{CP} e^{i\alpha_{CP}} a_{13,12} e^{i\delta_{13,12}} g_-(t). \tag{2.21}$$

Taking the squared modulus of the amplitude gives the decay probability

$$
\begin{aligned}
|A_{D^0\to K^0_S\pi^+\pi^-}(t, m^2_{12}, m^2_{13})|^2 = \; & a^2_{12,13}|g_+(t)|^2 + a^2_{13,12}r^2_{CP}|g_-(t)|^2 \\
& + 2\mathcal{R}e\left[a_{12,13}a_{13,12}r_{CP}e^{i(\delta_{12,13}-\delta_{13,12}-\alpha_{CP})}g_+(t)g^*_-(t)\right]. \\
= \; & a^2_{12,13}e^{-\Gamma t} + a_{12,13}a_{13,12}r_{CP}\Gamma t e^{-\Gamma t} \; \cdot \\
& [y_D\cos(\delta_{12,13}-\delta_{13,12}-\alpha_{CP}) + x_D\sin(\delta_{12,13}-\delta_{13,12}-\alpha_{CP})] \\
& + \mathcal{O}(x^2_D) + \mathcal{O}(y^2_D) + \mathcal{O}(x_D y_D),
\end{aligned}
\tag{2.22}
$$

where x_D and y_D are the D mixing parameters defined in Equations 2.8 and 2.9. Terms involving x^2_D, y^2_D and $x_D y_D$ are ignored in what follows. It is clear that this expression depends on the magnitudes and the difference between the phases of both the D^0 and $\overline{D}^0$ decay amplitudes at a given point in the Dalitz plot. By examining the density of the Dalitz plot in localised regions, one can obtain information about the absolute value of the squared amplitude, but it is not possible to extract the strong phases by doing this alone. Mixing analyses in $D^0 \to K^0_S\pi^+\pi^-$ have previously adopted the approach of building an amplitude model, as described in the previous section, to extract the strong phase information. Once a suitable model is found then it can be incorporated into a maximum likelihood fit using the above expression to extract the mixing and CPV parameters. However, it is possible to remove this model dependence and its associated systematic uncertainty by using strong phase measurements in bins of the Dalitz plot carried out at an e^+e^- collider as inputs, as described in the following section.

2.2 Removing Model Dependence

2.2.1 The CKM angle γ

The Dalitz plot binning formalism used in the analysis presented in this thesis follows that originally proposed by Giri et al. [57] as a model-independent approach to measuring the

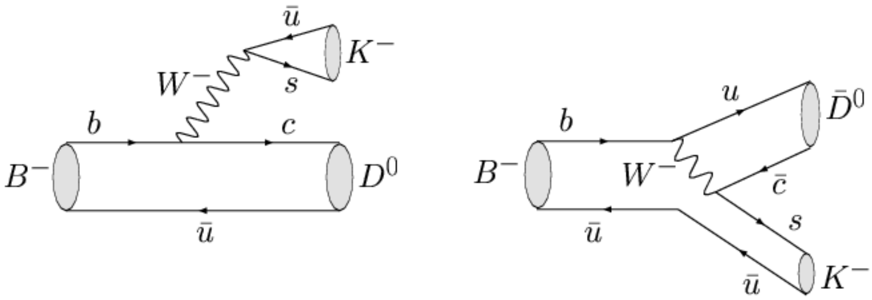


Figure 2.3: Feynman diagrams contributing to the $B^- \to \tilde{D}K^-$ decay amplitude.

CKM angle $\gamma = \arg[-V_{ud}V_{ub}^*/V_{cd}V_{cb}^*]$ (see Section 1.8) in $B^\mp \to \tilde{D}K^\mp$, $\tilde{D} \to K_S^0\pi^+\pi^-$ decays (where $\tilde{D}$ can be D^0 or $\overline{D}^0$). Accurate measurement of γ is a vital consistency check of the SM and achieving this has been one of the key goals of recent heavy flavour experiments including LHCb. $B^\mp \to \tilde{D}K^\mp$ decays propagate by two dominant processes (see Figure 2.3); a $b \to c\bar{u}s$ transition with a resulting D^0 and a colour suppressed $b \to u\bar{c}s$ transition with a resulting $\overline{D}^0$. If both of the D mesons decay to $K_S^0\pi^+\pi^-$, the Dalitz plot distribution is dependent on the interference between the D^0 and $\overline{D}^0$ amplitudes. Neglecting D^0-$\overline{D}^0$ mixing and CP-violating effects in the D decay, the decay amplitude as a function of the D Dalitz plot position for the B^- decay can be written as

$$\begin{aligned}\mathcal{A}_{B^-\to\tilde{D}K^-}(m_{12}^2, m_{13}^2) &\propto \mathcal{A}_{D^0\to K_S^0\pi^+\pi^-}(m_{12}^2, m_{13}^2) + \lambda r_B e^{i(\delta_B-\gamma)}\mathcal{A}_{\bar{D}^0\to K_S^0\pi^+\pi^-}(m_{12}^2, m_{13}^2)\\ &\propto a_{12,13}e^{i\delta_{12,13}} + \lambda r_B e^{i(\delta_B-\gamma)} a_{13,12}e^{i\delta_{13,12}},\end{aligned} \tag{2.23}$$

where δ_B is the relative strong phase between the two diagrams in Figure 2.3 and $r_B \approx c_F|V_{cs}V_{ub}^*|/|V_{us}V_{cb}|$ is the ratio of their amplitudes, where c_F is a colour suppression factor. The equivalent expression for the B^+ decay is obtained by interchanging $\gamma \to -\gamma$ and $\mathcal{A}_{D^0\to K_S^0\pi^+\pi^-}(m_{12}^2, m_{13}^2) \to \mathcal{A}_{\bar{D}^0\to K_S^0\pi^+\pi^-}(m_{12}^2, m_{13}^2)$. As these are tree level processes they are unlikely to be affected by new physics and thus provide a very clean probe of γ. If there is significant CP-violation in the B decay then the resulting Dalitz plot distributions for B^- and B^+ will differ due to the interference between the D^0 and $\overline{D}^0$ amplitudes being rotated by the phase γ and $-\gamma$, respectively. The squared amplitude for the B^- decay at a given point in the D Dalitz plot can be written as

$$\begin{aligned}|\mathcal{A}_{B^-\to\tilde{D}K^-}(m_{12}^2, m_{13}^2)|^2 &= a_{12,13}^2 + r_B^2 a_{13,12}^2 + 2r_B[a_{12,13}a_{13,12}^* e^{i(\delta_{12,13}-\delta_{13,12})}e^{-i(\delta_B-\gamma)}] \\ &= a_{12,13}^2 + r_B^2 a_{13,12}^2 + 2r_B a_{12,13}a_{13,12}\times \\ &\quad [\cos(\delta_{12,13}-\delta_{13,12})\cos(\delta_B-\gamma) + \sin(\delta_{12,13}-\delta_{13,12})\sin(\delta_B-\gamma)],\end{aligned} \tag{2.24}$$

where the interference term is assumed symmetric under the transition $m_{12}^2 \to m_{13}^2$ and subsequent complex conjugation. Dividing the Dalitz plot up into $\mathcal{N}$ bins and integrating over the phase-space of the i^{th} bin one can define the following quantities:

$$T_i = \int_i a_{12,13}^2 \mathrm{d}m_{12}^2 \mathrm{d}m_{13}^2, \tag{2.25}$$

$$c_i = \frac{1}{\sqrt{T_i T_{-i}}}\int_i a_{12,13}a_{13,12}\cos(\delta_{12,13}-\delta_{13,12})\mathrm{d}m_{12}^2 \mathrm{d}m_{13}^2, \tag{2.26}$$

$$s_i = \frac{1}{\sqrt{T_i T_{-i}}}\int_i a_{12,13}a_{13,12}\sin(\delta_{12,13}-\delta_{13,12})\mathrm{d}m_{12}^2 \mathrm{d}m_{13}^2., \tag{2.27}$$

where T_i is the integral of the squared amplitude and c_i (s_i) is the integral of the amplitude weighted by the cosine(sine) of the strong phase difference between D^0 and $\overline{D}^0$ across bin i. Due to the symmetry of the interference term, Giri et al. suggested dividing the Dalitz plot in bins that are symmetric about the $m_{12}^2 = m_{13}^2$ axis as indicated in Figure 2.4. The bins are labelled such that the upper half ($m_{12}^2 > m_{13}^2$) are indexed i and the lower half ($m_{13}^2 > m_{12}^2$) are indexed $-i$ where i runs from 1 to n and the total number of bins is $\mathcal{N} = 2n$. The relationships between the variables c_i and s_i when transitioning from i to $-i$ are given by

$$c_{-i} = c_i \quad \text{and} \quad s_{-i} = -s_i. \tag{2.28}$$

There is no relationship between T_i and T_{-i}. It is clear that if the bins were centred along the axis of symmetry that s_i would equal zero. By constructing a binning scheme in this manner and integrating the amplitude expression given in Equation 2.24 and its conjugate

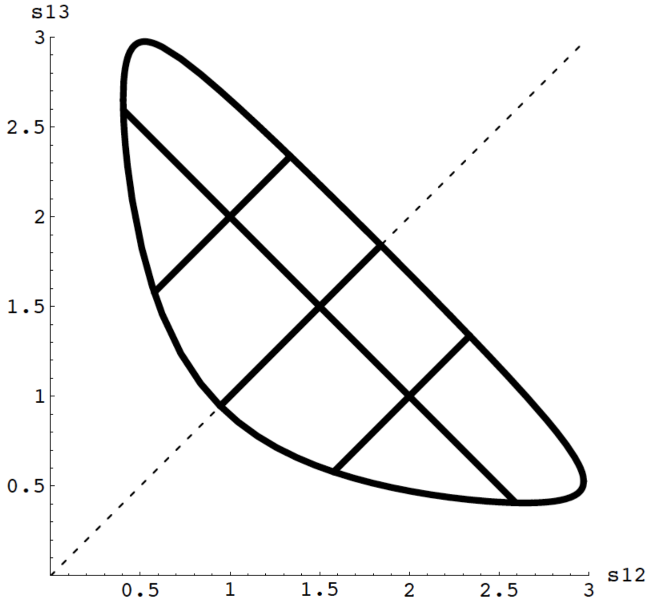


Figure 2.4: Example of symmetric binning where $s_{12} = m^2_{K^0_S\pi^+}$ and $s_{13} = m^2_{K^0_S\pi^-}$. Figure taken from [57].

across bins i and $-i$ one ends up with a set of $4n$ equations with $2n + 3$ unknowns,

$$c_i, \;\; s_i, \;\; r_B, \;\; \delta_B, \;\; \gamma \tag{2.29}$$

that are solvable for the cases where $n \geq 2$. Giri et al. also noted that one can measure the values of c_i and s_i for a binning scheme with $n \geq 2$ bins by looking at pairs of D^0 $\overline{D}^0$ mesons produced in $\Psi(3770)$ decays (a measurement that was performed by the CLEO collaboration [58] and is discussed in the following section) and use the results as input to a measurement of γ that requires no model of the resonant structure of the D decay. The technique was further examined by Bondar and Poluektov [59] who noted that the sensitivity to γ can be greatly increased if the binning is optimised using an amplitude model. A particularly interesting feature of their study was that optimising the bins in such a manner does not introduce a model-dependence, but that a poor choice of binning will reduce the statistical sensitivity to γ.

2.2.2 Mixing and CPV in $D^0 \to K^0_S\pi^+\pi^-$

It has been suggested that the same approach to removing model dependence in the measurement of γ can be applied to mixing and CPV in $D^0 \to K^0_S\pi^+\pi^-$ decays [60]. Starting from Equation 2.22 and integrating over a region i of the Dalitz plot, one arrives at

$$\begin{aligned}\int_i |A_{D^0\to K^0_S\pi^+\pi^-}(t, m^2_{12}, m^2_{13})|^2 dm^2_{12}dm^2_{13} = \ & e^{-\Gamma t}\int_i a^2_{12,13}dm^2_{12}dm^2_{13} + r_{CP}\Gamma t e^{-\Gamma t} \ \times \\ & \left[y_D \int_i a_{12,13}a_{13,12}\cos(\delta_{12,13} - \delta_{13,12} - \alpha_{CP})dm^2_{12}dm^2_{13} \right. \\ & \left. + x_D \int_i a_{12,13}a_{13,12}\sin(\delta_{12,13} - \delta_{13,12} - \alpha_{CP})dm^2_{12}dm^2_{13}\right]. \end{aligned} \tag{2.30}$$

Using the trigonometric identities $\cos(a \pm b) = \cos a \cos b \mp \sin a \sin b$ and $\sin(a \pm b) = \sin a \cos b \pm \cos a \sin b$ and substituting in Equations 2.25, 2.26 and 2.27 one arrives at an expression for the time-dependent decay rate in a given Dalitz plot bin,

$$\begin{aligned}\mathcal{P}_{D^0}(t)_i = e^{-\Gamma t}[T_i + r_{CP}\Gamma t\sqrt{T_i T_{-i}} \times \{&y_D(c_i \cos\alpha_{CP} + s_i \sin\alpha_{CP}) \\ &+ x_D(s_i \cos\alpha_{CP} - c_i \sin\alpha_{CP})\}].\end{aligned} \tag{2.31}$$

The equivalent expression for $\overline{D}^0$ is given by

$$\begin{aligned}\mathcal{P}_{\overline{D}^0}(t)_i = e^{-\Gamma t}[T_{-i} + \frac{1}{r_{CP}}\Gamma t\sqrt{T_i T_{-i}} \times \{&y_D(c_i \cos\alpha_{CP} + s_i \sin\alpha_{CP}) \\ &- x_D(s_i \cos\alpha_{CP} - c_i \sin\alpha_{CP})\}],\end{aligned} \tag{2.32}$$

where direct CP-violation is assumed to be zero such that T_i and T_{-i} are interchanged when moving from D^0 to $\overline{D}^0$. Again, it is clear that for $\mathcal{N} = 2n$ bins one has $4n$ equations with $2n + 4$ unknowns,

$$c_i,\ \ s_i,\ \ x_D,\ \ y_D,\ \ \alpha_{CP},\ \ r_{CP} \tag{2.33}$$

that are solvable for $n \geq 2$. Note that Γ is also technically an unknown. However, the average decay width is obtainable without performing a Dalitz plot fit and the world average for this parameter is known with relatively high precision such that a binning scheme consisting of

$\mathcal{N} > 4$ bins would suffice for this type of measurement. Applying exactly the same principle as Giri et al. proposed for measuring γ offers a framework to measure charm mixing in $K^0_\mathrm{S}\pi^+\pi^-$ decays without any knowledge of the resonant structure of the Dalitz plot, thus removing any model dependence and its associated systematic uncertainty. Analyses that utilise this binned technique require input values of c_i and s_i. The CLEO experiment has provided this information and developed several binning schemes for both $K^0_\mathrm{S}\pi^+\pi^-$ and $K^0_\mathrm{S}K^+K^-$ that were originally intended for use in the measurement of γ, but that can also be used in the context of charm mixing.

2.3 Measurement of the D^0-$\overline{D}^0$ Strong Phase Difference at CLEO

The following section summarises the techniques used to extract the strong phase information at CLEO and how the binning scheme used in this analysis was developed. For a more detailed description of the analysis performed by CLEO please refer to [58].

The relative strong phase between D^0 and $\overline{D}^0$ can be accessed by studying D mesons produced at the $\Psi(3770)$ resonance. The $\Psi(3770)$ decays to a $D\bar{D}$ pair that is in an entangled state until one of them decays. After one of the D mesons decays, the state of the other meson is then well defined, but until this point the two are indistinguishable in terms of their CP and flavour content. This is a useful phenomenon as, by examining the CP content of one of the D decays, known as the tag decay, one can obtain information about the CP content of the other D decay. The CLEO experiment ran at the $\Psi(3770)$ resonance for a period between 2003 and 2006 amassing a few fb^{-1} of data and performed an analysis to extract values of c_i and s_i from $\tilde{D} \to K^0_{S,L}h^+h^-$ (here, $\tilde{D}$ represents D^0 and $\overline{D}^0$) decays. Examining event yields in binned Dalitz plots for both CP-tagged and flavour tagged $\tilde{D} \to K^0_\mathrm{S}h^+h^-$ decays offers sensitivity to c_i and s_i. As the $\Psi(3770)$ has CP $= -1$, by looking at events where one of the D mesons decays to a CP eigenstate, for instance $D^0 \to \pi^+\pi^-, K^+K^-$, gives information about the CP content of the opposite decay. As $K^0_\mathrm{S}h^+h^-$ is a CP mixed state, the CP content of this channel cannot be accessed directly in other D production modes. Flavour tagging works in a similar way, but where the tagging D decays to a flavour eigenstate, for instance $D^0 \to K^-\pi^+$. The amplitude for a CP-tagged $\tilde{D} \to K^0_S h^+h^-$ decay is given by

$$\mathcal{A}_{CP\pm}(m^2_{12}, m^2_{13}) = \frac{1}{\sqrt{2}}[\mathcal{A}_D(m^2_{12}, m^2_{13}) \pm \mathcal{A}_D(m^2_{13}, m^2_{12})], \tag{2.34}$$

where $(+)$ indicates CP-even and $(-)$ CP-odd states of the decay. Taking the modulus squared of the amplitude and integrating across a given region in the Dalitz plot gives,

$$M^{\pm}_i = n_{CP\pm}(T_i \pm 2c_i\sqrt{T_iT_{-i}} + T_{-i}), \tag{2.35}$$

where $M^{\pm}_i$ is the number of CP-tagged events in bin i, $n_{CP\pm} = N^{\pm}/2N_f$ is a normalisation factor and T_i and T_{-i} are the fractional event yields in bin $\pm i$ of the flavour-tagged Dalitz plot. Here $N^{\pm}$ is the total number of CP-tagged D^0 mesons in the dataset regardless of how the associated D decays and N_f is the number of flavour-tagged decays. This ratio can also be expressed in terms of the branching fractions of the D, $\mathcal{B}_{\mathcal{CP}\pm}(\mathcal{B}_f)$ to CP$_\pm$ (flavour) eigenstates. It is clear from Equation 2.35 that measuring the yields in individual bins from both the flavour-tagged and CP-tagged Dalitz plots gives sensitivity to c_i. Sensitivity to s_i comes from examining $\Psi(3770)$ decays where both the D mesons decay to $K^0_S h^+h^-$. The amplitude for this mode can be written as

$$\mathcal{A}(m^2_{12}, m^2_{13}, m'^2_{12}, m'^2_{13}) = \frac{\mathcal{A}_{D^0}(m^2_{12}, m^2_{13})\mathcal{A}_{D^0}(m'^2_{13}, m'^2_{12}) - \mathcal{A}_{D^0}(m'^2_{12}, m'^2_{13})\mathcal{A}_{D^0}(m^2_{13}, m^2_{12})}{\sqrt{2}}, \tag{2.36}$$

where the primed and unprimed invariant mass squared terms are from the opposite $\tilde{D} \to K^0_S h^+h^-$ decays. Again, taking the modulus squared of this amplitude and integrating across a region in the Dalitz plot one arrives at,

$$M_{ij} = h_{corr}(T_iT_{-j} + T_{-i}T_j - 2\sqrt{T_iT_{-j}T_{-i}T_j}(c_ic_j + s_is_j)), \tag{2.37}$$

where the index i (j) is the index of the first (second) $\tilde{D} \to K^0_S h^+h^-$ decay in the event. The

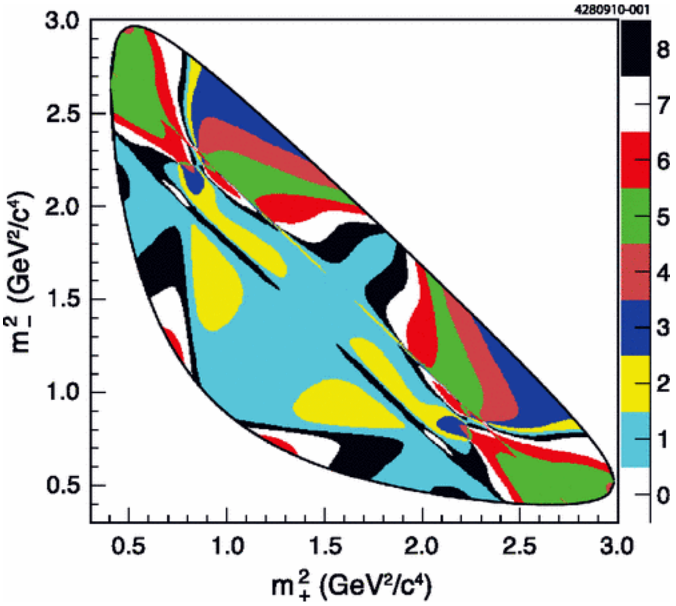


Figure 2.5: Equal-$\Delta\delta$ binning of the $K^0_{\rm S}\pi^+\pi^-$ Dalitz plot developed by CLEO [58].

normalisation term is given by $h_{corr} = N_{D\bar{D}}/2N_f^2 = N_{D\bar{D}}/8N_{D\bar{D}}\mathcal{B}_f^2$ where $N_{D\bar{D}}$ is the number of $D\bar{D}$ pairs both decaying to $K^0_{\rm S}h^+h^-$. It is clear that analysing the CP-tagged, double $K^0_{\rm S}h^+h^-$-tagged and flavour-tagged events together allows c_i and s_i to be determined. In addition to the terms for $D^0 \to K^0_{\rm S}h^+h^-$ one can write analogous expressions for the decay $D^0 \to K^0_{\rm L}h^+h^-$ where the equivalent parameters c_i and s_i are denoted c'_i and s'_i,

$$M_i'^{\pm} = h_{CP_\pm}(T'_i \mp 2c'_i\sqrt{T'_iT'_{-i}} + T'_{-i}), \tag{2.38}$$

$$M'_{ij} = h_{corr}(T_iT'_{-j} + T_{-i}T'_j + 2\sqrt{T_iT'_{-j}T_{-i}T'_j}(c_ic'_j + s_is'_j)). \tag{2.39}$$

Here the CP-tagged expression is purely dependent on the $K^0_{\rm L}h^+h^-$ decay, but the $D\bar{D}$ pair is tagged as one side decaying to $K^0_{\rm S}h^+h^-$ and the other to $K^0_{\rm L}h^+h^-$. This offers extra information about the range that c_i and s_i can occupy in phase-space and constraints of the type $\Delta c_i = c'_i - c_i$ and $\Delta s_i = s'_i - s_i$ are imposed in the determination of their value.

The CLEO analysis made measurements of $c_i^{(\prime)}$ and $s_i^{(\prime)}$ in both $K^0_{\rm S}\pi^+\pi^-$ and $K^0_{\rm S}K^+K^-$ decays using several different binning schemes. But, for the purposes of this mixing analysis, the $K^0_{\rm S}\pi^+\pi^-$ binning scheme selected was the one optimised such that the variation of the strong phase difference across the bin is minimised. The binning scheme was optimised using

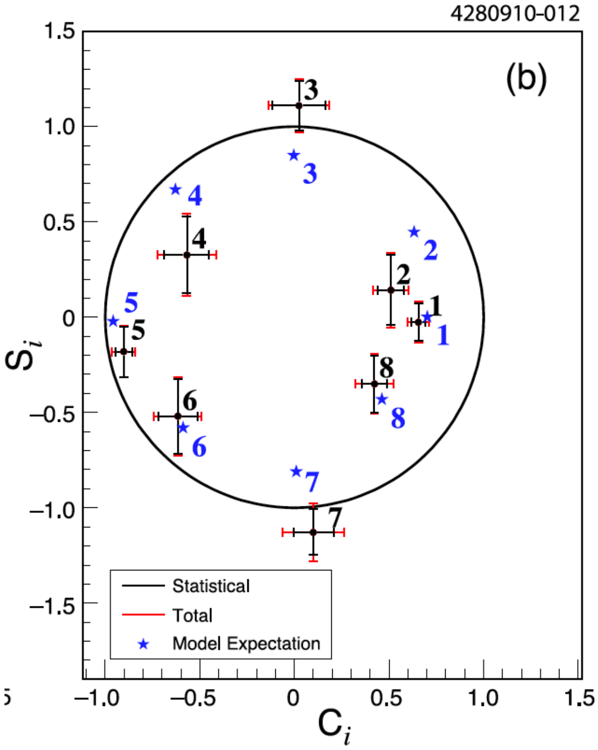


Figure 2.6: Measured values of c_i and s_i for the equal-$\Delta\delta_D$ binning [58]. The circle indicates the $c_i^2 + s_i^2 = 1$ boundary.

a 2008 BaBar amplitude model described in [53]. Splitting the Dalitz plot up into $2n$ bins one can define a bin boundary that minimises the strong phase variation within each bin such that

$$2\pi(i - 3/2)/n < \Delta\delta_D(m_{12,}^2, m_{13}^2) < 2\pi(i - 1/2)/n, \tag{2.40}$$

where $\Delta\delta_D = \delta_{12,13} - \delta_{13,12}$ is taken by evaluating the amplitude model at each data point within the bin. The $-i^{\mathrm{th}}$ bin is defined symmetrically about the $m_{12}^2 = m_{13}^2$ axis. Bondar and Poluektov noted that this provided better sensitivity to γ, but that selecting too fine a binning choice would result in s_i tending to zero and that too coarse would result in large amplitude variations within the bin and violation of the condition $c_i^2 + s_i^2 = 1$. The equal-$\Delta\delta$ binning that CLEO developed, shown in Figure 2.5, consists of $n = 8$ bins and is the binning which is primarily used in the mixing measurement presented in this thesis. The values of T_i, T_{-i}, c_i and s_i that were measured by CLEO are given in Tables 2.2 and 2.3 and Figure 2.6[1].

---

[1] It should be noted that there is a mistake in [58] such that, in the text, they define the $-i^{\mathrm{th}}$ region as $m_{12}^2 < m_{13}^2$ when in fact it is defined as $m_{12}^2 > m_{13}^2$ in the actual binning scheme they utilised in the analysis.

i	$T_i(\%)$		$T_{-i}(\%)$	
	Measured	Predicted	Measured	Predicted
1	17.0 ± 0.5	17.2	8.3 ± 0.4	8.3
2	8.4 ± 0.4	8.4	2.4 ± 0.2	1.9
3	7.2 ± 0.3	6.9	2.3 ± 0.2	1.7
4	2.4 ± 0.2	2.5	1.6 ± 0.2	2.0
5	7.6 ± 0.4	8.6	4.8 ± 0.3	5.3
6	5.9 ± 0.3	5.9	1.3 ± 0.2	1.5
7	12.8 ± 0.5	12.4	1.6 ± 0.2	1.4
8	13.0 ± 0.5	13.0	3.1 ± 0.2	2.8

Table 2.2: Extracted values of T_i from CLEO for the BABAR 2008 equal-$\Delta\delta$ binning [58]. Both the measured values and those predicted from the model are given.

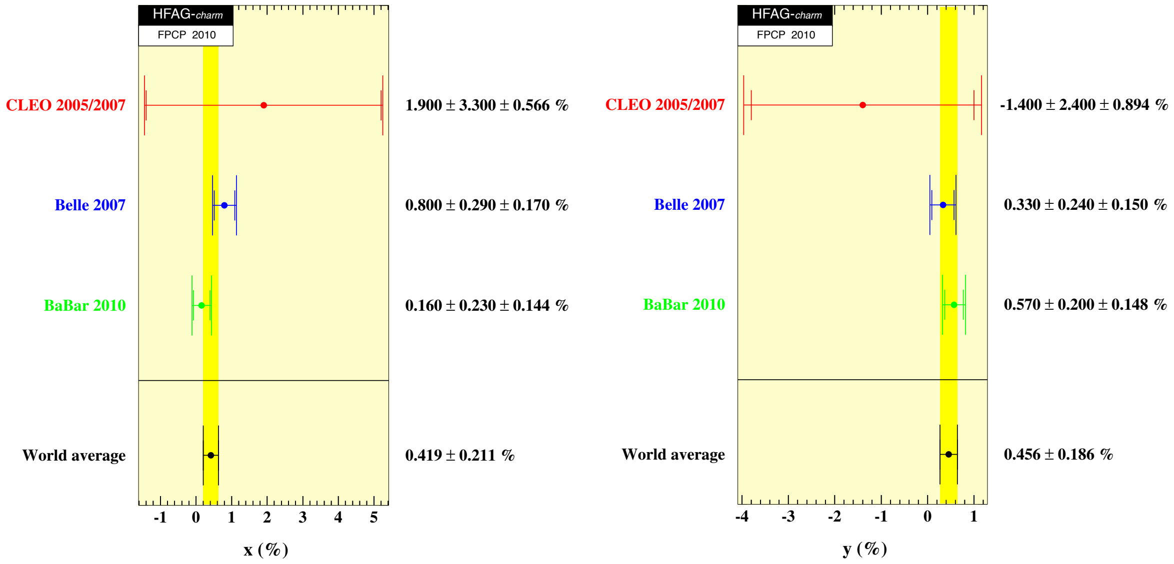
i	c_i	s_i
1	0.655 ± 0.036 ± 0.042	-0.025 ± 0.098 ± 0.043
2	0.511 ± 0.068 ± 0.063	0.141 ± 0.183 ± 0.066
3	0.024 ± 0.140 ± 0.080	1.111 ± 0.131 ± 0.044
4	-0.569 ± 0.118 ± 0.098	0.328 ± 0.202 ± 0.072
5	-0.903 ± 0.045 ± 0.042	-0.181 ± 0.131 ± 0.026
6	-0.616 ± 0.103 ± 0.072	-0.520 ± 0.196 ± 0.059
7	0.100 ± 0.106 ± 0.124	-1.129 ± 0.120 ± 0.096
8	0.422 ± 0.069 ± 0.075	-0.350 ± 0.151 ± 0.045

Table 2.3: Measured values of c_i and s_i from CLEO for the BABAR 2008 equal-$\Delta\delta$ binning where the first uncertainty is statistical and the second is systematic [58].

These values are used as input to the fit to extract the mixing parameters presented in this thesis. A complete description of how this is done is given in Chapter 5. One should note that the constraints imposed on c_i and s_i only enforce the condition $c_i^2 + s_i^2 = 1$ within uncertainties. It is clear from Figure 2.6 that the central values for bins 3 and 7 sit outside of this boundary which may well bias any measurement that uses these values as input. The effect of this and of the uncertainty on the measurements made by CLEO are investigated as a systematic uncertainty in Section 6.

2.4 Status of Mixing and CPV in Charm and Future Prospects

There has been a flurry of analyses aimed at measuring mixing and CP-violation in charm decays in recent years. The high luminosity of the B-factory experiments and the large charm production cross-sections at hadron colliders like the Tevatron [61] and the LHC have enabled them to amass vast datasets offering measurements of unrivalled precision compared with the previous generation of particle physics experiments. Recently, LHCb, the CDF Collaboration at the Tevatron [62] and the Belle Collaboration [63] have published mixing analyses examining the time-dependent ratio of Cabbibo favoured and doubly Cabbibo suppressed $D^0 \to K^{\mp}\pi^{\pm}$ decays that each independently rule out the no mixing hypothesis to $> 5\sigma$ [43, 64, 65]. LHCb has pursued this further and performed a search for direct and indirect CPV in the same channel, but finds that the rate of CPV is consistent with zero [66]. Searches for mixing in other channels such as decays to CP eigenstates $D^0 \to h^+h^-$ where $h = \pi, K$ offer sensitivity to the value of y_{CP} which is equivalent to y_D in the limit of no CP-violation. The current best measurements in this mode from LHCb and BaBar show evidence for charm mixing at the 3.3σ level [67, 68]. The analyses mentioned so far have been performed on channels that have high production rates at high luminosity e^+e^- and hadron colliders and are sensitive to the magnitude and sign of y_D, but only the magnitude of x_D. For example, the $D^0 \to K^{\mp}\pi^{\pm}$ mode is sensitive to $x_D'^2 = (x_D \cos\delta_{K\pi} + y_D \sin\delta_{K\pi})^2$ and $y_D' = y_D \cos\delta_{K\pi} - x_D \sin\delta_{K\pi}$ which are the regular mixing parameters rotated by the strong phase difference, $\delta_{K\pi}$ between the Cabbibo favoured $D^0 \to K^-\pi^+$ and doubly Cabbibo suppressed $D^0 \to K^+\pi^-$ modes, and is thus insensitive to the relative sign between x_D and y_D. Measuring the rate of mixing in $D^0 \to K^0_S\pi^+\pi^-$ decays offers sensitivity to the magnitude and the relative sign between the regular mixing parameters. This mode is also sensitive to indirect CP-violation. The same is true of the $D^0 \to K^0_S\pi^+\pi^-\pi^0$ and $D^0 \to K^0_S K^+K^-$ modes. However, multi-body channels containing one or more π^0 are extremely difficult to reconstruct at hadron collider experiments such as LHCb and there are currently no strong phase measurements available for this mode. There are strong phase



(a) (b)

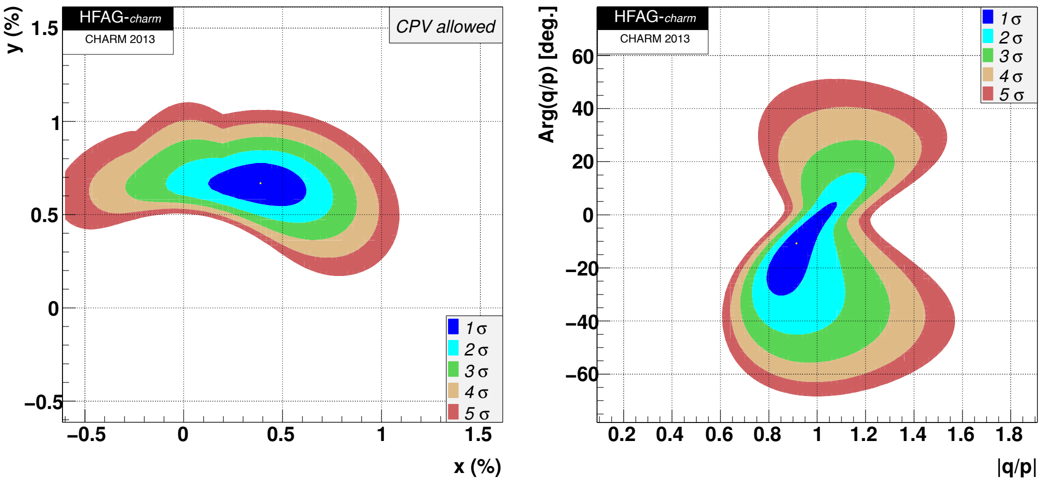
Figure 2.7: World averages for x (a) and y (b) from $D^0 \to K^0_S\pi^+\pi^-$ decays provided by HFAG [48].

measurements available for $D^0 \to K^0_S K^+K^-$, but the decay rate is kinematically suppressed with respect to $K^0_S\pi^+\pi^-$. Therefore, $D^0 \to K^0_S\pi^+\pi^-$ offers a unique high statistics probe of the rate of mixing in charm and the relative sign between x_D and y_D. There are several published mixing analyses in $D^0 \to K^0_S\pi^+\pi^-$ decays from BaBar, Belle and CLEO [47, 55, 56], but none use this model-independent approach. Summaries of all charm mixing and CPV results as well as computation of the world averages are provided by the Heavy Flavour Averaging Group (HFAG) [48]. The current world averages for x and y obtained from analyses of $D^0 \to K^0_S\pi^+\pi^-$ decays are

$$x = (0.419 \pm 0.211)\% \qquad y = (0.456 \pm 0.186)\%, \tag{2.41}$$

where the input measurements and averages for x and y are given in Figures 2.7 (a) and (b), respectively. Incorporating information from all of the available charm mixing and CPV results, the world averages are calculated as

$$x = (0.490 \pm 0.150)\% \qquad y = (0.617 \pm 0.079)\%. \tag{2.42}$$



(a) (b)

Figure 2.8: Contour plots for (a) x, y and (b) $|q/p|$, Arg(q/p) provided by HFAG [48].

The HFAG group produces contour plots that combine the significance of every measurement to help understand the overall picture. Figures 2.8 (a) and (b) show the x_D, y_D and $|q/p| = r_{CP}$, Arg$(q/p) = \alpha_{CP}$ contour plots with coloured regions representing the 1–5σ bands. It is clear from these plots that the no-mixing hypothesis is ruled out with a very large significance, but much is still to be done in reducing the uncertainty on the absolute value of the mixing parameters. This is particularly important for x for which the uncertainty on the current world average is a factor of 2 larger than for y. CP-violation in charm has eluded detection, as expected if it occurs at the small rate predicted by the SM. The smallness of both mixing and CP-violation in charm coupled with the large theoretical uncertainties means that our knowledge can be greatly improved by high precision experimental measurements in the charm sector. With LHCb due to take more data and Belle II being commissioned [69] much more $K^0_S\pi^+\pi^-$ data will become available in the coming years. However, as the available data mounts, so does the necessity to reduce systematic errors associated with any measurement. Analyses that follow the traditional model building approach may struggle to reduce their model systematic as data samples grow. Adding more data will improve the models, but it is not clear that the model systematic uncertainty will scale with the statistical uncertainty when modelling very large datasets. By contrast, the sensitivity of the model-independent method described in this chapter can be improved

Parameter	Stat, 0.5 M	Stat, 10M	Stat, 100M	Syst (c_i, s_i)
$\sigma(x_D)$ (%)	0.2513	0.0541	0.0170	0.0761
$\sigma(y_D)$ (%)	0.2724	0.0611	0.0190	0.0871
$\sigma(\alpha_{CP})$	0.1752	0.0371	0.0120	0.0240
$\sigma(r_{CP})$ (°)	12.46	2.423	0.741	0.882

Table 2.4: Expected statistical and systematic uncertainties from c_i and s_i using the current CLEO strong-phase measurements for a range of sample sizes as determined in [60].

by analysing larger samples of $D^0 \to K^0_S h^+ h^-$ decays produced at the $\Psi(3770)$ resonance, something that experiments such as BES-III [70] are actively pursuing [71]. Additional improvements on the CLEO measurements can be made by developing binning schemes which are optimised to increase their sensitivity to charm mixing. These binnings can be further improved by optimising them with a highly accurate amplitude model. Thus, it is crucial to pursue both the model-dependent and model-independent routes to gain the greatest insights from the data. However, it has been shown by Thomas and Wilkinson [60] that the systematic error on extracted values of the mixing and CPV parameters due to the uncertainty on the current strong phase measurements from CLEO will not be a leading uncertainty for some years of data taking at LHCb. Table 2.4 summarises their findings. The estimated systematic uncertainties due to the current measured values of c_i and s_i from CLEO are listed with estimates of the statistical uncertainties for a range of sample sizes. It is clear that the systematic uncertainty from the current CLEO strong phase measurements will not begin to dominate until data samples exceed several million signal events, almost an order of magnitude larger than the current datasets from Belle and BaBar. The work presented in this thesis is a vital step in the development of this new, powerful tool for measuring charm mixing in $D^0 \to K^0_S h^+ h^-$ decays that will likely play an extensive role in mixing analyses in the near future and in years to come.

Chapter 3

The LHCb Detector

The data analysed in this thesis was recorded during 2011 by the LHCb detector, one of the four main experiments at the Large Hadron Collider (LHC), CERN. The following chapter describes both the LHC and the LHCb detector with a summary of their performance during 2011 data taking.

3.1 The LHC

The Large Hadron Collider (LHC) is the largest particle accelerator in the world [72]. It makes up part of the vast accelerator complex at the European Centre for Nuclear Research (CERN) which straddles the Franco-Swiss border near the city of Geneva. The LHC is a synchrotron accelerator with a circumference of around 27 km housed in a tunnel 100 m underground (see Figure 3.1). It is designed to collide protons at a centre-of-mass energy of $\sqrt{s} = 14$ TeV and an instantaneous luminosity of 1×10^{34} cm^{-2} s^{-1} with a bunch crossing rate of 40 MHz. The protons begin as a hydrogen gas which is then enveloped in an electric field to strip the protons of their electrons. This process gives about 70 % pure protons which are subsequently accelerated using the linear LINAC2 accelerator to an energy of 50 MeV. The 50 MeV beam of protons is then injected into the first of a sequence of 4 synchrotrons the last of which is the LHC. The Proton Synchrotron Booster (PSB) receives protons from LINAC2 and accelerates them to 1.4 GeV at which point they are injected into the Proton Synchrotron (PS). The PS accelerates the protons to 25 GeV before injecting them into the Super Proton Synchrotron (SPS) which accelerates them to 450 GeV, the injection energy for the LHC.

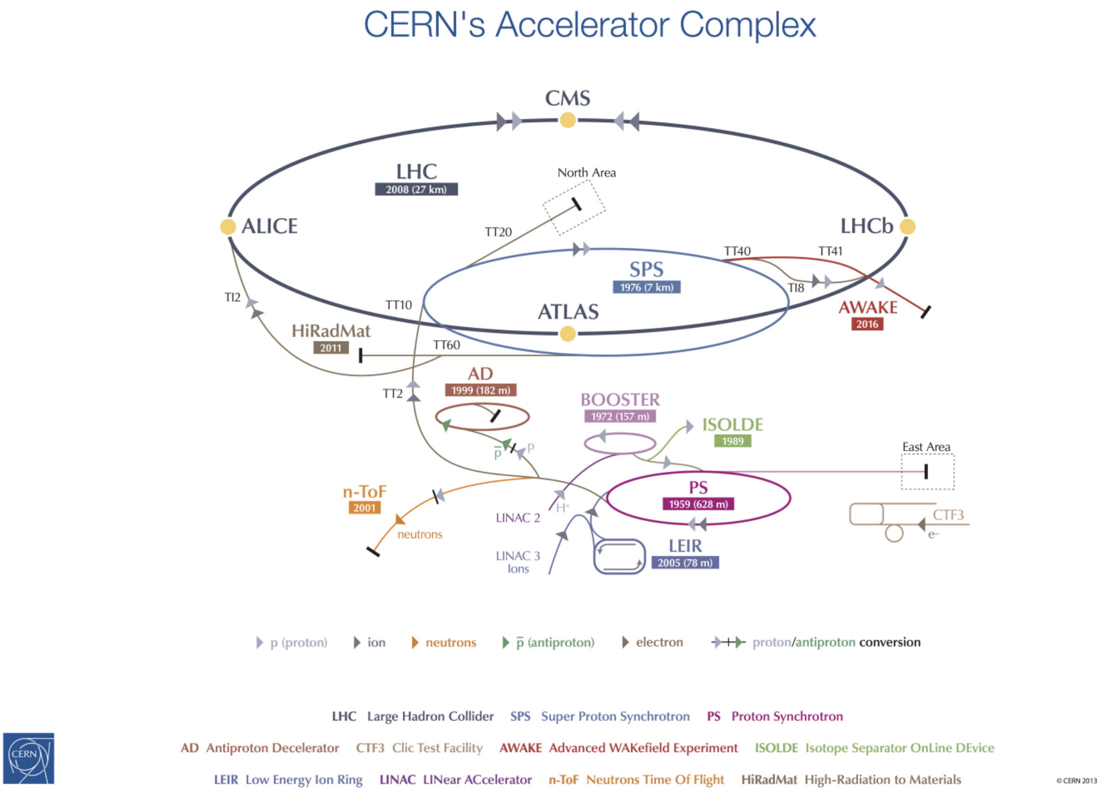


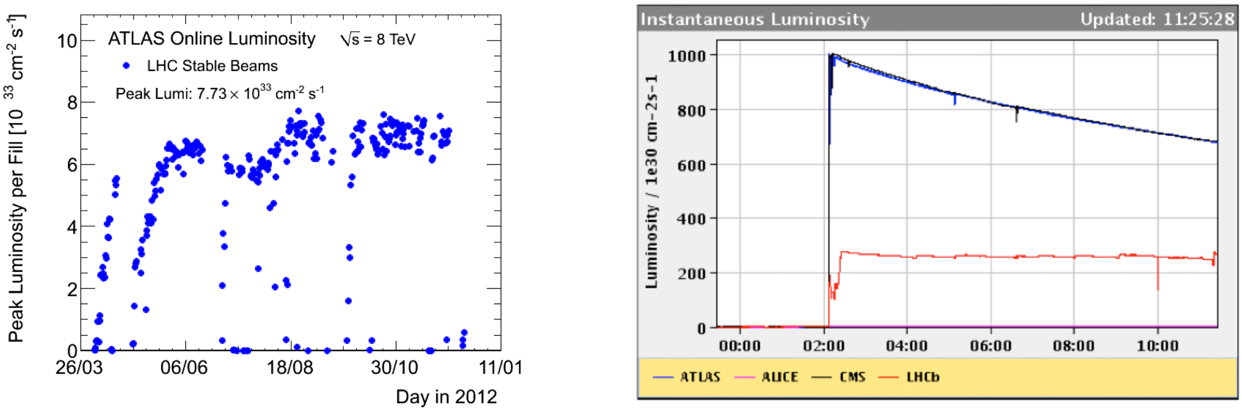
Figure 3.1: The CERN accelerator complex. Image provided by CERN.

Protons from the SPS are then injected into the twin beam pipes of the LHC. This is the first mode of operation for the LHC when preparing to record physics data and is known as a FILL. Once both beam pipes have been filled the LHC begins to accelerate the beams in opposite directions until they reach their colliding energy, a state known as RAMP. Once the protons have been accelerated to their collision energy the LHC moves to SQUEEZE and the beams are collimated to make them as bright as possible before bringing them into collision at the 4 main interaction points. Once the beams are focused and colliding, the machine is in the STABLE BEAMS state at which point the majority of the detectors begin to take data. Once stable beams has been reached only minor adjustments to the beam parameters can be made. The beam currents gradually diminish through interactions as the protons continue to cycle, reducing the number of particles crossing per bunch and the instantaneous luminosity begins to decrease (see Figure 3.2 (b)). Once it drops too low the beams are dumped and the cycle starts again. Fills can last up to 14 hours with the detectors recording data for > 90 % of the fill. Following the accelerator commissioning in 2009, the

LHC ran at a centre-of-mass energy of $\sqrt{s} = 7$ TeV achieving a maximum luminosity of $\sim 3.65 \times 10^{33}$ cm^{-2}s^{-1} during 2010 and 2011. In 2012 the centre-of-mass energy was increased to $\sqrt{s} = 8$ TeV and a peak luminosity of $\sim 7.73 \times 10^{33}$ cm^{-2}s^{-1} was achieved (Figure 3.2 (a)).

The LHC ring provides proton collisions at four large detectors. The two largest, ATLAS [73] and CMS [74] are General Purpose Detectors (GPDs) situated opposite each other on the LHC ring (Figure 3.1). They are both hermetic detectors constructed with cylindrical layers of precision detection systems surrounding a section of the LHC where the beams can be focussed to collide. As general purpose experiments they both have broad research programs and extensive physics reach. However, one of the main physics goals of the GPDs was to perform direct searches for the Higgs Boson, and in July 2012 a $> 5\sigma$ excess of events in several of the Higgs' decay channels was observed by both detectors at a mass of ~ 125 GeV [22, 21]. The GPDs have also been employed in direct searches for particle states that are predicted in extensions to the SM [75, 76, 77], precision tests of the Standard Model [78] and examining the substructure of the proton [79] among many other things [80, 81].

The LHC is also capable of colliding Lead (Pb) ions and short runs with Lead-Lead and Lead-Proton collisions took place at the LHC over 2010 to 2013. The acceleration chain for lead is slightly different to that for protons. In particular, the source for lead ions is isotopically pure lead-208 (^{208}Pb) which is heated until it is a vapour and then ionised by passing an electric field across the vapour at which point the lead ions are extracted to be fed to the accelerator system. The majority of the acceleration chain is the same as for protons, but the initial acceleration is performed by the LINAC3 accelerator which is specifically designed to accelerate lead ions which are then accumulated in a storage ring, the Low Energy Ion Ring (LEIR). This acceleration and storage continues the ionisation process ensuring that all electrons have been removed from the lead. Once enough ions have been accumulated they are injected into the PS accelerator and follow the same acceleration chain to the LHC as for protons. The ion beams are accelerated to a centre-of-mass energy of around $\sqrt{s} = 2.76$ TeV per colliding nucleon pair. The GPDs can be used to record and analyse



(a) (b)

Figure 3.2: (a) Peak instantaneous luminosity delivered to the ATLAS detector during 2012 data taking. (b) Instantaneous luminosity as a function of time delivered to the four large experiments by the LHC during a typical proton-proton fill. Figures taken from [87].

these collisions [81], but another of the four large detectors, the ALICE experiment [82, 83], is specifically designed to record heavy ion data. ALICE is constructed with a similar geometry to the GPDs, but its design is optimised to detect the large numbers of tracks produced in ion collisions. ALICE and the GPDs have conducted searches for Quark-Gluon Plasma, a state of matter close to what might have occurred shortly after the big bang [84] and explored the origin of mass within hadronic bound states [85, 86].

The fourth large experiment on the LHC ring is LHCb [88, 89]. Both ALICE and LHCb have limits on the instantaneous luminosity at which they can operate, $\sim 5 \times 10^{29}$ cm^{-2}s^{-1} and $\sim 4 \times 10^{32}$ cm^{-2}s^{-1}, respectively [88, 82], which are lower than the LHC design luminosity of 10^{34} cm^{-2}s^{-1}. Therefore, special measures have to be taken to ensure that the luminosity provided at the GPDs is maximal while keeping the luminosity at a lower level for ALICE and LHCb. During low luminosity running the LHC can simply limit the number of colliding bunches at ALICE and LHCb. However, running at the design luminosity of the GPDs would require many more colliding bunches at the other two experiments than they can handle. To overcome this the beams are vertically shifted away from each other to reduce the cross-sectional overlap, thus reducing the number of collisions per bunch crossing. This can be actively adjusted during a fill to maintain a constant luminosity at the interaction points of ALICE and LHCb, a process known as luminosity levelling. Figure 3.2 shows the

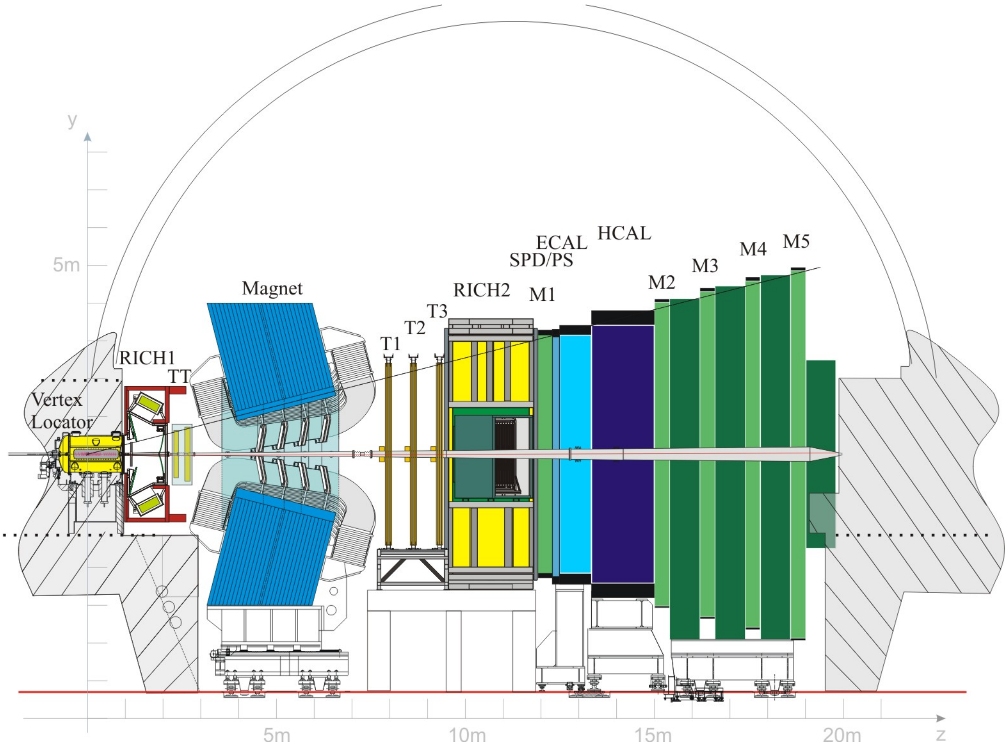
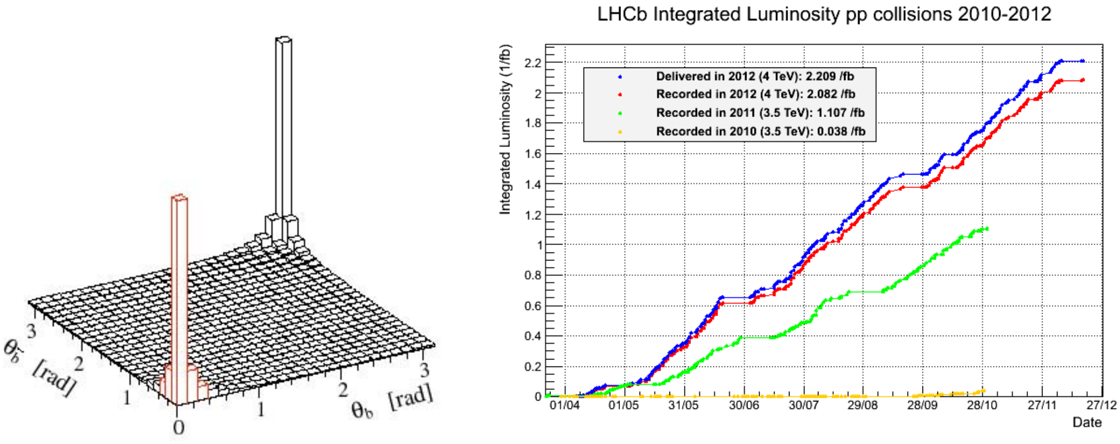


Figure 3.3: Cutaway of the LHCb detector. Figure taken from [89].

luminosity levelling in action during a typical proton-proton fill. The delivered luminosity to the GPDs is at $\sim 1 \times 10^{34}$ cm^{-2}s^{-1} at the start of the fill and gradually falls while it is kept approximately constant at LHCb. As LHCb was used to record the data analysed in this thesis, a more detailed description of the LHCb detector is given in the following sections.

3.2 The LHCb Detector

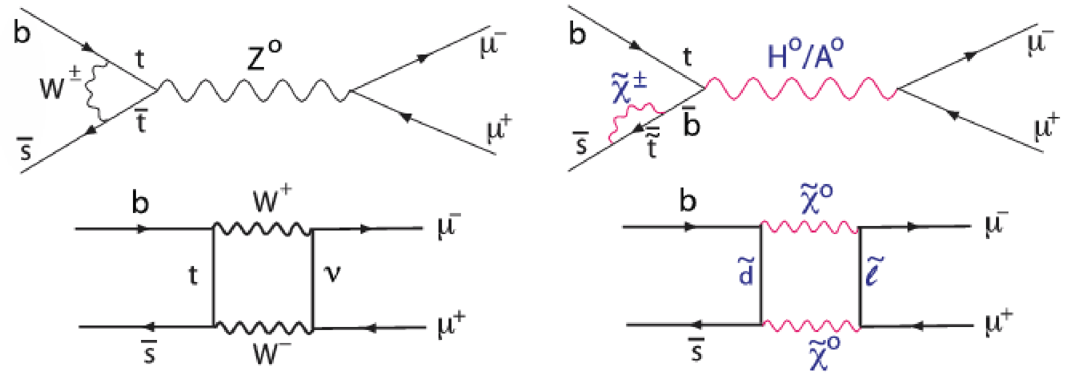
LHCb is a dedicated heavy flavour experiment situated at interaction point 8 of the LHC proton ring [88]. The broad physics goals of LHCb are to probe mixing and CP-violation in the B and D sectors, utilise the high collision rate at the LHC to measure rare processes, and perform precision tests of the Standard Model as well as a plethora of direct and indirect searches for new physics at the TeV scale and beyond. The detector is a single arm spectrometer which sits forward of the interaction point. A cutaway of the detector is shown in Figure 3.3. The interaction region is centred at 0 m along the z axis towards the left hand edge of the Vertex Locator with the various sub-systems positioned to the right. This type of detector geometry is akin to what is commonly used at fixed target experiments, but is motivated by the rapidity distribution of b and c quarks produced at the LHC. Figure 3.4 shows the distribution of b quarks produced at the LHC with LHCb's acceptance highlighted in red and clearly demonstrates the motivation behind the choice of detector geometry. The integrated production cross-sections for B mesons with transverse momentum $0 < p_T < 40$ GeV/c and rapidity $2.0 < y < 4.5$ at $\sqrt{s} = 7$ TeV in LHCb have been measured to be $\sigma(pp \to B^+X) = 38.9 \pm 0.3\,(\text{stat.}) \pm 2.5\,(\text{syst.}) \pm 1.3\,(\text{norm.})$ μb, $\sigma(pp \to B^0X) = 38.1 \pm 0.6\,(\text{stat.}) \pm 3.7\,(\text{syst.}) \pm 4.7\,(\text{norm.})$ μb and $\sigma(pp \to B^0_sX) = 10.5 \pm 0.2\,(\text{stat.}) \pm 0.8\,(\text{syst.}) \pm 1.0\,(\text{norm.})$ μb, where the third uncertainty is due to existing branching fraction measurements [90]. The charm cross section at LHCb is two orders of magnitude larger, with the integrated cross-section for the production of $c\bar{c}$ pairs with $p_T < 8$ GeV/c and $2.0 < y < 4.5$ at $\sqrt{s} = 7$ TeV in LHCb measured to be $\sigma(c\bar{c}) = 1.419 \pm 0.012\,(\text{stat.}) \pm 0.116\,(\text{syst.}) \pm 0.065\,(\text{frag.})$ mb, where the third uncertainty is due to the fragmentation functions used [91].



(a) (b)

Figure 3.4: (a) Polar angle distribution of b quark pairs produced at the LHC. It is clear that the majority of the quarks are produced in the forward and backward regions with the region of LHCb's acceptance highlighted in red. (b) Cumulative integrated luminosity delivered to (2012) and recorded by (2010-2012) LHCb.

LHCb takes advantage of such high production rates of B and D mesons to make very high precision measurements of processes across the heavy flavour sector. Similar physics can be done at the GPDs allowing for direct comparisons between multiple independent measurements all performed with proton collisions provided by the LHC. One example is measuring the rate of the very rare decay $B_s^0 \to \mu^+\mu^-$. The branching fraction of this decay has a very accurate SM prediction, but is heavily suppressed. The current theoretical estimate of the branching fraction is $\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.23 \pm 0.27) \times 10^{-9}$ [92]. The SM decay can proceed by a flavour changing neutral current (FCNC) process via virtual W exchange and subsequent $t\bar{t}$ annihilation to produce a Z^0 which decays to two muons, or by two virtual W mesons in a box diagram (see Figures 3.5 (a)). However, there are other diagrams that can be drawn involving the exchange of non-SM particles. For instance, in supersymmetric (SUSY) extensions to the Standard Model this decay can also proceed via a chargino and stop loop or via the exchange of two neutralinos in a box diagram which can alter the decay rate [93] (see Figure 3.5 (b)). Therefore, any deviation from the predicted SM rate would be an indication of new physics beyond the SM. Both LHCb and CMS have performed searches for this decay which have been combined to give the first measurement of the branching fraction of $B_s \to \mu^+\mu^-$ [94, 95].



(a) (b)

Figure 3.5: Feynman diagrams contributing to $B_s^0 \to \mu^+\mu^-$ decays. (a) depicts the SM processes that contribute to the decay amplitude and (b) depicts potential BSM processes allowed in supersymmetry. Image taken from [96].

Starting from the interaction point at z = 0 in Figure 3.3, particles that are moving in the forward (right) direction pass through a series of sub-detectors designed to measure various properties. The first of these is the Vertex Locator (VELO) which is designed to reconstruct charged tracks and vertices produced by particle interactions (Section 3.2.1). They then pass through the first of two RICH detectors and the upstream TT tracking station (Sections 3.2.2 and 3.2.4), through the magnet (Section 3.2.3), the downstream tracking, the second RICH detector, the calorimeters and the muon detectors (Sections 3.2.4, 3.2.2, 3.2.5 and 3.2.6). A description of each sub-detector is given in the following sections along with a summary of the software and data acquisition (DAQ) used to operate the detector and analyse the recorded data.

3.2.1 The VELO

The Vertex Locator (VELO) is the sub-detector closest to the primary interaction region. The detector is designed to measure the trajectory of charged particles as well as pinpoint their origin. Charged particles can emanate from a primary vertex (PV) i.e. a proton-proton interaction within the detector acceptance, or from the decay of any long lived particles that produce secondary decay vertices (DVs) which are distinct from any PVs in the event. Providing high precision measurements of these secondary vertices is crucial to any analysis of time-dependent decay rates, which is the case for a large fraction of the physics analyses performed at LHCb including the work presented in this thesis. A second parameter that is crucial to these and other analyses is the impact parameter (IP) of a track, defined as the perpendicular distance of closest approach of the track to a primary vertex. The impact parameter and a related variable, the impact parameter chi-square χ^2_{IP}, are used extensively to identify and remove background events.

The VELO detector itself is a silicon (Si) strip detector composed of 42 semicircular detector modules which are mounted on two halves each with 21 modules. A schematic of one side of the VELO is shown in Figure 3.6. During the FILL, RAMP and SQUEEZE modes of LHC operation the two halves of the VELO are kept 35 mm (radial) away from the beam axis (open) as the aperture of the beam is large and are then brought together to within 7 mm of the beam axis (closed) once stable beams are reached. If the modules were left in the closed position during injection they would degrade very quickly. The two halves of the detector are mounted inside a vessel which maintains a vacuum around the detection modules that is separate to the beam vacuum. Each side of the VELO vacuum vessel is enclosed in a thin aluminium *RF-foil* which is corrugated to allow for the alternating layers

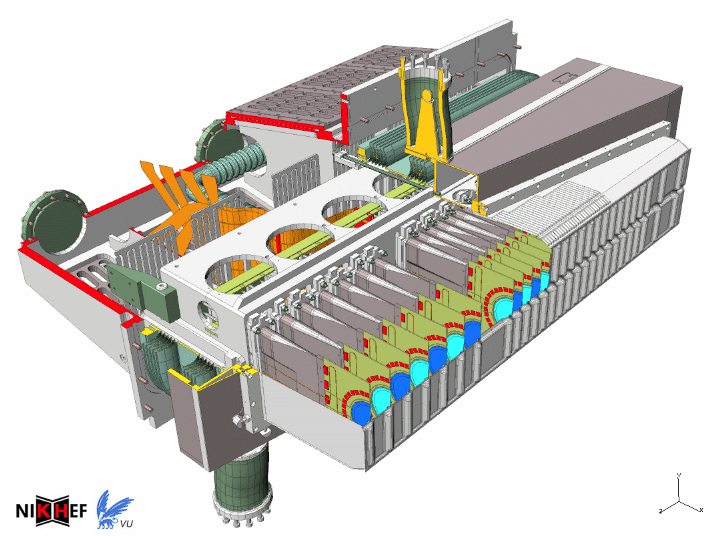
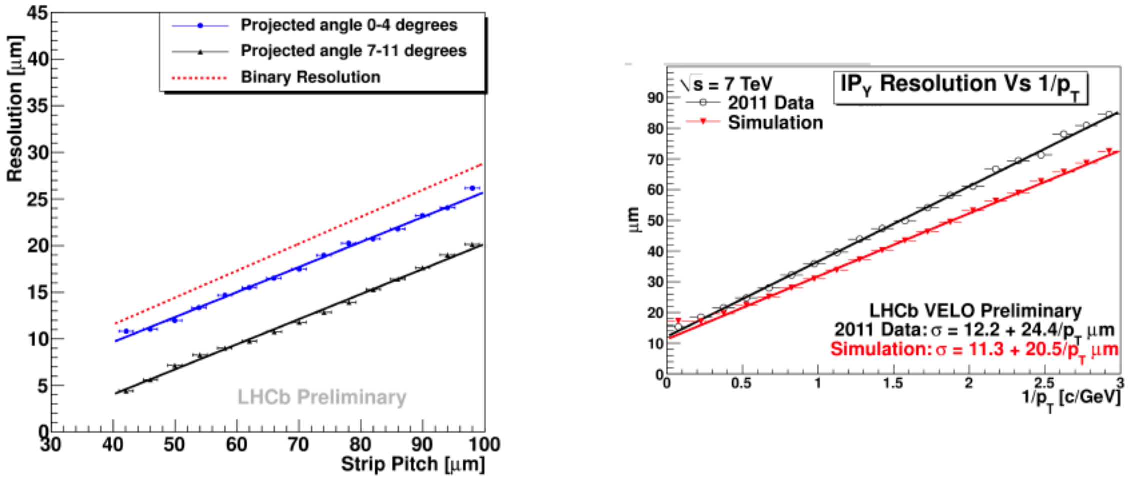


Figure 3.6: Schematic showing one side of the VELO detector in the closed position. Image provided by the LHCb VELO group.

of detection modules to overlap. This ensures that there are no gaps in the sensor coverage at the point where the two halves meet and is used in measuring the relative alignment of the two halves. The detector modules are more densely packed around the primary interaction point to improve the primary vertex resolution and reduce the track extrapolation distance to PVs and DVs. Each detector module is composed of two sensors: an r sensor that measures the radial distance to the beam line, and a ϕ sensor which measures the azimuthal angle in the x-y plane. Both types of sensor consist of 300 μm thick silicon that has been specially treated to work in the high radiation environment of the LHC with 2048 individual strips per sensor. The r sensors are oriented in concentric rings emanating from from the centre with the ϕ strips oriented perpendicular to the beam line. The pitch of the r and ϕ strips ranges from 38 μm in the regions closest to the beam line and increases approximately linearly up to 101.6 μm at the outer radius. Each sensor has 16 × 128 readout channels and is designed to give a single hit resolution of 4 μm. In addition to the 42 regular modules there are 4 pile-up veto modules consisting of just an r sensor situated upstream of the primary interaction point. These modules are only used to identify events with large numbers of primary interactions and are not used in the track reconstruction. In addition to track and vertex reconstruction the VELO also plays a role in the high level trigger at LHCb. It is used to select events that contain at least one high p_T track that has a large impact

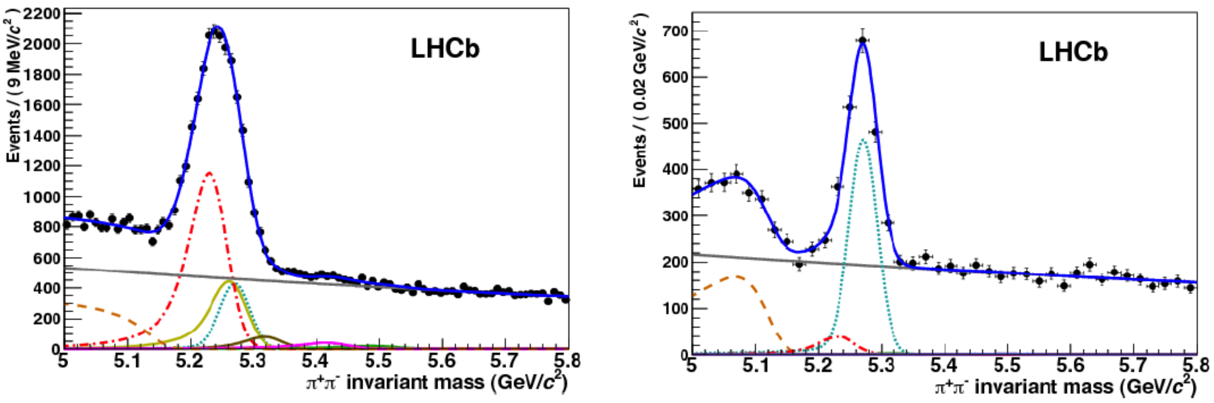


(a) (b)

Figure 3.7: (a) Single hit resolution of the VELO r sensors as a function of strip pitch at different projected angles. (b) Impact parameter resolution at optimal angle. Figures taken from [98].

parameter with respect to all primary vertices in the event, a signature that is common to events that contain heavy particles. This places strict requirements on the performance of the VELO. Specifically, the signal to noise ratio of the readout must be high, the channel readout efficiency must be at least 99 % and the point resolution must remain constant as a function of the radiation dosage [97].

The performance of individual modules varies according to the projected angle and the sensor pitch. During 2011 the VELO provided a point resolution of 4 μm in the 40 μm pitch region at an optimal angle of $\sim 10^{\circ}$ and a momentum-dependent impact parameter resolution of $(11.5 + 24.5/p_T(\mathrm{GeV}/c))$ μm (Figures 3.7 (a) and (b)) [98]. The track finding efficiency was measured to be > 98 % [99] using a tag and probe method from $J/\psi \to \mu^+\mu^-$ decays. Candidate pairs of muons are selected such that one is fully reconstructed (tag) and the other is only partially reconstructed using information from the other tracking stations (probe). The track finding efficiency is then measured in bins of track momentum by taking the ratio of the number of probes successfully reconstructed in the VELO to the total number of probes. Due to the VELO's excellent impact parameter resolution the decay time resolution for $B_s^0 \to J/\psi\,\phi$ was measured from data to be 45 fs. A similar value was measured for $D^0 \to K_{\mathrm{S}}^0\pi^+\pi^-$ using Monte Carlo which is described in Section 6.2.



(a) (b)

Figure 3.8: Invariant mass distribution for $B \to \pi^+\pi^-$ (cyan) decays with (a) kinematic and vertex cuts only, and (b) with PID information included [100]. $B^0 \to \pi^+\pi^-$ (dotted cyan), $B^0 \to$ 3-body (red dot-dash), $B_s^0 \to KK$ (yellow), $B_s^0 \to K\pi$ (brown), $\Lambda_b \to pK$ (purple) and $\Lambda_b p\pi(green)$. Figures taken from [100].

3.2.2 The RICH Systems

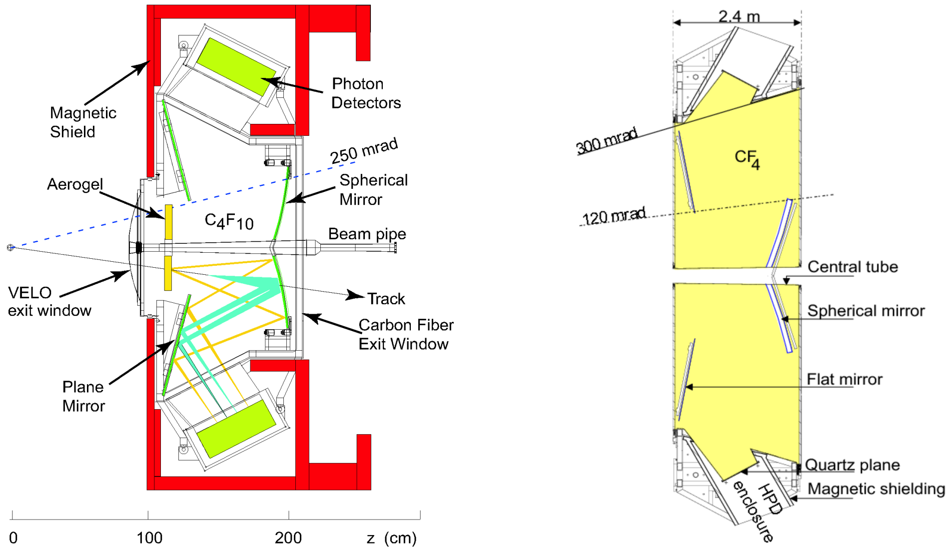
The decay products of heavy flavour particles produced in LHCb very often contain charged pions, kaons and protons. Reconstructing these decays requires these particle types to be identified and their properties measured. In the absence of particle identification it would be very difficult for the calorimetry and tracking alone to separate these particle types. Take, for example, pions and kaons. A charged pion will leave hits in the tracking detectors and energy deposits in both the electromagnetic and hadronic calorimeters. A charged kaon will behave in exactly the same way. Given the relatively small mass difference between a pion and a kaon ($\Delta(m_K - m_\pi) \approx 350$ MeV/ c^2) and the limited resolution of the calorimeters ($\sigma(E) \approx$ 250 MeV/c^2 for a track momentum of 100 GeV/c^1), it would be essentially impossible to distinguish between the two cases making measurements at LHCb susceptible to a high level of contamination from misidentified particle combinations. This effect can be clearly seen in $B^0 \to \pi^+\pi^-$ decays. The B^0 is reconstructed by combining pairs of charged particles that appear to originate from a common decay vertex that is significantly displaced from the primary vertex. However, $B^0 \to \pi^+\pi^-$ is not the only B^0 decay mode to exhibit this particular decay topology, nor the only B meson. Figure 3.8 (a) shows the invariant mass

---

[1] It is assumed that the resolution is dominated by the Hadronic calorimeter which is described in Section 3.2.5.

distribution of $B \to \pi^+\pi^-$ decays where the selection only utilises kinematic and vertex-related variables. It is clear that the contribution from other decay modes of the B^0 and decays of other types of B-hadron completely swamp that from the mode of interest (cyan peak). Figure 3.8 (b) shows the same invariant mass distribution, but with PID information incorporated into the selection. Many modes are completely removed, such as the Λ_b, $B_s \to K\pi$ and $B_s \to KK$ modes, and the remainder are heavily suppressed with respect to the signal. PID information is applied to the selection of all basic particle types in LHCb. The RICH also plays a crucial role in flavour identification at LHCb. When studying mixing or CP violation, information about the initial state of a heavy particle is often required. This is true of the analysis presented in this thesis and can be achieved by selecting D^0 that come from $D^{*+} \to D^0\pi^+$ decays (see Chapter 4). However, the situation is more difficult for B decays and the RICH plays a vital role in the flavour identification of B mesons using a technique known as flavour tagging [101]. B mesons are produced in the fragmentation of $b\bar{b}$ pairs which hadronise into a $B\bar{B}$ meson pair or a $B\bar{B}$ meson pair in addition to a b baryon anti-baryon pair. Information about the flavour of one of the pairs comes from reconstructing part of the cascade decay of its partner. This type of decay often produces pions and kaons with a mean momentum of 10 GeV. Therefore, to achieve a high tagging efficiency, the RICH system is required to separate between various particle types across a broad range of momenta.

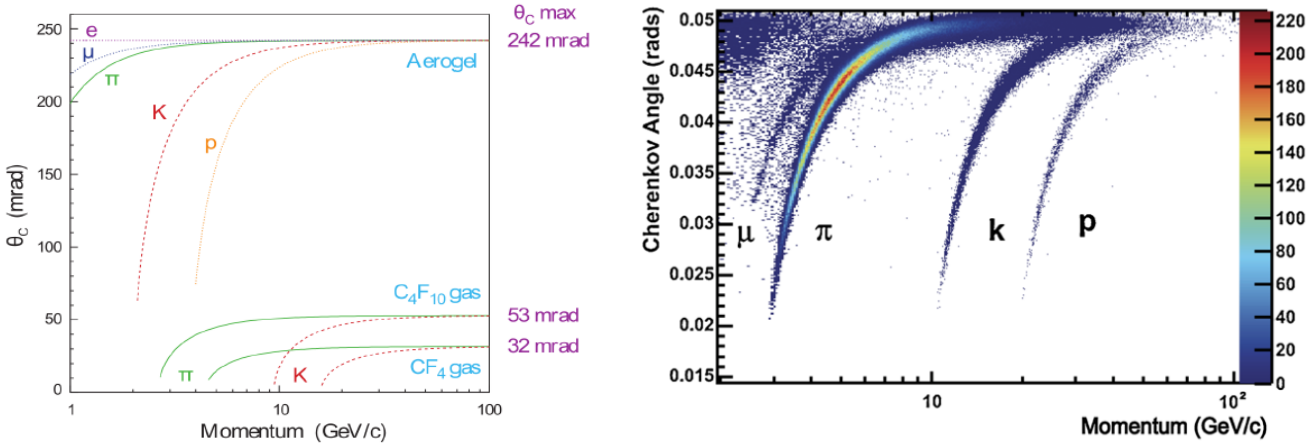
The LHCb RICH system consists of two Ring Imaging Cherenkov (RICH) detectors [88]. RICH1 (Figure 3.9 (a)) is placed before the magnet, sandwiched between the VELO and the TT tracking stations and covers the lower momentum region 2–40 GeV/c over the full LHCb acceptance. To ensure that all low momentum particles produced within the LHCb acceptance are captured by RICH1 it is sealed directly to the exit window of the VELO vacuum chamber. Placing it before the magnet prevents these low momentum tracks from being bent outside of this acceptance. RICH2 (Figure 3.9 (b)) is placed further downstream between the OT tracking stations and the first of the muon stations. It covers the higher momentum region of 15–100 GeV/c in the angular range of 15–120 mrad. As high momentum tracks are less affected by the bending magnet, RICH2 can be placed downstream of the OT,



(a) (b)

Figure 3.9: (a) Schematic of RICH1. (b) Schematic of RICH 2. Figures provided by the LHCb RICH group.

reducing the material budget for charged tracks and so improving the momentum resolution. Together they cover the entire geometric acceptance of LHCb across a momentum range of 2–100 GeV/c. The lower limit of 2 GeV is due to the fact that any particles produced below this momentum will not pass through the LHCb magnet. The RICH detectors measure the amount of Cherenkov light produced by a charged particle passing through a scintillating gas, known as the radiator, as well as the angle of the light with respect to the particle momentum direction. Cherenkov light is produced when a charged particle enters a medium at a velocity larger than the phase velocity of light in that medium. The Cherenkov light produced by an incident charged particle is reflected from a spherical mirror producing a ring of photons which are then reflected from a plane mirror onto a bank of Hybrid Photon Detectors (HPDs) developed especially for the LHCb RICH system. Each HPD is a tube with an active surface that is 75 mm in diameter. Incident photons release photoelectrons from the surface which are subsequently accelerated across a potential difference of $\sim$ 16 kV and onto an array of pixel detectors. Each HPD has 32 × 32 pixels with a pixel size of 2.5 × 2.5 mm with RICH 1 and RICH 2 having 196 and 288 HPDs, respectively, giving a total of 484. Fluorocarbon gases at room temperature and pressure are used as the radiators in both detectors; C_4F_{10} in RICH1 and CF_4 in RICH2. Roughly 5 % CO_2 is added to the



(a) (b)

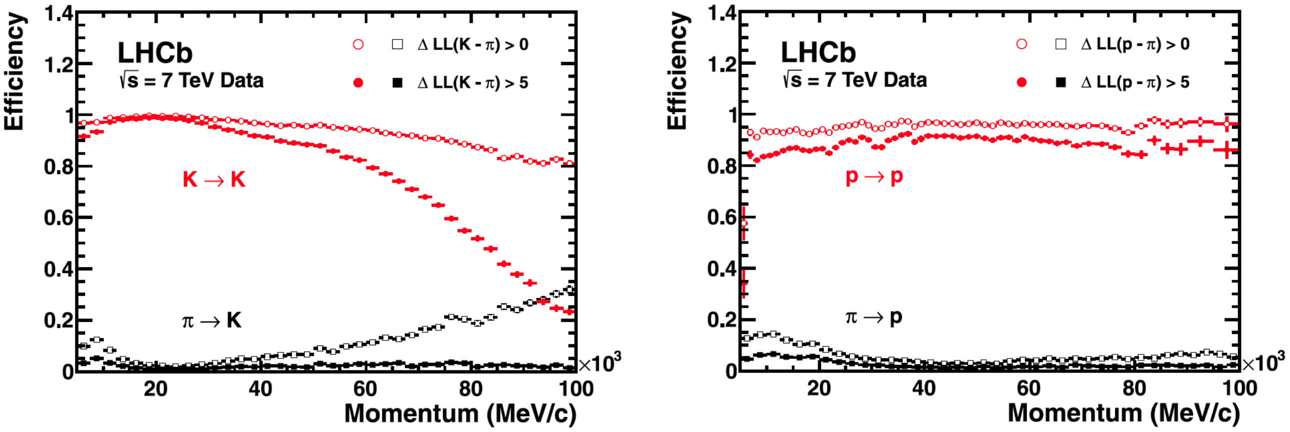
Figure 3.10: (a) Expected Cherenkov angle as a function of track momentum for various particle types in each of the 3 radiators. Figure provided by the LHCb RICH group. (b) The same plot for the C_4F_{10} radiator using isolated tracks. Figure taken from [100].

CF_4 in RICH2 to suppress the scintillation for high momentum particles. In addition to the C_4F_{10} RICH1 has a second radiator in the form of 16 silica aerogel tiles [102]. The momentum threshold for kaons to produce Cherenkov light in C_4F_{10} is 9.3 GeV/c. Kaons with momentum less than this threshold can be separated from pions by vetoing charged tracks that do not produce any light, but separation between protons and kaons at low momentum can only be done using the aerogel which has a much lower emission threshold. Figure 3.10 (a) shows the expected Cherenkov angle as a function of track momentum for various particle types in each of the 3 radiators. Figure 3.10 (b) shows the same plot for the C_4F_{10} radiator using isolated tracks (tracks that produce rings which do not overlap with any other from the same radiator) from real data recorded during 2011 [100] demonstrating the separation power of the RICH.

Information from the RICH and tracking systems is used as input to a log-likelihood algorithm which provides a set of variables that can be used to distinguish between various particle types. As the vast majority of particles produced at the LHC are pions, the overall event likelihood is initially calculated assuming the pion mass hypothesis for all particles. Following this, the likelihood for each track is recomputed using the mass hypothesis for each of e, μ, π, K and p individually with the likelihood for all other tracks fixed to the initial estimate. The particle is then identified by the mass hypothesis that gives the largest increase in the overall likelihood and the process repeated until the total event log-likelihood converges. The contribution to the likelihood from background photons that do not emanate from a charged particle is estimated before each likelihood calculation by comparing the expected photon distribution due to the track parameters and current best mass hypothesis with the photon hits observed and adding a penalty term to the likelihood. Suitable convergence is usually found after just 2 iterations. The change in the total log-likelihood, ΔLL with respect to the pion hypothesis is saved for each particle type and can be used to separate them in data. A quantitative measure of the performance of these PID variables is crucial for analyses that use PID and wish to measure the absolute rate of any process. The performance is determined by using a series of control channels that have high production rates at LHCb: $K_{\mathrm{S}}^{0} \rightarrow \pi^{+}\pi^{-}$, $\Lambda \rightarrow p\pi^{-}$ and $D^{\star +} \rightarrow D^{0}\pi^{+}$ where $D^{0} \rightarrow K^{-}\pi^{+}$. These samples contain copious $K^{\pm}$, $\pi^{\pm}$ and p particles and a very high purity can be achieved using a kinematic selection that is independent of any PID variables. These samples are then used to estimate the PID efficiency for each particle type. For example, the fraction of kaons passing the cut $\Delta\mathrm{LL}(K-\pi) > 0$ will give an estimate of the kaon ID efficiency, and the fraction of pions with $\Delta\mathrm{LL}(K-\pi) > 0$ gives the pion mis-ID probability for the same cut. These are plotted as a function of track momentum for this cut and for a much tighter requirement of $\Delta\mathrm{LL}(K-\pi) > 5$ in Figure 3.11(a) and clearly demonstrate the discriminating power of this variable. The average efficiency for kaon identification across the entire momentum range for $\Delta\mathrm{LL}(K-\pi) > 0$ is 95 % for 10 % pion mis-ID. However, the mis-ID rate can be reduced to just 3 % for a kaon ID efficiency of 85 % when the tighter requirement is applied. A similar plot for the proton ID efficiency is shown in Figure 3.11(b).

3.2.3 The LHCb Magnet

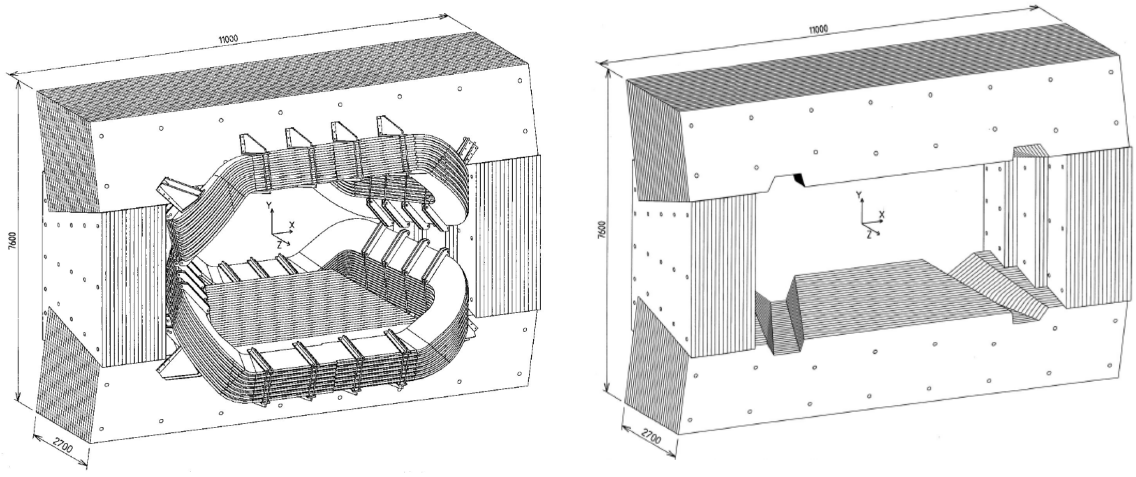
The LHCb dipole magnet is a 1,600 ton warm electromagnet that bends charged particles as they fly through the detector. This enables the momentum of charged tracks to be determined by measuring the curvature of their trajectory. The LHCb magnet consists of



(a) (b)

Figure 3.11: (a) Kaon ID and pion mis-ID efficiency for $\Delta\mathrm{LL}(K-\pi) > 0$ and $\Delta\mathrm{LL}(K-\pi) > 5$ as a function of track momentum from the 2011 data control samples. (b) Proton ID and pion mis-ID efficiency for $\Delta\mathrm{LL}(p-\pi) > 0$ and $\Delta\mathrm{LL}(p-\pi) > 5$ as a function of track momentum from the 2011 data control samples. Figures taken from [100].

two trapezoidal coils with each end bent at a 45° angle in the transverse plane (Figure 3.12). Each coil is constructed from 15 Aluminium “pancake” mono-layers, essentially concentric rings of conductor wound into the trapezoidal shape with each layer containing around 290 m of conductor. The gap between the coils is wedge shaped in both the horizontal and vertical directions and follows the total detector acceptance. The original design of LHCb was to incorporate a race-track like superconducting magnet [103]. However, due to the complicated geometry required by LHCb a warm magnet design was finally settled upon. The warm magnet design has other advantages as it can be ramped up quickly (at the same rate as the LHC magnets), cost significantly less to build and was faster to construct. It is also much easier to invert the magnetic field. This last feature is extremely useful in studying any charge detection asymmetry due to imperfections in the detector. Such asymmetries can fake CP-violation if not properly understood and corrected for. This is because the magnetic field bends tracks with different electric charges in opposite directions. At low momentum this results in all positively charged tracks seeing one half of the detector and negatively charged tracks the other. Therefore, the measured rates of charge conjugate modes can differ significantly due to differences in the detector response of the two halves. To reduce this effect, LHCb regularly flips its magnet polarity throughout the data taking period such that the total dataset is split roughly 50:50 across each magnet polarity. The



(a) (b)

Figure 3.12: (a) Schematic of the LHCb magnet coils and yoke. (b) With coils removed. Figure taken from [103].

magnet provides a bending power of approximately 4 Tm. However, to measure particle momenta accurately the magnetic field variation throughout the detector must be known precisely. During 2011 a survey of the magnet field strength was performed in situ using a series of Hall probes. The magnetic field strength was determined to an accuracy of $\sim$ 0.2 mT with a magnetic field strength of $\sim$ 0.8 T in the y plane (the bending plane). The results of this survey were used to calibrate the tracking systems to achieve the excellent momentum resolution discussed in the following section.

3.2.4 The Tracking System

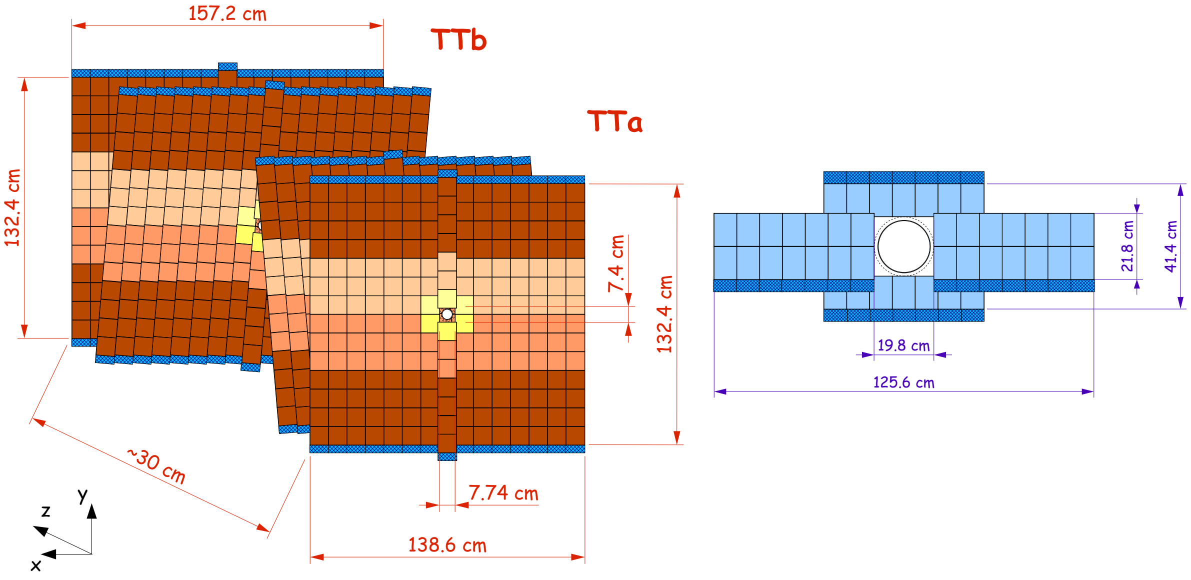
The LHCb tracking system [88] is used to follow the trajectory of charged particles through the detector and measure their momenta. It consists of the Tracker Turicensis[2] (TT) which is located upstream of the magnet and the downstream tracking stations T1–T3. Each of the three T-stations consists of an Inner Tracker (IT) which covers the region closest to the beam pipe and an Outer Tracker (OT) which covers the remaining detector acceptance. The tracking system (including the VELO) provides an overall momentum resolution varying from $\sigma_p/p = 0.4$ % @ 5 GeV/c to 0.6 % @ 100 GeV/c.

---

[2] Also known as the Trigger Tracker.

Silicon Tracker

The Tracker Turicensis (TT), and the Inner Tracker (IT) make up the Silicon Tracker (ST) system [89]. Both detectors use silicon microstrip sensors with a strip pitch of approximately 200 μm. Each tracking station has four detector layers with vertical strips in the first and last layer and strips rotated by a stereo angle of -5° and $+5^{\circ}$ in the second and third layer in a (x-u-v-x) arrangement. The ST system was designed to give an overall single-hit resolution of about 50 μm such that the momentum resolution had no significant contribution from multiple scattering across almost the entire range of particle momenta. It was also built to withstand the expected occupancy and exposure to radiation, maintain a high single-hit efficiency and conform to the material budget of LHCb.



(a) (b)

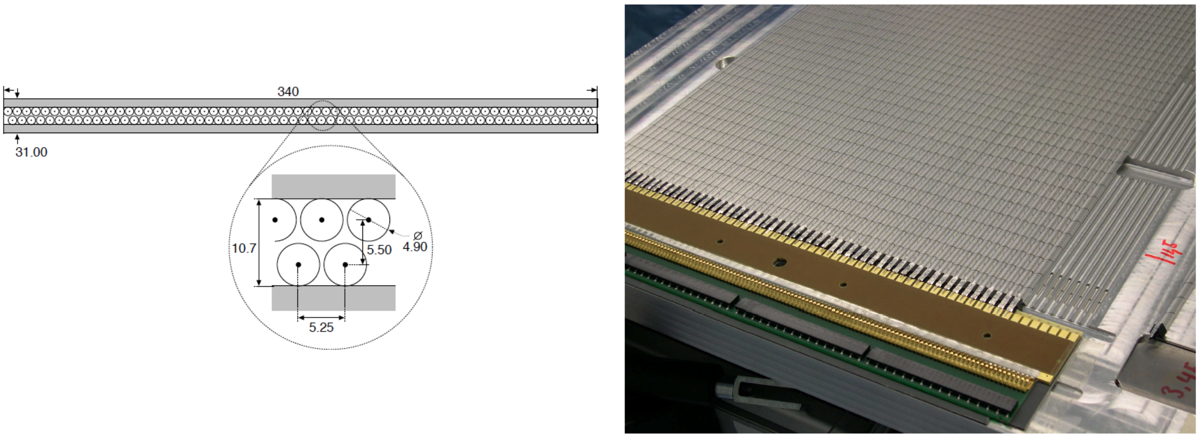
Figure 3.13: (a) Schematic of the TT detector. (b) Schematic of an IT detector layer. Figures taken from [89].

Tracker Turicencis The TT is a planar tracking system 150 cm wide and 130 cm high that covers the entire LHCb acceptance. It has an active area of 8.4 m^2 with 143,360 readout strips. Each of the four detector layers is made up of two half modules containing 7 rows of sensors as in Figure 3.13 (a). The detector layers are housed in a large, light-proof, thermally and electrically insulated container. A steady temperature of 5°C is maintained within the detector volume which is continuously flushed with Nitrogen to prevent the build up of condensation on the surface of the detection layers. Individual modules within a detector layer are staggered by about 1 cm along the beam direction and overlap by a few millimetres in the x-direction to ensure total detector coverage and for the purposes of alignment. Hybrid front-end electronics are built on the end of each row of sensors allowing for the read-out, cooling and support structure to be built on the upper and lower sides of the detector outside of the LHCb acceptance. The TT silicon sensors are 500 μm thick, and carry 512 readout channels with a strip pitch of 183 μm.

Inner Tracker Each of the three Inner Tracker (IT) stations consists of four individual boxes arranged around the beam pipe as shown in Figure 3.13 (b). Each box is light-proof, thermally and electrically insulated, and maintains the same atmospheric conditions as for the TT. Each box contains four detector layers with each detector layer consisting of seven individual detector modules. The modules are staggered in a similar way as for the TT, but they are separated by just 4 mm in the beam direction and 3 mm in the x direction. This finer resolution in the beam direction is necessary to measure the curvature of lower momentum tracks that are significantly bent by the magnet. The modules in the upper and lower boxes are built of a single silicon sensor and readout, but modules in the left and right boxes (the magnet bending plane) contain 2 sensors with a single readout. The silicon sensors in the single-sensor boxes are 320 μm thick with a strip pitch of 198 μm. Those in the double-sensor modules have the same strip pitch, but are 410 μm thick. The relative thicknesses were chosen to ensure the signal to noise ratio met the detector efficiency requirements while minimising the material budget.

Outer Tracker

The Outer Tracker (OT) [88] is a gaseous straw-tube detector that is capable of measuring the momentum of charged particles over a large acceptance area. It consists of 53,760 straw



(a) (b)

Figure 3.14: (a) Schematic of straws in each OT detector layer. (b) Photo of one layer of straws encased in an aluminium template. Images taken from [89].

tubes with an active area of approximately 5 × 6 m^2 and 12 detection layers each with two layers of straw tubes. Each tube is 2.4 m long and 4.9 mm in diameter, and is filled with an Argon (Ar), Carbon Dioxide (CO_2) and Oxygen (O_2) gas mixture. A 25 μm thick gold-plated tungsten wire anode runs down the centre of each tube with an inner cathode made from electrically conducting carbon foil doped with a polyimide film and an outer cathode foil of polyimide film laminated with a layer of aluminium. The straws in each layer are glued together using a precision aluminium template as shown in Figures 3.14 (a) and (b). Each half of an OT station is built up of 11 modules, 7 long and 4 short. The long modules sit either side of the region closest to the beam line which is covered by the IT. They have readout electronics at both ends of the module, contain 256 straws and have an active length of 485 cm. The shorter modules are about half the length and sit above and below the region closest to the beam pipe covered by the IT. They contain 128 straws and a single readout at the end furthest from the beam line. Each module consists of four layers which are oriented in the same way as for the IT. The size of the inner region occupied by the IT is determined by the requirement that the detector occupancy of the straw-tubes must not exceed 10 % at the LHCb design luminosity.

3.2.5 The Calorimeters

The calorimetry at LHCb plays a crucial role in electron identification, the measurement of photon energies and in the trigger. Due to the high volume of energetic charged particles produced in proton collisions at the LHC many high energy photons are produced in an event. Therefore, an electromagnetic calorimeter is a necessity for analyses that look to reconstruct radiative decays or final states that contain π^0. The LHCb calorimeter system consists of a Scintillator Pad Detector (SPD), a Pre-shower Detector (PRS), an Electromagnetic Calorimeter (ECAL) and an Hadronic Calorimeter (HCAL), and is sandwiched between the muon stations M1 and M2 (see Figure 3.3). It is used to distinguish between electrons, photons and hadrons by tracing their path through the calorimeter system. Information from all four of the calorimeter sub-systems is used in the identification process. Figure 3.15 shows the characteristic energy deposits left by photons, electrons and hadrons in the calorimeters. Particles initially enter the SPD which can detect electrons and charged hadrons. They then pass through a layer of Lead (Pb) at which point electrons and photons typically produce dense electromagnetic showers which are subsequently detected by the PRS and the energy absorbed by the ECAL. Incident hadrons will also sometimes begin to shower in the Lead, but will usually deposit the majority of their energy in the HCAL. It is clear that by combining information from each sub-system a distinctive energy signature is left in the calorimeters for these three types of particle. Each detector works by collecting scintillation light and transporting it to a Photo-Multiplier Tube (PMT) along fibre-optic cables. As the amount of scintillation light collected is proportional to the energy of the shower produced by an incident particle, the calorimeters can also be used to measure the energy of the particle types it identifies. The calorimeters are also used in the first level trigger to select highly inelastic events by searching for large transverse energy deposits in the ECAL and HCAL. A more detailed discussion of the calorimeters' role in the trigger decision making is given in Section 3.2.7.

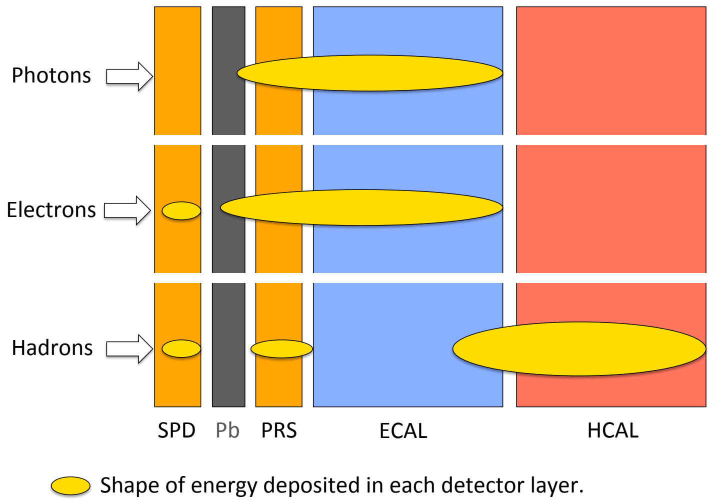


Figure 3.15: Characteristic energy deposits for photons, electrons and hadrons traversing the LHCb calorimeter system.

Scintillator Pad (SPD) and Pre-shower (PRS) Detectors

The SPD and PRS detectors are both 15 mm thick scintillator pads that are separated by a 12 mm thick sheet of Lead corresponding to 2.5 radiation lengths (X_0). Both detectors have 6016 readout channels that are split into three segments increasing in cross-sectional area the further away from the beam line they are positioned (Figure 3.16). The same readout geometry is used in the SPD, PRS and ECAL detectors giving a 1 to 1 spatial correspondence with the readout for the first three subsystems. Each scintillating pad contains a helicoidal fibre optic cable which collects the scintillation light. This light is then sent from both ends of each fibre to a bank of Multi-Anode Photomultiplier Tubes (MaPMT) that are located on the upper and lower sides of the detector. Each fibre is read out on an MaPMT that contains 64 readout channels from 2×2 mm^2 pixel photo-anodes and collects on average 25 photoelectrons in response to a minimum ionising particle.

The Electromagnetic Calorimeter (ECAL)

The electromagnetic calorimeter incorporates a shashlik style calorimeter that consists of alternating layers of absorber and scintillating material. The total area covered by the ECAL

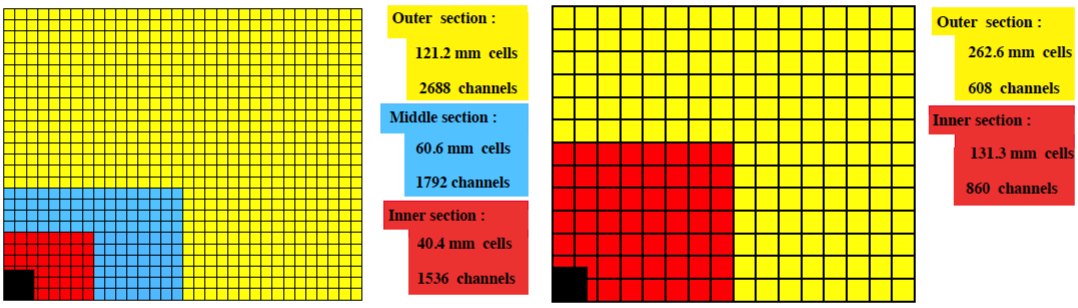


Figure 3.16: Schematic of the calorimetry readout geometry for the SPD/PRS and ECAL (left) and HCAL (right). The cell dimensions listed in the left hand image are for the ECAL [89].

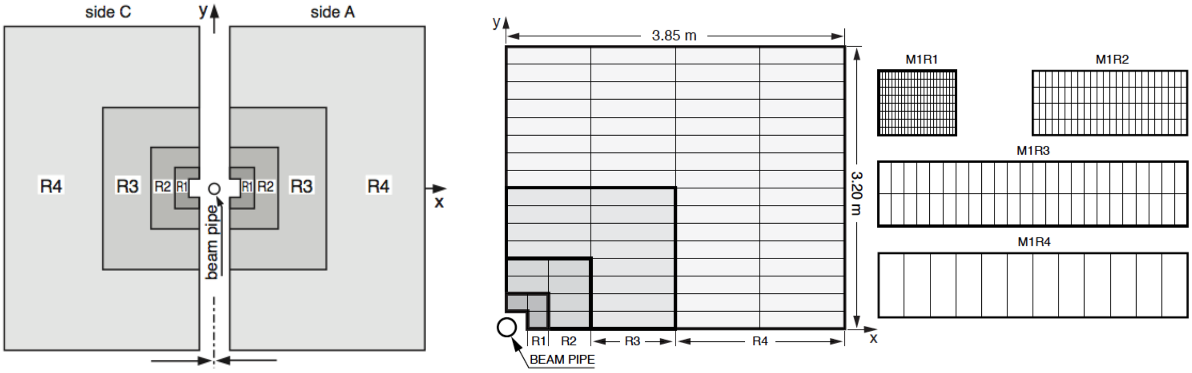
is the same as the projection of the tracking system and consists of three regions containing banks of detector modules each with different readout cell sizes. The geometry of the ECAL readout is shown in Figure 3.16. Each module is constructed from 66 alternating layers of 2 mm thick lead, 120 μm reflecting paper and 4 mm thick scintillator tiles corresponding to 25 X_0. The cells are punctured with precisely positioned holes that house fibre optic cables which transport the scintillation light to a bank of photomultiplier tubes attached to the back of the module. The light from each cell is captured by an individual PMT with the density of cells and fibres being larger in the inner regions of the detector to improve the resolution where the track density is high. The design energy resolution of $\sigma_E/E = 10\%/\sqrt{E} \oplus 1\%$ (E in GeV) results in a B mass resolution of 65 MeV/c^2 for the $B \rightarrow K^*\gamma$ decay with a high E_T photon and 75 MeV/c^2 for $B \rightarrow \rho\pi^0$ with a π^0 mass resolution of $\sim$ 8 MeV/c^2 [89]. The energy resolution for the modules in each region was found to vary from around 3 % down to 1 % as the particle momentum increases from 15 to 100 GeV/c [89].

The Hadronic Calorimeter (HCAL)

The LHCb hadronic calorimeter is a sampling calorimeter with cells made from alternating layers of 2 cm thick iron and polystyrene scintillating tiles. The HCAL is designed to stop essentially all the remaining hadrons. The scintillating tiles are oriented parallel to the beam axis with the iron absorber positioned perpendicular to the beam direction. The light from each cell is read out individually to a PMT and the density of cells is greater in the inner part of the detector. The readout geometry for the HCAL is shown in Figure 3.16. The material in the HCAL corresponds to 5.6 λ_I where λ_I is the hadronic interaction length. The energy resolution is ~ 12 % for a track momentum of 100 GeV/c.

3.2.6 The Muon Detectors

The muon detector system plays a crucial role in both the trigger and in muon identification at LHCb. It is designed to detect charged particles that penetrate through the other sub-detectors and aid in the measurement of their momenta. The detector system itself consists of five stations that measure the x, y position of track hits perpendicular to the direction of the beam. The first of these stations (M1 in Figure 3.3) is sandwiched in-between RICH 2 and the calorimeters and is only used to improve the momentum resolution for muon candidates recorded by the trigger. The remaining stations (M2–M5) are situated downstream of the rest of the detector and are used in the trigger and for offline muon identification. Each station consists of 276 multi-wire proportional chambers (MWPCs), but with the central part of the first station equipped with 12 Gas Electron Multiplier (GEM) detectors accounting for about 1% of the total muon detector coverage. The downstream muon stations are interspersed with 80 cm thick iron absorbers for removing hadronic particles that survive the calorimeters. Each station is split into two halves, the A and C sides, which are in turn split into four panelled regions, R1 to R4, which increase in area the further away from the beam line they are positioned (see Figure 3.17 (a)). Each of these four regions is then built up of rectangular chambers which are in turn split into logical pads with the chambers shaped such that the spatial resolution is finer in the vertical plane compared to the horizontal to account for the bending of the LHCb magnet (see Figure 3.17 (b)). The granularity of the pads in each region decreases the further they are from the beam line and are sized according to the requirements of the trigger and the offline muon identification. The granularity in the y direction scales with the area of the chambers and is the same for all stations. The granularity in the x direction is finer in stations M2 and M3 with double the number of pads and coarser in M4, M5 and in the outer regions of M1 with half the number of pads.



(a) (b)

Figure 3.17: (a) Head-on view of one of the muon stations with regions R1-4 labelled. (b) Left: a quadrant of the M1 station with the four regions R1 - R4 highlighted. The rectangles represent chambers of logical pads. Right: the relative size and shape of the logical pads within the chambers for each region. Images taken from [104].

This allows for measurement of the muon momentum in stations M2 and M3 before it has passed through much of the iron absorbers achieving an overall momentum resolution of $\sim$ 20 %. Due to multiple scattering, the momentum resolution is not as good in stations M4 and M5 which are primarily used to identify tracks that survive. Each logical pad has its own readout channel in the front-end electronics with 55,296 pads in total.

Multi-wire Proportional Chambers (MWPC)

MWPCs generally consist of a grid of wires surrounded by a gas mixture and encased in a chamber which is earthed. Incoming charged particles ionise the gas molecules in the chamber setting off a cascade of drift electrons which are collected in the wire. Each layer of LHCb's MWPC detectors are built of two gas chambers stacked together. Each chamber contains a gas mixture of 40 % Ar, 55 % CO_2 and 5 % CF_4 contained within two cathode plates which are separated by a 5 mm gap. A potential difference of 2.5–2.8 kV is applied between the plates. Within the gas volume are a series of 30 μm diameter gold-plated Tungsten wires that are spaced 2 mm apart. M2–M5 have two of the double chambers stacked together and M1 has just one. The LHCb muon system consists of 1,386 MWPCs in total.

Gas Electron Multiplier (GEM) Detectors

GEM detectors function in a similar way to MWPCs, but there are some key differences. GEMs generally consist of thin foils surrounded by a gas mixture and sandwiched between two copper plates. The foils are perforated with very fine holes that are precisely spaced at regular intervals. A potential difference is applied across the plates creating large electric fields in the holes. Incident muons ionise the gas molecules and create an avalanche of electrons and ions which are drawn towards the holes and projected onto the two copper plates. The detector is then read out by placing conducting readout strips along each copper plate. LHCb's GEM detectors are built of 3 foil layers sandwiched within a gas volume that consists of 45 % Ar, 15 % CO_2 and 40 % CF_4. The increased amount of CF_4 gives a better detector time resolution compared to the MWPCs. The potential applied across the three foils ranges from 3.5 kV/cm up to 5 kV/cm. The total GEM detector module consists of 12 layers each with two triple-layer GEMs that are stacked together with a total active area of 20×2.4 cm^2. The GEM detectors are placed in the centre of the M1 station, the region that contains the highest flux of charged particles. GEMs were chosen in place of MWPCs as they are more radiation hard, function better with high occupancy, and achieve a superior spatial resolution due to the high gains achieved by stacking them together.

Muon Identification

The offline muon identification combines information from the muon stations and the tracking stations. The offline identification can be done at two levels. The first is a simple check to see whether a track passes through the calorimeters, iron absorbers and a number of the downstream muon stations. This is known as the isMuon selection and involves counting hits in the muon stations within a region of interest about the direction of the track extrapolation. The size of the region of interest and the number of muon stations the track is required to traverse to satisfy the isMuon condition both depend on the track momentum and are computed separately for each muon station. The second method is a likelihood computation for the muon and non-muon hypotheses based on the track hits left in the muon stations and information from the trackers. This is then used in a global likelihood calculation for all the different particle hypotheses that incorporates information from the RICH and calorimeter systems. The muon likelihood is taken as a cumulative probability distribution of the average squared significance of the distance between the track hits seen in the muon chambers and those that would be expected given the extrapolation of the track from the upstream tracking detectors. The difference between the likelihood calculation for the muon hypothesis and the not-muon hypothesis ΔLL_{μ} can then be used as a discriminating variable for identifying muons.

3.2.7 The Trigger

The LHCb trigger [89, 105] is necessary to reduce the rate of visible *pp* collisions at the LHC down to a manageable level that can be sent to storage. The trigger is split into two levels. The first level trigger (L0) is implemented in hardware and uses information from the tracking, calorimetry and the muon detectors to reduce the event rate from 40 MHz (the LHC bunch crossing rate) down to 1 MHz at which the rest of the detector can be read out. The High Level Trigger (HLT) is implemented in software and uses information from all of the sub-detectors to reduce the readout rate of 1MHz to a few kHz which can then be saved to disk for analysis. The trigger system is designed to select signal events with the highest possible efficiency ensuring that the majority of the data recorded is useful to analysts. The fact that the bulk of the trigger is implemented in software is an excellent design feature. It means the trigger can be easily updated to take advantage of improvements throughout the data taking period and allows analysts to rerun the trigger offline in exactly the same configuration as used for recording the data. The latter feature is exploited in this analysis and others [106] to determine the trigger efficiency as a function of D^0 decay time using a data driven technique which is described in Section 4.6. The trigger is run using the MOORE application and the entire configuration is determined by just 4 parameters:

1. The version of MOORE used to take the data

2. The Trigger Configuration Key (TCK)

3. The Detector Database (DDDB) tag

4. The LHCb Conditions Database (LHCBCOND) tag

where (1) and (2) fully describe the decision making logic and cut values used in the trigger and (3) and (4) provide a database of parameters which describe the geometry and conditions of the detector at the time of data taking, respectively. A schematic of the trigger decision making process for 2012 data taking is given in Figure 3.18. A more in depth description of the LHCb trigger is given in the following sections, however, for further details on the trigger and its performance during 2011 please refer to [105].

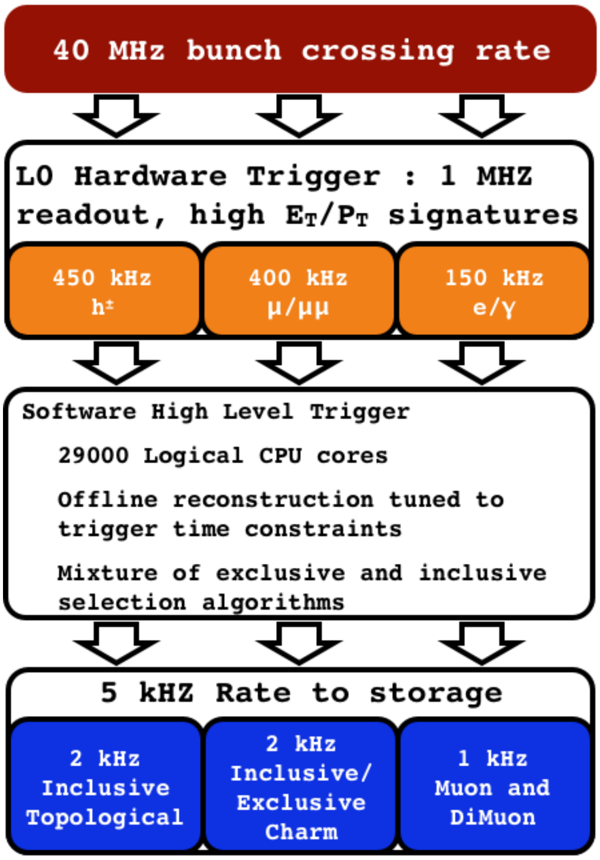


Figure 3.18: The LHCb trigger during the 2012 data taking period. Note: the trigger setup was slightly different in 2011 with the output rate to disk at 2–3 kHz and no inclusive charm lines. Figure provided by the LHCb Trigger group.

First Level Trigger (L0)

Decays of heavy particles such as B and D mesons often contain decay products that have very high transverse momentum, p_T, or transverse energy, E_T. The L0 trigger reduces the bunch crossing rate of the LHC down to the 1 MHz readout rate of LHCb by selecting events that have these characteristics. L0 itself is divided up into several independent triggers using information from the calorimeters and muon systems. The calorimeter triggers are designed to accept high E_T hadron, electron and photon clusters utilising information from the ECAL and HCAL about the energy deposited and information from the SPD and PS about the number of tracks in the event. The muon trigger searches for the 2 highest p_T muons in a given quadrant of the muon stations. This information is passed to the L0 decision unit which takes the logical OR of each trigger decision as well as applying any pre-scaling. Pre-scaling involves randomly rejecting events to reduce the acceptance rate of the trigger. The whole decision making process takes around 2.5 μs.

High Level Trigger (HLT)

The High Level Trigger (HLT) is entirely software and is run on a batch of 29,000 logical CPU cores. Events that pass the L0 trigger are sent to the first software level, HLT1, which reduces the 1 MHz rate down to about 40 kHz. Events that then pass HLT1 are passed to the final trigger level, HLT2, which performs a fast version of the full event reconstruction using information from all the sub-detectors to select decays of interest.

HLT1 The HLT1 trigger performs a partial event reconstruction based on information from the VELO and tracking stations. The VELO track reconstruction is fast enough to be run on a sub-set of the track objects that leave a large number of hits in the VELO modules. Although this step misses a fraction of the track hits, the reconstruction quality is sufficient to identify tracks and vertices to within a few μm with respect to the offline version that uses all track hits. Several different lines are run in parallel and the final HLT1 decision is the logical OR of their outputs. The main HLT1 line used in this analysis, `Hlt1TrackAllL0`, searches for tracks with high p_{T}, large impact parameter with respect to all PVs in the event, and good track fit quality as determined by the fit χ^2 and by comparing the number of observed hits with the number expected for a given track. These tracks are also required to leave hits in at least 5 VELO modules. Events that pass the muon L0 triggers are still required to contain a track with a small IP and large p_T, but are only required to leave hits in 4 VELO modules. However, the tracks reconstructed as muons are extrapolated and combined with hits left in the muon station, M3. Assuming a fixed muon momentum a search window in the $x-y$ plane is defined and the event is accepted by the muon HLT1 trigger if sufficient hits are left in the remaining muon stations within this window. Events that are accepted by HLT1 are then sent to HLT2 for a more complete event reconstruction and selection.

HLT2 HLT2 runs a fast version of the full offline track reconstruction. The track finding efficiency is 1–2 % lower per-track than the full offline reconstruction. The full muon reconstruction is also run and tracks are associated with clusters in the ECAL and HCAL to identify electrons, photons and hadrons. The RICH information is also used in a subset of the HLT2 decisions. Once the event reconstruction has been applied a search for particular decays of interest is performed by running a series of roughly 200 selection algorithms (trigger lines) and taking the logical OR of their decisions. The trigger algorithms can be split into two categories: exclusive and inclusive. Exclusive triggers search for composite particles and particular decays by attempting to reconstruct the entire decay chain and applying a selection to the final particle objects. This type of trigger is predominantly used for selecting charm decays. Inclusive triggers attempt to partially reconstruct the decay, reducing the computation time, and accept events that pass more generalised selection criteria than for exclusive lines. The LHCb trigger uses an inclusive trigger for selection of B decays known as the topological trigger. The trigger searches for combinations of 2, 3 and 4 good quality tracks that fall within a given window in mass. As missing particles push the reconstructed mass lower than the true B mass, the selection uses a corrected mass which accounts for this based on the missing momentum transverse to the flight direction of the B candidate.

Additional selection criteria are applied using a set of variables which are fed to a boosted decision tree (BDT). The use of a BDT improves the background rejection of the topological trigger while keeping the signal efficiency near optimal.

Trigger Decisions

The triggering decisions at LHCb can be considered in three classes. Events can belong to more than one class. These are the overall trigger decision (DEC), whether the event was Triggered On the Signal candidate (TOS) and whether the event was Triggered Independently of the Signal candidate (TIS). The overall decision is simply whether a candidate in an event fired the trigger or not. A trigger decision is TOS if the trigger line was fired using only tracks or other objects belonging to the signal candidate (i.e. the signal candidate alone was sufficient to fire the trigger). A decision is TIS if something else in the event fired the trigger (i.e. the signal candidate was not necessary for the trigger). This could be any of the other tracks in the event not including the tracks or other objects left by the signal candidate or particles produced in its decay. Candidates can be both TIS and TOS. They can also be neither TIS nor TOS, but such events are usually fired by some mis-reconstructed candidate and should not be used in a physics analysis. The idea behind employing the TIS/TOS system is to enable trigger efficiencies to be estimated from data. Taking the ratio of the number of candidates which are TOS to the sum of those that are TIS and TOS in bins of a particular variable gives an estimate of the trigger efficiency as a function of that variable. However, this is not the case for all variables. For example, the trigger efficiency as a function of decay time cannot be extracted this way and a more novel technique for extracting the trigger efficiency from data has to be employed.

3.2.8 Stripping and Offline Data Processing

Data that has passed HLT2 is then permanently saved to disk and a copy made for further processing and offline event reconstruction. The event reconstruction is performed by the Brunel package. It also runs the full pattern recognition, track and primary vertex reconstruction as well as reconstructing the Cherenkov rings in the RICH. Combining information from the calorimeters, tracking system, RICH and muon detectors it re-computes the particle likelihood differences ΔLL $(\pi - X)$ where $X = K, \mu, e, p$ (described in Sections 3.2.2 and 3.2.6) for all tracks in the event. The time allowance for the Brunel step is greater than for the HLT and better quality fit routines are used in the reconstruction. A further selection known as the "stripping" is then applied using the DaVinci software package. The stripping searches for decays of interest, according to criteria custom-written for each analysis, typically by combining tracks of a given particle hypothesis into composite particle candidates. The selection process works in a similar way to the HLT, but since the full offline information is available, the addition of the RICH and muon information to the particle likelihoods and improved reconstruction increases the background rejection power of the selections. The output of the stripping is divided into streams which are groups of individual stripping lines. Generally, each stripping line is designed for a specific analysis and selects just the mode of interest although more generic selections with more than one mode are also possible. Stripping decisions can also be categorised using the TIS/TOS method discussed in Section 3.2.7. Each stripping line saves its final particle candidates in a unique Transient Event Store (TES) container within a given output stream and is made available for use by analysts in the format of a DST file. As the raw event data is also saved, periodic "re-strippings" are performed to take advantage of improvements in the selection and reconstruction software or to apply different selection criteria.

3.2.9 LHCb Simulation

Simulated events produced using a Monte Carlo (MC) method can be an invaluable tool for investigating detector efficiencies and controlled consistency checks. In the high statistics realm of the LHC, precision modelling of detector efficiencies and accurate representation of the real data are crucial. Therefore, it is important that the MC is simulated in a way that accurately represents a particle's production, decay and detection, requiring a complete model of the underlying particle interactions that take place at the LHC and an accurate description of the detector geometry and its response. The generation of simulated data at LHCb is performed using a GAUDI based software package called GAUSS which interfaces with several other packages developed to generate particular processes. The first step is to generate an initial proton-proton collision, any additional pile up collisions, the result of multiple proton-proton collisions occurring within the same recorded event, and the resulting hadronisation of the quarks and gluons produced. This and the subsequent decay of short lived particles is done using the PYTHIA [107] event generator package. The decay of long lived particles is controlled by EVTGEN [108] with the PHOTOS [109] and TAUOLA [110] packages used to correct for radiative decays and simulate τ decays, respectively. Although, events are generated in the fully allowed 4π phase space region around the interaction region, generated particles that do not remain within the LHCb acceptance are not kept for the next stage of processing to avoid waste of CPU and disk resources. Generated events are also usually forced to contain at least one decay of the particular mode under study. Once the particle interactions have been generated they are propagated through a digital model of the detector. Implemented in GEANT4, the geometry of the detector and information on its material composition as well as running conditions such as the alignment are stored in XML databases which encode all of the relevant information. The detector response and subsequent digitisation is simulated using LHCb's BOOLE application. These latter stages of the event generation are the most CPU intensive. GAUSS has the option to apply kinematic cuts at the generator level to remove events that would likely not pass the trigger or stripping, reducing wasted CPU time. The raw simulated data can then be passed through the reconstruction, trigger and stripping using the same framework as real data. For simulated data the trigger and stripping can be run in flag mode, retaining candidates that fail the selection. The software retains the truth information about the generated particles so that efficiencies and resolutions can be measured from simulated data. An alternative type of simulated data known as Toy Monte Carlo is used in fit diagnostics and systematic cross-checks. Toy Monte Carlo is simply a set of kinematic variables that are generated according to a set of probability density functions (PDFs). As toy MC is used extensively in this analysis, details of how each dataset is generated are given in Chapters 5 and 6.

Chapter 4

Dataset and Selection

The following chapter describes the dataset used in the analysis and the selection applied. The swimming technique for correcting decay time biases is introduced and its implementation at LHCb described. Specific sources of background and their characteristic signatures as a function of the kinematic variables used in the analysis are also discussed.

4.1 Introduction

This analysis uses 1.0 fb^{-1} of data collected by the LHCb detector in 2011. Due to the small rate of mixing in charm, the statistical uncertainty is expected to dominate so the yield (and hence efficiency) should be as large as possible. At the same time, the purity needs to be high in order to minimise uncertainties (both statistical and systematic) associated with the background. It is also essential to know the flavour of the D meson (D^0 or $\overline{D}^0$) at production[1]. The flavour is tagged by requiring it is produced via the decays:

$$D^{*+} \to D^0 \pi^+_{slow}, \tag{4.1}$$

$$D^{*-} \to \bar{D}^0 \pi^-_{slow}. \tag{4.2}$$

The low-momentum pion π^+_{slow} is commonly known as the *slow pion*. As the D^{*+} decay

---

[1] Note that D^0 is used as shorthand for D^0 and $\overline{D}^0$ unless explicitly stated otherwise.

proceeds via a strong interaction there are no flavour changing effects and the charge of the soft pion is used to identify the D^0 flavour. The Q-value of the D^{*+} decay (i.e. the difference in mass of the system before and after the decay, which is given by $Q = m(D^{*+}) - m(D^0) - m(\pi^+_{slow})$) is approximately 5.5 MeV [23]. Therefore, the mass difference between D^{*+} and $(D^0 + \pi^+_{slow})$ mesons will peak at around this value for D^0 that are correctly selected as coming from D^{*+} decays. This mass difference, denoted Δm, can then be used to estimate the amount of correctly and incorrectly tagged signal events in the data. Often the value of Δm is defined in a different manner to the Q-value such that the pion mass is not subtracted giving $\Delta m = m(D^{*+}) - m(D^0)$. This is the definition used throughout what follows unless specified otherwise.

This analysis studies the D^0 resonant structure by means of a Dalitz plot analysis dependent on the D^0 proper lifetime. The decay

$$D^0 \to K^0_s \pi^+ \pi^-, \tag{4.3}$$

is reconstructed, where $K^0_s \to \pi^+\pi^-$. The following sections describe the methods used to select decays of this type from the data as well as investigating the level of contamination from potential background events that are selected in addition to the signal mode.

4.2 Trigger Selection

All signal D^0 candidates used in the final fit are required to pass a set of trigger lines designed specifically to select $K^0_\mathrm{S}\pi^+\pi^-$ decays. A description of the LHCb trigger is given in Section 3.2.7. The L0 selection is loose such that events are not required to have fired any specific line thus avoiding the efficiency penalty of making an L0 TOS requirement. The L0 trigger is of limited importance in this analysis as it does not have any lifetime-biasing effects. It does affect the efficiency variation across the Dalitz plot but, as will be shown in Section 6.5, this has a limited systematic effect on the mixing parameter measurements. At HLT1, all associated D^0 candidates are required to be TOS with respect to the single track

Property	Cut
D^0 daughter π p_T	> 1000.0
D^0 daughter π momentum	> 1500.0 MeV
D^0 daughter π track χ^2/DoF	< 5.0
D^0 daughters' vertex χ^2/DoF	< 10.0
D^0 daughters' vertex displacement w.r.t PV	> 2.0 mm
K_s daughter π momentum	> 2000.0 MeV
K_s daughter π χ^2_{IP} w.r.t PV	> 9.0
K_s daughter π track χ^2/DoF	< 20.0
K_s mass window (w.r.t PDG value)	± 50.0 MeV (before vertexing) ± 11.4 MeV (after vertexing)
K_s vertex displacement w.r.t PV	< 650.0 mm
K_s vertex displacement χ^2 w.r.t PV	> 100.0
K_S DIRA	> 0.9999
D^0 decay vertex χ^2/DoF	< 20.0
D^0 p_T	> 2000.0 MeV
D^0 mass window (w.r.t PDG value)	± 100.0 MeV
D^0 DIRA	> 0.0
`Hlt1`* TOS	`true`
Maximum number of tracks in an event	120

Table 4.1: List of cuts a candidate has to pass to be TOS on the exclusive $K^0_S\pi^+\pi^-$ trigger line. Hlt1* refers to the fact the HLT2 decision is taken as the logical AND of the HLT2 line decision and any HLT1 line TOS decision. The DIRA is the scalar product of a particle's momentum and displacement-from-PV vectors.

trigger line, `Hlt1TrackAllL0` (described briefly in Section 3.2.7 and in more detail in [111]). At HLT2, the associated D^0 must be TOS with respect to an exclusive $K^0_S h^+h^-$ line which is designed to select $K^0_S\pi^+\pi^-$, $K^0_S K^+K^-$ and $K_s K^\mp\pi^\pm$ decays, but does not attempt to reconstruct the D^{*+} candidate. K^0_S candidates are constructed from tracks that leave hits in both the VELO and the downstream tracking stations (so-called *long tracks*), are consistent with the pion hypothesis and pass the relevant cuts. The full list of cuts applied in HLT2 is given in Table 4.1. The trigger selection was kept stable throughout the data taking period, but different versions of the reconstruction software and detector alignment were used. This is of particular importance for the swimming which is described in Section 4.6. Table 4.2 contains a list of all combinations of MOORE version, TCK (see Section 3.2.7) and database tags used to collect the data used in this analysis. No additional trigger requirements are placed on the D^{*+} candidates.

Moore version	TCK	CondDB tag	DDDB tag	Run range
v12r5	0x5A0032	head-20110331	head-20110302	89333–90899
v12r5	0x5B0032	head-20110331	head-20110302	
v12r5	0x5D0032	head-20110331	head-20110302	
v12r6p1	0x6D0032	head-20110512	head-20110302	91631–94386
v12r6p1	0x6D0032	head-20110524	head-20110302	
v12r6p1	0x700034	head-20110512	head-20110302	
v12r6p1	0x710035	head-20110524	head-20110302	
v12r6p1	0x730035	head-20110524	head-20110302	
v12r6p1	0x740036	head-20110524	head-20110302	
v12r6p1	0x710035	head-20110622-Reco10	head-20110302	
v12r8	0x760037	head-20110622	head-20110302	95929 - 101011
v12r9p1	0x790037	head-20110622	head-20110302	101012 - 104486
v12r9p1	0x790037	head-20110722	head-20110722	
v12r9p1	0x790037	head-20110901	head-20110722	
v12r9p1	0x790038	head-20110901	head-20110722	

Table 4.2: Every combination of MOORE version, Trigger Configuration Key (TCK), conditions and database tags (CondDB and DDDB) used during 2011 data taking. The MOORE version determines the version of the reconstruction code used, the TCK controls the trigger line configuration used and the conditions and database tags determine the detector alignment constants used. The HLT cuts remained constant throughout 2011 data taking but the reconstruction and alignment were altered over time to improve the detector performance.

4.3 Stripping Selection

The final D^{*+} candidate reconstruction is done via the stripping (described in Section 3.2.8). A dedicated stripping line is used which is designed to select $D^{*+} \to (D^0 \to K^0_\mathrm{S}\pi^+\pi^-)\,\pi^+_{slow}$, $D^{*+} \to (D^0 \to K^0_\mathrm{S}K^+K^-)\,\pi^+_{slow}$ and $D^{*+} \to (D^0 \to K^0_s K^\mp\pi^\pm)\,\pi^+_{slow}$ decays. In addition to performing the D^{*+} reconstruction, the stripping is also used to further refine the data by applying additional selection criteria to the K^0_S and D^0 candidates. The list of cuts applied in the stripping to select the $K^0_\mathrm{S}\pi^+\pi^-$ mode are listed in Table 4.3.

Property	Cut
D^0 daughter π momentum	> 1500 MeV
D^0 daughter π track χ^2/DoF	< 4.0
D^0 daughter π ($\Delta\mathrm{LL}_e$ - $\Delta\mathrm{LL}_\pi$)	< 10.0
D^0 daughter π ($\Delta\mathrm{LL}_p$ - $\Delta\mathrm{LL}_\pi$)	< 15.0
D^0 daughter π ($\Delta\mathrm{LL}_K$ - $\Delta\mathrm{LL}_\pi$)	< -1.0
K_s DIRA	> 0.9997
K_s mass window (w.r.t PDG)	$\pm$ 11.4 MeV
K_s vertex displacement χ^2 w.r.t PV	> 100.0
D^0 decay time	> 0.3 ps
D^0 vertex χ^2/DoF	< 13.0
D^0 DIRA	> 0.0
D^0 p_T	> 1500 MeV
D^0 mass window (w.r.t PDG)	$\pm$ 130.0 MeV (before vertexing) $\pm$ 110.0 MeV (after vertexing)
D^{*+} p_T	> 2200 MeV
D^{*+} vertex χ^2/DoF	< 20.0
π^+_{soft} (DLL_e - DLL_π)	< 5.0
Δm window	> 0.0 MeV < 15.0 MeV
Maximum number of tracks in an event	150

Table 4.3: List of cuts applied to $K^0_\mathrm{S}\pi^+\pi^-$ candidates in the stripping. These cuts were stable throughout the data taking period in 2011. The DIRA is the scalar product of a particle's momentum and displacement-from-PV vectors.

4.4 Offline Processing

Ntuples for use in offline analysis were produced using a custom software package specifically developed for the purposes of this analysis. In addition to saving all of the relevant particle information in the ntuple, two forms of kinematic re-fitting are applied using LHCb's implementation of the `DecayTreeFitter` (DTF) algorithm [112]. The first instance constrains the π^+_{slow} to originate from the primary vertex. This gives an improvement in Δm resolution of a factor of ~ 2.5, the effect of which can be clearly seen in Figure 4.1. The second instance is to constrain the invariant mass of the 3-body decay to the average D^0 mass taken from the PDG. This brings events that lie outside of the kinematically allowed region of the Dalitz plot back inside and gives the Dalitz distribution a sharp cut-off at the kinematic limit. Plots of the Dalitz distribution from data with and without the DTF mass constraint are shown in Figures 4.2 (a) and (b), respectively. In both instances any kinematic variables that are recalculated during the re-fit are kept in the resulting ntuples in addition to their original values. Despite constraining the D^0 mass around 1.7 % of events passing the selection remain outside of the kinematic boundary of the Dalitz plot. The Dalitz plot binning scheme is provided by CLEO in [58] in the form of a grid of micro-bins 0.0054×0.0054 GeV$^2/c^4$ in area. To maximise the number of events in the final sample a bin-finding algorithm is used to capture those events which are not associated with any micro-bin. It starts by scanning north, south, east and west until it finds the closest bins in each direction. It then triangulates the closest micro-bin and appends the associated CLEO index to that event. The effect of including and excluding these events is explored as a systematic uncertainty in Section 6.9.

4.5 Offline Selection

Although the output from the stripping is of high purity, further reduction of background is achieved by applying a cut-based selection to the data. A list of the cuts and their values is given in Table 4.4. This set of cuts helps suppress background to a very low

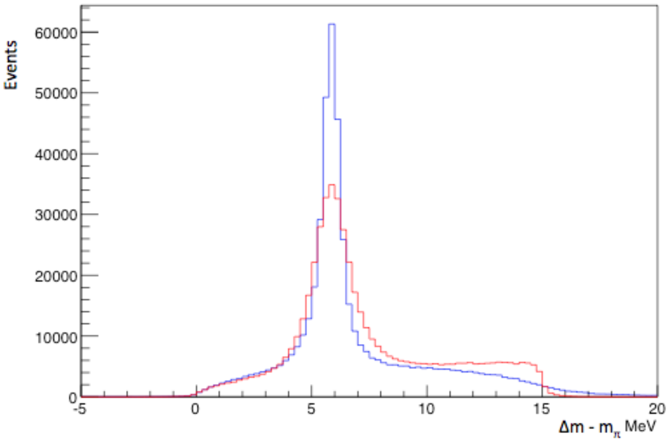


Figure 4.1: The $(\Delta m - m_\pi)$ distributions before (red) and after (blue) applying `DecayTreeFitter`. Data used to produce this plot are required to be TOS on the relevant HLT1 and HLT2 lines, and DTF is required to have converged with a $\chi^2/\mathrm{DoF} < 10.0$. Note that the red line has a cut-off at $\Delta m = 15$ MeV corresponding to the stripping cut, whereas the blue line instead slopes off. This is due to the difference in resolution between the DTF and non-DTF distributions. The red distribution also has a small tail at 15 MeV due to a slight difference in the value of D^0 mass used in the stripping compared with what is calculated in the offline processing.

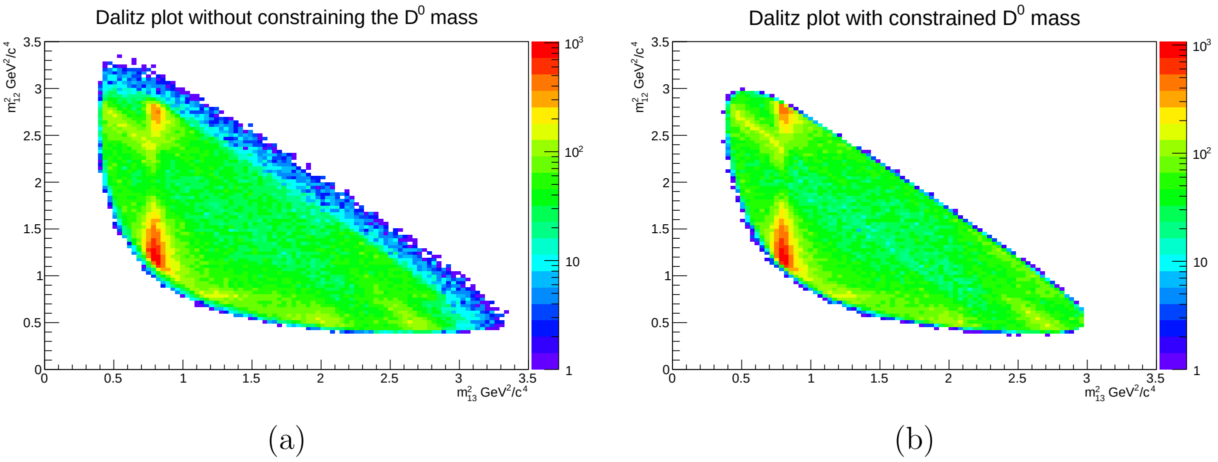


Figure 4.2: Dalitz plot distributions (a) before and (b) after applying `DecayTreeFitter`. Data used to produce this plot are required to be TOS on the relevant HLT1 and HLT2 lines, and DTF is required to have converged.

level while keeping a large proportion of signal. In addition to the cuts in Table 4.4, signal and background regions in D^0 mass and a signal region in Δm are defined. The D^0 mass sidebands are used to determine the characteristics of the combinatoric background whereas the signal windows in mass and Δm are used to maximise the purity of the final data sample used in the fit. These regions are listed in Table 4.5. The average multiplicity of candidates in events across the full range in mass and Δm that are retained after imposing all the selection cuts is 1.05. A multiple candidate rejection procedure is applied. For events that contain multiple D^0 candidates a single D^0 is selected at random[2]. For events where there are multiple D^{*+} candidates sharing the same D^0 candidate, the one with the best D^{*+} decay vertex fit χ^2 is chosen. After the full selection and multiple candidate procedure, the resulting 2D D^0 mass vs. Δm distribution is shown in Figure 4.3 for the full data sample.

Property	Cut
D^0 decay time	< 10.0 ps
$\sigma(t)$ D^0 decay time uncertainty	< 10.0 ps
χ^2/ndf of D^0 proper time fit	< 20.0
D^0 flight distance (w.r.t PV)	> 2.0 mm
K_s flight distance (w.r.t D^0 end vertex)	> 10.0 mm
π_{soft} PID DLL(e - π)	< 2.0
π_{soft} ghost probability (from Neural Net)	< 0.7
`DecayTreeFitter` ndf	> 0
n_{TP} Number of turning points from the swimming	2

Table 4.4: List of rectangular cuts used in the offline selection. This does not include mass windows which are listed in Table 4.5. The cut on N_{DoF} is a requirement that DTF converged successfully. The cut on n_{TP} is discussed in Section 4.6.

Region	Cut
D^0 mass signal (w.r.t to PDG)	± 20.0 MeV
Δm signal	$144.2 < \Delta m < 146.4$ MeV
D^0 mass lower sideband	$1.785 < m < 1.810$ GeV
D^0 mass upper sideband	$1.920 < m < 1.945$ GeV

Table 4.5: Signal and background regions in m and Δm.

---

[2] A major advantage of this requirement is that it simplifies the swimming code.

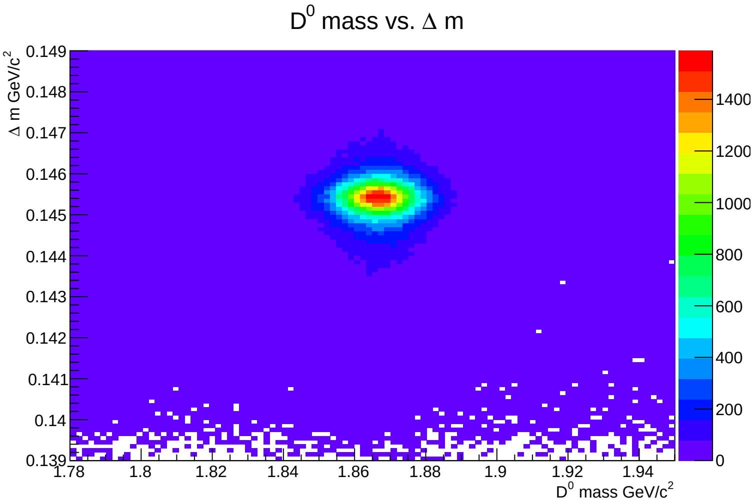
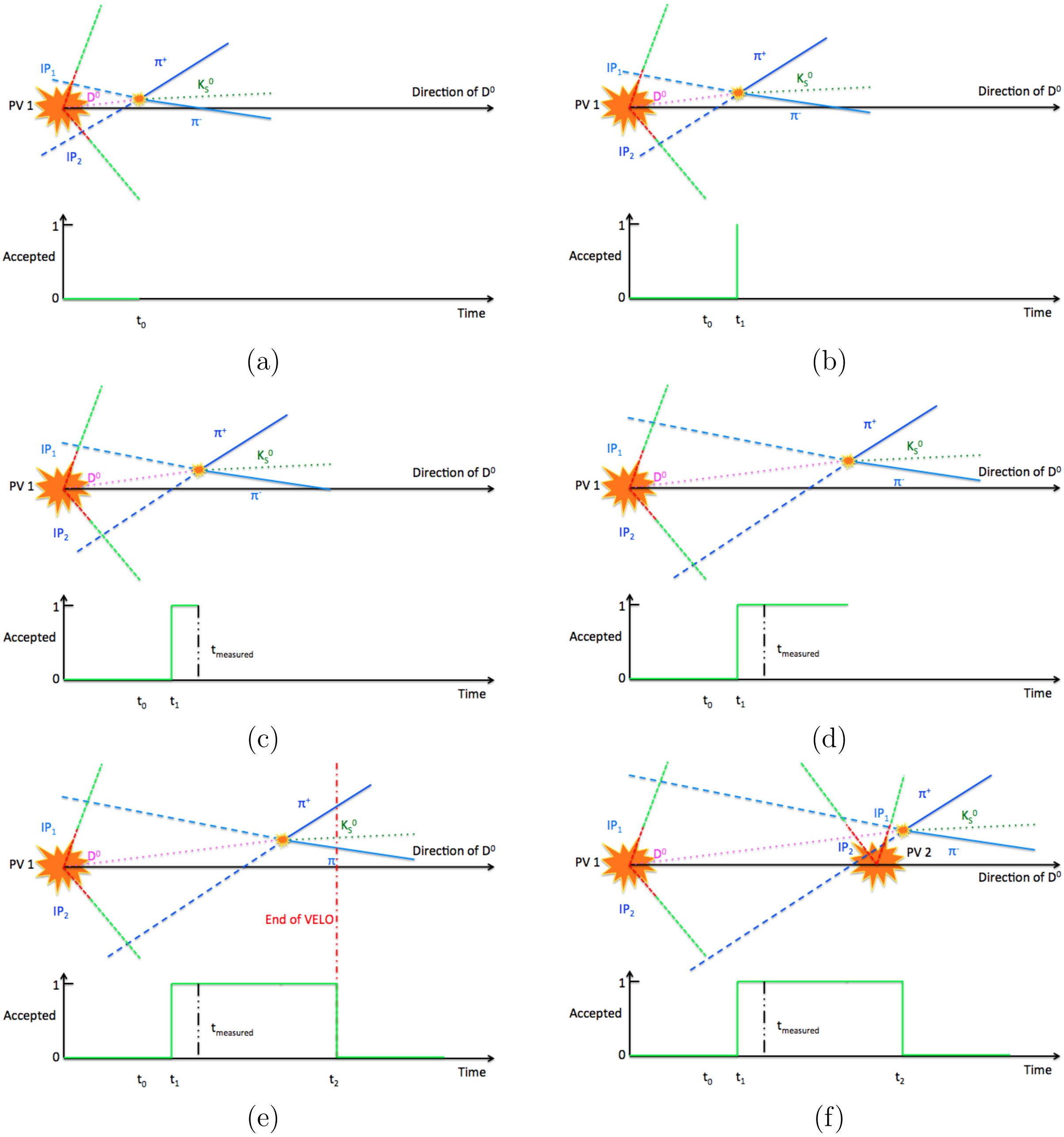


Figure 4.3: 2D D^0 mass vs. Δm distribution from data after full selection has been applied. Re-fitted variables have been used in the calculation of Δm where DTF is required to converge.

4.6 Swimming

The swimming is a data-driven technique for correcting decay time biases induced by cuts applied in the trigger, stripping and offline selection that are correlated with the reconstructed decay time. The swimming is used to determine the per-event decay time acceptance i.e. the limits in decay time within which a candidate would pass the selection cuts given the kinematics and geometry of that particular candidate in a given event. This information can then be used to correct for the decay time bias on a per-event basis in the fit (details on how this is done are given in Section 5.2). The swimming technique was first used by the NA11 experiment at CERN [113], was further developed by the CDF collaboration [114] and then further refined by LHCb [115, 116, 117]. The fundamental technique used by each experiment involves re-evaluating the acceptance decision for the candidate across a range of decay times until the regions of positive acceptance are defined to a certain precision. Figures 4.4 (a)–(f) demonstrate how this works in the simplified case where $D^0 \to K^0_S\pi^+\pi^-$ candidates pass the selection if at least one of the D^0 daughter pions has a large impact parameter (IP) with respect to all primary vertices in the event. In each figure a schematic of the decay is given along with the acceptance as a function of reconstructed decay time. In Figure 4.4 (a) it is clear that at t_0 neither of the daughter pions pass the IP cut and the event is not accepted. Figure 4.4 (b) shows that at t_1 the D^0 daughter π^+ passes the IP cut and the event is accepted. Figure 4.4 (c) and (d) indicate the measured decay time and that the event is accepted for all decay times where the IP of each of the D^0 daughters is larger than the cut. Figures 4.4 (e) and (f) demonstrate two possible causes of upper acceptance effects. Figure 4.4 (e) demonstrates that at very large decay times, the tracks may not be reconstructed due to the physical length of the VELO. Figure 4.4 (f) demonstrates the effect of multiple primary vertices in the event. As the IP cut is defined with respect to all primary vertices in the event, if there is a significant separation between multiple PVs such that they can be resolved, then a candidate may pass the IP cut when measured from its origin vertex, but may then fail the cut when measured from a secondary PV further downstream corresponding to a longer decay time as calculated with respect to its original PV.

The simplified example discussed above demonstrates the fundamental concept behind the swimming algorithm. However, in general, data selection is much more complicated with many decay time biasing cuts implemented at different stages of the reconstruction process. The swimming involves re-running any selection that has decay time biasing cuts (for LHCb this means the trigger, stripping and, for this analysis, the offline selection) many times. The implementation of swimming at LHCb is greatly simplified compared with that of NA11 and CDF due to the vast majority of the trigger being implemented in software rather than hardware. As mentioned in Section 3.2.7, this makes re-running the trigger offline a very simple operation that only requires a handful of parameters to ensure that the same configuration is run offline as was used to record the data. To simplify the procedure further, rather than moving the decay vertex and daughter tracks (as demonstrated in Figures 4.4 (a)–(f)) to vary the decay time, the primary vertices are moved instead. At LHCb, this method of varying the decay time is legitimate as the full VELO track reconstruction and PV association is run each time the trigger decision is evaluated. However, there are some caveats to this technique which are discussed in Sections 4.6.4 and 4.6.6. The total proper time acceptance of a given event can be constructed as the product of the acceptances of sequentially applied requirements—trigger, stripping, offline cuts—and for each of these the



(a) (b) (c) (d) (e) (f)

Figure 4.4: Schematics demonstrating the effect of impact parameter cuts on decay time acceptance in $D^0 \to K^0_S\pi^+\pi^-$ decays.

acceptance is either 0 or 1 at each value of the proper time. This can be expressed as a sum of Heaviside functions. Since the combined acceptance must be zero at both very large positive and very large negative proper times, it takes the form of a top-hat function or a sum of non-overlapping top-hat functions. The limits of the accepted region in decay time for each stage of the selection are described using a vector of turning points, where each turning point is defined as a point in decay time where the accepted decision for a given line has changed. As one can swim the logical OR of trigger and stripping lines, turning points are saved where the decision of any of these lines changes regardless of whether the overall accepted decision has or not. It is not necessary to utilise this extra information for the purposes of this analysis, but it is kept for diagnostic purposes.

In practice, the decay time acceptance is determined separately for (a) the trigger, (b) the stripping, and (c) the offline selection. The product of these acceptances is then obtained by merging the turning points. Steps (a) and (b) are performed using the swimming algorithm as implemented in the LHCb software. This is discussed briefly in the following sections. For a more detailed description of swimming at LHCb please refer to [118]. The swimming of the offline selection is done analytically and is described in Section 4.6.3. There are additional acceptance effects that are not included in these three components, due to the VELO acceptance and the dependence of the track reconstruction efficiency on impact parameter. These effects are discussed in Sections 4.6.4 and 4.6.6.

4.6.1 Swimming the Trigger

The HLT TOS decisions for a D^0 candidate depend on its kinematics and decay time. The LHCb trigger software can be re-run offline with exactly the same configuration as used for taking the data, enabling the trigger decision to be re-evaluated across a range of decay times. This provides information about the trigger acceptance in the region of decay time scanned. The swimming of the trigger for data is performed within the central production framework provided by the LHCb collaboration via the MOORE application. The trigger configuration is set on a per-run basis by cross-referencing the run number of the input data with a pre-compiled database containing the trigger configurations (in the form of TCK and database tags as discussed in Section 3.2.7) for every run in 2011.

The per-event acceptance is initially determined by moving the position of all primary vertices in the event[3] along the direction of the D^0 momentum. This has the effect of varying the decay time of the candidate, but without the need to move the position of the track hits left in the VELO. This approach is preferable as it is much simpler to implement and is insensitive to any potential differences in the track objects created in HLT1 at each swimming step due to imperfect extrapolation of the track trajectory/hits, but does have some shortcomings (addressed in Section 4.6.4). This simplified technique is only viable at LHCb due to the fact that the VELO track reconstruction is run each time the HLT1 decision is evaluated with the resulting track objects being used throughout the rest of the trigger decision making process. To make efficient use of computing resources an iterative procedure is used to determine the position of the turning points. A rough search for the turning points is initially performed with a granularity of 4 mm between ±200 mm of the associated primary vertex position and a granularity of 40 mm between ±600 and ±200 mm. The finer granularity is used in the region corresponding to $\pm$ 10 D^0 lifetimes as the acceptance in this range must be known to a high precision if one is to recover the correct distribution in that range. For upper acceptance effects at very large decay times the impact of getting the turning point position wrong has less of an effect on the shape of the recovered distribution and thus does not need to be determined as accurately. There is a small chance that, in the region of interest, the finer granularity of 4 mm could miss very narrow top hats. However, this type of acceptance is primarily driven by overlapping (yet resolved) primary vertices which do not occur frequently under nominal running conditions at LHCb and so are unlikely to have a significant effect. Once the rough turning points have been determined the iteration procedure starts by scanning in 1 mm steps ±4 mm about the turning point, then 250 μm and so on with the final turning point position determined to a precision of $15.625/\sqrt{12}$ μm after 4 iterations. In addition to evaluating the trigger decisions the primary vertex association is also re-run at every swimming step to ensure that any acceptance effects

---

[3] All PVs must be moved to account for effects similar to that demonstrated in Figure 4.4 (f).

due to the candidate being associated with another primary vertex are accounted for. The HLT1 and HLT2 decisions at each turning point are saved in the output from this swimming step and the resulting DST file fed to the next stage, which is to swim the stripping.

4.6.2 Swimming the Stripping

The swimming of the stripping is done using essentially the same algorithm as for swimming the trigger, but with the stripping decision for the D^0 evaluated at each swimming step. Since the decay time acceptance depends purely on the D^0 and not on the D^{*+}, only the D^0 candidate selection is rerun with information about the slow pion being ignored[4].

4.6.3 Swimming the Offline Selection

The only decay time biasing cut in the offline selection is the D^0 flight distance cut at 2 mm. As this cut is not applied at the trigger or stripping level, candidates that pass the cut may have to have their total decay time acceptance altered to account for the fact this cut has been applied. To include this in the resulting acceptance for candidates that pass the cut, the D^0 decay time corresponding to a flight distance of 2 mm w.r.t the associated PV is calculated and is treated as an additional turning point. If this turning point intersects a region of positive acceptance then it replaces the low decay time edge for the top hat it intersects. In all cases any top hats below this value are excluded. This ignores a potential complication that can occur if there are further top-hats at short lifetimes due to a change in the PV association. However, as we reject events with more than 2 turning points (Table 4.4), any additional acceptance that is below the stripping decay time cut will have no effect and can safely be ignored.

---

[4] At a technical level, for each D^0 candidate any associated D^{*+} candidates are retained in the swimming output and recombined with the swum D^0 candidate during the offline reconstruction step (see Section 4.4).

4.6.4 Swimming the VELO Acceptance

As the swimming moves the primary vertices rather than the tracks, it does not take into account that the tracks themselves may move out of the geometric acceptance of the VELO (as demonstrated in Figure 4.4 (e) for the D^0 daughter pions). At high decay times there is a possibility that some or all of the signal tracks may not be reconstructed due to the fact they may not traverse enough of the VELO tracking stations. For the HLT1 track reconstruction the minimum number of VELO stations that a hadron track must leave hits in is 5. The equivalent value for a track to be reconstructed offline is 3. This is of particular importance in this analysis as the K^0_S candidates are made up of long tracks (tracks that leave hits in the VELO and the downstream tracking stations) and K^0_S particles in LHCb typically fly a significant distance before decaying. To account for this effect, the origin vertex of each of the four tracks is swum along the D^0 momentum direction. At each step the number of VELO stations that the track passes through is calculated with a simple model of the VELO geometry, and a vector of VELO turning points for both the HLT1 and offline acceptance for each track is calculated and merged with those from the trigger, stripping, and offline selection. The offline VELO acceptance is applied for all of the daughter tracks (logical AND). The HLT1 VELO acceptance requirement is only applied for tracks that are used in the HLT1 decision of the D^0 at the decay time under consideration, requiring that at least one pass (logical OR).

4.6.5 Merging the Acceptances

The merging step takes the product of the acceptances from the trigger, stripping offline selection and the VELO, each of which is a sum of top-hats. The final acceptance is then simply another vector of turning points that describe the total acceptance of the D^0 candidate. Only events with exactly two turning points after merging (i.e. an acceptance function that is a single top-hat with a switch-on and a switch-off) are considered in this analysis.

4.6.6 Parameterising the Tracking Efficiency

The track reconstruction imposes an additional acceptance effect that is not accounted for in the VELO acceptance. The track-finding algorithm employs a method of clone-killing that is designed to prevent the same track being reconstructed many times from subtly different hit patterns. The clone-killing has the effect of the track reconstruction preferring “on-axis” tracks i.e. tracks that originate from the beam line. As D^0 and especially K_s mesons fly significant distances before decaying, their daughter tracks will tend to be “off-axis” and are less likely to be reconstructed leading to a depletion of events at larger D^0 flight distances (and hence larger decay times). It has been shown previously [119] that for long tracks this efficiency is quadratic as a function of the track radial impact parameter (DOCAZ) and may be parameterised as

$$\epsilon(\text{DOCAZ}) = \alpha + \beta(\text{DOCAZ})^2. \tag{4.4}$$

The radial impact parameter is defined as the track’s distance of closest approach to the beam line i.e.

$$\text{DOCAZ} = \frac{\boldsymbol{x}}{|\boldsymbol{p}|} \cdot (\boldsymbol{p} \times \hat{\boldsymbol{b}}) \tag{4.5}$$

where $\boldsymbol{x}$ is the position vector of the track origin vertex, $\boldsymbol{p}$ is the track 3-momentum and $\hat{\boldsymbol{b}}$ is the beam axis unit vector. The beam axis position in the $x - y$ plane is taken as the average over all primary vertices in the event. The values of α and β are determined for all four daughter tracks independently using signal Monte Carlo. Taking the ratio of the number of reconstructed tracks to the number of those that are reconstructible in bins of DOCAZ^2 and fitting a straight line to the resulting distribution enables the parameters α and β to be determined. This is done using 1.6M truth matched signal MC events that are produced without any generator level cuts and are taken with no additional selection applied such that any structure in the ratio is due to the effect of the track reconstruction efficiency. The fitted values for each daughter track are given in Table 4.6 and the fit projections are shown in Figure 4.5. Parameterising this efficiency as a function of decay time for use in the fit requires a relationship between the D^0 decay time and DOCAZ. This is linear, but the parameterisation is unique to every event as it is dependent on the kinematics of the D^0.

Daughter Track	α	β (mm^{-2})
D^0 daughter π^+	0.9617 ± 0.0003	-0.0685 ± 0.0121
D^0 daughter π^-	0.9619 ± 0.0003	-0.0623 ± 0.0121
K_s daughter π^+	0.9447 ± 0.0007	-0.0055 ± 0.0002
K_s daughter π^-	0.9427 ± 0.0008	-0.0050 ± 0.0002

Table 4.6: Fitted values parameterising the track reconstruction efficiency as a function of DOCAZ^2.

The origin vertex of each D^0 daughter track is swum along the D^0 momentum direction and the value of DOCAZ and the corresponding D^0 decay time are calculated at each point. A straight line is fitted to the resulting distribution such that

$$\text{DOCAZ}(t)_i = \gamma_i + \delta_i t \tag{4.6}$$

where γ_i and δ_i are unique to event i. The resulting time-dependent efficiency for this event can be written as

$$\epsilon(t)_i = \alpha + \beta(\gamma_i + \delta_i t)^2. \tag{4.7}$$

The values of γ_i and δ_i are evaluated for each event and saved in the resulting ntuple. A description of how the efficiency is treated in the fit is given in Section 5.2.5. The effect of an imperfect description of this efficiency on the resulting mixing parameters is explored as a systematic uncertainty in Section 6.8.

4.7 Monte Carlo Studies

Several Monte Carlo samples were generated for use in this analysis and are described in the following section. All of the MC events produced contain at least one signal event, i.e. at least one ($D^{*+} \to D^0\pi^+_{soft}$, $D^0 \to K^0_S\pi^+\pi^-$, $K^0_s \to \pi^+\pi^-$) decay is contained within the simulated event. Table 4.7 lists the MC samples and the number of events generated. The *signalflat* sample contains signal events that are generated according to a phase-space distribution (i.e. a uniform, flat Dalitz plot). The *signalmix* and *signalmixfiltered* samples contain signal events produced with an isobar decay model with the amplitude composed of a

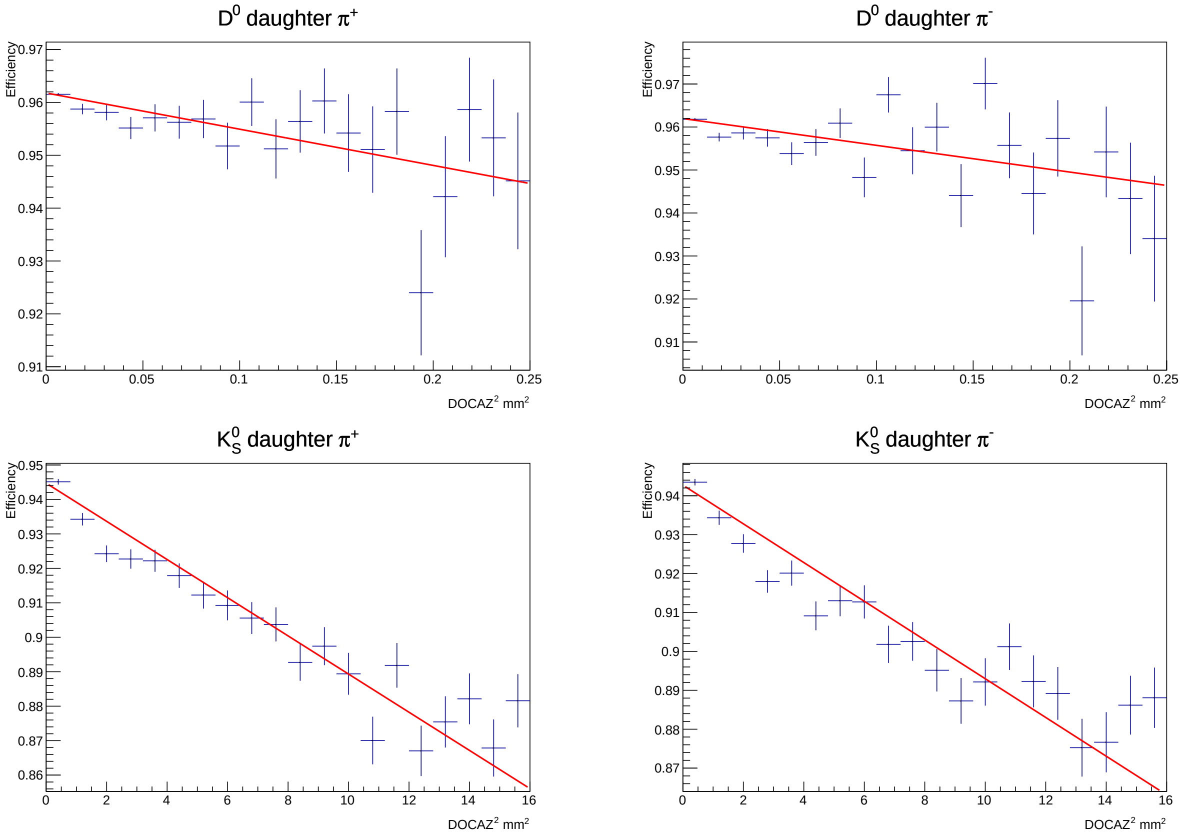


Figure 4.5: Fits to track reconstruction efficiency from MC as a function of DOCAZ^2 for each of the four daughter tracks.

Sample	Events Generated	Events Reconstructed
signalflat	10M	197918
signalmix	10M	197014
signalmixfiltered	600k	303206

Table 4.7: Number of generated Monte Carlo events for each sample of $K^0_S\pi^+\pi^-$ signal MC. N.B. Reconstructed yields are the number of events output from the stripping with no further selection applied.

Variable	Cut
D^0 p_T	> 2.0 GeV/c
D^0 t	> 0.25 ps
D^0 decay vertex z-pos	< 8 m
At least one $\pi^\pm$ $\|\boldsymbol{p}\|$	> 2.7 GeV/c
At least one $\pi^\pm$ p_T	> 1.4 GeV/c
K^0_S decay vertex z-pos	< 600 mm
D^{*+} p_T	> 1.5 GeV/c

Table 4.8: List of cuts applied in the generator level filtering.

linear combination of relativistic Breit-Wigner propagators and with decay time dependence corresponding to mixing parameters $x_D = y_D = -1$ % and no CP-violation. The standard LHCb generation procedure, as described in Section 3.2.9, was used. The resulting samples are processed in the same way as the real data. For the samples that have resonant Dalitz plot decays, the implementation of the model used to generate these decays with `EvtGen` is based on the 2008 BaBar amplitude model which is described in [53]. Due to the low efficiency of the trigger and stripping, most of the events generated will not contain a reconstructed, triggered, and selected signal candidate. Since CPU and disk resources were limited, two types of event filtering are used to improve the output rate of useful events. The first is at the generator level, i.e. before the detector response has been simulated, and is referred to as generator level filtering. The cuts used for this are looser than those in the trigger and stripping selections to allow for resolution and are listed in Table 4.8. This type of filtering is applied to all the MC samples listed in Table 4.7. The second type of filtering is applied at the stripping level after the full detector response has been simulated and the reconstruction and trigger applied, and is referred to as stripping filtering. This is only applied to the *signalmixfiltered* sample.

4.8 Dalitz Plot Efficiency

In the limit that the strong phase is constant within a Dalitz plot bin (DP), the sensitivity of the mixing and CPV parameters to efficiency variation with DP position[5] vanishes. Although the bins were chosen so as to minimise the strong phase variation, there will be some finite effect in practice. The baseline results for the mixing parameters are obtained under the assumption that the efficiency is symmetric about the axis $m^2_{12} = m^2_{13}$ and that it is slowly varying across the Dalitz plot bins. In order to determine the systematic effects due to these assumptions (described in Section 6.5), a model of the relative efficiency will be required.

The *signalflat* MC sample is generated with a uniform Dalitz Plot distribution. Therefore the relative efficiency of the complete selection (trigger, reconstruction/stripping, offline cuts) as a function of position in the DP is simply given by the distribution of truth-matched *signalflat* events after it has been applied. An empirical efficiency model, outlined below, is fitted to this distribution. There are clearly limitations to the accuracy of this model, mainly due to (a) the limited amount of phase space signal MC available, and (b) systematic differences in efficiency between data and MC. However, a highly accurate model of the efficiency variation with DP position is not needed for this analysis as it is not used to correct for any systematic effects (unlike, e.g., the BaBar and Belle analyses which fitted amplitude models); it merely needs to be good enough to determine the systematic uncertainties.

The model to parameterise the efficiency (Equation 4.8) was chosen from several candidate polynomials of different orders by comparing the χ^2 per degree of freedom. The function that best describes the data is

$$\begin{aligned} \epsilon(m^2_{12}, m^2_{13}) &= au^3 + bu^2 + cu^2v + du + euv + fv + gv^2 + hv^3 \qquad (4.8) \\ u &= (m^2_{12} + m^2_{13}),\ v = m^2_{12} - m^2_{13} \end{aligned}$$

where the Dalitz plot coordinates are transformed to preserve the kinematic limits. The

---

[5] Variation as a function of decay time is much more important to this analysis and was discussed in Section 4.6.

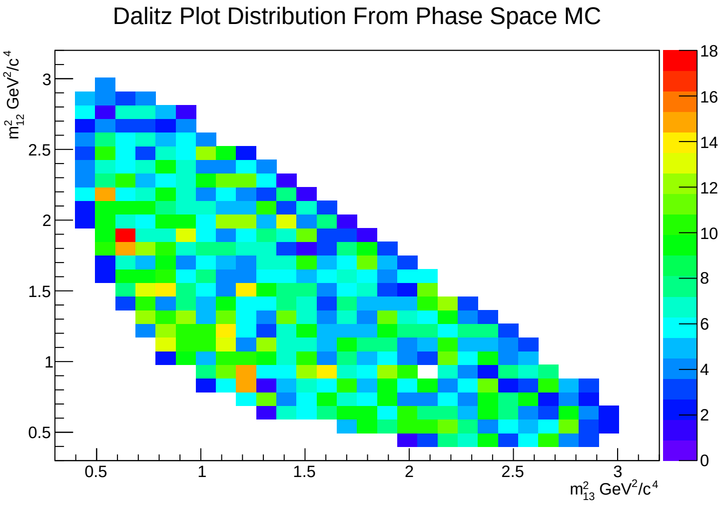
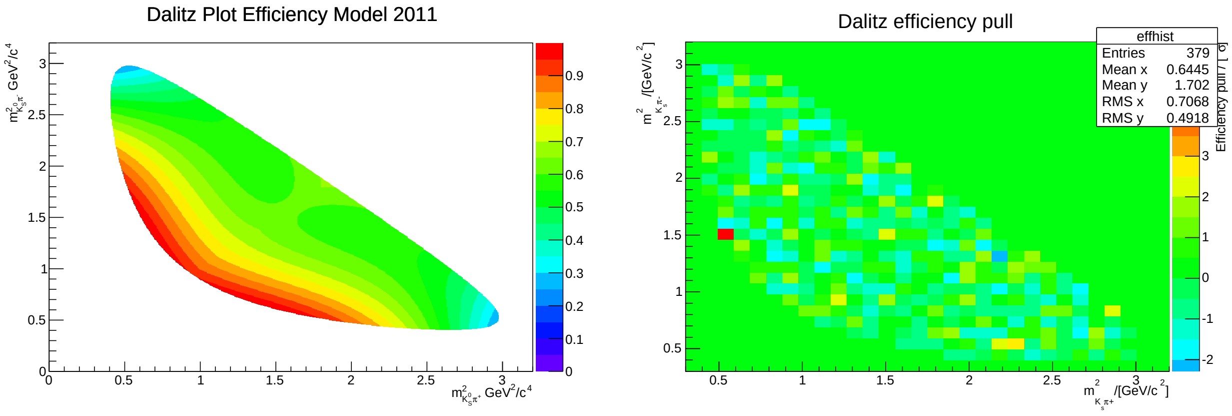


Figure 4.6: Dalitz plot distribution of *signalflat* MC after the complete selection.

resulting parameters from the best fit are listed in Table 4.8 and plots of the selected MC events, model and resulting pull distribution are given in Figures 4.6 and 4.7. There are a handful of bins with large pulls at the edge of the DP where the sample size is small leading to non-Gaussian errors. This model is denoted the 2011 model. A selection of other models are used to investigate the various systematic effects associated with the efficiency model itself. These are described in Appendix B.



(a) (b)

Figure 4.7: (a) The 2011 efficiency model as fitted to the distribution in Figure 4.6. (b) Pulls between the model and the data in Dalitz Plot bins.

Parameter	Value
a	-0.408 ± 0.105
b	0.730 ± 0.161
c	0.067 ± 0.134
d	0.249 ± 0.125
e	-0.166 ± 0.153
f	-0.224 ± 0.166
g	-0.100 ± 0.035
h	-0.204 ± 0.140

Table 4.9: Fitted parameter values for the Dalitz plot efficiency model defined in Equation 4.8.

4.9 Background Categories

In addition to the signal, there are several background components present in the data. These are considered in the following subsections. In the fit they will be separated from the signal by their distributions in D^0 mass m, Δm, the logarithm of the D^0 impact parameter chi-square $ln(\chi^2_{IP})$ and decay time t. Consequently, we can group together classes of background that have the same distributions in these variables; those which have different behaviour in any of the discriminating variables may warrant individual treatment if present in non-trivial quantities. If so, they will be treated as separate, additional components in the fit.

4.9.1 Specific Sources of Background

$D^0 \to K^0_S h^+ h^-$

One of the more dangerous potential backgrounds is from other $D^0 \to K^0_S h^+ h^-$ decays, especially $D^0 \to K^0_S K^\mp \pi^\pm$ and $D^0 \to K^0_S K^+ K^-$ where one or both of the D^0 daughter kaons are mis-identified as pions. These types of event have a highly non-trivial decay time distribution as they can exhibit mixing. They also have resonant structure in the Dalitz plot. The latter would be particularly tricky to describe as the distribution will be warped by the fact one or more of the daughters are reconstructed under the wrong mass hypothesis. However, this last fact also means that the reconstructed peak in m for these events will be

Decay	Magnet Polarity	Generated	Reconstructed	Efficiency
$D^0 \to K^0_S K^- \pi^+$	Down	1,431,200	27	$(1.89 \pm 0.36) \times 10^{-5}$
$D^0 \to K^0_S K^- \pi^+$	Up	1,431,200	28	$(1.96 \pm 0.37) \times 10^{-5}$
$D^0 \to K^0_S K^+ \pi^-$	Down	1,431,200	26	$(1.82 \pm 0.36) \times 10^{-5}$
$D^0 \to K^0_S K^+ \pi^-$	Up	1,431,200	28	$(1.96 \pm 0.36) \times 10^{-5}$

Table 4.10: Number of reconstructed $K^0_S K^{\mp}\pi^{\pm}$ decays passing the $K^0_S\pi^+\pi^-$ stripping line.

shifted down, away from the signal peak and into the lower sideband. This shift is larger for $K^0_S K^+ K^-$ events as both kaons would be reconstructed under the wrong mass hypothesis. The $K^0_S K^+ K^-$ events will also be suppressed by two pion PID cuts. Therefore, it is fair to assume[6] that if there is no significant contribution from $K^0_S K^{\mp}\pi^{\pm}$ decays then this will also hold for $K^0_S K^+ K^-$.

To assess if there is a sizeable contribution from either of the $K^0_S K^{\mp}\pi^{\pm}$ modes, the trigger, stripping and offline selection for $K^0_S\pi^+\pi^-$ is applied to a set of $K^0_S K^{\mp}\pi^{\pm}$ Monte Carlo events and a value for the "efficiency" for $K^0_S K^{\mp}\pi^{\pm}$ to fake $K^0_S\pi^+\pi^-$ is estimated. It should be noted that the MC is generated using the generator level cuts listed in Table 4.8, but where one of the pions is replaced with a kaon. Four samples each consisting of 1,415,200 events split by $K^0_S K^{\mp}\pi^{\pm}$ mode and magnet polarity were generated using the full LHCb MC generation chain (described in Section 3.2.9). Table 4.10 shows the number generated, number reconstructed and corresponding efficiency for each decay mode and magnet polarity. The weighted average of the reconstruction efficiency across each polarity is used in the following calculations.

The $K^0_S\pi^+\pi^-$ MC reconstruction efficiency is calculated to be $\epsilon(K^0_S\pi^+\pi^-) = (0.117 \pm 0.002)\%$ using the *signalmix* sample listed in Table 4.7. Using Equations 4.9, the total MC reconstruction efficiency for $K^0_S\pi^+\pi^-$, and branching fractions for each mode taken from the PDG [23] one can estimate the total contamination from these decays in the real data. $\mathcal{N}$ is a normalisation factor that scales the number of observed $K^0_S\pi^+\pi^-$ events to the yield

---

[6] Bearing in mind that $\mathcal{BF}(D^0 \to K^0_S K^+ K^-) \sim \mathcal{BF}(D^0 \to K^0_S K^{\mp}\pi^{\pm})$ since one is phase-space suppressed and the other is Cabibbo-suppressed.

Decay	Branching Fraction
$D^0 \to K^0_S\pi^+\pi^-$	(2.83 ± 0.20)%
$D^0 \to K^0_S K^-\pi^+$	(0.35 ± 0.05)%
$D^0 \to K^0_S K^+\pi^-$	(0.21 ± 0.04)%
$D^0 \to K^0_S K^+K^-$	(0.45 ± 0.03)%

Table 4.11: Branching fractions for $D^0 \to K^0_S\pi^+\pi^-$, $D^0 \to K^0_S K^\mp\pi^\pm$ and $D^0 \to K^0_S K^+K^-$ modes taken from the PDG [23].

Mode	Events after selection	Contamination
$K^0_S\pi^+\pi^-$	172605 ± 472	N/A
$K^0_S K^-\pi^+$	232 ± 73	(0.13 ± 0.04)%
$K^0_S K^+\pi^-$	137 ± 46	(0.08 ± 0.03)%
$K^0_S K^\mp\pi^\pm$	369 ± 86	(0.21 ± 0.05)%

Table 4.12: Estimated contamination from $K^0_S K^\mp\pi^\pm$ decays from signal Monte Carlo.

obtained from the mass fits (see Section 5.4.1).

$$\begin{aligned}
\mathcal{N} \times \epsilon_{K^0_S\pi^+\pi^-} \times \mathcal{BF}(D^0 \to K^0_S\pi^+\pi^-) &= N_{K^0_S\pi^+\pi^-} = 172605 \pm 472 \\
\mathcal{N} \times \epsilon_{K^0_S K^-\pi^+} \times \mathcal{BF}(D^0 \to K^0_S K^-\pi^+) &= N_{K^0_S K^-\pi^+} \\
\mathcal{N} \times \epsilon_{K^0_S K^+\pi^-} \times \mathcal{BF}(D^0 \to K^0_S K^+\pi^-) &= N_{K^0_S K^+\pi^-}
\end{aligned} \tag{4.9}$$

Re-arranging the top expression in Equation 4.9 to extract $\mathcal{N}$ and substituting into the bottom 2 equations allows $N_{K^0_S K^-\pi^+}$ and $N_{K^0_S K^-\pi^+}$ to be estimated. The branching ratios for each mode are given in Table 4.11 and the expected number of each mode after trigger, stripping and selection are given in Table 4.12. The rate of $K^0_S K^\mp\pi^\pm$ contamination within the signal mass window is approximately 2 per mille and is comparable to the statistical error on the extracted yield for $K^0_S\pi^+\pi^-$. As a result, this component can be ignored in the signal fit. However, this may not be the case for the D^0 mass sidebands. As this is a real D^0 decay reconstructed under the wrong mass hypothesis one would expect some reflection in the lower mass sideband as the mass difference between the kaon and the pion is lost during the daughter combination. This can be clearly seen in Figure 4.8 where the mass distributions for the each of the four $K^0_S K^\mp\pi^\pm$ samples reconstructed under the $K^0_S\pi^+\pi^-$ mass hypothesis are given. As the mass sidebands are used to estimate the shape of the decay time distribution for the combinatoric background, contamination from $K^0_S K^\mp\pi^\pm$ decays could result in a bias.

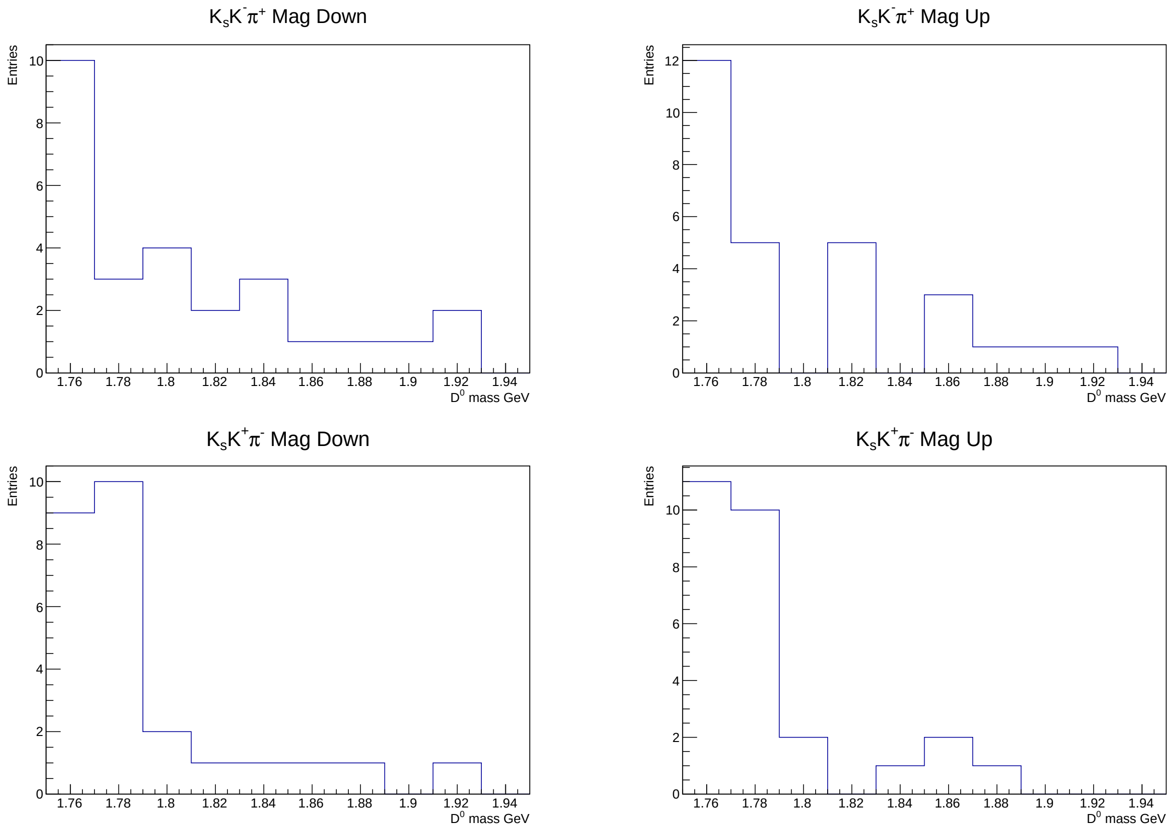


Figure 4.8: $K^0_{\rm S}K^{\mp}\pi^{\pm}$ MC events reconstructed under the $K^0_{\rm S}\pi^+\pi^-$ mass hypothesis for each of the four MC samples studied.

The effect of this is investigated in Section 6.10.3[7]

$D^0 \to K^0_{\rm S}\pi^+\pi^-\pi^0$

Mis-reconstructed $D^0 \to K^0_{\rm S}\pi^+\pi^-\pi^0$ decays where the π^0 is missed are another potentially dangerous background. Again, as it is a real D^0 decay it can mix and, as we have many bodies in the final state, it can have a non-trivial Dalitz plot distribution. It also has a larger branching than that of the signal mode: $\mathcal{BF}(D^0 \to K^0_{\rm S}\pi^+\pi^-\pi^0) = 5.4 \pm 0.6$ % compared to $\mathcal{BF}(D^0 \to K^0_{\rm S}\pi^+\pi^-) = 2.83 \pm 0.20$ %. There is also missing mass in the form of the π^0 resulting in the reconstructed D^0 mass for these events being shifted into the lower sideband. As this shift is of the order 140 MeV it is unlikely that there is any significant contribution from these decays within the mass signal window but there could be a significant amount

---

[7] Note that there may be differences between the data and MC efficiencies, especially for PID and χ^2_{IP} cuts. However, these are neglected here since the rate of $K^0_{\rm S}K^{\mp}\pi^{\pm}$ contamination is so low.

in the lower sideband. As the combinatoric background shape is determined using the mass sidebands contamination from these decays could lead to mis-modelling of the combinatoric component within the signal window. The effect of this is investigated in Section 6.10.3.

$D^0 \to \pi^+\pi^-\pi^+\pi^-$

The decay $D^0 \to \pi^+\pi^-\pi^+\pi^-$ produces the same final state as $D^0 \to K^0_S\pi^-\pi^+$, though without an intermediate K^0_S. This decay is suppressed by two cuts. The first is the K^0_S mass requirement in the stripping; this helps but there will still be a small fraction that sits under the reconstructed K^0_S mass. The second is the requirement that the D^0 and the K^0_S vertex are separated by a significant distance, which will further suppress any contribution from these decays. The VELO can achieve a vertex resolution down to 11 μm (around 50 μm in the z-direction) and K^0_S particles at LHCb have a typical lab-frame flight length in the tens of cm with a large proportion leaving the VELO completely before decaying. Taking into account these factors a fairly conservative cut of 10 mm was placed on the K^0_S flight distance with respect to the D^0 decay vertex. This along with cuts placed in the quality of the K^0_S decay vertex and the mass requirements in the trigger and stripping are expected suppress any background from $D^0 \to \pi^+\pi^-\pi^+\pi^-$ decays to a negligible level.

4.9.2 Combinatoric Background

The combinatoric background consists of mis-reconstructed events. These can be from mis-reconstructed D decays as described in the previous sections, but also from mis-reconstructed B decays and random clusters of tracks emanating from the primary vertex. The contributions from mis-reconstructed D and B decays are suppressed by requiring that the vertex fit is of good quality and that the D candidate points back to the primary vertex. The random clustering is suppressed by requiring the D candidate to fly a significant distance from the primary vertex as well as requiring that the daughter tracks have a large χ^2_{IP}. These requirements remove a large quantity of the combinatoric background, but there is a component that remains and must be modelled in the fit. The shape of the combinatoric background in any given variable can be estimated by examining the relevant distribution plotted from data in the mass sidebands. However, as discussed previously, the lower sideband in particular can be contaminated with events that exhibit signal-like behaviour and, more importantly, which are not representative of the background in the signal window. To see whether this has any effect on the final result the fit will be run with the combinatoric description taken from both sidebands and then from each of the upper and lower sidebands separately. The results of this test are described in Section 6.10.3.

4.9.3 Wrong π^+_{slow} Decays

Wrong-π^+_{slow} decays are real $D^0 \to K^0_S\pi^+\pi^-$ decays that are recombined with a random slow pion. As the D decay is correctly reconstructed these events will peak in D^0 mass but, as the π^+_{slow} does not originate from a D^{*+}, will follow a smooth distribution in Δm that increases from threshold. These events will be some admixture of D^0 and $\overline{D}^0$ decays and will require special treatment in the fit. The full treatment of these events in the fit is described in Sections 5.4.2 and 5.5.2.

4.9.4 Secondary D Decays

Finally, there is contamination from mis-reconstructed B/Λ_b decays where $B/\Lambda_b \to D^0X$ and $D^0 \to K^0_S\pi^+\pi^-$ for which the D^0 is correctly reconstructed. These events will peak in D^0 mass (and, for the subset that proceed via a D^{*+}, in Δm), but will have a non-trivial decay time distribution due to the presence of the B. Therefore, there can be secondary components to both prompt signal and wrong-π^+_{slow} decays that behave differently in Δm, but are indistinguishable in D^0 mass. An additional variable is, therefore, required to distinguish between the prompt and secondary components in the data. As secondary decays generally fly further than prompt decays and are missing momentum due to the lost tracks they tend not to point back towards the primary vertex. The $ln(\chi^2_{IP})$ of the D candidate can be used to distinguish between the prompt and secondary events. The distributions in $ln(\chi^2_{IP})$ as a function of decay time from truth-matched signal Monte Carlo are given for both prompt and secondary decays in Figure 4.9. It is clear that there is a region of overlap between the prompt and secondary events at low decay times that cannot be totally removed with a cut. The secondary decays are then modelled as additional components in the fit, the full treatment of which is described in Sections 5.4.2 and 5.5.2.

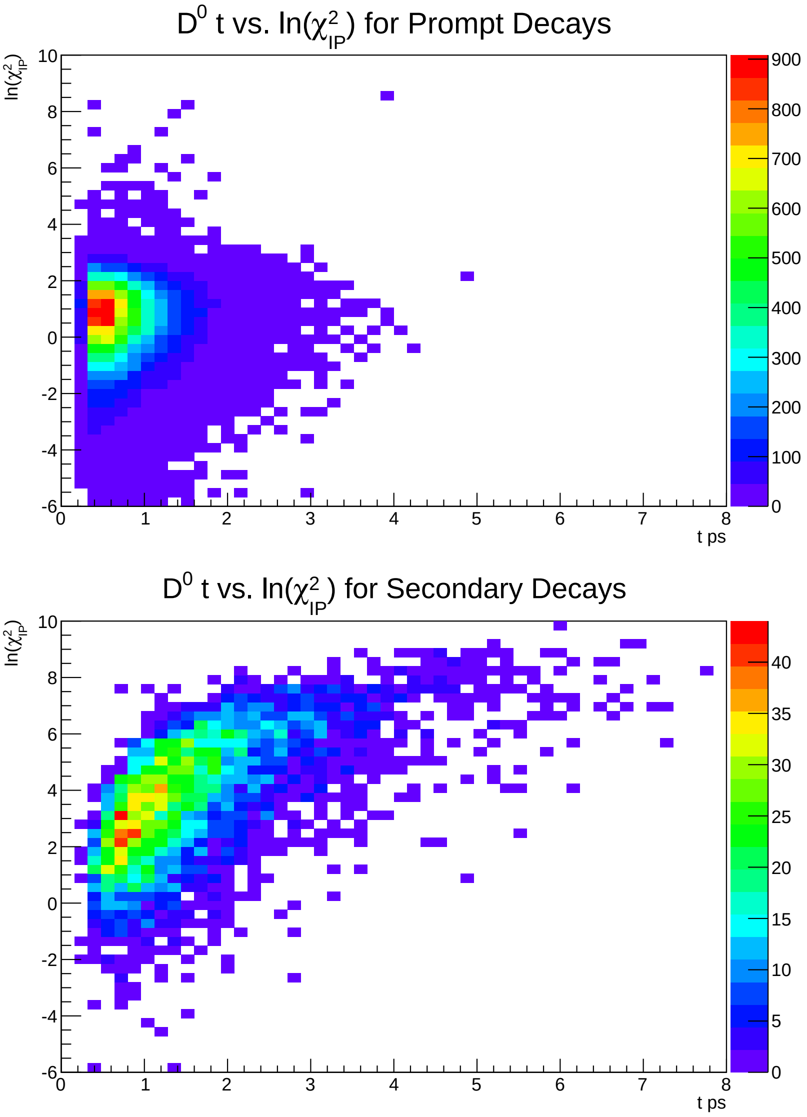


Figure 4.9: t vs. $ln(\chi^2_{IP})$ distributions for prompt (top) and secondary (bottom) decays taken from truth matched signal Monte Carlo.

Chapter 5

Fit

The aim of the fit procedure is to extract the mixing parameters from the available data. This requires accurately modelling the background components outlined in the previous chapter and correcting the signal components for detector acceptance effects. The total procedure consists of two fits; a fit to the mean D^0 lifetime integrated across the Dalitz plot (hereby referred to as the "lifetime fit"), and a simultaneous fit to data split by D^0 flavour across each Dalitz plot bin to extract the mixing parameters (hereby referred to as the "mixing fit"). The lifetime fit is performed to check the integrity of the generic fit procedure using real data and to extract the shapes of the $ln(\chi^2_{IP})$ distributions for prompt and secondary events which are then fixed in the mixing fit. As the lifetime fit is performed to a mixture of D^0 and $\overline{D}^0$ candidates it was not necessary to blind the fitted D^0 lifetime as it is not sensitive to the rate of mixing[1]. Both the lifetime and mixing fits are conducted in a similar way and follow closely the work done in [67]. Each fit models a different set of components that are labelled by *class* and *subclass* as in Table 5.1. The definition and individual treatment of each *class* and *subclass* for the lifetime and mixing fits are described in Sections 5.4 and 5.5, respectively. The input variables for each fit are divided into two categories which are assumed to be independent and uncorrelated; mass-dependent variables

---

[1] The average lifetime of the CP-eigenstate components is in principle sensitive to y, but since we have a mixture of CP-even, CP-odd, and non-CP-eigenstate components there is cancellation and dilution to the point that we are insensitive to it, given the uncertainties on the world average D^0 lifetime and on our lifetime value.

and time-dependent variables. Each fit is then split into two stages. The first stage is a fit to the mass-dependent distributions, which determines the relative contribution from each class. The second stage is then a fit to the time-dependent distributions, which determines the relative contribution from each subclass and extracts the physics observables. The mass-dependent variables are the D^0 mass for the lifetime fit, and both the D^0 mass and Δm for the mixing fit. In both fits the time-dependent variables are the D^0 decay time, t, the logarithm of the D^0 impact parameter chi-squared, $ln(\chi^2_{IP})$, the distance in decay time of the first turning point, TP_1, and the distance in decay time between the first and the second turning points, TP_{diff}. The combined fit to t and $ln(\chi^2_{IP})$ is required to separate the prompt and secondary contributions. The additional turning point variables are required due to the per-event acceptance that is applied to correct for the decay time biasing effects from the selection. A complete description of how the decay time acceptance is included in the fit is given in Section 5.2. The PDFs that are used to describe the mass-dependent distributions are implemented in the fitting package RooFit [120] and an extended maximum likelihood fit is performed. The PDFs that are used to describe the time-dependent distributions are implemented in custom code and a regular maximum likelihood fit is performed. The minimisation in all cases is performed using the MINUIT minimisation algorithm [121]. The following sections describe the lifetime and the mixing fits in detail.

Fit	Class	Subclass
Lifetime Fit	*signal*	*prompt* *secondary*
	combinatoric	
Mixing Fit	*right*-π	*prompt* *secondary*
	wrong-π	*prompt* *secondary*
	combinatoric	

Table 5.1: List of components in each fit. For the mixing fit, separate components are included for correctly tagged real D^0 (*right*-π) and incorrectly tagged real D^0 (*wrong*-π).

5.1 PDF Definitions

A subset of the PDFs common to both fits that are used to describe the m, Δm and $ln(\chi^2_{IP})$ distributions are defined here for convenience.

5.1.1 Parametric PDFs

Gaussian Distribution The Gaussian distribution for a random variable x can be written as

$$G(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\left(\frac{x-\mu}{\sigma}\right)^2} \tag{5.1}$$

where μ and σ are the mean and width of the distribution, respectively. All shape parameters can be allowed to float in the fit.

Bifurcated Gaussian Distribution The bifurcated Gaussian distribution for a random variable x can be written as

$$Bif(x;\mu,\sigma_L,\sigma_R) = \begin{cases} A \cdot G(x;\mu,\sigma_R) & x \geq \mu \\ B \cdot G(x;\mu,\sigma_L) & x < \mu \end{cases} \tag{5.2}$$

where μ is the line of bifurcation and σ_L (σ_R) is the width of the distribution to the left (right) of the mean. A and B are normalisation terms such that the amplitudes of both Gaussians either side of the mean are equal at $x = \mu$. All shape parameters can be allowed to float in the fit.

Crystal Ball The Crystal Ball PDF [122] is utilised in modelling the signal component in the fit to the D^0 mass distribution. For a random variable x the Crystal Ball PDF can be written as

$$CB(x;\mu,\sigma,\alpha,n) = N \cdot \begin{cases} e^{-\left(\frac{x-\mu}{\sqrt{2}\sigma}\right)^2} & \text{for } \frac{x-\mu}{\sigma} > \alpha \\ A \cdot (B - \frac{x-\mu}{\sigma})^{-n} & \text{for } \frac{x-\mu}{\sigma} \leq \alpha \end{cases} \tag{5.3}$$

where,

$$A = \left(\frac{n}{|\alpha|}\right)^n e^{-\frac{|\alpha|^2}{2}}, \quad \text{and} \quad B = \frac{n}{|\alpha|} - |\alpha|, \tag{5.4}$$

and N is a normalisation term. The shape is approximately Gaussian, but with a tail that is described by a power law that effectively models radiative tails in the signal mass component. The parameters μ and σ describe the Gaussian shape with α and n describing the tail. For use in this analysis n is restricted to positive integer values and is fixed prior to any fit. All other shape parameters can be allowed to float.

Chebyshev Polynomial A Chebyshev polynomial of the first kind is used to model the smooth background in Δm. These polynomials for a random variable x are described using the recurrence relation

$$T_0(x) = 1 \tag{5.5}$$

$$T_1(x) = x \tag{5.6}$$

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x). \tag{5.7}$$

Therefore, the total n^{th} order Chebyshev polynomial can be written as

$$f(x; c_0, ..., c_n) = \frac{1}{N}\left(T_0(x) + \sum_{k=1}^{n} c_k T_k(x)\right) \tag{5.8}$$

where N is a normalisation factor.

Single Exponential Distribution A single exponential distribution is used to describe the combinatoric background shape in D^0 mass and the prompt signal shape in decay time. The single exponential for a random variable x with decay constant τ can be written as

$$f(x) = \frac{1}{\tau} e^{-x/\tau}. \tag{5.9}$$

Convolution of Two Exponentials A convolution of two exponentials is used to describe the decay time distribution of secondary decays. This PDF for a random variable x can be written as

$$f(x) = \frac{1}{\tau_2 - \tau_1}(e^{-x/\tau_2} - e^{-x/\tau_1}) \tag{5.10}$$

where the decay constants, τ_1 and τ_2, are floated in the fit.

5.1.2 Kernel PDFs

Several distributions in both the mixing and lifetime fits are not easily described using empirical, parametric PDFs. Such distributions are modelled using non-parametric PDFs built up of Gaussian kernels. Kernel PDFs are designed to approximate a smooth, continuous probability density from a discrete data distribution. Each data point is represented by a PDF that is non-divergent and normalised to 1 across the available phase-space. The total PDF is then taken as the sum of the kernels for each data point. A graphical example of this can be seen in Figure 5.1. Kernels can take on many shapes depending on the distribution one is trying to smooth. For the purposes of this analysis a Gaussian kernel was chosen following by example [67]. All kernel PDFs are evaluated for a discrete range of values and saved in a finely binned histogram prior to any fit.

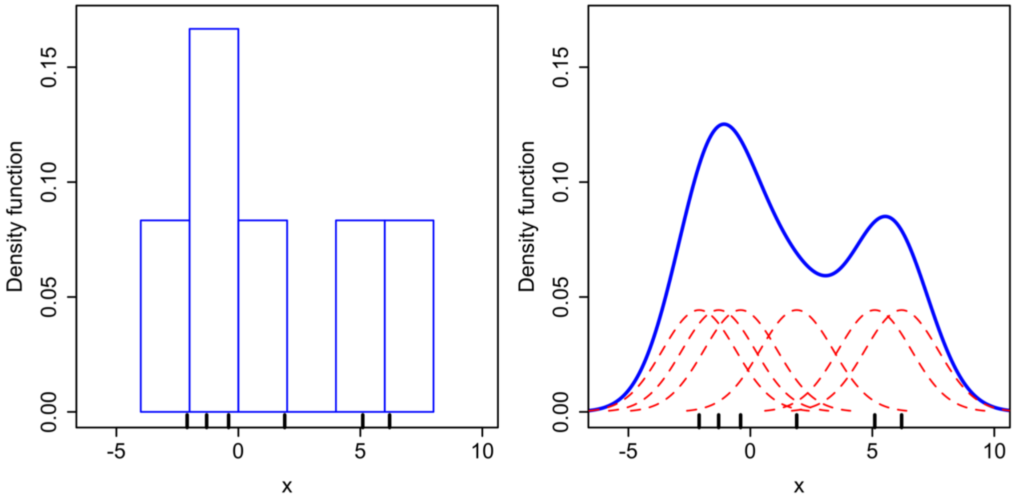


Figure 5.1: The right hand plot is an example of the Gaussian kernel density PDF applied to the binned data on the left. Image taken from [123].

1D Gaussian Kernel PDF

The 1-dimensional Gaussian kernel PDF as a function of a random variable, x for n data points can be expressed as

$$f_{kernel}(x) = \frac{1}{\sqrt{2\pi h^2}n} \sum_{i}^{n} e^{-\left(\frac{x-x_i}{\sqrt{2}h}\right)^2} \tag{5.11}$$

where h is known as the bandwidth of the distribution and controls the rate of smoothing. The ability of the kernel method to accurately represent a distribution heavily depends on estimating the optimum bandwidth for the data. For the case where the distribution one is trying to model is a normal distribution the optimum bandwidth can be defined [124] as

$$h = \left(\frac{4\sigma^5}{3n}\right)^{\frac{1}{5}} \tag{5.12}$$

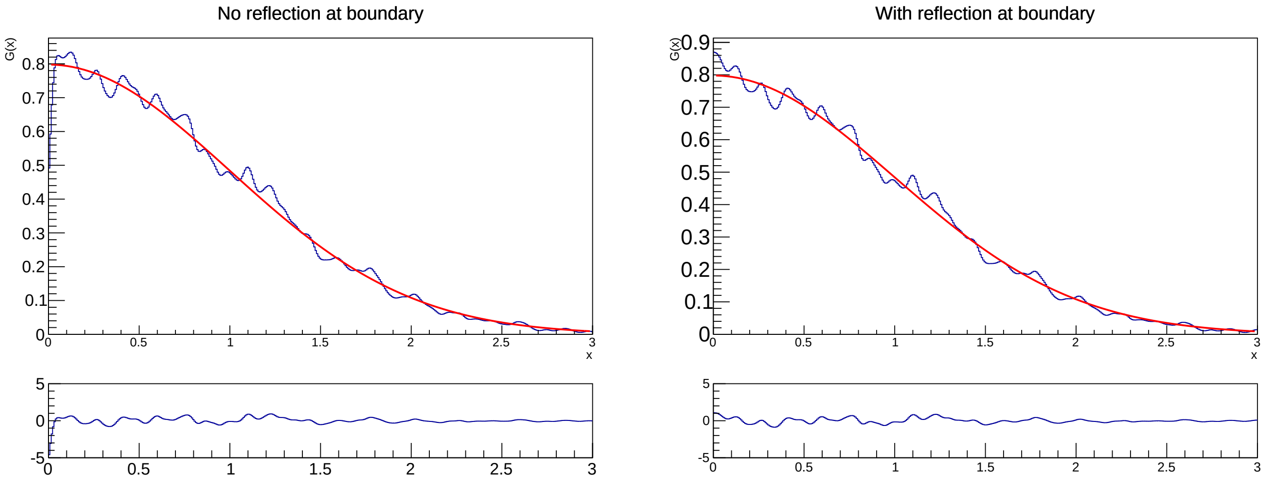
where σ is the standard deviation of the data. More generally, the optimum bandwidth for a given kernel shape, K can be written as

$$h^* = \left(\frac{R(K)}{\sigma_K^4 R(f''_{kernel}(x))n}\right)^{\frac{1}{5}} \tag{5.13}$$

where σ_K^2 is the variance of the kernel, $f''_{kernel}(x)$ is the second derivative of the distribution that is being modelled and

$$R(f) = \int_{-\infty}^{+\infty} f(x)^2 \mathrm{d}x \tag{5.14}$$

is defined as the "statistical roughness" of the data. To get a reliable estimate of h^* the distribution is initially drawn assuming it behaves normally with the optimum bandwidth set to Equation 5.12. Once a smooth estimate of the distribution has been drawn its first and second derivatives are calculated using the finite-difference technique and plugged into Equation 5.13. The distribution is then drawn again using the new optimum bandwidth estimate. For the case of 1-dimensional kernel PDFs this procedure reliably reproduces the measured distributions and good agreement between the data and kernel PDFs is observed. However, these kernel PDFs can struggle to correctly model regions close to a kinematic boundary or distribution cut-off due to over smoothing at the boundary edge and the tails of the distribution "leaking" out. This generally results in the kernel PDF undershooting the data in regions close to the boundary. This can be overcome by reflecting the leaked PDF back from the boundary, but this can lead to an overestimate by the kernel PDF. This effect can be clearly seen in Figures 5.2 (a) and (b) where the same distribution has been plotted with and without a reflection at the lower boundary. Both are produced from 15k events generated using a normal distribution with $\mu = 0$ and $\sigma = 1$. A cut is placed at $x = 0$ to create a boundary and the remaining data is kernelised with and without the reflection applied. It is clear that including the reflection does force the kernel PDF to sit above the true distribution close to the boundary. However, the deviation is smaller compared with not including the reflection. One should note that the pull distributions are normalised to arbitrary units to visualise the effect more clearly. The reflection is applied to all kernel PDFs used in the fit. As the kernel PDFs are predominantly used to model components that contribute only a few percent of the data any systematic effect from mis-modelling such regions will be small.



(a) (b)

Figure 5.2: Comparison between 2 kernel PDFs (blue) and the true distribution (red). Figure (a) has had no reflection of the kernel PDF applied at the cut boundary whereas (b) has. The pull histogram is in arbitrary units and is included to compare the deviation from the kernel PDF at the lower boundary when the reflection is not applied.

2D Gaussian Kernel PDF

The 2-dimensional Gaussian kernel PDF as a function of random variables x and y for n data points can be expressed as

$$f(x,y)_{kernel} = \frac{1}{(2\pi)h_x^2 h_y^2 n}\sum_i^n e^{-\left(\frac{x-x_i}{\sqrt{2}h_x}\right)^2} e^{-\left(\frac{y-y_i}{\sqrt{2}h_y}\right)^2}. \tag{5.15}$$

The bandwidth of the kernels in each dimension, h_x and h_y, respectively, are assumed independent. This may not necessarily be the case and one can extend this formalism to account for this. However, the implementation described here is sufficient for the purposes of this analysis. To determine the optimum bandwidth in each dimension the 1D projections are drawn and the optimum bandwidth evaluated in the same way as for the 1D kernel PDF described in the previous section. The final 2D PDF is then evaluated using Equation 5.15 with h_x and h_y set to the values calculated from the 1D projections.

5.2 Inclusion of Decay Time Acceptance and Efficiency

5.2.1 Decay Time Acceptance

The per-event decay time acceptance of the trigger, stripping and offline selection is parameterised as a series of top-hat functions that are 1 within the accepted decay time limits for that event and 0 everywhere else. This series of top-hats is described using turning points that are determined using the swimming technique as outlined in Section 4.6. To simplify the analysis, events with more that two turning points are rejected ensuring that the per-event decay time acceptance can be described using a single top-hat. Assuming a true exponential decay with some per-event acceptance, $A(t)$, which is parameterised by turning points $t_{min,i}$ and $t_{max,i}$ the per-event probability density as a function of decay time can be written as

$$f(t|A) = \frac{\frac{1}{\tau}e^{-t/\tau}A(t)\Theta(t)}{\int \frac{1}{\tau}e^{-t'/\tau}A(t')\Theta(t')\mathrm{d}t'} = \frac{\frac{1}{\tau}e^{-t/\tau}\Theta(t-t_{min,i})\Theta(t_{max,i}-t)}{\int_{t_{min,i}}^{t_{max,i}} \frac{1}{\tau}e^{-t'/\tau}\mathrm{d}t'}, \tag{5.16}$$

for t that lies within the interval $t_{min,i} < t < t_{max,i}$ and $f(t|A) = 0$ elsewhere. $\Theta(t)$ is the Heaviside function describing the region of positive acceptance. This can be generalised to give the per-event probability density for a particular class such that

$$f(t|A, class) = \frac{f_{unbiased}(t|class)\Theta(t - t_{min,i})\Theta(t_{max,i} - t)}{\int_{t_{min,i}}^{t_{max,i}} f_{unbiased}(t'|class)\mathrm{d}t'} \tag{5.17}$$

where $f_{unbiased}(t|class)$ is the unbiased PDF that describes that particular class. This per-event acceptance correction formalism is applied to the decay time PDFs for all components and used in the likelihood evaluation for the fit.

5.2.2 Average Acceptance Function

In addition to evaluating the per-event PDF given some acceptance, knowledge of the average acceptance as a function of decay time is required if one wants to make comparisons between the fitted distribution and that measured from data. The average acceptance function for a given class that has a true decay time distribution $f_{unbiased}(t|class)$ is defined as

$$\bar{A}(t|class) = \sum_i \left[\frac{\Theta(t - t_{min,i})\Theta(t_{max,i} - t)}{\int_{t_{min,i}}^{t_{max,i}} f_{unbiased}(t|class)\mathrm{d}t} \bigg/ \sum_j \frac{1}{\int_{t_{min,j}}^{t_{max,j}} f_{unbiased}(t'|class)\mathrm{d}t'} \right]. \tag{5.18}$$

To reproduce the biased distribution one simply takes the product of the unbiased PDF with the average acceptance function measured from the data.

5.2.3 Decay Time Acceptance for Kernel PDFs

A kernel PDF is used to describe the combinatoric decay time component in the fit and, therefore, must incorporate the per-event decay time acceptance. The kernel PDF that is produced from the data is modelled on the distribution that has been sculpted by the trigger, stripping and selection. The inclusion of the per-event decay time acceptance requires knowledge of the "true", unbiased distribution, such that the biased distribution is once again recovered after the acceptance has been applied in the fit. The decay time distribution for combinatoric events is taken from the D^0 mass sidebands and kernelised using the formalism described in Section 5.1.2. Using the turning points in the sideband data it is possible to produce an unbiased distribution from the biased data using an iterative procedure.

The biased distribution for a given class is simply the product of the true distribution with the average acceptance function such that,

$$f_{biased}(t|class) = f_{unbiased}(t|class)\bar{A}(t|class). \tag{5.19}$$

The form of $f_{unbiased}(t|class)$ is initially estimated by producing a kernel PDF from data generated using a linear function. The average acceptance function for this unbiased PDF is then drawn using the turning points from the sideband data. Re-arranging Equation 5.19 and inserting the estimated acceptance function provides a new estimate of the unbiased PDF. A second acceptance function is drawn from this new estimate and the process repeated. At each iteration the product of the estimated unbiased PDF and average acceptance is compared with the biased distribution measured from the data. The iteration process ceases when the mean of the absolute difference between the two distributions sampled at a large number of points is less than 1×10^{-6}. An example of this process in action is shown in Figure 5.3. It is clear that the biased distribution is successfully recovered after a small number of iterations.

5.2.4 Turning Point PDFs

The turning points are treated as additional input variables to the fit. It is not clear that the probability of an event having a pair of turning points at $t_{min,i}$ or $t_{max,i}$ is equal for all classes and subclasses. Therefore, the turning point distributions are required to be modelled for each component in the fit. Equivalently, one can argue that Equation 5.16 describes the probability density of an event given it has some acceptance, but does not describe the probability density of a given class of event having a particular acceptance, $f(A|class)$. This PDF is approximated to be the product of the distributions of $TP_1 = t_{min,i}$

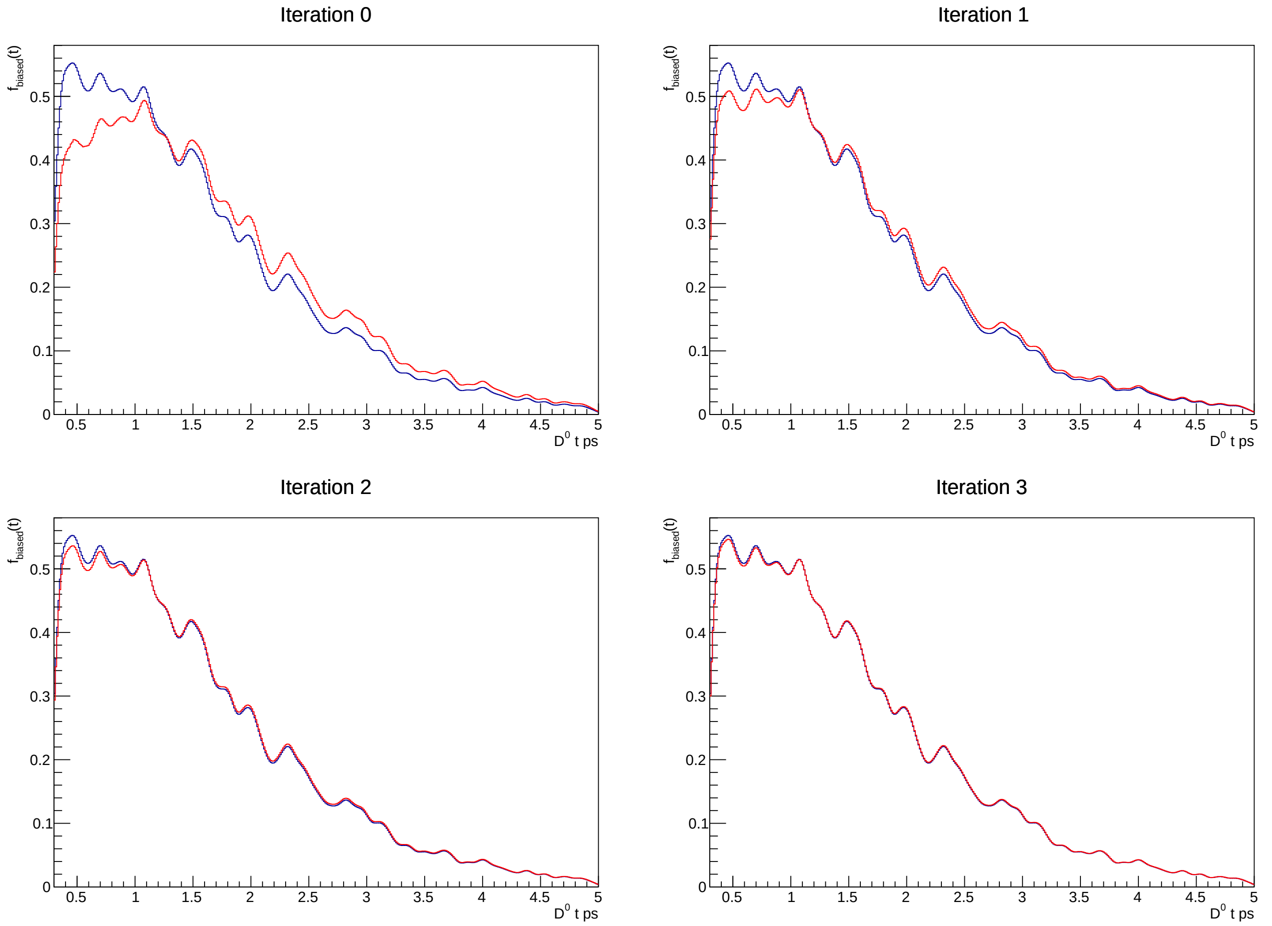
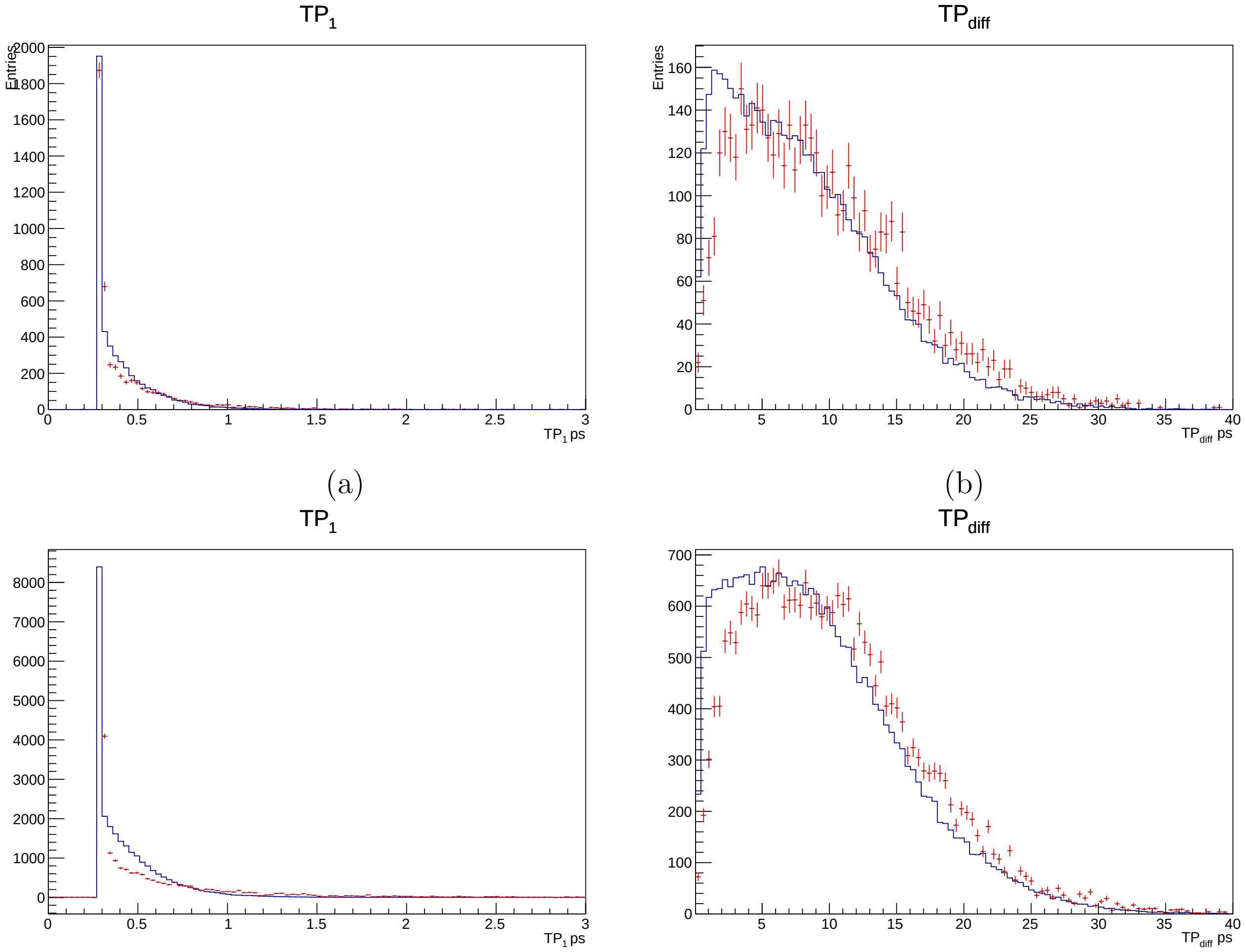


Figure 5.3: Comparisons of recovered biased kernel distribution (red) and that measured from data (blue) for iterations 1 → 4 (top left to bottom right). It is clear that the recovered distribution from the unbiased estimates quickly converges on the measured distribution.

and of $TP_{diff} = t_{max,i} - t_{min,i}$ such that

$$f(A|class) = f(TP_1|class)f(TP_{diff}|class). \tag{5.20}$$

This approximation is based on the assumption that, although the absolute position of a pair of turning points in decay time will be correlated, there is no such correlation between the position of the first turning point and the distance between subsequent turning points. For prompt $K^0_S\pi^+\pi^-$ decays the position of the first turning point is mainly driven by the kinematics of the D^0 whereas the difference between the first and second is mainly driven by the position of the K^0_S decay vertex, the K^0_S daughter kinematics and if there are any other primary vertices in the event. Although the kinematics of the D^0 and the K^0_S are correlated, the absolute position of the K^0_S decay vertex is predominantly driven by the decay time of the K^0_S which is uncorrelated with both the D^0 decay time and the kinematics of the D^0 decay. It is also true that the underlying event properties are uncorrelated with the kinematics of any signal candidates. For secondary $K^0_S\pi^+\pi^-$ decays the same rule of thumb applies, but it is not guaranteed the turning point distributions will be the same for prompt and secondary decays. In addition, the distribution of turning points for combinatoric events may well differ to that for the prompt and secondary signal classes. Figures 5.4 (a) and (b) show the distributions of TP_1 and TP_{diff} for matched prompt and secondary signal MC. There is clearly a difference between the prompt and secondary D^0. To see if this difference is modelled well by the MC a similar set of plots is produced from the data. Figures 5.4 (c) and (d) show the same plots, but for different regions of $ln(\chi^2_{IP})$. The prompt-rich distributions are from data that have $ln(\chi^2_{IP}) < 1.0$ and the secondary-rich distributions from data with $ln(\chi^2_{IP}) > 3.0$. In both cases all the data is required to sit within the D^0 mass and Δm signal windows, and the combinatoric background component is removed using a subtraction from the D^0 mass sidebands. A similar shape is found for both the TP_1 and TP_{diff} distributions in data and in the MC. However, the distributions for both components from data will still have some remaining contamination from the other component i.e. some prompt in the secondary distribution and vice versa. A re-weighting technique can be used to extract the true distributions for these components as described in [67]. The technique involves starting with initial estimates for the distributions, running the fit and then re-weighting the distributions using per-event Bayesian weights calculated from the fit result. This process is then repeated with the estimates of the turning point distributions being updated with the re-weighted version from the previous fit at the start of each iteration. The distributions are said to converge once the sum of all of the component distributions closely resemble the total distribution from data. An implementation of this was tried, but the method failed to converge and the sum of the component distributions began to diverge from the data. Therefore, the turning point distributions for prompt and secondary events are taken from all events that sit within the D^0 mass and Δm signal window and are assumed to be the same for both components. The turning point PDFs for combinatoric events are taken from the D^0 mass sidebands. A systematic uncertainty due to this simplified treatment of the turning point distributions is investigated in Section 6.4.



(a) (b)

(c) (d)

Figure 5.4: (a) TP_1 distribution for prompt (blue - solid line) and secondary (red - crosses) events from signal MC. (b) Equivalent plot for TP_{diff} from signal MC. (c) Background subtracted TP_1 distribution for prompt-rich (blue - solid line) and secondary-rich (red - crosses) events from data. (d) Equivalent plot for TP_{diff} from data.

The turning point PDFs describing the distributions for each component are implemented in the same way. The TP_1 distribution is dominated by a very sharp, narrow peak at the stripping decay time cut of 0.3 ps with a tail that extends up to 2–3 ps. The probability density of the TP_1 distribution is implemented as a histogram PDF with 100 bins. This implementation was chosen as it gave the best agreement between the model and the data. The serial nature of the PDF evaluation is not ideal, but given the majority of events all sit within one or two bins around the cut of 0.3 ps, the discriminating power mainly comes from the difference between the PDF at the peak and in the tail which the histogram implementation achieves. The distribution for TP_{diff} is much more slowly varying and can be implemented using a 1D Gaussian kernel PDF as described in Section 5.1.2. Figure 5.5 shows an example of each distribution as modelled in the fit. Note that the structure in the pulls for the TP_1 plot is mainly driven by the inclusion of the combinatoric background. As the prompt and secondary distributions use all data within the mass and Δm signal window there is still a small combinatoric component contributing to their shapes. The combinatoric distribution is then taken from the mass sidebands and included meaning it is somehow double-counted causing disagreement between the sum of all components and the distribution from data. The deviation in the TP_{diff} distribution at low decay times is due to the reflection applied to the kernel PDFs for each component which is discussed in Section 5.1.2. Despite the inconsistency between the model and the data in both plots the size of the systematic uncertainty associated with the mistreatment of these distributions is small (see Section 6.4).

5.2.5 Tracking Efficiency Correction

The evaluation of the input parameters that describe the efficiency as a function of decay time for a single track is described in Section 4.6.6. There are 4 final state tracks emanating from the D^0 decay. Each track has a reconstruction efficiency that has a quadratic dependence on decay time. Taking the product of the efficiencies of all four tracks leads to an 8th order polynomial in decay time such that the total per-event efficiency is given by,

$$\epsilon(t) = \prod_i (\alpha_i + \beta_i(\gamma_i + \delta_i t)^2) \tag{5.21}$$

where the product i runs over all four daughter tracks. The parameters α_i and β_i are set according to the type of daughter track (D^0 daughter π^+/π^- or K^0_S daughter π^+/π^-) and are constant for all events (the values of which are listed in Table 4.6). Both γ_i and δ_i are dependent on the kinematics of the D^0 decay and are unique for each event. This correction is applied to all components that are sensitive to the physics observables.

5.3 Parameter Blinding

The mixing fit is initially performed such that it is blind to the physics observables. The blinding is implemented so as to allow the comparison of parameters extracted using different fit configurations while still blind, enabling systematic checks to be performed without risk of biasing the result. Fitted values are encoded in a random string which is unique for any fit even if the values that are output are exactly the same. The key to decoding the value is nested within the string allowing values to be saved and unblinded at any time deemed appropriate. An example of the blinding of two identical values is given below.

Blind	Unblinded
`BLaPqcy2j6nMto7Mz`	100
`BDy6QCw8ukAsPLs1L`	100

5.4 Lifetime Fit

The lifetime fit considers 3 components. The prompt signal component describes real $D^0 \to K^0_S\pi^+\pi^-$ decays that originate from the primary vertex and is sensitive to the mean D^0 decay time. The secondary signal component describes $D^0 \to K^0_S\pi^+\pi^-$ decays which are real but originate from secondary decays of B mesons. Finally, the combinatoric component describes the contribution from mis-reconstructed candidates. The combinatoric component is not sensitive to the mean D^0 lifetime. The secondary component does depend on the D^0 decay time, but as the full decay chain is not reconstructed, there is no sensitivity to the the D^0 lifetime in the fit. The fit proceeds by first performing a fit to the D^0 mass distribution to get the relative contribution from real D^0 (prompt and secondary) and combinatoric events. Information from the first fit is then used in the second which is a 2D fit to the decay time and $ln(\chi^2_{IP})$ distribution to separate the prompt and secondary signal components, parameterise their $ln(\chi^2_{IP})$ distributions for use in the mixing fit and extract the mean D^0 lifetime. Information about the turning point distributions is also used in the second stage fit. The first stage fit is described in Section 5.4.1 and the second stage in Section 5.4.2.

5.4.1 D^0 Mass Fit

The first stage is the mass-dependent part which consists of a 1-dimensional fit to the m distribution. Two distinct components are modelled in this fit; correctly reconstructed D^0 (*signal*) and combinatoric background (*comb.*). The signal component will have a peaking structure centred about the true D^0 mass whereas the combinatoric shape is expected to follow a smooth slowly varying distribution. To get a handle on the combinatoric background shape the fit is performed across a broad mass window of ± 80 MeV about the nominal PDG value. Note that, although Δm is not considered in the fit, all data used throughout the lifetime fit is required to sit within the Δm signal window. The total PDF for the first stage of the lifetime fit can be written as

$$f(m) = \sum_{class} f(m|class)P(class) \tag{5.22}$$

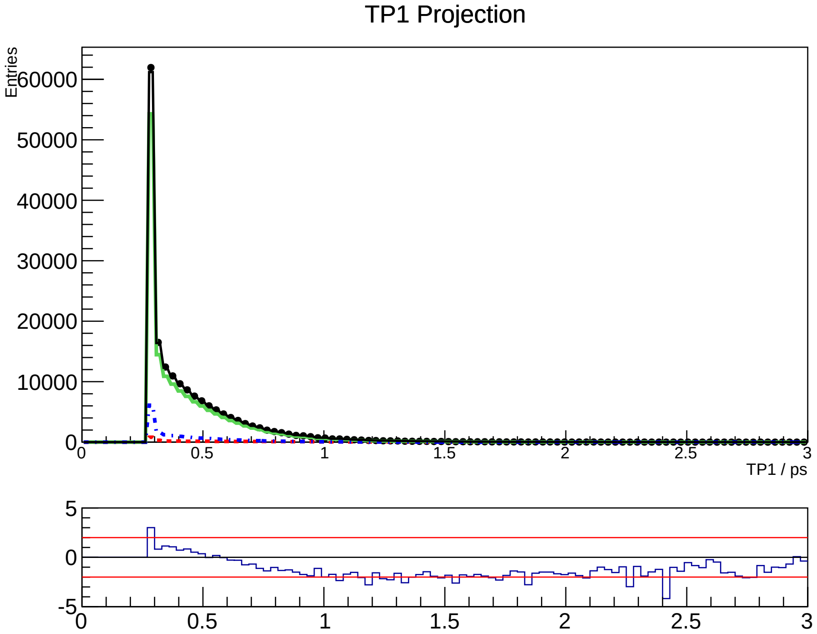
where $P(class)$ are the signal and background fractions and $\sum_{class} P(class) \equiv 1$. The signal shape is described by the sum of a single Gaussian, a bifurcated Gaussian and a Crystal Ball such that

$$f(m|signal) = f_1 G(m; m_{D^0}, \sigma_1) + f_2 Bif(m; m_{D^0}, \sigma_L, \sigma_R) + (1 - f_1 - f_2) CB(m; m_{D^0}, \sigma_2, \alpha, n). \tag{5.23}$$

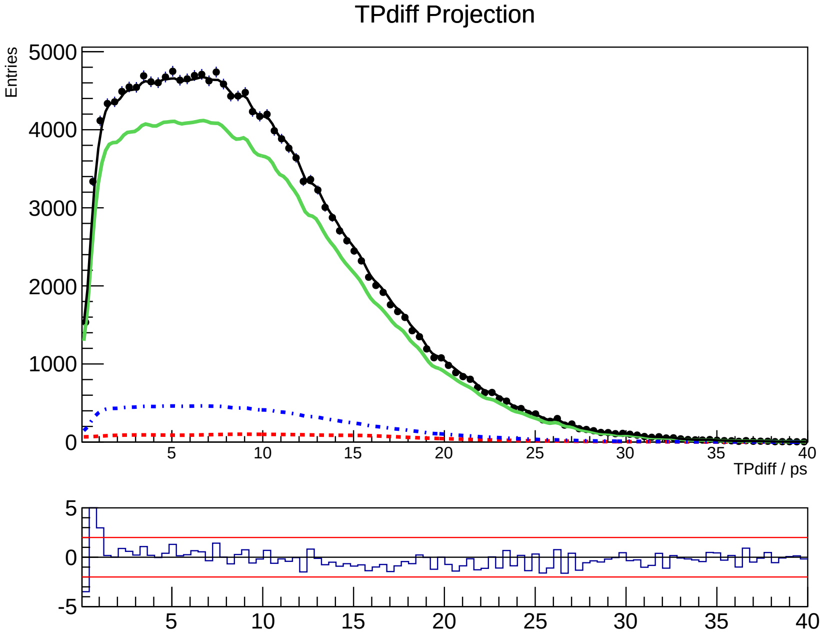
where m_{D^0} is the mean D^0 mass and the order of the Crystal Ball is fixed to $n = 3$. The combinatoric shape is described using a single exponential. The minimisation is performed across the wide mass window and the parameters describing each class are determined along with the signal and background fractions. The signal mass window cut is then applied to the data and the signal and background fractions recalculated within the new limits for use in the next stage. In addition to the signal and background fractions the probability for each event belonging to a given class is calculated using Bayes' theorem [125] such that

$$P(class|m) = \frac{f(m|class)P(class)}{f(m)} \tag{5.24}$$

and saved for use in the following fit stage. Figure 5.6 shows the mass fit projection for all of the data. The choice of PDFs shows good agreement with the data with a fit χ^2/ndf = 1.26. The total yield across the accepted region in D^0 mass is 172605 ± 472. The signal and background fractions within the mass signal window for this fit are 97.42 ± 0.26 % and 2.58 ± 0.04 %, respectively. This corresponds to a yield of 168151 ± 460 signal events containing a real D^0.

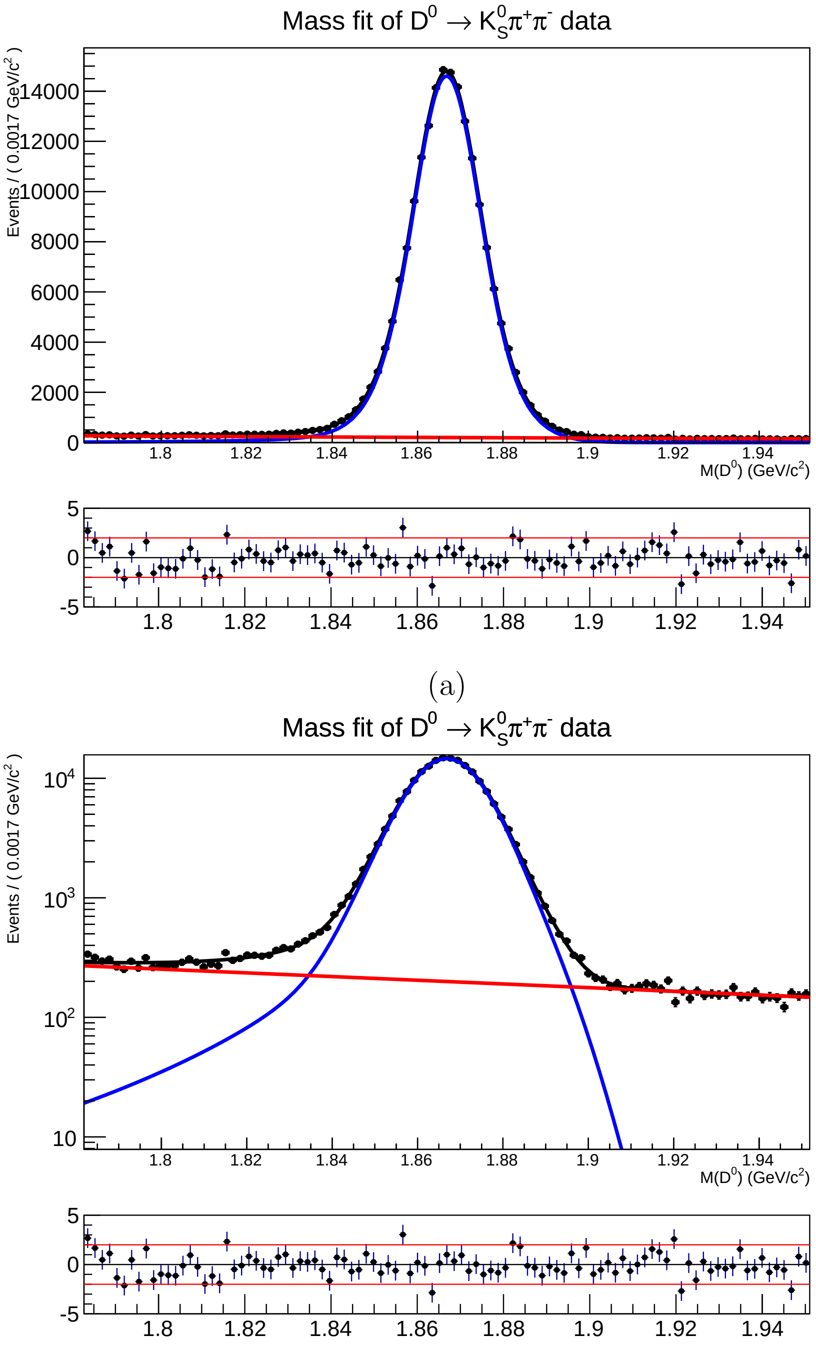


(a)



(b)

Figure 5.5: (a) TP_1 distribution in data (blue points), showing the fit components for prompt signal (green), secondary signal (blue) and combinatoric (red) events, and their sum (black). (b) Equivalent plot for TP_{diff}.



(a)

(b)

Figure 5.6: (a) Mass fit projection of all data passing the selection. Both real D^0 decays (blue) and combinatoric events (red) are displayed. (b) With $\log(y)$ axis.

5.4.2 Decay Time vs. $ln(\chi^2_{IP})$ Fit

The second stage fit is the time-dependent part and is a 2-dimensional fit to the D^0 decay time, t, and $ln(\chi^2_{IP})$. It is used to extract the fractions of each subclass as well as fit for the mean D^0 lifetime, τ. The fit is performed on data that passes the full selection and that sits within the D^0 mass and Δm signal windows. For each component in the fit, one is fitting the decay time and $ln(\chi^2_{IP})$ of the candidates given they are accepted by the trigger, stripping and selection. A conditional PDF is then constructed to describe the distributions in t and $ln(\chi^2_{IP})$ and any efficiency corrections applied to the PDFs that parameterise the decay time distribution for each component. Therefore, the total PDF that describes the 2D distribution in t and $ln(\chi^2_{IP})$ given a decay time acceptance, A, and a reconstructed D^0 mass, m, can be written as

$$f(ln(\chi^2_{IP}), t|A, m) = \sum_{class} f(ln(\chi^2_{IP}), t|A, class)P(class|m). \tag{5.25}$$

where the classes run over the signal and combinatoric components. The value of $P(class|m)$ is calculated from the previous fit using Equation 5.24. In the case of the signal class, the PDF on the right hand side can be separated into subclasses such that

$$f(ln(\chi^2_{IP}), t|A, sig.) = \sum_{subcl.} f(ln(\chi^2_{IP}), t|A, sig, subcl.)P(subcl.|sig.) \tag{5.26}$$

where the subclasses run over the prompt and secondary signal components and $P(subcl.|sig.)$ is the fraction of signal events that belong to each of the prompt and secondary subclasses and is allowed to float in the fit. The total $ln(\chi^2_{IP})$ vs. t PDF given some acceptance can then be expressed as

$$\begin{aligned} f(ln(\chi^2_{IP}), t|A, m) = &\sum_{prompt,sec.} f(ln(\chi^2_{IP}), t|A, sig., subcl.)P(subcl.)P(sig.|m) \\ &+ f(ln(\chi^2_{IP}), t|A, comb.)\, P(comb.|m). \end{aligned} \tag{5.27}$$

Including the probability density for a given class having a particular acceptance, $f(A|class) = f(TP_1|class)f(TP_{diff}|class)$, the full PDF used in the fit becomes

$$\begin{aligned} f(ln(\chi^2_{IP}), t|m) = \sum_{\substack{prompt,\\ sec.}} & f(ln(\chi^2_{IP}), t|A, sig., subcl.)f(TP_1|sig., subcl.) \\ & \cdot f(TP_{diff}|sig., subcl.)P(subcl.)P(sig.|m) \\ & + f(ln(\chi^2_{IP}), t|A, comb.)f(TP_1|comb.)f(TP_{diff}|comb.)P(comb.|m). \end{aligned} \tag{5.28}$$

The following sections describe the individual treatment of each component in the fit. Note that the abuse of notation $f(\mathbf{X}|A) = f(\mathbf{X}|TP_1, TP_{diff})$ is used throughout what follows.

Prompt Signal

The total prompt signal PDF can be written as

$$\begin{aligned} f(ln(\chi^2_{IP}), t|sig., prompt) = & f(ln(\chi^2_{IP})|t, A, sig., prompt)f(t|A, sig., prompt) \\ & \cdot f(TP_1|sig., prompt)f(TP_{diff}|sig., prompt). \end{aligned} \tag{5.29}$$

The decay time part follows a simple exponential with decay constant τ equal to the mean D^0 lifetime. As this component is sensitive to the physics observable it must have both the per-event decay time acceptance correction and the tracking efficiency correction applied. The decay time resolution is sufficiently small with respect to the total D^0 lifetime that it is ignored throughout. The decay time part of the PDF can then be written as

$$f(t|A, sig., prompt) = \frac{\frac{1}{\tau}e^{-t/\tau}\epsilon(t)}{\int_{t_{min,i}}^{t_{max,i}} \frac{1}{\tau}e^{-t'/\tau}\epsilon(t')\mathrm{d}t'}. \tag{5.30}$$

The $ln(\chi^2_{IP})$ distribution for prompt signal is modelled as a Gaussian plus a bifurcated Gaussian that share a common mean. Although, in the ideal case, it is not expected for the mean of the $ln(\chi^2_{IP})$ distribution for prompt signal to have any dependence on the D^0 decay time a small shift is observed in the data. The time evolution of the mean for prompt signal is modelled as a linear function. The $ln(\chi^2_{IP})$ PDF for prompt signal can then be written as

$$f(ln(\chi^2_{IP})|t, sig., prompt) = f_1 G(ln(\chi^2_{IP}); \mu(t), \sigma_1) + (1 - f_1) Bif(ln(\chi^2_{IP}); \mu(t), \sigma_L, \sigma_R), \tag{5.31}$$

where

$$\mu(t) = \mu(0) + At. \tag{5.32}$$

The turning point distributions for prompt signal are taken from data that lies within the D^0 mass and Δm signal window and are modelled as described in Section 5.2.4.

Secondary Signal

The total PDF for secondary signal is constructed in a similar way as for the prompt component as in Equation 5.29. The true decay time distribution is described using the convolution of two exponentials such that,

$$f(t|A, sig., sec.) = \frac{\frac{1}{\tau_2-\tau_1}(e^{-t/\tau_2} - e^{-t/\tau_1})}{\int_{t_{min,i}}^{t_{max,i}} \frac{1}{\tau_2-\tau_1}(e^{-t'/\tau_2} - e^{-t'/\tau_1})\mathrm{d}t'}. \tag{5.33}$$

The two-exponential model (beauty then charm) is reasonable given the similar lifetimes of the dominant b-hadrons ($B^{\pm}$, B^0, B^0_s, and Λ^0_b). However, the reconstructed decay time for the B meson component will not be correct as there is missing momentum from the decay products that are not reconstructed. This will have the effect of broadening the B decay time distribution and giving an incorrect value for the mean B lifetime, but should not affect the overall shape of the distribution rendering this a safe approximation. This will not affect the D^0 decay time as all daughters are properly reconstructed. Both of the decay constants τ_1 and τ_2 are allowed to float in the fit.

The $ln(\chi^2_{IP})$ distribution for secondaries is described using a single Gaussian plus a bifurcated Gaussian. The width parameters are shared with the prompt component, but scaled up by a factor α (which is also allowed to float in the fit) to account for a broadening of the distribution due to the missing momenta. The secondary $ln(\chi^2_{IP})$ PDF is then given by

$$f(ln(\chi^2_{IP})|t, sig., sec) = f_1 G(ln(\chi^2_{IP}); \mu_{sec}(t), \alpha{\cdot}\sigma_1) + (1-f_1) Bif(ln(\chi^2_{IP}); \mu_{sec}(t), \alpha{\cdot}\sigma_L(t), \alpha{\cdot}\sigma_R). \tag{5.34}$$

Both the mean, $\mu_{sec}(t)$, and the left hand sigma of the bifurcated Gaussian, $\sigma_L(t)$, evolve with time. The mean time dependence is given by

$$\mu(t)_{sec} = \mu_{sec} + B(1 - e^{Ct}) \tag{5.35}$$

with the sigma having a linear time-dependence such that

$$\sigma_L(t) = \sigma_L + Dt \tag{5.36}$$

where all parameters are allowed to float in the fit. The turning point distributions are treated in exactly the same way as for the prompt signal component.

Combinatoric

The distributions in decay time and $ln(\chi^2_{IP})$ for the combinatoric component will not necessarily conform to a slowly varying distribution that can easily be described using some empirical formula. It is also true that a parametric description is not required as this background has no sensitivity to the physics observable. Therefore, a 2D Gaussian kernel function as described in Section 5.1.2 is used to model this component in both of these variables. The shapes of the distributions are taken from the D^0 mass sidebands, kernelised and fixed in the fit. The model produced by the kernel PDF and the pulls with respect to the data are shown in Figures 5.7 (a) and (b), respectively. To enable the inclusion of the per-event decay time acceptance the evaluation of the 2D PDF is split into two following a similar procedure as for the parametric components such that

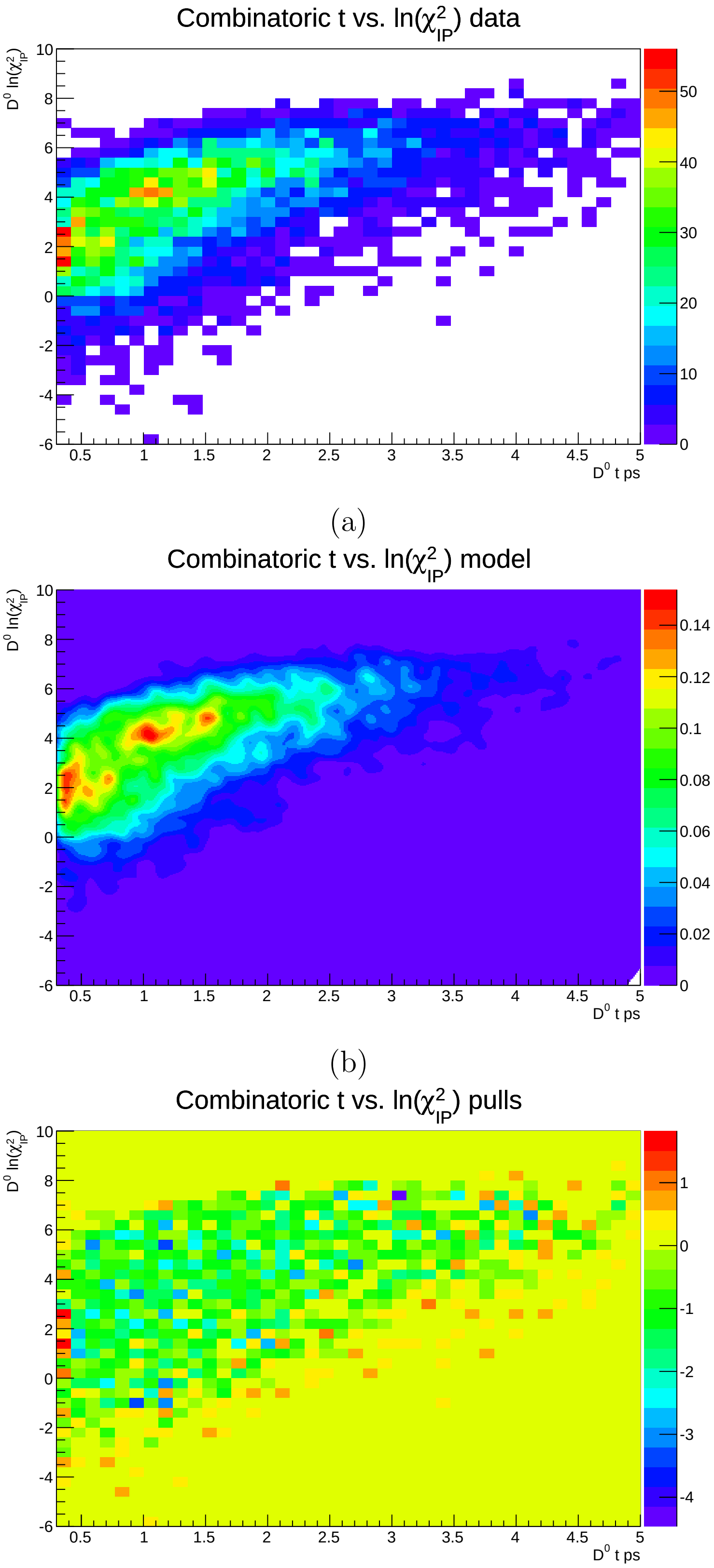
$$f(ln(\chi^2_{IP}), t|A, comb.) = f(ln(\chi^2_{IP})|t, A, comb.)\, \frac{f_{unbiased}(t|A, comb.)}{\int_{t_{min,i}}^{t_{max,i}} f_{unbiased}(t'|A, comb.)\mathrm{d}t'} \tag{5.37}$$

where $f_{unbiased}(t|A, comb.)$ is the unbiased kernelised decay time distribution as determined using the method described in Section 5.2.3 and $f(ln(\chi^2_{IP})|t, A, comb.)$ is the projection of $ln(\chi^2_{IP})$ at time t normalised to 1 across the $ln(\chi^2_{IP})$ fit range. The distributions for TP_1 and TP_{diff} are also taken from the mass sidebands and fixed in the fit. Modelling of both turning point distributions is performed in the same way as for all other components. The total PDF for the combinatoric component can then be written as

$$f(ln(\chi^2_{IP}), t, TP_1, TP_{diff}|comb.) = f(ln(\chi^2_{IP})|t, A, comb.)\,\frac{f_{unbiased}(t|A, comb.)}{\int_{t_{min,i}}^{t_{max,i}} f_{unbiased}(t'|A, comb.)\mathrm{d}t'} \cdot f(TP_1|comb.)f(TP_{diff}|comb.). \tag{5.38}$$

5.4.3 Fit Validation

Although the lifetime fit is not the main focus of this analysis, the fact that it is used to verify that the generic fit procedure and, in particular, that the decay time acceptance correction works correctly, requires any potential biases from the fit procedure to be understood. The fit validation is performed by running many fits to toy Monte Carlo data that is generated such that it resembles the real data as closely as possible. The generation procedure for the prompt and secondary signal components is essentially the same. A pair of turning points is generated from kernelised distributions of TP_1 and TP_{diff} which are taken from real data that sit within the D^0 mass signal window. A decay time is then generated according to the unbiased distribution and the process repeated until the generated decay time sits within the accepted region. Note that the tracking efficiency is not included in the generation and fit to simplify the procedure. Once the decay time and turning points are generated a value of $ln(\chi^2_{IP})$ is generated and the event is saved. The process is the same again for combinatoric events, but the decay time is generated using the recovered unbiased distribution that is determined using the procedure outlined in Section 5.2.3 and the turning points are generated from kernelised distributions taken from data that sit in the D^0 mass sidebands. The contribution from each component is also representative of the



(a)

(b)

(c)

Figure 5.7: (a) Distribution from sideband data. (b) 2D kernel PDF model. (c) Pulls between the sideband data and kernel PDF model.

data with $\sim$ 89% prompt signal, $\sim$ 8% secondary signal and $\sim$ 3% combinatoric. A total of 500 generation and fit steps are performed with each toy dataset containing 180k events. The generated mean D^0 lifetime is 410 fs and 495/500 fits converge successfully. A Gaussian is fitted to the resulting distribution of fitted values for τ and the pull distribution resulting in $\tau = (409.92 \pm 0.06)$ fs, $\mu_{pull} = (0.016 \pm 0.049)$ and $\sigma_{pull} = (1.03 \pm 0.04)$ indicating that the fit is stable and does not have any significant bias. The fitted distributions of τ and the pulls are given in Figures 5.8 (a) and (b), respectively.

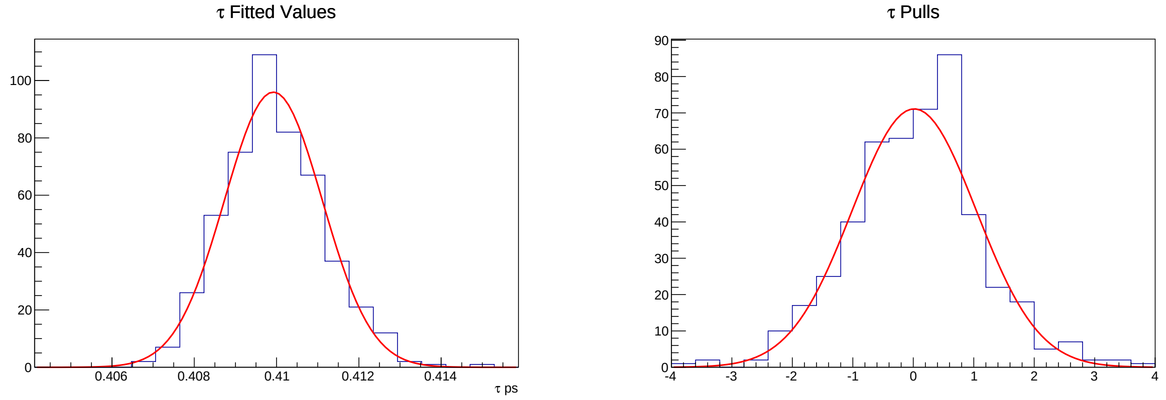


Figure 5.8: (a) Distribution of fitted values of τ from 495 successful toys. (b) Resulting pull distribution.

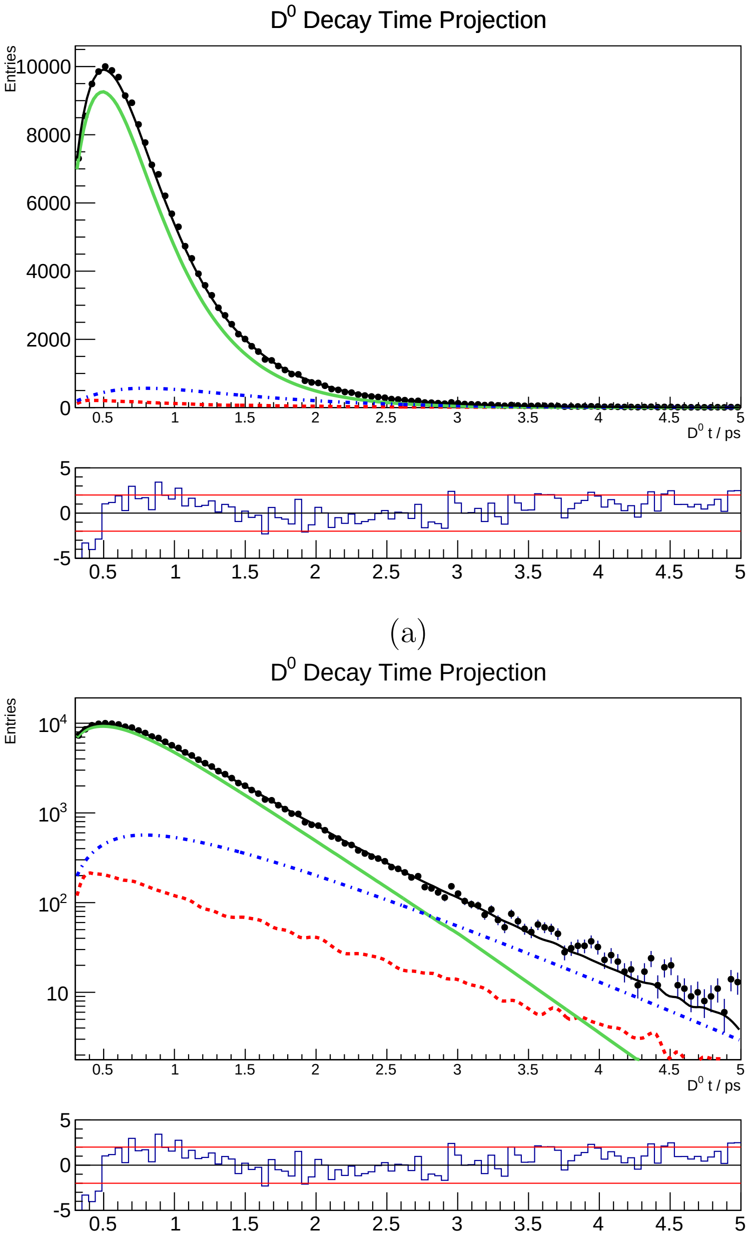
5.4.4 Fit Results

A selection of the fitted parameters from the second stage fit are listed in Table 5.2. The D^0 lifetime is measured to be $\tau = (410.87 \pm 1.14)$ fs which is well within 1σ of the world average of (410.1 ± 1.5 fs) [23]. The 1D decay time and $ln(\chi^2_{IP})$ projections for data within the D^0 mass and Δm signal windows are given in Figures 5.9 and 5.10. The reader should note that the time binning chosen for the $ln(\chi^2_{IP})$ projections is not used in the fit, but merely for display purposes.

The pulls for the decay time projection are not totally satisfactory with the overall χ^2/ndf = 3.17. However, similar behaviour is observed in the projections from the toy studies performed in Section 5.4.3 which indicates that, although the pulls are undesirable, the fit converges to the correct answer. There are also regions in $ln(\chi^2_{IP})$ close to the intersection between the prompt and secondary components in the range $0.45 < t < 0.75$ ps where the fit struggles to describe the data. Although this behaviour is not observed in the toys, the crucial point is that the relative contribution from the prompt and secondary components is correctly estimated and that the parameterisation taken from this fit adequately describes the $ln(\chi^2_{IP})$ distributions observed in the individual Dalitz plot bins for the mixing fit. These assumptions are explored as a systematic uncertainty in Section 6.11.

Parameter	Fitted Value
τ	410.87 ± 1.14 fs
μ	1.203 ± 0.011
σ_1	0.916 ± 0.007
σ_L	1.986 ± 0.010
σ_R	0.770 ± 0.014
f_1	0.457 ± 0.010
α	1.092 ± 0.016
μ_{sec}	1.803 ± 0.092
f_{sec}	10.106 ± 0.151 %
τ_1	606.97 ± 32.21 fs
τ_2	607.63 ± 32.19 fs
A	0.042 ± 0.008
B	5.500 ± 0.076
C	-0.654 ± 0.031
D	-0.092 ± 0.028

Table 5.2: Table of fitted parameters from the second stage of the lifetime fit.



(a)

(b)

Figure 5.9: (a) Decay time projection from the second stage lifetime fit with prompt signal (green), secondary signal (blue) and combinatoric (red) components identified. (b) The same plot, but with a log(y) axis.

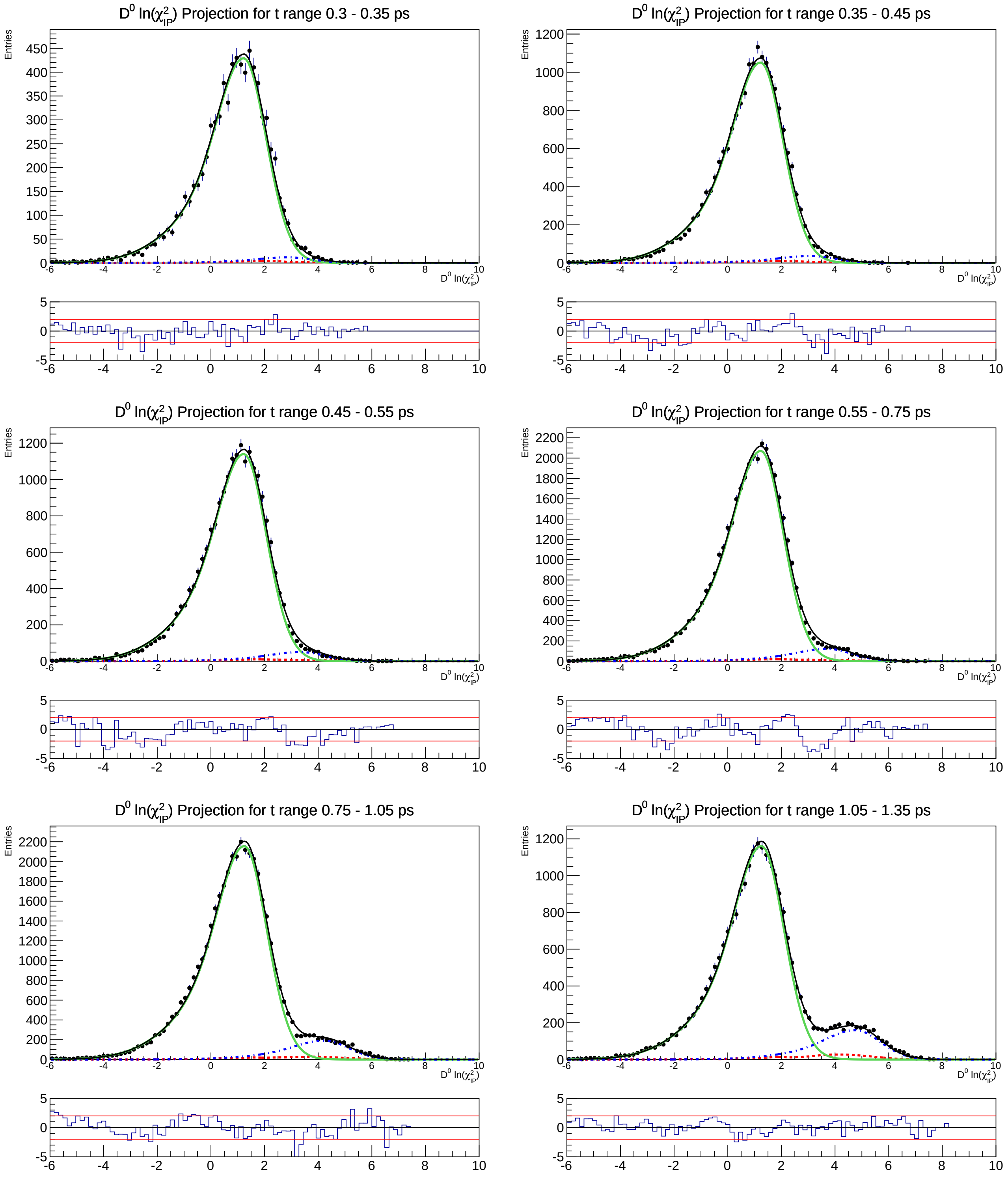


Figure 5.10: $ln(\chi^2_{IP})$ projections from the second stage lifetime fit in bins of decay time with prompt signal (green/solid line), secondary signal (blue/dot-dash line) and combinatoric (red/dash line) components identified.

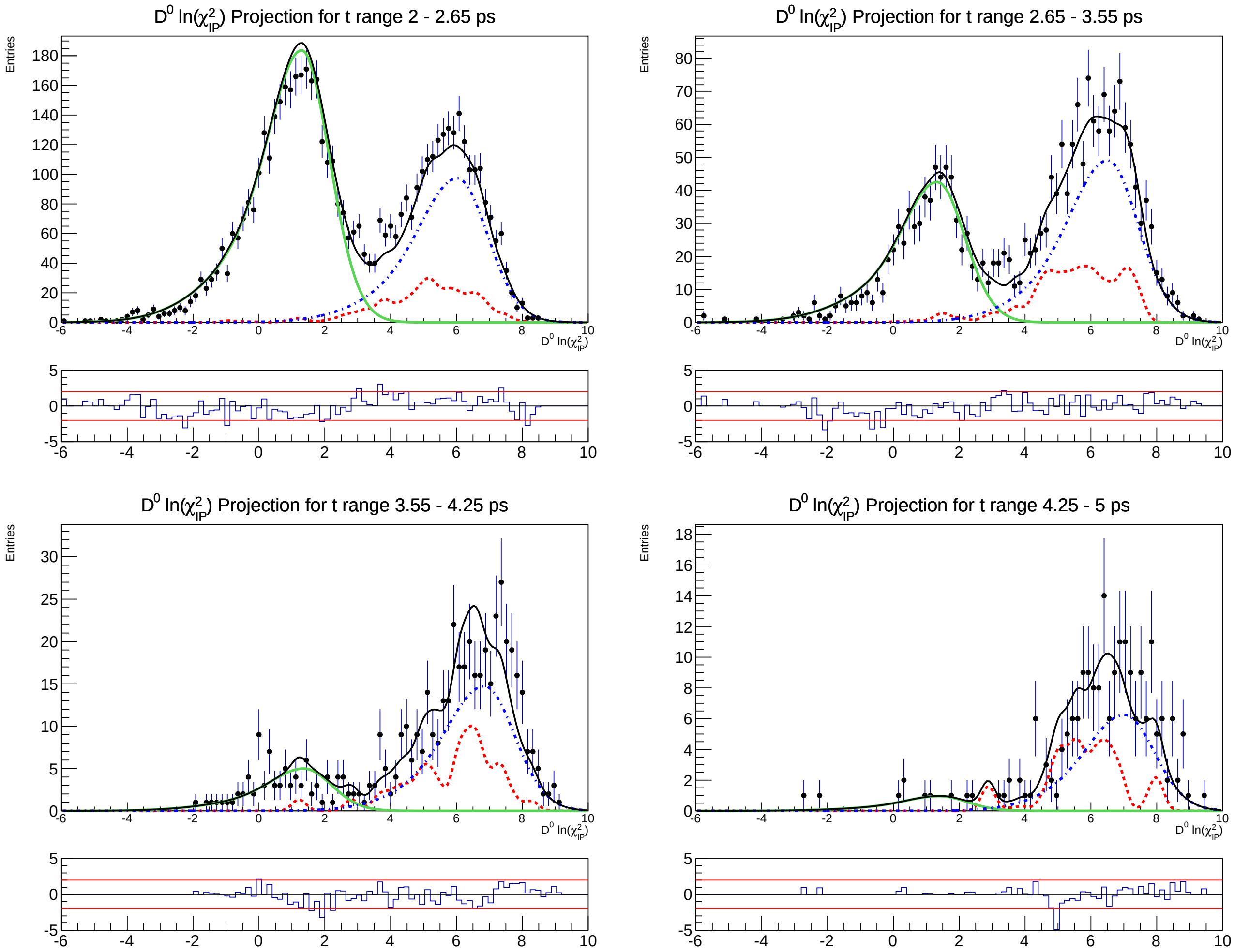


Figure 5.11: $ln(\chi^2_{IP})$ projections from the second stage lifetime fit in bins of decay time with prompt signal (green/solid line), secondary signal (blue/dot-dash line) and combinatoric (red/dash line) components identified.

5.5 Mixing Fit

The mixing fit is used to extract the mixing parameters and is split into two stages. The first stage is a 2D fit to the D^0 mass vs. Δm distribution in each of the individual Dalitz plot bins for D^0 and $\overline{D}^0$ data separately. This provides information about the relative contribution from the three classes of event considered in each Dalitz plot bin; correctly tagged real D^0 ($right\text{-}\pi^+_{slow}$), incorrectly tagged real D^0 ($wrong\text{-}\pi^+_{slow}$) and combinatoric background (*comb.*). The second stage is then a simultaneous fit to both the D^0 and $\overline{D}^0$ samples which is in turn a simultaneous fit to the t vs. $ln(\chi^2_{IP})$ distribution across each of the Dalitz plot bins which is used to extract the mixing parameters. Many of the components in the mixing fit use the same parameterisations in one or more kinematic variables as those obtained in the lifetime fit. Due to the lower statistics in the individual Dalitz plot bins the fit would become unstable if many parameters were floated in each bin. Therefore, a selection of parameters are fixed from the lifetime fit to improve the stability and a systematic uncertainty is assigned. The following sections describe each stage in further detail.

5.5.1 D^0 Mass vs. Δm Fit

The mass vs. Δm fit is performed to extract the relative contributions from each class of event in each Dalitz plot bin for D^0 and $\overline{D}^0$ separately. The total PDF is constructed in a similar way as for the first stage of the lifetime fit such that

$$f(m, \Delta m) = \sum_{class} f(m, \Delta m | class) P(class) \tag{5.39}$$

The right-π^+_{slow} component will have a peaking structure in both D^0 mass and Δm. The D^0 mass distribution for this component is described using the sum of two Gaussians and a Crystal Ball with shared mean. The Δm distribution is described using the sum of two Gaussians and a bifurcated Gaussian with shared mean. The PDF for this component can then be written as

$$
\begin{aligned}
f(m, \Delta m | right\text{-}\pi^+_{slow}) = {} & [f_1 G(m; m_{D^0}, \sigma_1) + f_2 G(m; m_{D^0}, \sigma_2) \\
& +(1 - f_1 - f_2) CB(m; m_{D^0}, \sigma_3, \alpha, n)] \times [f_3 G(\Delta m; \mu_{\Delta m}, \sigma_4) \\
& +f_4 G(\Delta m; \mu_{\Delta m}, \sigma_5) + (1 - f_3 - f_4) Bif(\Delta m; \mu_{\Delta m}, \sigma_L, \sigma_R)]
\end{aligned}
\tag{5.40}
$$

where the power of the Crystal Ball tail is fixed to $n = 3$. The wrong-π^+_{slow} component describes real D^0 that have been combined with a random π^+_{slow}. Therefore, this component is expected to peak in D^0 mass, but be smooth in Δm. The D^0 mass distribution for this component is shared with that for the right-π^+_{slow} whereas the Δm distribution is described using a 2$^{\text{nd}}$ order Chebyshev polynomial. The PDF for this component can then be written as

$$
\begin{aligned}
f(m, \Delta m | wrong\text{-}\pi^+_{slow}) = {} & [f_1 G(m; m_{D^0}, \sigma_1) + f_2 G(m; m_{D^0}, \sigma_2) + \\
& (1 - f_1 - f_2) CB(m; m_{D^0}, \sigma_3, \alpha, n)] \times (T_m(\Delta m))
\end{aligned}
\tag{5.41}
$$

Finally, the combinatoric component is not expected to peak in either D^0 mass or Δm. Therefore, the D^0 mass distribution is described using a single exponential and the Δm distribution shares the second order Chebyshev polynomial for the wrong-π^+_{slow} component.

As some of the Dalitz plot bins contain very few events the fit is initially performed to each of the D^0 and $\overline{D}^0$ data samples integrated across the Dalitz plot. It is expected that the signal shapes in both mass and Δm and the background shape in Δm will not vary as a function of Dalitz plot position. However, the shape of the background in D^0 mass and the relative yields will differ across each of the Dalitz plot bins. Therefore, the signal shapes and Δm background shape are fixed to that extracted from the global fits to the D^0 and $\overline{D}^0$ samples in the fits to each Dalitz plot bin with just the yields and D^0 mass background parameters floating. This improves the fit stability in bins with low statistics while still accurately describing the data. The mass vs. Δm fit results are summarised in Table 5.3 and the projections for D^0 and $\overline{D}^0$ in each bin are given in Figures 5.15–5.22.

Dalitz Bin	χ^2/ndf Mass Fit	χ^2/ndf Δm Fit	$P(right\text{-}\pi^+_{slow})$	$P(wrong\text{-}\pi^+_{slow})$	$P(comb.)$
D^0 bin -1	1.26	1.16	0.9009 ± 0.0105	0.0577 ± 0.0024	0.0414 ± 0.0007
D^0 bin -2	0.79	1.14	0.8704 ± 0.0232	0.0687 ± 0.0059	0.0609 ± 0.0017
D^0 bin -3	0.72	0.96	0.8872 ± 0.0233	0.0744 ± 0.0056	0.0384 ± 0.0014
D^0 bin -4	0.80	0.72	0.8975 ± 0.0257	0.0617 ± 0.0059	0.0409 ± 0.0016
D^0 bin -5	1.14	0.97	0.9019 ± 0.0157	0.0551 ± 0.0036	0.0430 ± 0.0010
D^0 bin -6	0.93	1.06	0.8568 ± 0.0261	0.0807 ± 0.0069	0.0630 ± 0.0020
D^0 bin -7	0.96	0.70	0.8194 ± 0.0244	0.1157 ± 0.0071	0.0649 ± 0.0019
D^0 bin -8	0.91	1.08	0.8662 ± 0.0173	0.0875 ± 0.0045	0.0463 ± 0.0011
D^0 bin 1	1.15	1.06	0.9423 ± 0.0076	0.0343 ± 0.0014	0.0234 ± 0.0004
D^0 bin 2	0.80	1.30	0.9509 ± 0.0122	0.0301 ± 0.0021	0.0189 ± 0.0005
D^0 bin 3	0.89	1.10	0.9476 ± 0.0151	0.0337 ± 0.0027	0.0187 ± 0.0007
D^0 bin 4	0.92	0.84	0.9159 ± 0.0238	0.0505 ± 0.0051	0.0336 ± 0.0014
D^0 bin 5	1.06	1.15	0.9291 ± 0.0133	0.0384 ± 0.0027	0.0323 ± 0.0007
D^0 bin 6	1.04	0.79	0.9463 ± 0.0137	0.0355 ± 0.0025	0.0181 ± 0.0006
D^0 bin 7	0.95	1.31	0.9624 ± 0.0088	0.0278 ± 0.0014	0.0098 ± 0.0003
D^0 bin 8	0.80	1.25	0.9556 ± 0.0088	0.0303 ± 0.0015	0.0142 ± 0.0003
$\overline{D}^0$ bin -1	1.01	1.07	0.9415 ± 0.0075	0.0357 ± 0.0014	0.0223 ± 0.0004
$\overline{D}^0$ bin -2	0.77	1.12	0.9487 ± 0.0122	0.0315 ± 0.0021	0.0198 ± 0.0006
$\overline{D}^0$ bin -3	0.87	0.81	0.9486 ± 0.0150	0.0333 ± 0.0027	0.0182 ± 0.0007
$\overline{D}^0$ bin -4	0.70	0.89	0.9194 ± 0.0235	0.0448 ± 0.0049	0.0358 ± 0.0014
$\overline{D}^0$ bin -5	1.05	0.83	0.9294 ± 0.0131	0.0406 ± 0.0026	0.0300 ± 0.0007
$\overline{D}^0$ bin -6	0.81	0.90	0.9462 ± 0.0136	0.0352 ± 0.0024	0.0187 ± 0.0006
$\overline{D}^0$ bin -7	0.84	0.98	0.9604 ± 0.0088	0.0303 ± 0.0014	0.0093 ± 0.0003
$\overline{D}^0$ bin -8	1.05	0.98	0.9544 ± 0.0088	0.0322 ± 0.0015	0.0134 ± 0.0003
$\overline{D}^0$ bin 1	0.82	1.37	0.9087 ± 0.0104	0.0507 ± 0.0023	0.0406 ± 0.0007
$\overline{D}^0$ bin 2	0.93	1.00	0.8780 ± 0.0225	0.0667 ± 0.0056	0.0553 ± 0.0016
$\overline{D}^0$ bin 3	0.71	1.09	0.8914 ± 0.0232	0.0682 ± 0.0055	0.0404 ± 0.0014
$\overline{D}^0$ bin 4	0.72	0.87	0.9046 ± 0.0255	0.0547 ± 0.0058	0.0407 ± 0.0016
$\overline{D}^0$ bin 5	0.87	1.03	0.9040 ± 0.0158	0.0526 ± 0.0036	0.0434 ± 0.0010
$\overline{D}^0$ bin 6	0.87	1.02	0.8512 ± 0.0264	0.0866 ± 0.0071	0.0621 ± 0.0020
$\overline{D}^0$ bin 7	0.93	1.01	0.8224 ± 0.0238	0.1224 ± 0.0068	0.0552 ± 0.0017
$\overline{D}^0$ bin 8	0.75	1.00	0.8826 ± 0.0170	0.0733 ± 0.0041	0.0439 ± 0.0011

Table 5.3: Summary of D^0 mass vs. Δm fits including relative contributions from each class of event.

5.5.2 Decay Time vs. $ln(\chi^2_{IP})$ Fit

The decay time vs. $ln(\chi^2_{IP})$ fit is a simultaneous fit to the D^0 and $\overline{D}^0$ samples which is in turn a simultaneous fit to each of the individual Dalitz plot bins. The total PDF for bin i can be written as

$$f(ln(\chi^2_{IP}), t|A, m, \Delta m) = \sum_{class} f(ln(\chi^2_{IP})|t, A, class) f(t|A, class)_i P(class|m, \Delta m)_i \quad (5.42)$$

where class runs over the right-π^+_{slow}, wrong-π^+_{slow} and combinatoric components and $f(t|A, class)_i$ is the decay time distribution for Dalitz plot bin, i. $P(class|m, \Delta m)_i$ is determined in each Dalitz plot bin from the mass vs. Δm fit described in the previous section. The right-π^+_{slow} and wrong-π^+_{slow} components are also split into prompt and secondary subclasses in a similar way as for the signal component in the lifetime fit such that

$$f(ln(\chi^2_{IP}), t|right\text{-}\pi^+_{slow}, A, m, \Delta m) = \sum_{\substack{prompt,\\ sec.}} f(ln(\chi^2_{IP})|t, A, right\text{-}\pi^+_{slow}, subcl.)$$
$$f(t|A, right\text{-}\pi^+_{slow}, subcl.)_i P(subcl.|right\text{-}\pi^+_{slow}) \quad (5.43)$$

for the right-π^+_{slow} component. A similar expression can be written for the wrong-π^+_{slow} component. The value of $P(prompt|right\text{-}\pi^+_{slow}) = P(prompt|wrong\text{-}\pi^+_{slow}) = 1 - f_{sec}$ and $P(sec|right\text{-}\pi^+_{slow}) = P(sec|wrong\text{-}\pi^+_{slow}) = f_{sec}$ where f_{sec} is fixed to the value extracted from the lifetime fit. The following sections describe the individual treatment of each component in the fit.

Prompt Right-π^+_{slow}

The prompt right-π^+_{slow} component describes correctly tagged real D^0 and is the most sensitive to the mixing and CP violation parameters. The decay time distribution (before decay time acceptance and Dalitz plot efficiency effects) for D^0 in Dalitz plot bin i is given by,

$$\begin{aligned} f_{D^0}(t)_i = e^{-\Gamma t}\Big(T_i + t\sqrt{T_i T_{-i}} r_{CP}\Gamma\,\big[\,y_D(c_i\cos(\alpha_{CP}) + s_i\sin(\alpha_{CP})) \\ +x_D(s_i\cos(\alpha_{CP}) - c_i\sin(\alpha_{CP}))\big]\Big)\,. \end{aligned} \tag{5.44}$$

Factoring out T_i and defining the mixing term as $M_{D^0} = y_D(c_i\cos(\alpha_{CP}) + s_i\sin(\alpha_{CP})) + x_D(s_i\cos(\alpha_{CP}) - c_i\sin(\alpha_{CP}))$ this can be re-written as

$$f_{D^0}(t)_i = e^{-\Gamma t}T_i\left(1 + t\sqrt{T_{-i}/T_i}\, r_{CP}\Gamma M_{D^0}\right). \tag{5.45}$$

Recall that T_i is defined as

$$T_i = \int_i a^2_{12,13}\mathrm{d}m^2_{12}\mathrm{d}m^2_{13}. \tag{5.46}$$

If there is some non-uniform Dalitz plot efficiency, $\epsilon(m^2_{12}, m^2_{13})$ then T_i becomes

$$T'_i = \int_i a^2_{12,13}\epsilon(m^2_{12}, m^2_{13})\mathrm{d}m^2_{12}\mathrm{d}m^2_{13}. \tag{5.47}$$

Assuming that the Dalitz plot efficiency is symmetric about the axis $m^2_{12} = m^2_{13}$ such that $\epsilon(m^2_{12}, m^2_{13}) = \epsilon(m^2_{13}, m^2_{12})$ then the equivalent expression for T_{-i} can be written as,

$$T'_{-i} = \int_i a^2_{13,12}\epsilon(m^2_{12}, m^2_{13})\mathrm{d}m^2_{12}\mathrm{d}m^2_{13}. \tag{5.48}$$

Plugging this into equation 5.45 gives

$$f_{D^0}(t)_i = e^{-\Gamma t}T'_i\left(1 + t\sqrt{T'_{-i}/T'_i}\, r_{CP}\Gamma M_{D^0}\right). \tag{5.49}$$

Based on the assumption that the Dalitz plot efficiency is indeed symmetric then one can assume that the ratio $T_i/T_{-i} = T'_i/T'_{-i}$. This assumption is checked and a systematic uncertainty assigned in Section 6.5. It is also true that T'_i is equivalent to the fraction of the prompt right-π^+_{slow} yield that lies in bin i. The total number of right-π^+_{slow} events is determined from the fit to the mass vs. Δm distribution for D^0 events and is equivalent to $P(right\text{-}\pi^+_{slow})_i$. The fraction of these events which are prompt is fixed to that determined from the lifetime fit such that $P(prompt|right\text{-}\pi^+_{slow}) = P(prompt|sig.)$ giving $T'_i = P(prompt|right\text{-}\pi^+_{slow})P(right\text{-}\pi^+_{slow})_i$. Applying these assumptions and inserting the expression for T'_i into Equation 5.49 gives the decay time component for D^0, prompt right-π^+_{slow} events (before decay time acceptance) such that

$$f_{D^0}(t)_i = P(prompt|right\text{-}\pi^+_{slow})P(right\text{-}\pi^+_{slow})_i e^{-\Gamma t}\left(1 + t\sqrt{T_{-i}/T_i}r_{CP}\Gamma M_{D^0}\right), \quad (5.50)$$

where this becomes Equation 5.44 in the case where the right-π^+_{slow} signal is the only component present in the data. The equivalent expression for $\overline{D}^0$ in bin i is given by,

$$f_{\overline{D}^0}(t)_i = P(prompt|right\text{-}\pi^+_{slow})P(right\text{-}\pi^+_{slow})_i e^{-\Gamma t}\left(1 + t\sqrt{T_i/T_{-i}}\frac{1}{r_{CP}}\Gamma M_{\overline{D}^0}\right), \quad (5.51)$$

where $M_{\overline{D}^0} = y_D(c_i\cos(\alpha_{CP}) + s_i\sin(\alpha_{CP})) - x_D(s_i\cos(\alpha_{CP}) - c_i\sin(\alpha_{CP}))$, $P(right\text{-}\pi^+_{slow})_i$ is determined from the fit to the mass vs. Δm distribution for $\overline{D}^0$ events in bin i and $P(prompt|right\text{-}\pi^+_{slow}) = 1 - f_{sec}$.

The $ln(\chi^2_{IP})$ distribution for prompt right-π^+_{slow} decays is modelled in the same way as for the prompt signal component in the lifetime fit and the parameters fixed to those listed in Table 5.2.

Secondary Right-π^+_{slow}

The secondary right-π^+_{slow} component for each Dalitz plot bin in both decay time and $ln(\chi^2_{IP})$ is treated in the same way as for the lifetime fit with the parameters fixed to those listed in Table 5.2. It is also assumed that the fraction of right-π^+_{slow} events that emanate from secondary decays is the same in each Dalitz plot bin. A systematic uncertainty associated with these assumptions is investigated in Section 6.10.5.

Prompt Wrong-π^+_{slow}

The prompt wrong-π^+_{slow} component in the mixing fit is assumed to be an admixture of the D^0 and $\overline{D}^0$ amplitudes. The mass vs. Δm fit gives the total contribution from wrong-π^+_{slow} events in each Dalitz plot bin, $P(wrong\text{-}\pi^+_{slow})_i$. Assuming that both the D^0 and $\overline{D}^0$ components follow the same Dalitz plot distribution, but mirrored about the axis $m^2_{12} = m^2_{13}$, then in bin i one would expect the fraction of D^0 events to be T_i and of $\overline{D}^0$ to be T_{-i}. Equivalently, for bin $-i$ one would expect fractions of T_i for $\overline{D}^0$ and T_{-i} for D^0 events. This also assumes that there is no charge asymmetry in the mis-reconstruction of pions emanating from the primary vertex such that the overall fraction of D^0 and $\overline{D}^0$ events contributing to the wrong- π^+_{slow} component is evenly split across the Dalitz plot i.e. the total fraction of D^0 events in the wrong-π^+_{slow} component integrated across the Dalitz plot is $p_{D^0} = 0.5$. The equivalent fraction for $\overline{D}^0$ is $1 - p_{D^0} = 0.5$. Therefore, the decay time component for prompt wrong-π^+_{slow} events in bin i can be written as

$$\begin{aligned} f(t)_i = & P(wrong\text{-}\pi^+_{slow})_i P(prompt|wrong\text{-}\pi^+_{slow}) \\ & \cdot \left(\frac{1}{p_{D^0}T_i + (1 - p_{D^0})T_{-i}} \right) \left[p_{D^0}T_i f_{D^0}(t)_i + (1 - p_{D^0})T_{-i} f_{\overline{D}^0}(t)_i \right] \end{aligned} \tag{5.52}$$

where $f_{D^0}(t)_i$ and $f_{\overline{D}^0}(t)_i$ are the normalised PDFs given in Equations 5.50 and 5.51, respectively, $P(wrong\text{-}\pi^+_{slow})_i$ is determined from the mass vs. Δm fit and $P(prompt|wrong\text{-}\pi^+_{slow}) = 1 - f_{sec}$. The assumption that $p_{D^0} = 0.5$ is investigated as a systematic uncertainty in Section 6.10.5. The equivalent expression for bin $-i$ is given by

$$\begin{aligned} f(t)_{-i} = & P(wrong\text{-}\pi^+_{slow})_{-i} P(prompt|wrong\text{-}\pi^+_{slow}) \\ & \cdot \left(\frac{1}{p_{D^0}T_{-i} + (1 - p_{D^0})T_i} \right) \left[p_{D^0}T_{-i} f_{D^0}(t)_{-i} + (1 - p_{D^0})T_i f_{\overline{D}^0}(t)_{-i} \right]. \end{aligned} \tag{5.53}$$

Note that the same parameterisation is used for both D^0 flavours as the distribution of this background is independent of the D^{*+} reconstruction, but does depend on the Dalitz plot bin being considered.

The $ln(\chi^2_{IP})$ parameterisation is fixed to that used to describe the prompt right-π^+_{slow}

component, is constant for all Dalitz plot bins and has parameters set to those listed in Table 5.2. The turning point distributions are also treated in the same way as for the right-π^+_{slow} signal component.

Secondary Wrong-π^+_{slow}

The secondary wrong-π^+_{slow} component is again treated in the same way as for the secondary right-π^+_{slow} component with the parameterisation fixed from the lifetime fit and the fraction of wrong-π^+_{slow} events emanating from secondaries fixed for all Dalitz plot bins.

Combinatoric

The combinatoric distributions in decay time and $ln(\chi^2_{IP})$ are modelled using Gaussian kernel PDFs in the same way as for the lifetime fit. However, the distributions in each Dalitz plot bin for each D^0 flavour are modelled individually based on data taken from the D^0 mass sidebands. The turning point distributions for the combinatoric component are also taken from the D^0 mass sidebands for each Dalitz plot bin and D^0 flavour individually. Treating each bin independently allows for accurate modelling of these distributions and accounts for differences as a function of position in the Dalitz plot that would otherwise be difficult to parameterise.

CLEO Parameter Uncertainty

The c_i, s_i and T_i parameters used in the mixing fit have uncertainties due to the measurements made by CLEO. These uncertainties need to be propagated to the mixing parameters. The uncertainties from c_i and s_i are handled by allowing them to float in the fit with a penalty added to the likelihood[2]. The penalty term includes information about the uncertainties on

---

[2] Another method considered was to perform an ensemble of fits, with the values of $\{c_i, s_i\}$ sampled randomly for each fit according to the CLEO values and error matrix.

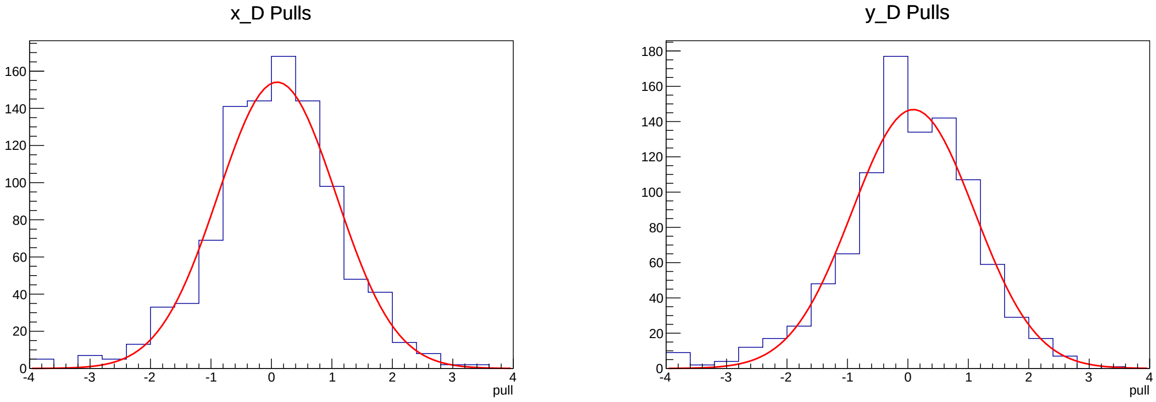
each parameter and the correlation between them. The PDF that describes this is given by

$$f(x_1, ..., x_n) = \frac{1}{(2\pi)^{n/2}\sqrt{\det(\Sigma)}}\exp(-\frac{1}{2}[x_1 - \mu_1, ..., x_n - \mu_n]\Sigma^{-1}[x_1 - \mu_1, ..., x_n - \mu_n]) \quad (5.54)$$

where Σ is the covariance matrix between parameters $\mu_1, ..., \mu_n$. In the fit, $x_1, ..., x_n$ runs over all 16 of the floating c_i and s_i parameters with each treated as a random variable and $\mu_1, ..., \mu_n$ are the central values measured by CLEO. By adding this term the uncertainties are automatically propagated to the statistical uncertainties estimated by MINUIT in the minimisation. Note that both the statistical and systematic uncertainties from the CLEO measurement are included in the same manner, but with the individual covariance matrices taken from [58]. In addition to the uncertainties from c_i and s_i, the fact that the ratio of T_i/T_{-i} is fixed in the fit to that measured by CLEO requires the uncertainties from these parameters to be included. One way to do this would be to allow these parameters to float in the fit, but with independent Gaussian constraints for each value of T_i. However, this would result in many floating parameters and may cause fit stability issues. Therefore, as these parameters are uncorrelated with each other, it was deemed appropriate to evaluate the uncertainty from these parameters as an additional systematic uncertainty which is discussed in Section 6.7.

5.5.3 Fit Validation

Any potential bias in the fit is investigated using a large number of fits to toy Monte Carlo. The toy MC is generated such that it has very similar properties to the real data. The first step is to generate a random number to decide whether the event is D^0 or $\overline{D}^0$. The data is split evenly between the two D^0 flavours. The second step is to decide which Dalitz plot bin the event is in. A weight is assigned to each Dalitz bin according to the total event yield seen in data that sit within the m and Δm signal window such that the generated event yield in each Dalitz plot bin is as close to that seen in real data as possible. Once the Dalitz bin is decided then the event is generated according to one of the components in the fit. The fractions of each class in each Dalitz plot bin listed in Table 5.3 are used as weights to decide

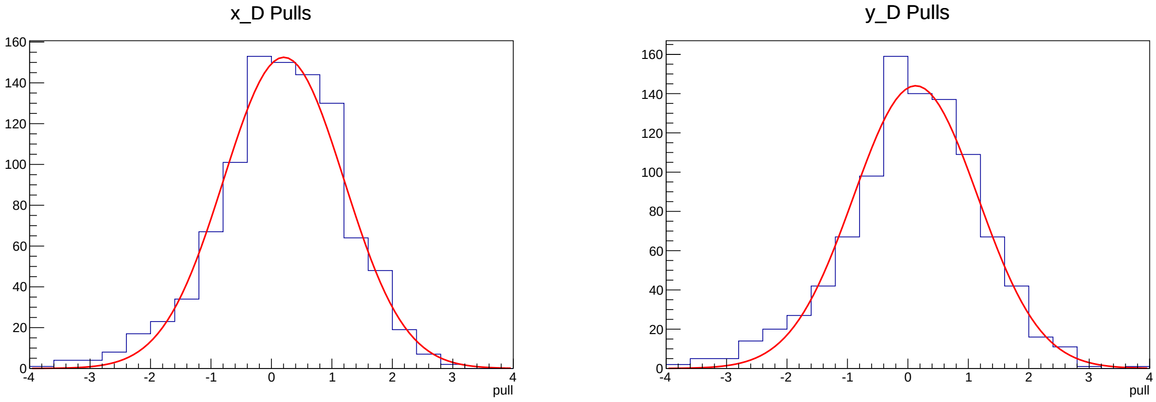


(a) (b)

Figure 5.12: (a) Pulls from 1000 generation and fit steps for x_D with c_i and s_i fixed. (b) Equivalent plot for y_D.

which class of event to generate. In the case an event is generated according to the right-π^+_{slow} or wrong-π^+_{slow} classes, then $1 - f_{sec}$ are generated as prompt and f_{sec} generated as secondary where f_{sec} is taken from Table 5.2. The event generation for each component follows a similar procedure as for the lifetime fit toys (described in Section 5.4.3). The decay time and $ln(\chi^2_{IP})$ for the signal components (prompt/sec right-π^+_{slow} and prompt/sec wrong-π^+_{slow}) are generated according to the parametric description appropriate for the class/subclass, Dalitz plot bin and D^0 flavour. Their turning points are generated using kernelised turning point distributions from data that sit within the m and Δm signal window in the appropriate Dalitz plot bin. The decay time, $ln(\chi^2_{IP})$ and turning points for combinatoric events are also generated according to kernelised distributions taken from real data, but which lie in the D^0 mass sidebands. An additional sideband sample consisting of roughly 5 % of the total generated event yield is also generated and used to model the combinatoric distributions in the fit. The number of events generated in each Dalitz plot bin follows the distribution seen in the mass sidebands. Each event is then generated using the same parameterisation as for the combinatoric component in a given Dalitz plot bin.

A total of 1000 generation and fit steps were performed with the CLEO parameters fixed. The parameters describing the decay time component for secondary right-π^+_{slow} and secondary wrong-π^+_{slow} are fixed to the generated values as are the $ln(\chi^2_{IP})$ parameters for all components. Therefore, the only floating parameters are Γ, x_D and y_D. The resulting pulls

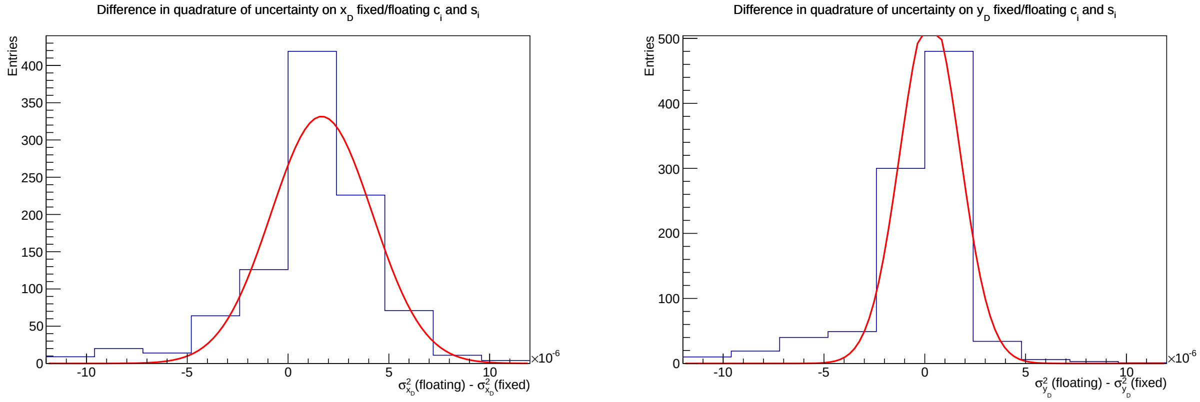


(a) (b)

Figure 5.13: (a) Pulls from 1000 generation and fit steps for x_D with c_i and s_i allowed to float with a constraint. (b) Equivalent plot for y_D.

for x_D and y_D in the 997 fits that completed successfully are given in Figures 5.12 (a) and (b), respectively. The mean and sigma of the pulls are $\mu = 0.083 \pm 0.035$ and $\sigma = 1.031 \pm 0.032$ for x_D, and $\mu = 0.083 \pm 0.036$ and $\sigma = 1.012 \pm 0.029$ for y_D. There is a bias of approximately +8 % of the statistical error for both x_D and y_D corresponding to a shift in their absolute values of $\Delta x_D \approx -0.035$ % and $\Delta y_D \approx -0.038$ %. There is also a slight bias observed in the extracted value of Γ which is not present in the lifetime fit. The fitted mean and sigma of the pull for Γ are $\mu = -0.200 \pm 0.032$ and $\sigma = 0.994 \pm 0.023$ and corresponds to a shift of $\Delta\Gamma \approx +0.0013$ MeV, approximately 20% of the statistical uncertainty. Although the bias on Γ is unsatisfactory, the bias on the mixing parameters is small enough to accommodate as a systematic uncertainty (the details of how this is evaluated are described in Section 6.1). The width of all the pull distributions are consistent with 1 indicating the statistical uncertainties are correctly estimated in the fit.

In the fit to data the values of c_i and s_i are allowed to float such that the uncertainty on these parameters from the CLEO measurements are correctly propagated to the extracted physics observables. Therefore, an additional 1000 fits are also run floating the c_i and s_i parameters using the constraint outlined in Section 5.5.2. The fits are run over the same data using the same configuration as in the previous case where c_i and s_i were fixed, with the only difference being that c_i and s_i are allowed to float. The resulting pulls for x_D and y_D in the 1000 fits that completed successfully are given in Figures 5.13 (a) and (b), respectively.



(a) (b)

Figure 5.14: (a) Difference in quadrature for fitted values of x_D when c_i and s_i are fixed and when they are allowed to float. (b) Equivalent plot for y_D.

The mean and sigma of the pulls are $\mu = 0.128 \pm 0.036$ and $\sigma = 1.064 \pm 0.032$ for x_D, and $\mu = 0.074 \pm 0.038$ and $\sigma = 1.057 \pm 0.033$ for y_D. The bias in y_D is consistent when moving from fixed to floating c_i and s_i. However, there is a slight increase in the size of the bias observed in x_D. Nonetheless, the increased bias on x_D corresponds to a systematic shift of $\Delta x_D \approx 0.056$ % which is bearable given it will likely not be a dominant systematic uncertainty. The pulls for Γ have fitted values $\mu = -0.188 \pm 0.031$ and $\sigma = 0.989 \pm 0.023$ which are consistent with the results in the case where c_i and s_i are fixed.

These toys are also used to estimate the uncertainty associated with floating the c_i and s_i parameters. A comparison between the statistical uncertainties on x_D and y_D from fits where c_i and s_i are allowed to float with those from fits where they are fixed is performed. The squared difference in quadrature (defined as $\Delta\sigma^2(1-2) = \sigma_1^2 - \sigma_2^2$) is computed for each parameter for each fit taking the uncertainty from the floating case from the fixed. The resulting distributions of $\Delta\sigma^2_{x_D}(float - fix)$ and $\Delta\sigma^2_{y_D}(float - fix)$ are each fitted with a Gaussian where the mean describes the squared uncertainty associated with floating c_i and s_i. Figures 5.14 (a) and (b) show the resulting fits for x_D and y_D, respectively. The fitted means of each distribution are $\mu = 1.667 \pm 0.089 \times 10^{-6}$ for x_D and $\mu = 0.262 \pm 0.067 \times 10^{-6}$ for y_D. Taking the square root of the sum of the mean and its error gives an estimate of the uncertainty associated with floating c_i and s_i for each parameter. These are found to be $\sigma_{x_D}(CLEO) = 0.133$ % and $\sigma_{y_D}(CLEO) = 0.057$ %. There is clearly a tail in both distributions that dips below zero indicating the uncertainty actually decreases when floating c_i and s_i for some of the fits. Although mathematically allowed, this is a little counter intuitive and will likely lead to an underestimation of the size of the uncertainty associated with floating c_i and s_i. However, the pulls from the fits where c_i and s_i are floated indicate uncertainties determined from the fit are correct. Therefore, it is assumed that the overall "experimental uncertainty" output from the fit to data where c_i and s_i are floated should correctly account for the CLEO uncertainties on c_i and s_i.

5.5.4 Fit Results

The fit to data is performed with the CP-violation parameters fixed[3] such that $\alpha_{CP} = 0$ and $r_{CP} = 1$ (i.e. no CPV) and with both mixing parameters initially blinded while systematic checks were performed. The selection, fit procedure and systematic uncertainties were all fixed before unblinding the results. The fit results after unblinding for x_D and y_D for the cases where c_i and s_i are fixed and when they are floating are summarised in Table 5.4. The nominal results that will be quoted come from the fit where c_i and s_i are allowed to float such that the uncertainties from the CLEO measurements are propagated to the statistical uncertainties extracted from the fit. The values of c_i and s_i extracted from the final fit are listed in Table 5.5 alongside the CLEO results for comparison. There is good agreement. The values of Γ extracted in both cases are consistent when moving from fixed c_i and s_i to letting them float. They are also compatible with the value of τ as measured from the lifetime fit and with the world average. Taking the value of Γ found in the fit where c_i and s_i are allowed to float gives $1/\Gamma = 1/(2.4351 \pm 0.0062) \times 10^{-3} = 410.66 \pm 1.05$ fs. The value of τ from the lifetime fit is 410.87 ± 1.14 fs. The slight bias in Γ seen in the mixing toys corresponds to an expected shift in τ extracted from the mixing fit compared to that from the lifetime fit of -0.48 fs. The shift in data is slightly smaller than this at -0.21 fs, but both are well within the statistical uncertainty on τ. The shifts between the extracted values of x_D and y_D for the fits where c_i and s_i are fixed and when they are floated are

---

[3] This is a safe assumption given the statistical precision of the dataset and the small rate of CPV observed in charm.

$\Delta x_D = -0.139\,\%$ and $\Delta y_D = +0.024\,\%$. The shift in the central value of x_D is comparable to the error associated with floating c_i and s_i as determined from the toy studies in the previous section. Finally, the correlation coefficient between x_D and y_D from the nominal mixing fit is -0.37.

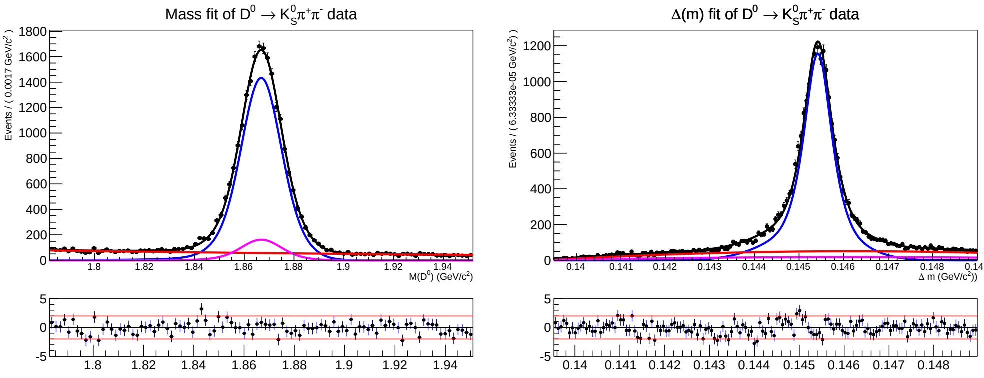
Parameter	Fixed c_i and s_i	Floating c_i and s_i	Difference
x_D (%)	-0.725 ± 0.478	-0.863 ± 0.527	-0.139
y_D (%)	-0.050 ± 0.446	-0.026 ± 0.463	$+0.024$
Γ (MeV)	2.435 ± 0.006	2.435 ± 0.006	0.000

Table 5.4: Fitted mixing and lifetime parameters from blind fits to real data.

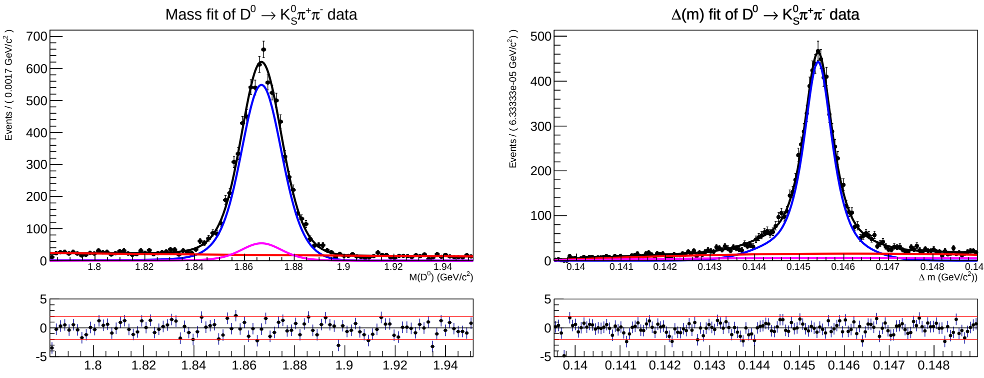
The decay time projections for each Dalitz plot bin and D^{*+} tag are given in Figures 5.23–5.30. Good agreement between the fit and the data is observed. The following chapter describes how the systematic uncertainties for x_D and y_D are determined. A detailed discussion of the results is then given in Section 6.12 and in Chapter 7.

Parameter	Mixing Fit	CLEO Results	Pull $(mix\text{ - }CLEO)/\sigma_{mix}$
c_1	0.658 ± 0.056	0.655 ± 0.036 ± 0.042	+0.055
c_2	0.516 ± 0.094	0.511 ± 0.068 ± 0.063	+0.053
c_3	0.079 ± 0.166	0.024 ± 0.140 ± 0.080	+0.331
c_4	-0.570 ± 0.159	-0.569 ± 0.118 ± 0.098	-0.006
c_5	-0.903 ± 0.063	-0.903 ± 0.045 ± 0.042	0.000
c_6	-0.612 ± 0.129	-0.616 ± 0.103 ± 0.072	+0.031
c_7	0.110 ± 0.167	0.100 ± 0.106 ± 0.124	+0.060
c_8	0.426 ± 0.104	0.422 ± 0.069 ± 0.075	+0.039
s_1	-0.068 ± 0.106	-0.025 ± 0.098 ± 0.043	-0.406
s_2	0.125 ± 0.181	0.141 ± 0.183 ± 0.066	-0.088
s_3	1.046 ± 0.142	1.111 ± 0.131 ± 0.044	-0.458
s_4	0.293 ± 0.208	0.328 ± 0.202 ± 0.072	-0.168
s_5	-0.186 ± 0.126	-0.181 ± 0.131 ± 0.026	-0.040
s_6	-0.593 ± 0.207	-0.520 ± 0.196 ± 0.059	-0.353
s_7	-1.184 ± 0.152	-1.129 ± 0.120 ± 0.096	-0.362
s_8	-0.362 ± 0.153	-0.350 ± 0.151 ± 0.045	-0.078

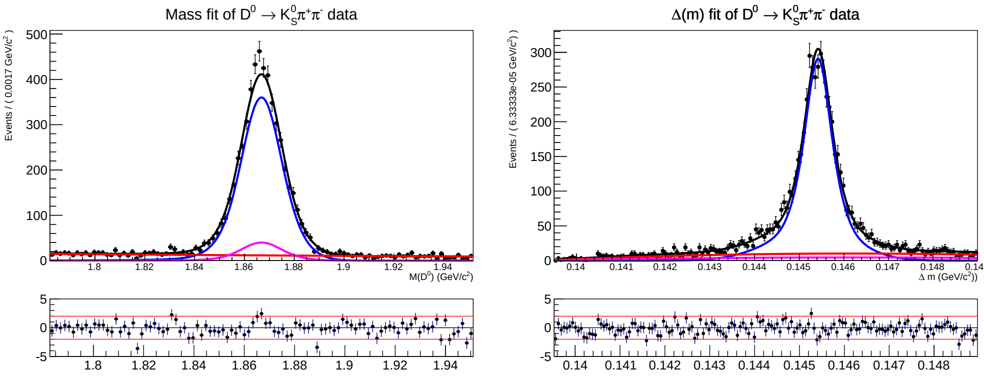
Table 5.5: Fitted central values for c_i and s_i from the mixing fit where the errors quoted are statistical only. The measured values from CLEO are also included where the first error is statistical and the second is systematic [58].



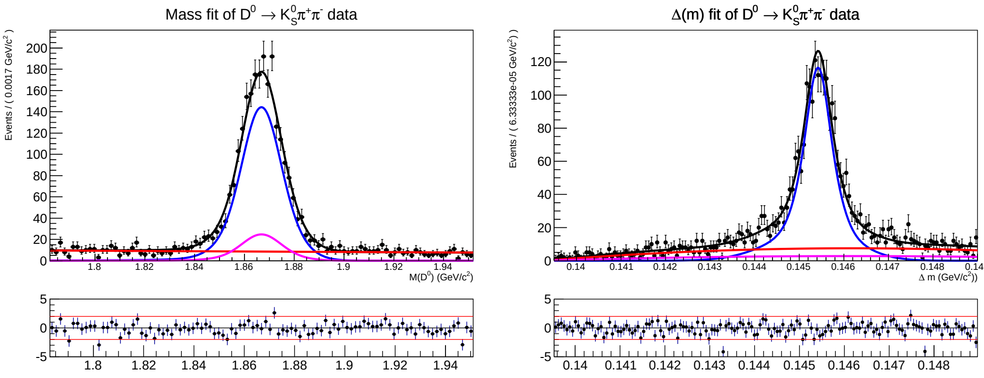
Bin 1 Bin 1



Bin 2 Bin 2

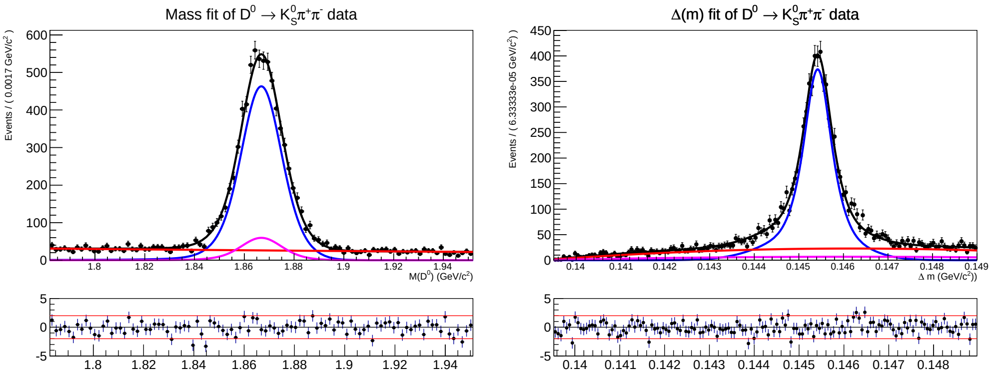


Bin 3 Bin 3

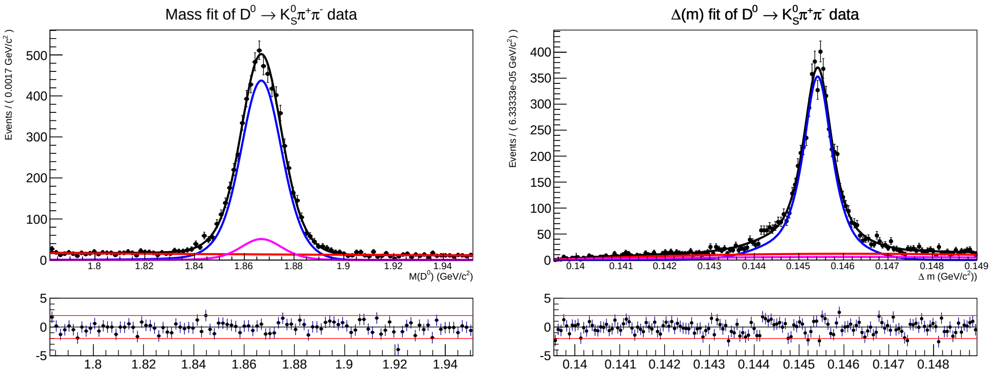


Bin 4 Bin 4

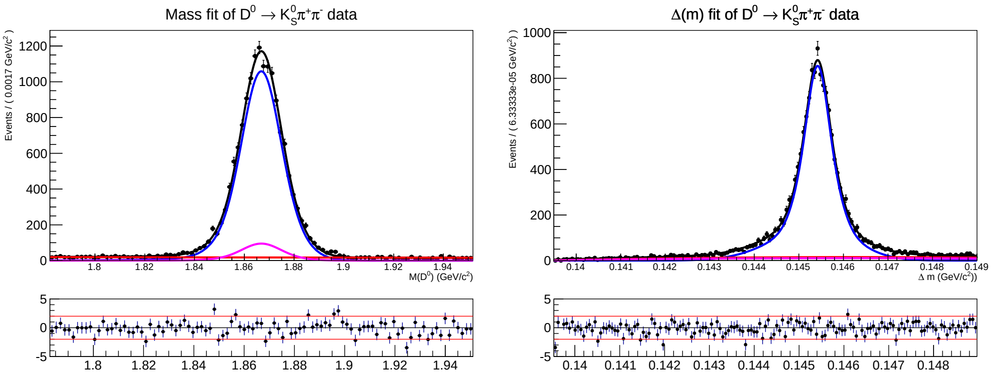
Figure 5.15: Mass and Δm fit projections for D^0 in Dalitz plot bins 1 → 4. The right-π^+_{slow} (blue), wrong-π^+_{slow} (magenta) and combinatoric (red) components are indicated.



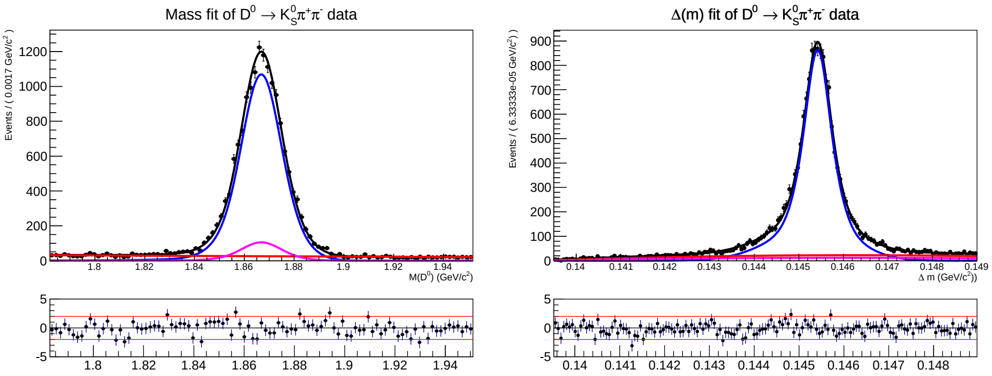
Bin 5 Bin 5



Bin 6 Bin 6

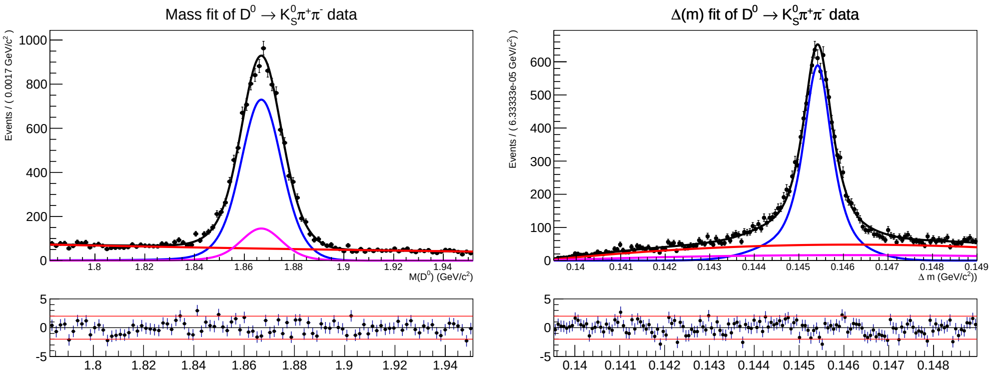


Bin 7 Bin 7

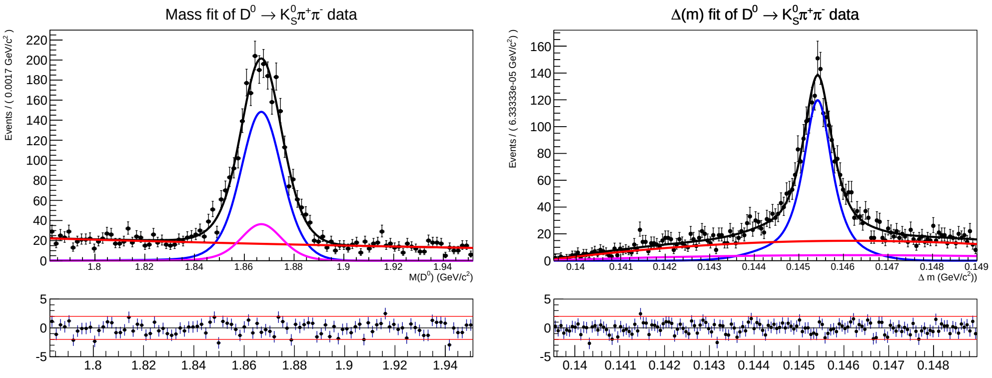


Bin 8 Bin 8

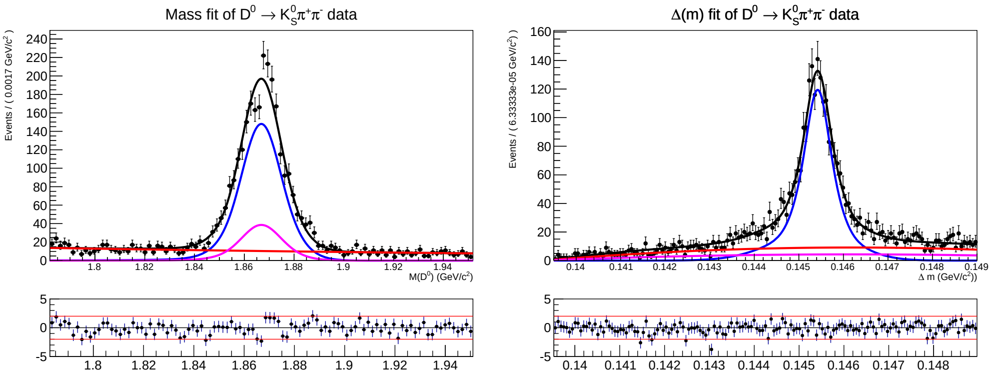
Figure 5.16: Mass and Δm fit projections for D^0 in Dalitz plot bins $5 \to 8$. The right-π^+_{slow} (blue), wrong-π^+_{slow} (magenta) and combinatoric (red) components are indicated.



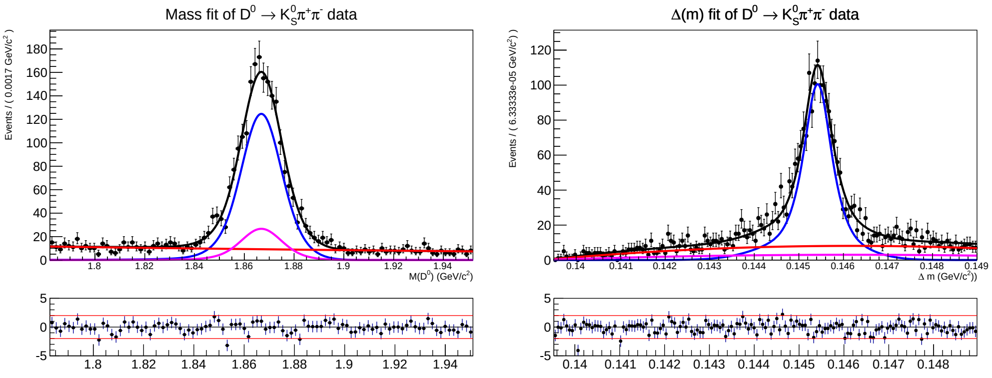
Bin -1 Bin -1



Bin -2 Bin -2

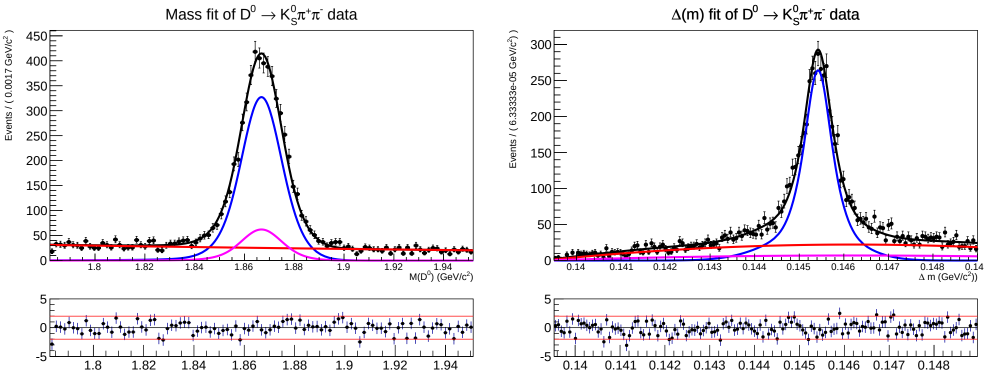


Bin -3 Bin -3



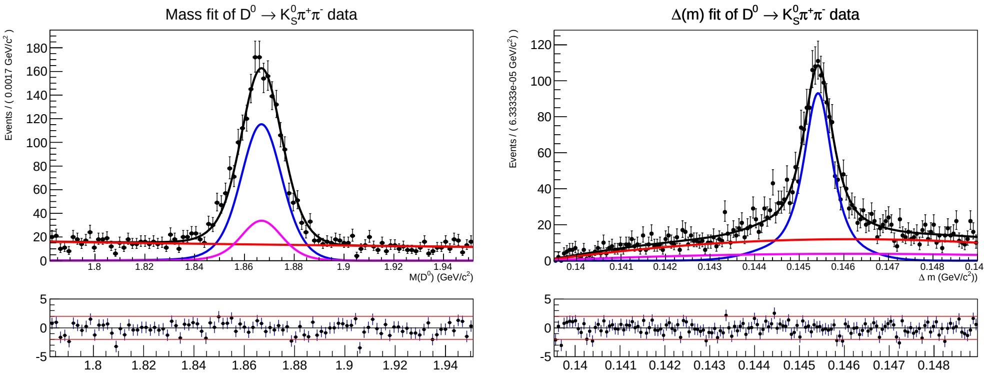
Bin -4 Bin -4

Figure 5.17: Mass and Δm fit projections for D^0 in Dalitz plot bins -1 →-4. The right-π^+_{slow} (blue), wrong-π^+_{slow} (magenta) and combinatoric (red) components are indicated.



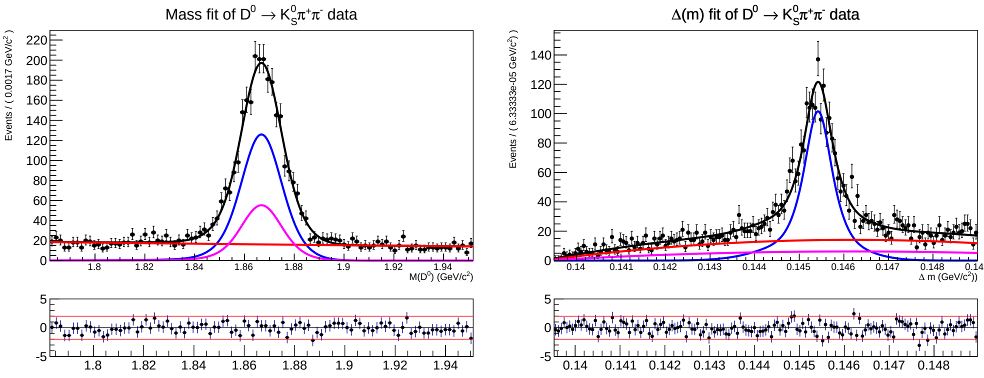
Bin -5

Bin -5



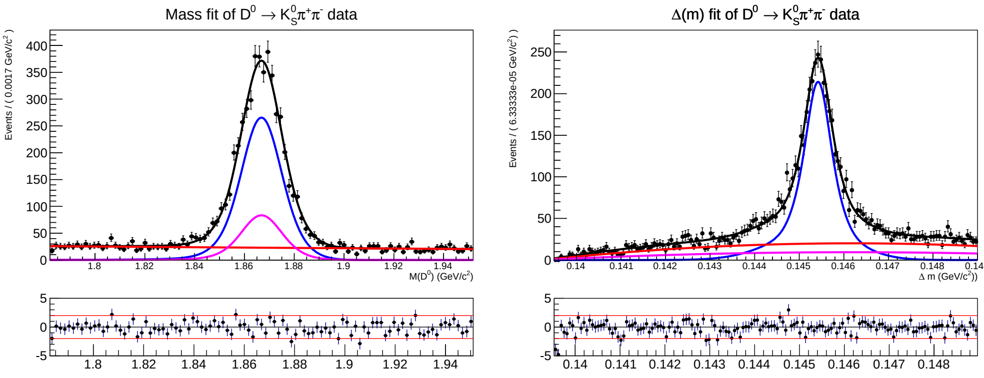
Bin -6

Bin -6



Bin -7

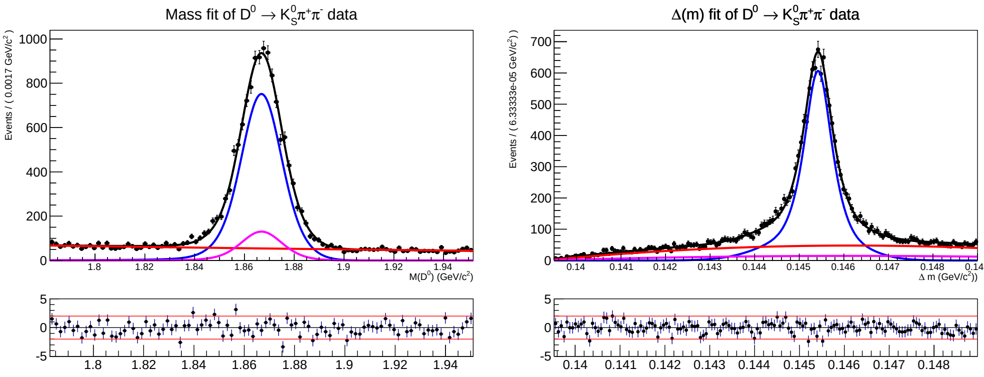
Bin -7



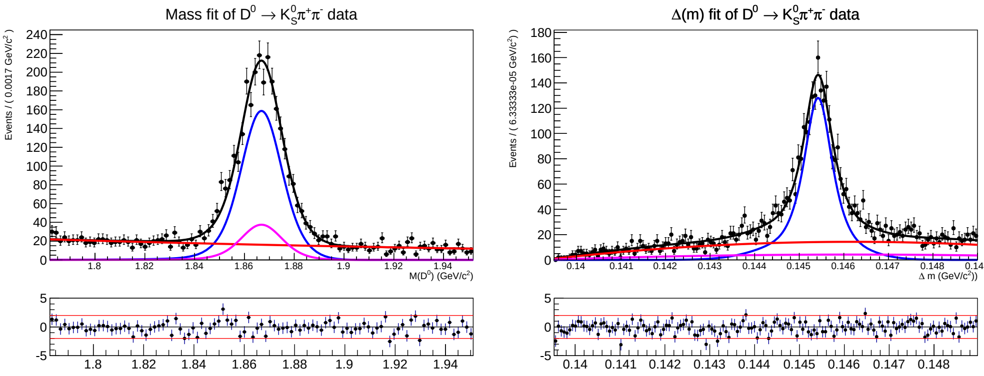
Bin -8

Bin -8

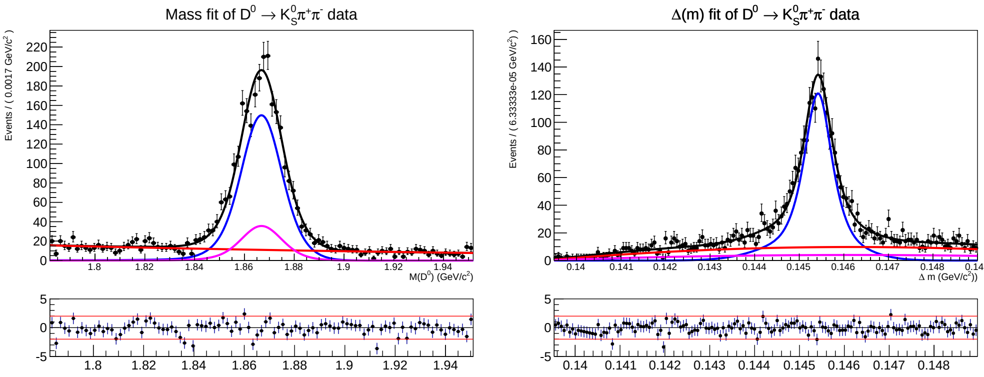
Figure 5.18: Mass and Δm fit projections for D^0 in Dalitz plot bins -5 →-8. The right-π^+_{slow} (blue), wrong-π^+_{slow} (magenta) and combinatoric (red) components are indicated.



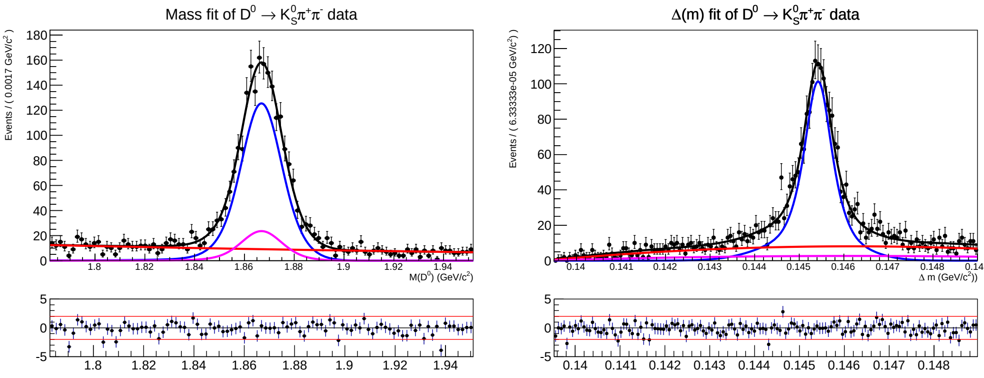
Bin 1 Bin 1



Bin 2 Bin 2

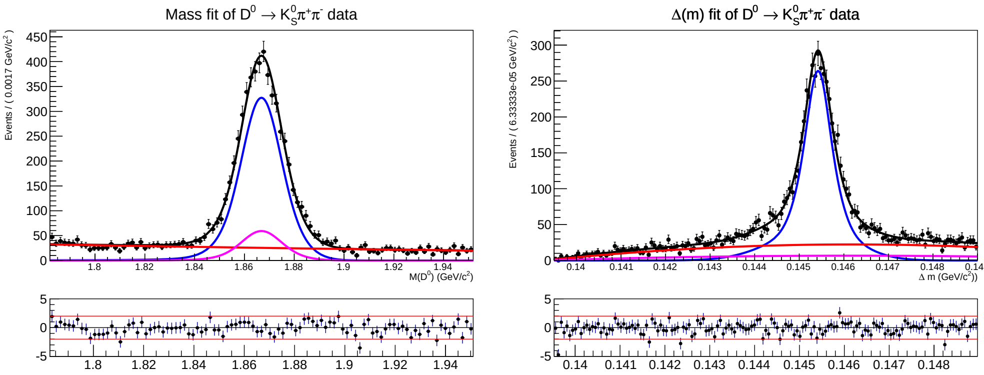


Bin 3 Bin 3

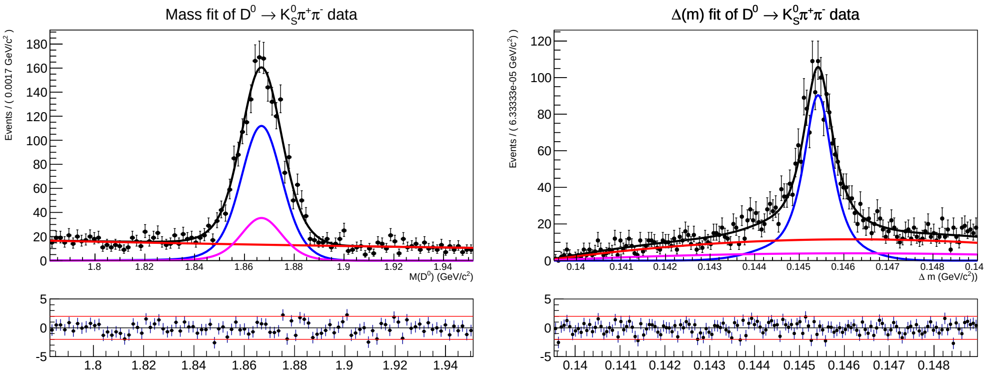


Bin 4 Bin 4

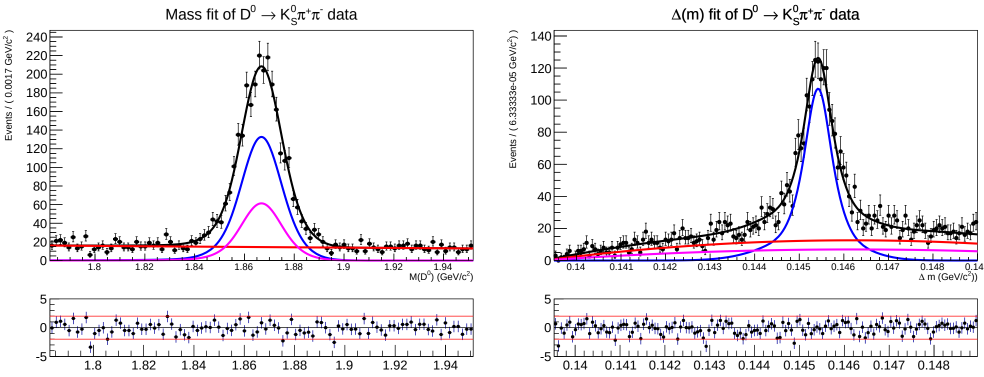
Figure 5.19: Mass and Δm fit projections for $\overline{D}^0$ in Dalitz plot bins 1 → 4. The right-π^+_{slow} (blue), wrong-π^+_{slow} (magenta) and combinatoric (red) components are indicated.



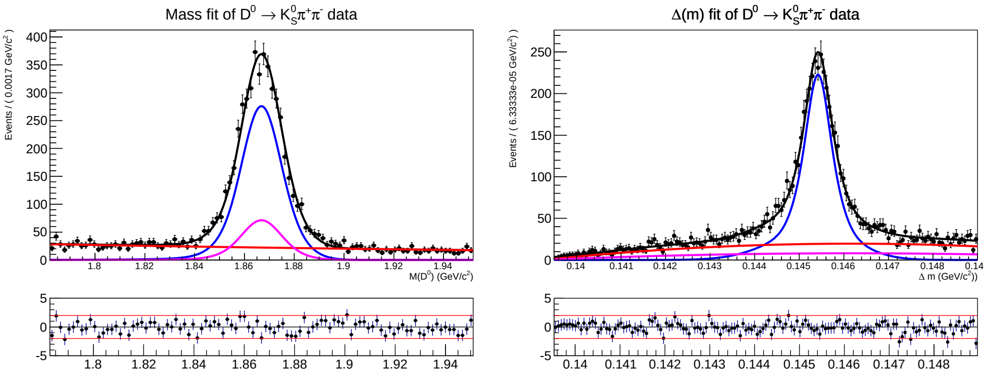
Bin 5 Bin 5



Bin 6 Bin 6

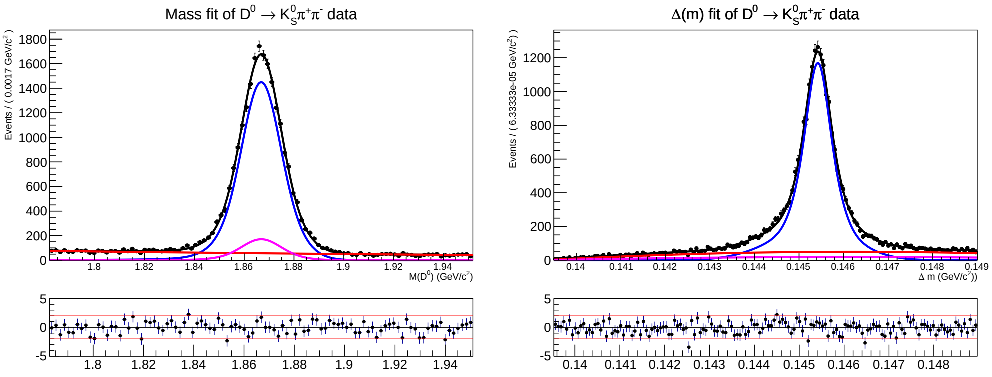


Bin 7 Bin 7



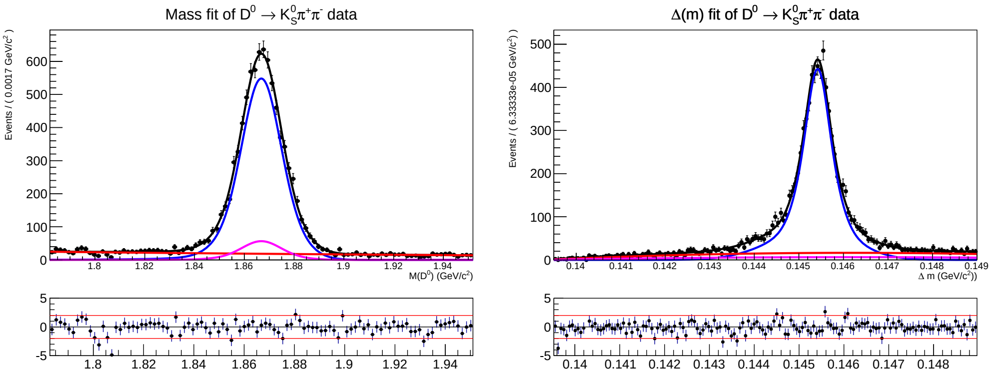
Bin 8 Bin 8

Figure 5.20: Mass and Δm fit projections for $\overline{D}^0$ in Dalitz plot bins 5 → 8. The right-π^+_{slow} (blue), wrong-π^+_{slow} (magenta) and combinatoric (red) components are indicated.



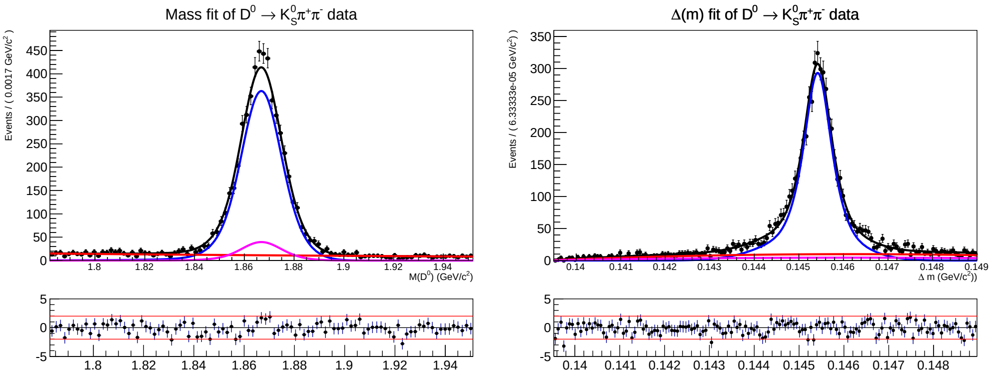
Bin -1

Bin -1



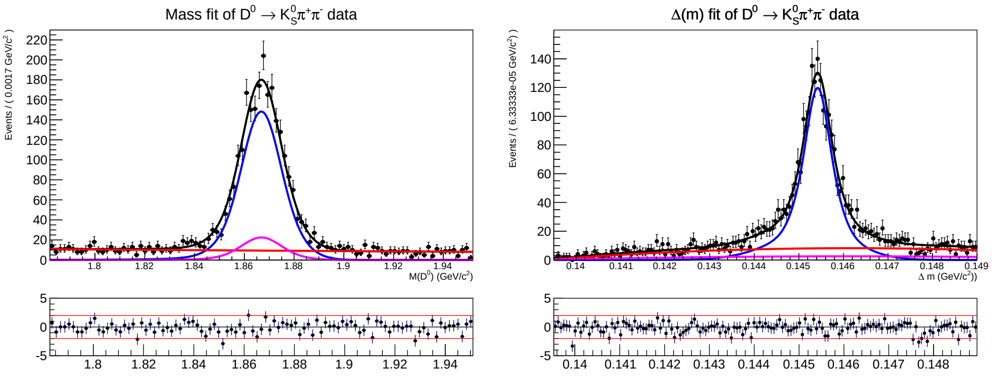
Bin -2

Bin -2



Bin -3

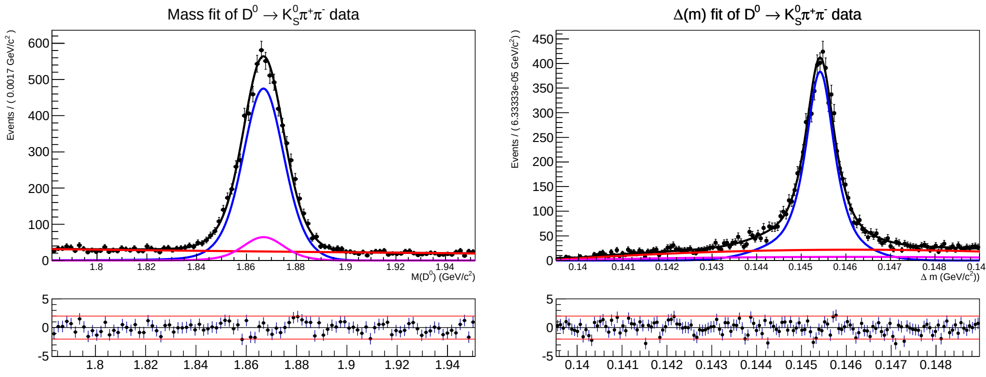
Bin -3



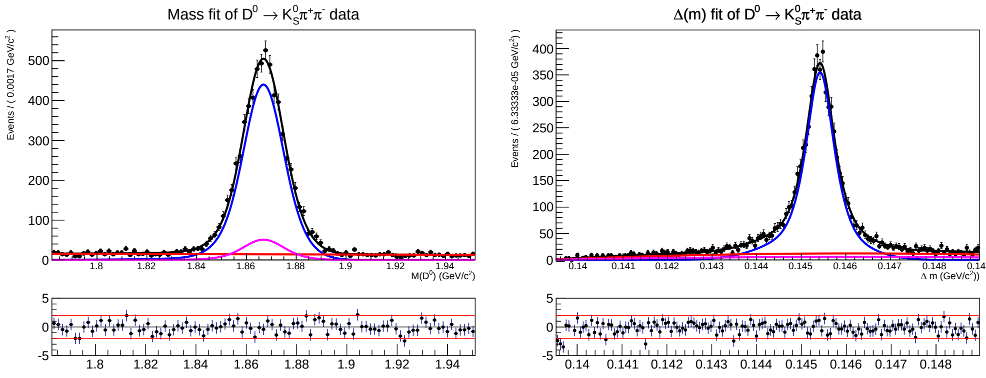
Bin -4

Bin -4

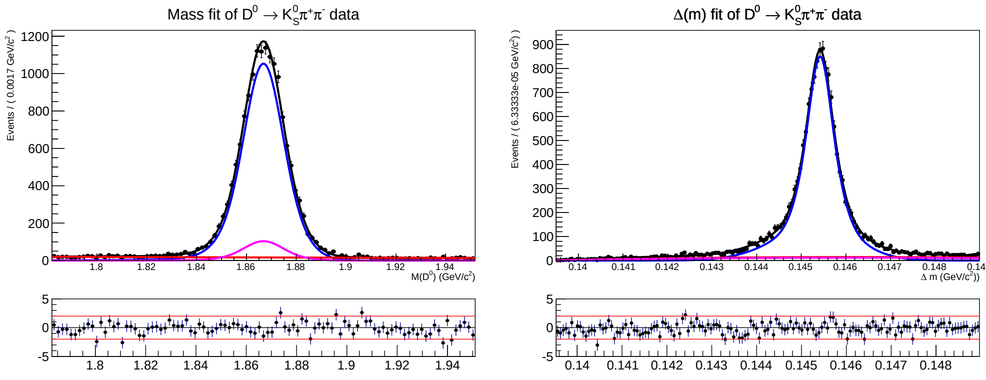
Figure 5.21: Mass and Δm fit projections for $\overline{D}^0$ in Dalitz plot bins -1 →-4. The right-π^+_{slow} (blue), wrong-π^+_{slow} (magenta) and combinatoric (red) components are indicated.



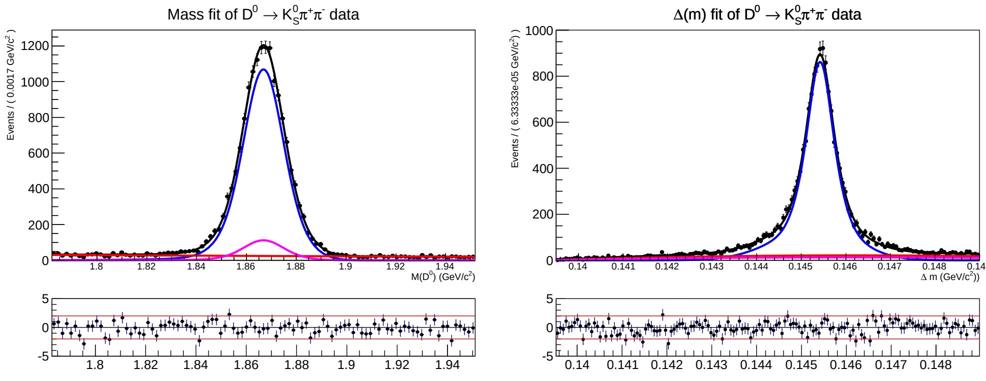
Bin -5 Bin -5



Bin -6 Bin -6

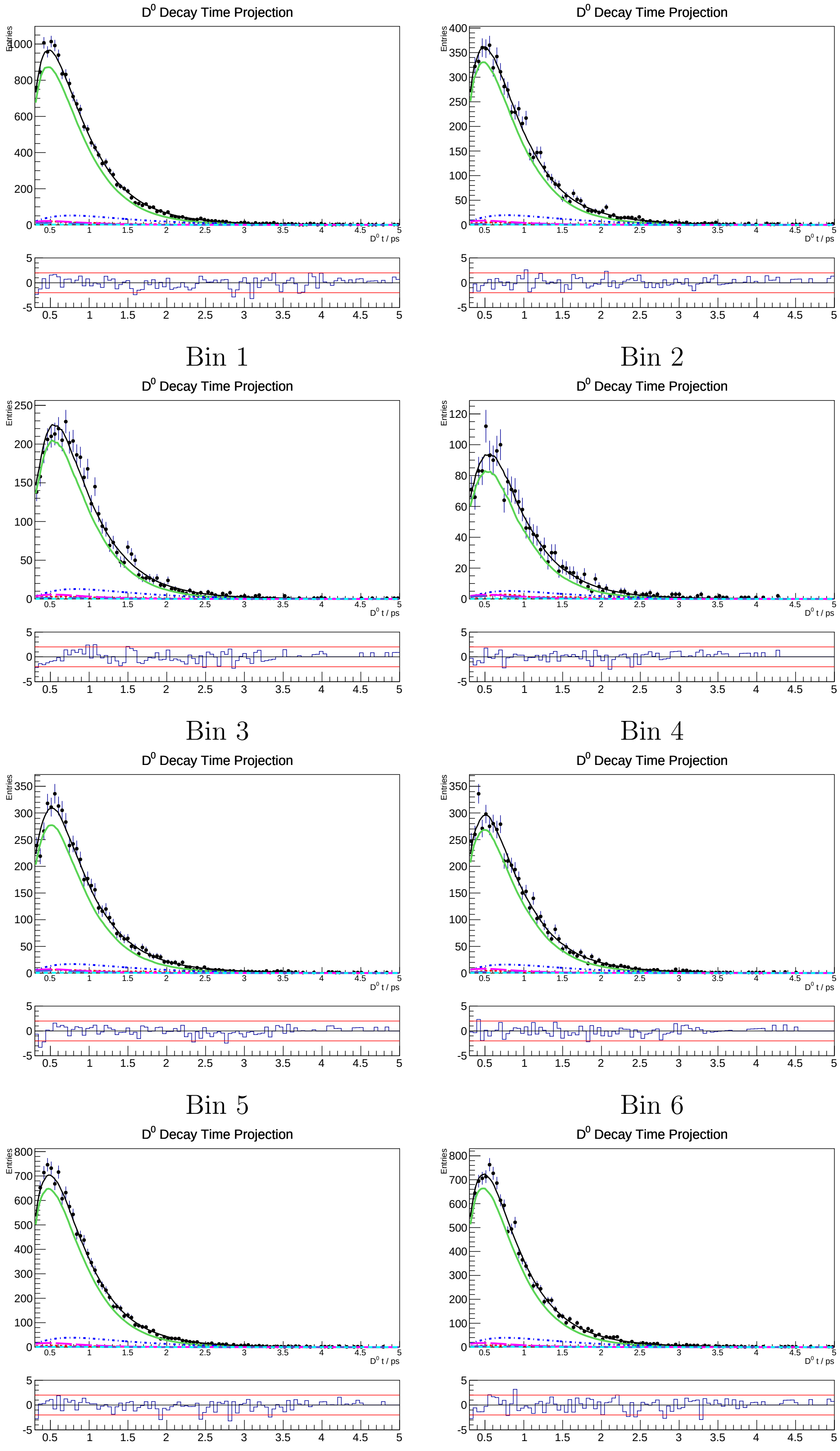


Bin -7 Bin -7



Bin -8 Bin -8

Figure 5.22: Mass and Δm fit projections for $\overline{D}^0$ in Dalitz plot bins -5 →-8. The right-π^+_{slow} (blue), wrong-π^+_{slow} (magenta) and combinatoric (red) components are indicated.



Bin 1

Bin 2

Bin 3

Bin 4

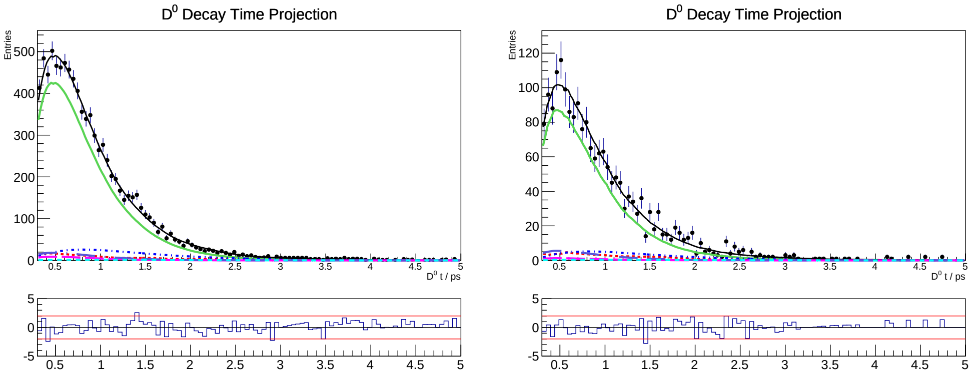
Bin 5

Bin 6

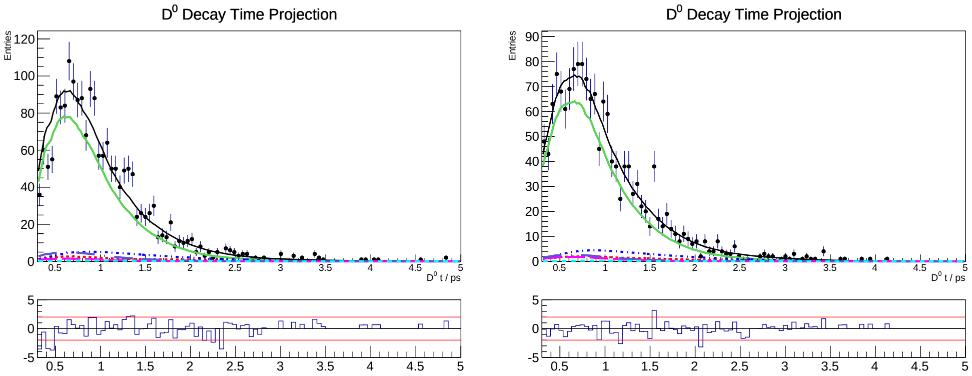
Bin 7

Bin 8

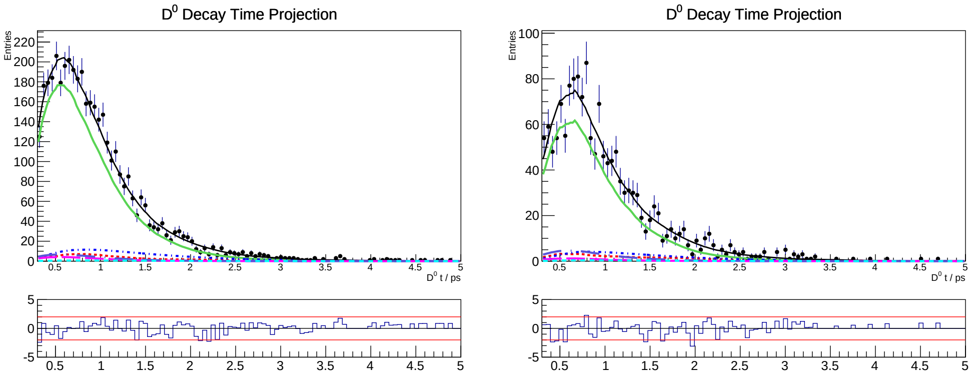
Figure 5.23: Decay time fit projections for D^0 in Dalitz plot bins 1 → 8. The prompt right-π^+_{slow} (green/solid), secondary right-π^+_{slow} (blue/short dot-dash), prompt wrong-π^+_{slow} D^0 (magenta/long dash), prompt wrong-π^+_{slow} $\overline{D}^0$ (navy/long dot-dash), secondary wrong-π^+_{slow} (cyan/short dash) and combinatoric (red/shorter dash) components are indicated.



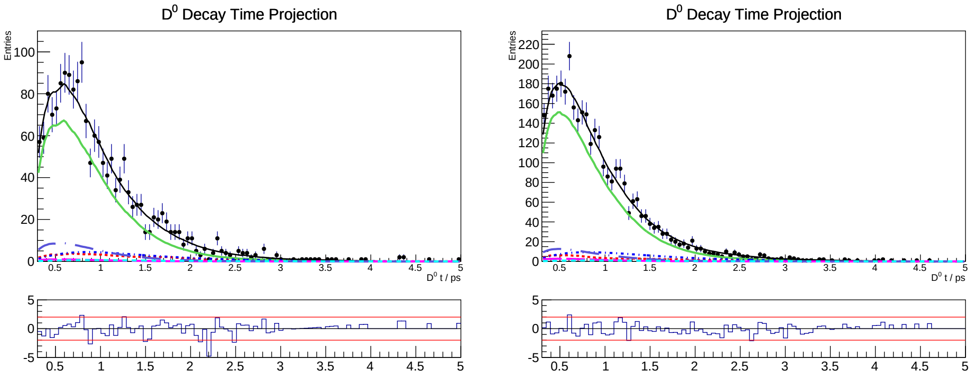
Bin -1 Bin -2



Bin -3 Bin -4

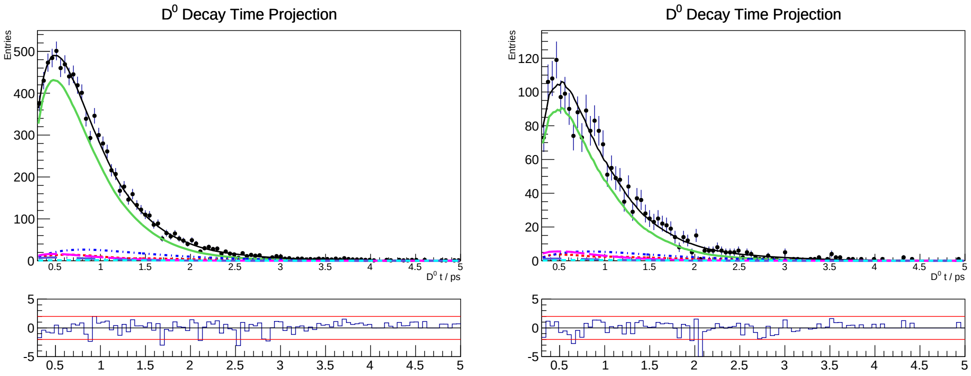


Bin -5 Bin -6

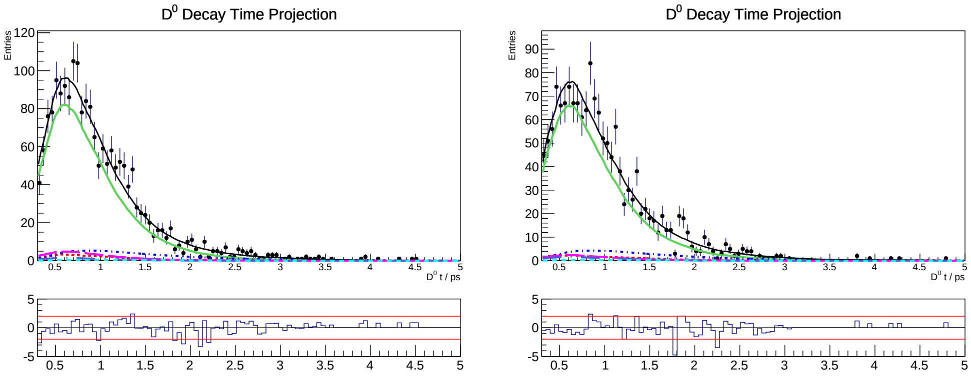


Bin -7 Bin -8

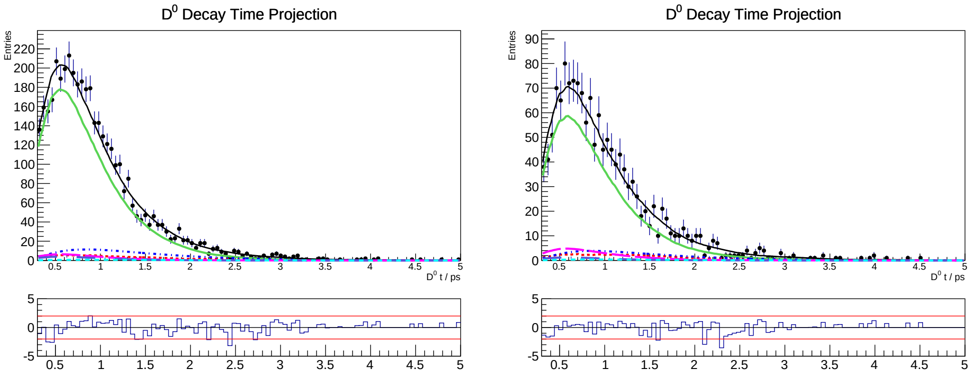
Figure 5.24: Decay time fit projections for D^0 in Dalitz plot bins -1 →-8. The prompt right-π^+_{slow} (green/solid), secondary right-π^+_{slow} (blue/short dot-dash), prompt wrong-π^+_{slow} D^0 (magenta/long dash), prompt wrong-π^+_{slow} $\overline{D}^0$ (navy/long dot-dash), secondary wrong-π^+_{slow} (cyan/short dash) and combinatoric (red/shorter dash) components are indicated.



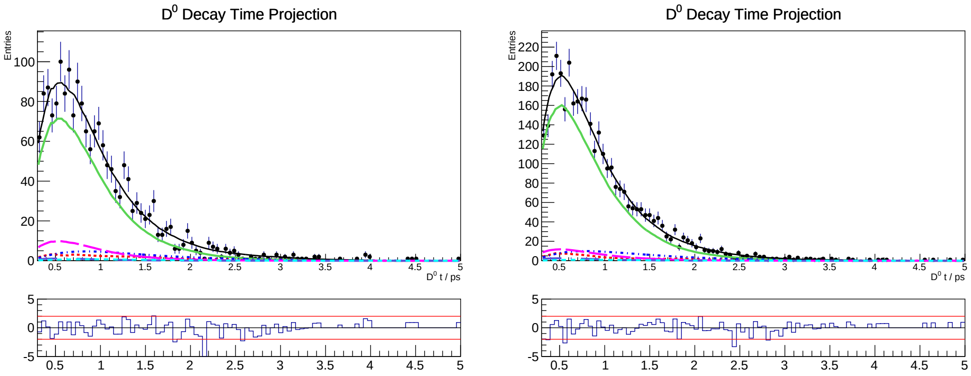
Bin 1 Bin 2



Bin 3 Bin 4

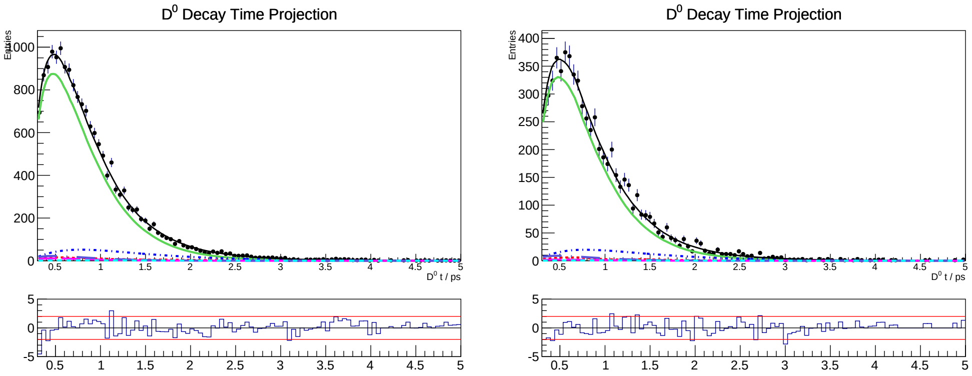


Bin 5 Bin 6

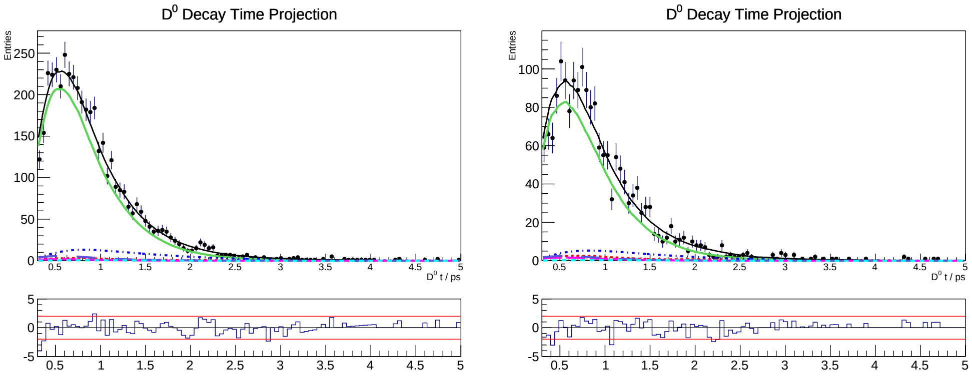


Bin 7 Bin 8

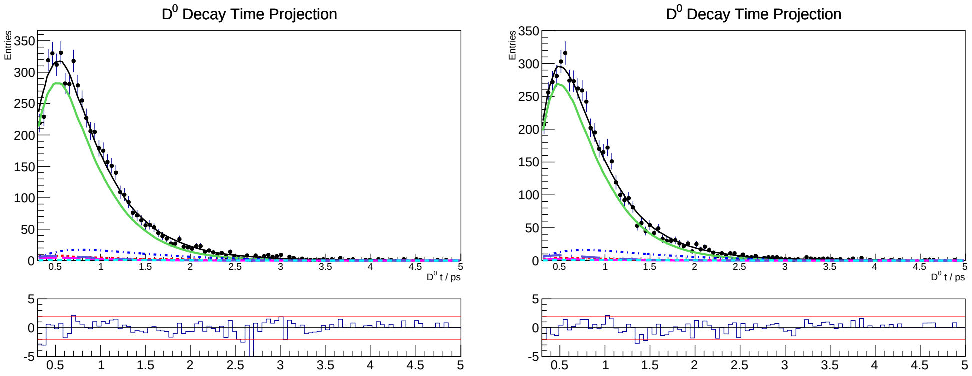
Figure 5.25: Decay time fit projections for $\overline{D}^0$ in Dalitz plot bins 1 → 8. The prompt right-π^+_{slow} (green/solid), secondary right-π^+_{slow} (blue/short dot-dash), prompt wrong-π^+_{slow} D^0 (magenta/long dash), prompt wrong-π^+_{slow} $\overline{D}^0$ (navy/long dot-dash), secondary wrong-π^+_{slow} (cyan/short dash) and combinatoric (red/shorter dash) components are indicated.



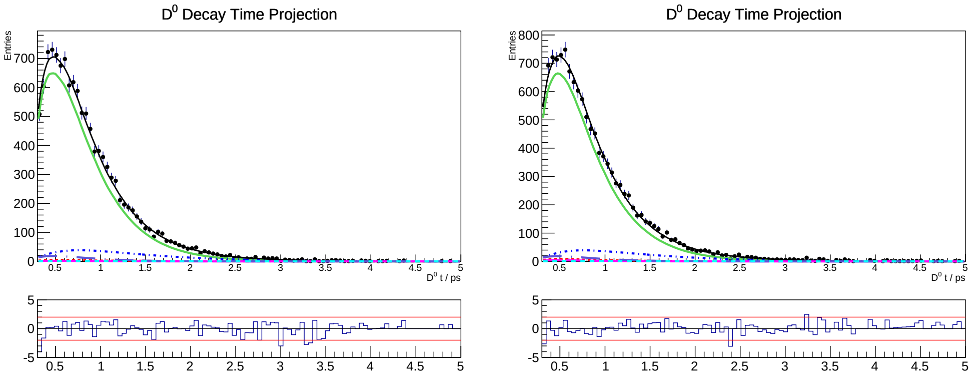
Bin -1 Bin -2



Bin -3 Bin -4

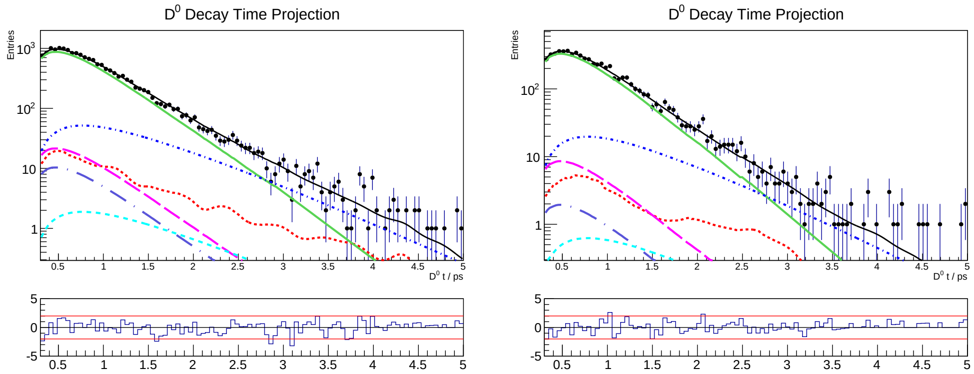


Bin -5 Bin -6

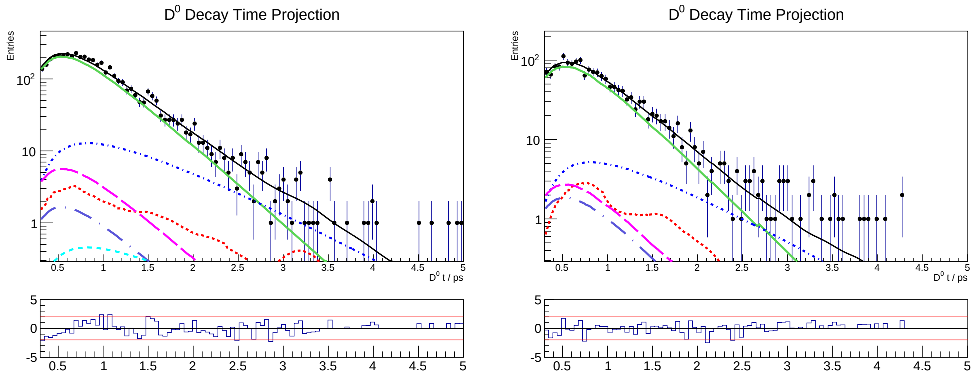


Bin -7 Bin -8

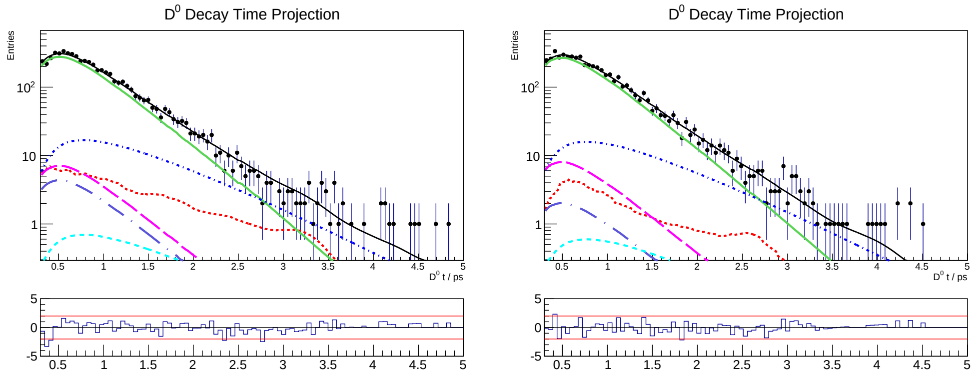
Figure 5.26: Decay time fit projections for $\overline{D}^0$ in Dalitz plot bins -1 →-8. The prompt right-π^+_{slow} (green/solid), secondary right-π^+_{slow} (blue/short dot-dash), prompt wrong-π^+_{slow} D^0 (magenta/long dash), prompt wrong-π^+_{slow} $\overline{D}^0$ (navy/long dot-dash), secondary wrong-π^+_{slow} (cyan/short dash) and combinatoric (red/shorter dash) components are indicated.



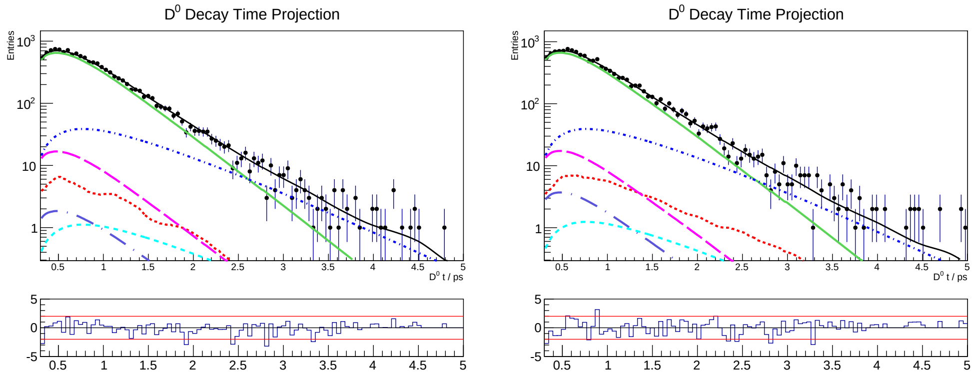
Bin 1 Bin 2



Bin 3 Bin 4

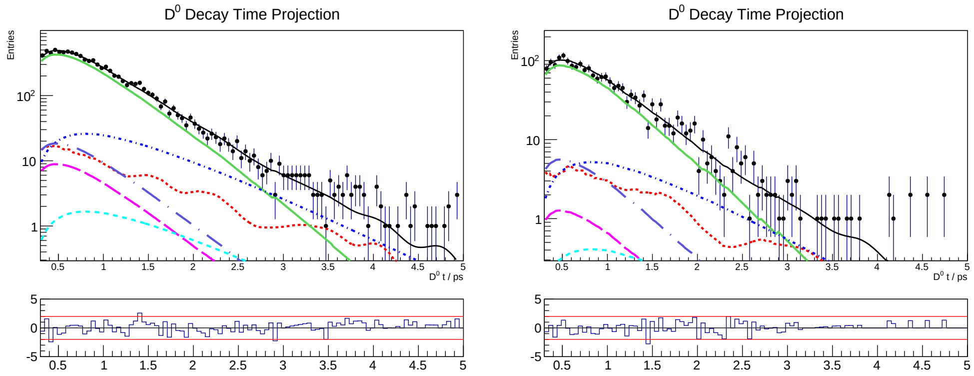


Bin 5 Bin 6

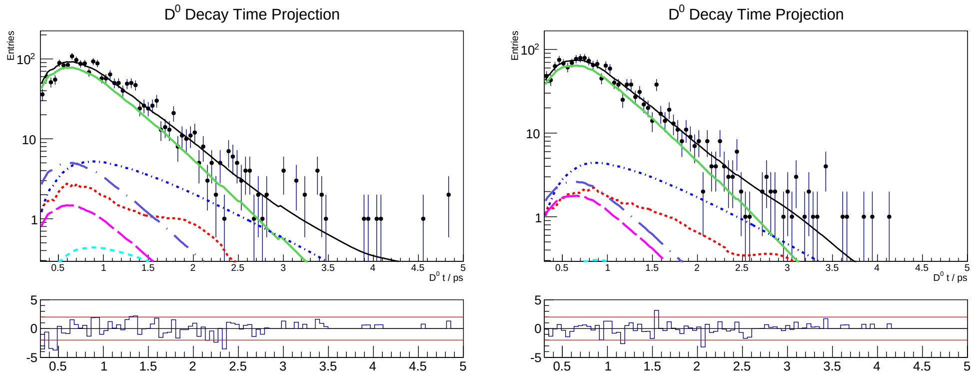


Bin 7 Bin 8

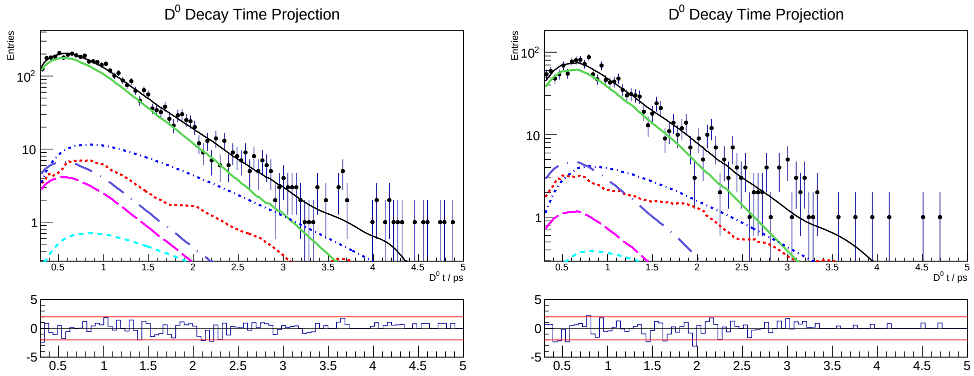
Figure 5.27: Decay time fit projections for D^0 in Dalitz plot bins $1 \rightarrow 8$. Log-y axis. The prompt right-π^+_{slow} (green/solid), secondary right-π^+_{slow} (blue/short dot-dash), prompt wrong-π^+_{slow} D^0 (magenta/long dash), prompt wrong-π^+_{slow} $\overline{D}^0$ (navy/long dot-dash), secondary wrong-π^+_{slow} (cyan/short dash) and combinatoric (red/shorter dash) components are indicated.



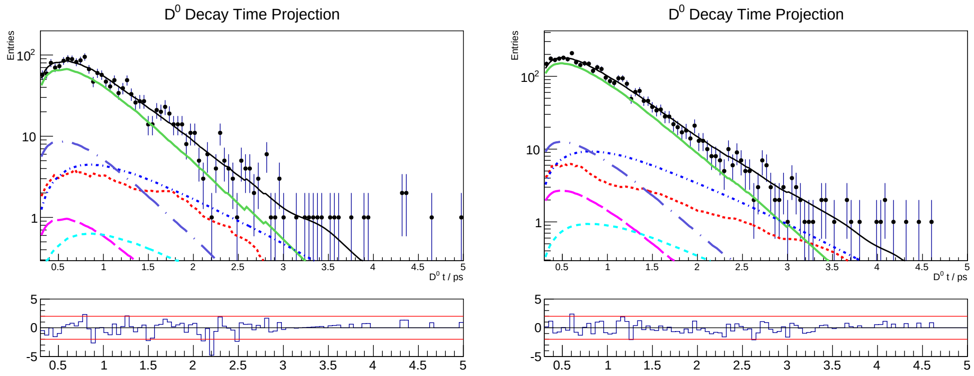
Bin -1 Bin -2



Bin -3 Bin -4

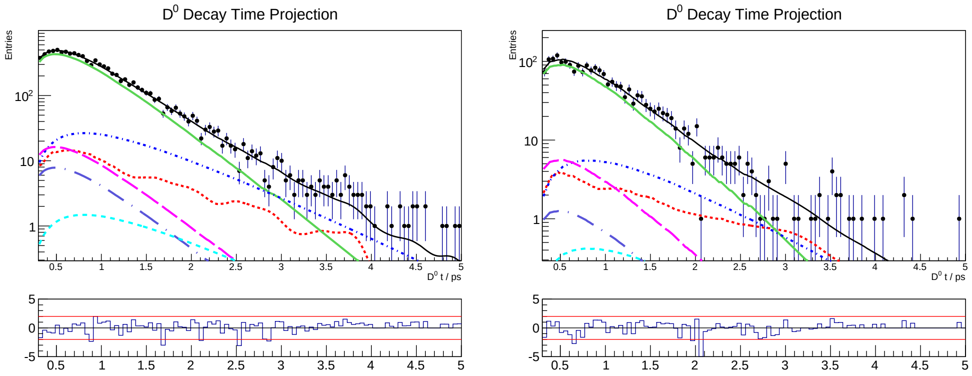


Bin -5 Bin -6

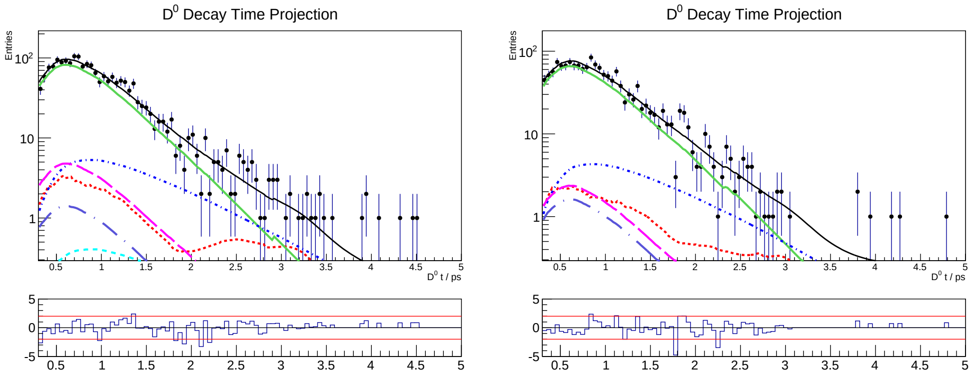


Bin -7 Bin -8

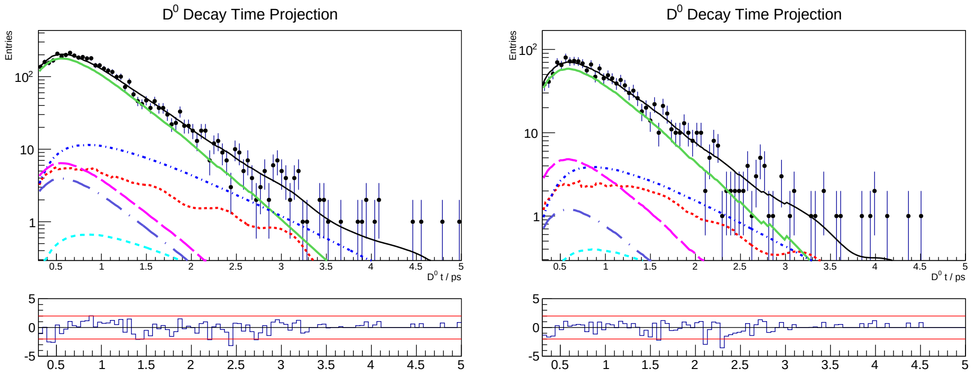
Figure 5.28: Decay time fit projections for D^0 in Dalitz plot bins -1 →-8. Log-y axis. The prompt right-π^+_{slow} (green/solid), secondary right-π^+_{slow} (blue/short dot-dash), prompt wrong-π^+_{slow} D^0 (magenta/long dash), prompt wrong-π^+_{slow} $\overline{D}^0$ (navy/long dot-dash), secondary wrong-π^+_{slow} (cyan/short dash) and combinatoric (red/shorter dash) components are indicated.



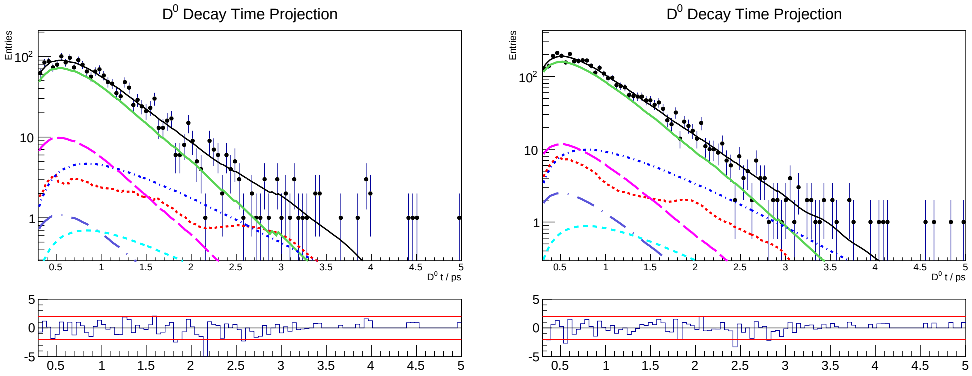
Bin 1 Bin 2



Bin 3 Bin 4

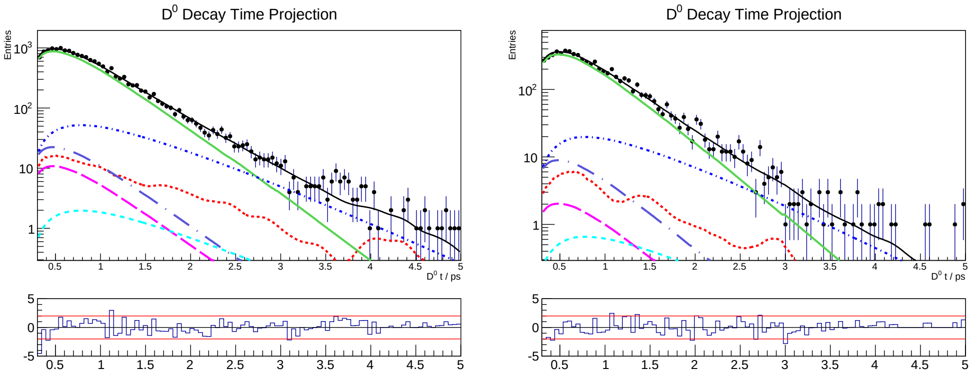


Bin 5 Bin 6

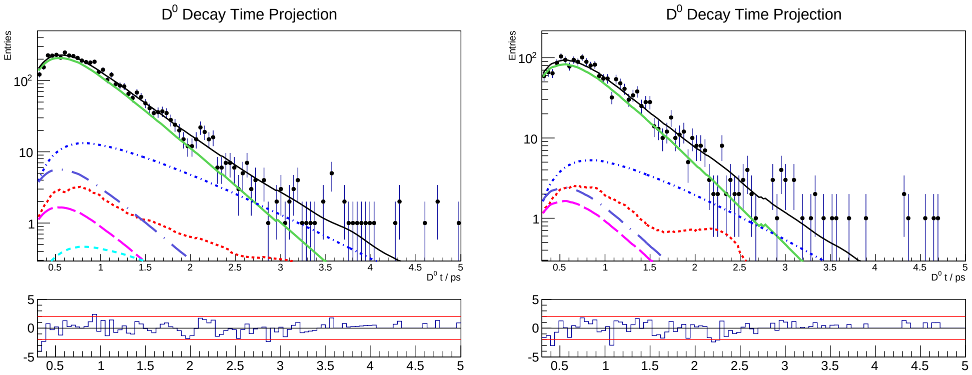


Bin 7 Bin 8

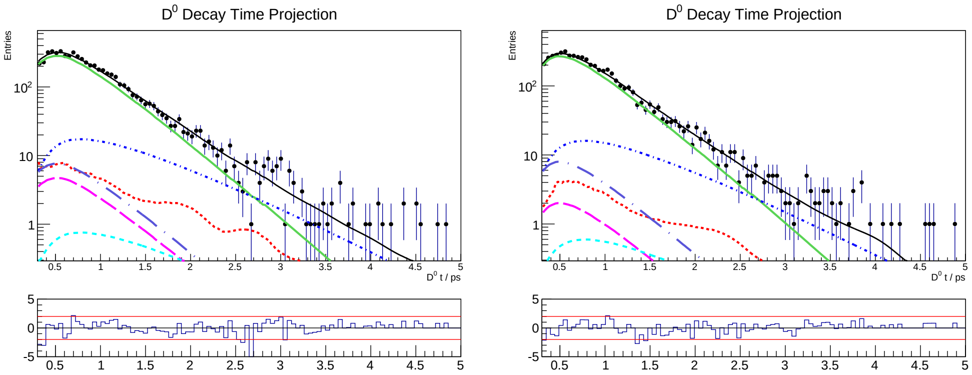
Figure 5.29: Decay time fit projections for $\overline{D}^0$ in Dalitz plot bins $1 \rightarrow 8$. Log-y axis. The prompt right-π^+_{slow} (green/solid), secondary right-π^+_{slow} (blue/short dot-dash), prompt wrong-π^+_{slow} D^0 (magenta/long dash), prompt wrong-π^+_{slow} $\overline{D}^0$ (navy/long dot-dash), secondary wrong-π^+_{slow} (cyan/short dash) and combinatoric (red/shorter dash) components are indicated.



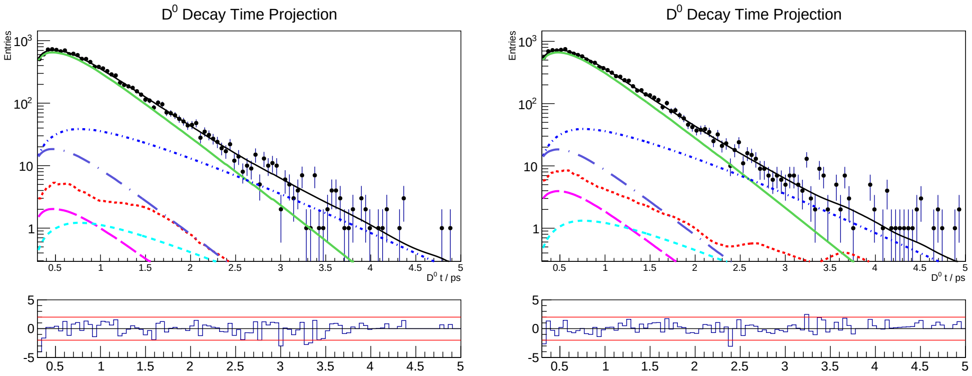
Bin -1 Bin -2



Bin -3 Bin -4



Bin -5 Bin -6



Bin -7 Bin -8

Figure 5.30: Decay time fit projections for $\overline{D}^0$ in Dalitz plot bins -1 →-8. Log-y axis. The prompt right-π^+_{slow} (green/solid), secondary right-π^+_{slow} (blue/short dot-dash), prompt wrong-π^+_{slow} D^0 (magenta/long dash), prompt wrong-π^+_{slow} $\overline{D}^0$ (navy/long dot-dash), secondary wrong-π^+_{slow} (cyan/short dash) and combinatoric (red/shorter dash) components are indicated.

Chapter 6

Systematic Uncertainties

The following chapter describes how each contribution to the systematic uncertainty has been evaluated for each of the mixing parameters. There are several distinct contributions to the overall systematic uncertainty. These stem from assumptions made about the behaviour of each category of event as a function of the fit variables, the propagation of any uncertainties on measured parameters used as input to the fit as well as from imperfections in the detector and the effect of the selection criteria on certain measured quantities. A summary of the total uncertainty on the measured parameters is given at the end of this chapter.

6.1 Fit Bias

As described in Section 5.5.3, a small, but significant, bias on the extracted values of x_D and y_D is observed in the pulls of fits to toy data. There is also a bias on the extracted value of Γ. The following section describes how the systematic uncertainties associated with these biases is evaluated. The bias for x_D is larger in the case where c_i and s_i are allowed to float. However, as the final fit is performed in this manner to propagate the CLEO uncertainties to the mixing parameters, the larger bias must be accommodated as a systematic uncertainty. The toys used in the fit validation studies were generated with $\Gamma = 2.44\,\mathrm{MeV}$, $x_D = y_D = -1\,\%$. The distribution of fitted values taken from the fit

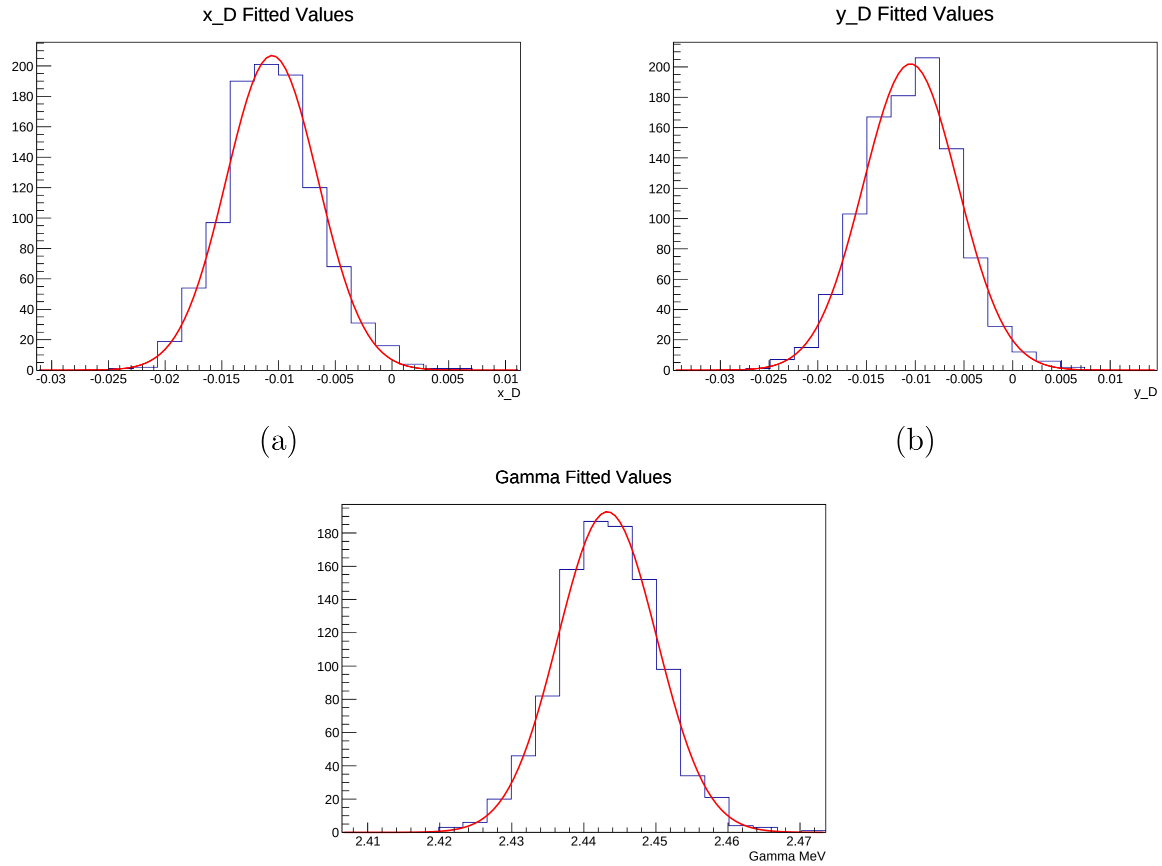
Parameter	Generated Value	Mean Fitted Value	Bias
Γ (MeV)	2.44	2.4412 ± 0.0002	$+0.0012 \pm 0.0002$
x_D (%)	-1.00	-1.0158 ± 0.0140	-0.0158 ± 0.0140
y_D (%)	-1.00	-1.0125 ± 0.0157	-0.0125 ± 0.0157

Table 6.1: Bias observed on the extracted mixing parameters from toys where c_i and s_i are allowed to float.

validation toys with c_i and s_i floated are plotted in Figures 6.1 (a)–(c). A Gaussian fit is performed to each distribution and the size of the bias is taken as the difference between the fitted mean and the generated central values for each parameter. The observed biases are summarised in Table 6.1. The keen eyed reader will notice that the biases quoted here are different from those in the Section 5.5.3. The biases quoted here are calculated from fits to the distributions of results for the toys (Figures 6.1(a)–(c)), whereas those in the previous section are estimated from fits to the pull distributions and the mean statistical uncertainty from the toys.

Taking the difference between the mean fitted values and the generator values for each of the mixing parameters shows the fit procedure is biased by $\Delta x_D(fit - gen) = -0.0158 \pm 0.0140\,\%$ and $\Delta y_D(fit - gen) = -0.0125 \pm 0.0157\,\%$. Correcting for the bias in the fit to data would result in systematic uncertainties equivalent to the size of the uncertainty on the measurement of the bias i.e. $\sigma_{x_D}(\text{syst.}) = 0.0140\,\%$ and $\sigma_{y_D}(\text{syst.}) = 0.0157\,\%$. However, although the toys are generated in such a way that they mimic the data as closely as possible, there may still be subtle inconsistencies between the two that could result in the correction being over or under estimated. Because of this and the fact that the size of the biases on x_D and y_D are much smaller than the statistical uncertainty in data, a systematic uncertainty that is equal to the magnitude of the bias and its uncertainty taken in quadrature is assigned to the mixing parameters. The resulting systematics on the mixing parameters due to the fit bias are $\sigma_{x_D}(\text{syst.}) = 0.0211\,\%$ and $\sigma_{y_D}(\text{syst.}) = 0.0201\,\%$.

The toy generation and fit procedure used here is also used to evaluate a subset of systematic effects in the following sections. Given that there is an inherent bias observed in the tests performed here, although small, it must be accounted for when evaluating the size



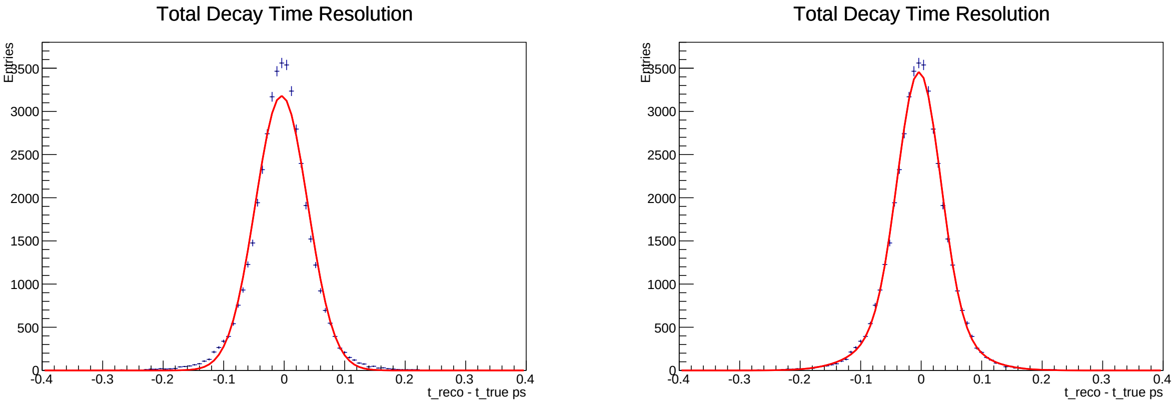
(a) (b) (c)

Figure 6.1: (a) Distribution of fitted values for x_D from validation toys where c_i and s_i are allowed to float. (b) Equivalent plot for y_D. (c) Equivalent plot for Γ.

of any systematic shift that is measured using toy experiments of this kind. Therefore, any shifts observed in the mixing parameters during these systematic checks are calculated with respect to the biased values listed in Table 6.1. This does not affect systematics that are evaluated with data or with toys that are generated using an amplitude model.

6.2 Decay Time Resolution

Decay time resolution is not accounted for in the lifetime fit nor the mixing fit. This is justified given the excellent resolution achievable with the VELO and the slow rate of mixing in comparison with the mean D^0 lifetime. Initial studies indicated that ignoring decay time resolution was likely to have very little effect on the mean lifetime of the D^0. However, the main concern is whether this has a significant impact on the extracted values of the mixing



(a) (b)

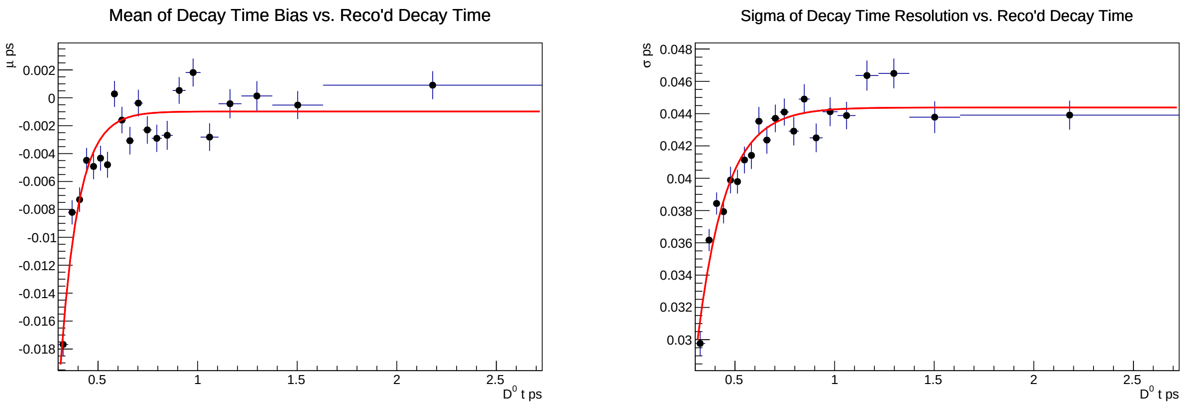
Figure 6.2: Decay time resolution for truth-matched signal candidates as measured from MC. (a) Fitted with a single Gaussian. (b) Fitted with a double Gaussian.

parameters.

The first ingredient is to get an estimate of the size of the decay time resolution for $D^0 \to K^0_S\pi^+\pi^-$ decays. The decay time resolution is measured using truth-matched $signalmixfiltered$ MC with a total sample size of roughly 40k events after the full selection has been applied. The difference between the true decay time and the reconstructed value is plotted and a Gaussian fitted to the resulting distribution. The width of this Gaussian corresponds to the decay time resolution as estimated from the signal MC. Figure 6.2 (a) shows the resulting distribution and fit for all truth-matched signal candidates. The Gaussian fitted parameters are $\mu = -4.31 \pm 0.20$ fs and $\sigma = 43.32 \pm 0.19$ fs. The actual resolution from the MC is probably a bit lower than the fitted value, but the tails either side of the distribution force the width to be larger. However, fitting a double Gaussian to the distribution gives much closer agreement with the data (Figure 6.2 (b)) indicating that there may also be an effect due to variation in the resolution as a function of some other variable. The average mean and width of the double Gaussian fit are $\mu = -4.22 \pm 0.36$ fs and $\sigma = 40.77 \pm 0.64$ fs, respectively. The global decay time resolution used in what follows is taken to be the larger of the two fitted values (i.e. the result of the single Gaussian fit).

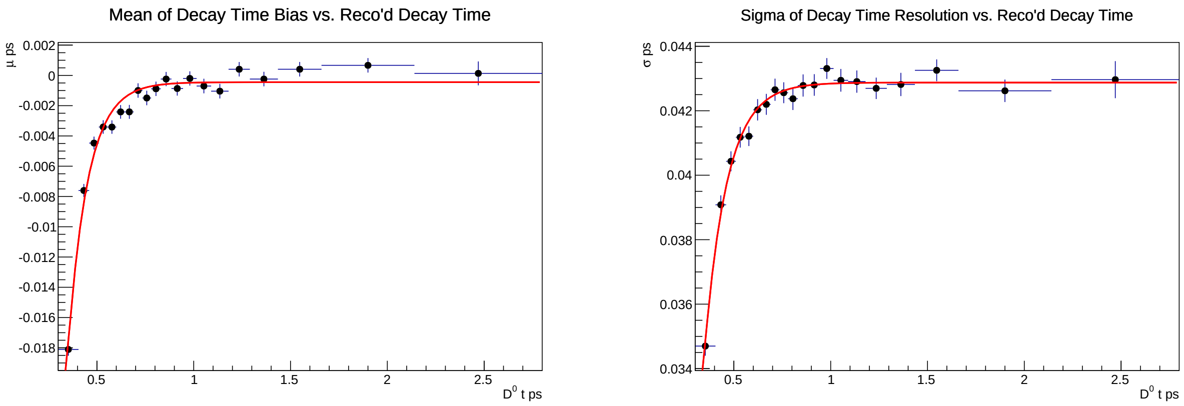
In addition to global decay time resolution effects, the parameters describing the resolution may also be dependent on decay time. To understand if this is an effect fits to the decay time resolution are performed in bins of decay time. A single Gaussian is used to model the decay time resolution in each bin. Events are binned according to their true decay time and the binning is chosen such that there are $\geq$ 2k events in each bin. Figure 6.3 (a) shows the measured decay time bias as a function of true decay time. It is clear that there is a threshold effect close to the stripping decay time cut of 0.3 ps where the decay time bias is large. This effect is likely due to the selection favouring events that are reconstructed with higher decay times close to their acceptance boundary. For example, for events that are produced at 0.3 ps only those that are smeared in an upward direction will pass the stripping cut. Above the stripping cut the size of the effect drops as other cuts in the trigger and stripping that are less correlated with decay time dominate the process. The width of the decay time resolution also narrows in the region close to the cut (see Figure 6.3 (b)) likely driven by the same mechanism. A simple test is devised to deduce if the stripping cut is in fact the primary cause of this effect and to estimate the size of any systematic effect on the physics observables. A total of 1000 toy datasets each consisting of 180k events were generated using the same setup as for the mixing fit validation studies. However, the generated decay time is subsequently smeared according to the width of the global decay time resolution i.e. $\sigma = 43.32 \pm 0.19$ fs. The smearing is done with no decay time bias. However, if the smeared, pseudo-reconstructed decay time falls outside of the generated acceptance then the event is not saved and another is generated in its place. The smearing is only applied to events that are generated as prompt right-π^+_{slow} and prompt wrong-π^+_{slow} and the results should closely mimic the time-dependent effect that is observed in the signal MC. Figures 6.4 show the same resolution plots as Figure 6.3, but taken from one of the toy datasets generated in the manner described. The plots are binned according to the true decay time, the binning chosen so there are roughly 8k events in each bin (with a minimum of 2k), and only events generated as prompt right-π^+_{slow} and prompt wrong-π^+_{slow} are included. It is clear that the effect observed in the signal MC is successfully reproduced in the toys by simply smearing the decay time and rejecting events that are smeared outside of their acceptance.

To determine the size of any systematic effect on the physics observables, 1000 fits were performed to the toy datasets generated in this manner, but with no attempt to account for the decay time resolution effects in the fit. The resulting distributions of fitted values of x_D



(a) (b)

Figure 6.3: (a) Fit to time evolution of the decay time bias. (b) Fit to time evolution of the resolution width.



(a) (b)

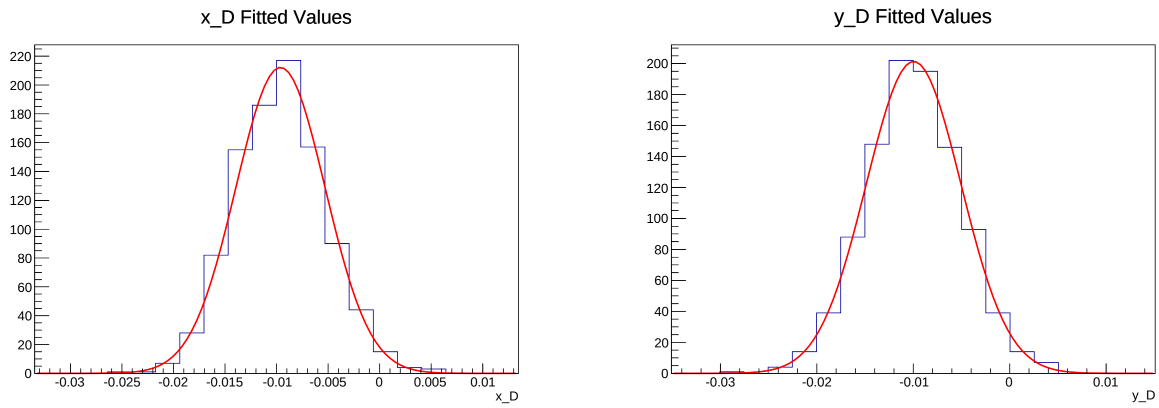
Figure 6.4: (a) Fit to time evolution of the pseudo-reconstructed decay time bias taken from toy MC. (b) Fit to time evolution of the resolution width taken from toy MC.

Parameter	Generated Value	Mean Fitted Value	Expected Value
x_D (%)	-1.00	-0.9532 ± 0.0140	-1.0158 ± 0.0140
y_D (%)	-1.00	-0.9802 ± 0.0156	-1.0125 ± 0.0157

Table 6.2: Mean of Gaussian fit to distributions of fitted values for x_D and y_D from toys that have been generated with decay time resolution effects. The last column is the biased value expected from the fit validation toys performed in Section 6.1.

and y_D are shown in Figures 6.5 (a) and (b), respectively. Gaussian fits are performed to both distributions and the results are summarised in Table 6.2. Listed in Table 6.2 are the generated mixing parameters, the mean fitted value from the fits to toys that have decay time resolution included in the generation and the expected value given the bias observed in the fit validation studies. The shifts away from the generated values for each parameter are $\Delta x_D(fit - gen) = +0.0468 \pm 0.0140\,\%$ and $\Delta y_D(fit - gen) = +0.0198 \pm 0.0156\,\%$. This shift is in the opposite direction to the bias observed in the fit validation studies and is of comparable size. If the toys are to be trusted, then it is clear that these biases begin to cancel each other out with respect to the generated value. However, the shifts in the fitted values compared to the expected values from the fit validation studies are much larger with $\Delta x_D(fit - exp) = +0.0626 \pm 0.0190\,\%$ and $\Delta y_D(fit - exp) = +0.0323 \pm 0.0221\,\%$. Despite their size, these shifts are still much smaller than the statistical uncertainty seen in data. The sum in quadrature of the magnitude of the bias with respect to the expected values and its associated uncertainty are assigned as a systematic uncertainty on both x_D and y_D resulting in $\sigma_{x_D}(\text{syst.}) = 0.0654\,\%$ and $\sigma_{y_D}(\text{syst.}) = 0.0391\,\%$.

Another potential issue is significant variation in the decay time resolution and bias across the Dalitz plot bins. If the reconstructed decay time of signal candidates is positively biased in bin i with respect to bin $-i$ then this could fake a mixing signal due to a higher population of signal events with long decay times in one bin compared to the other. To estimate the size of any potential effect, fits to the decay time resolution are performed in each Dalitz plot bin. A single Gaussian is used to model the resolution in each bin and the extracted mean and sigma saved. The fitted parameters for each bin are plotted in Figures 6.6 (a) and (b), where (a) is the decay time bias in each bin and (b) is the resolution. The fitted parameters



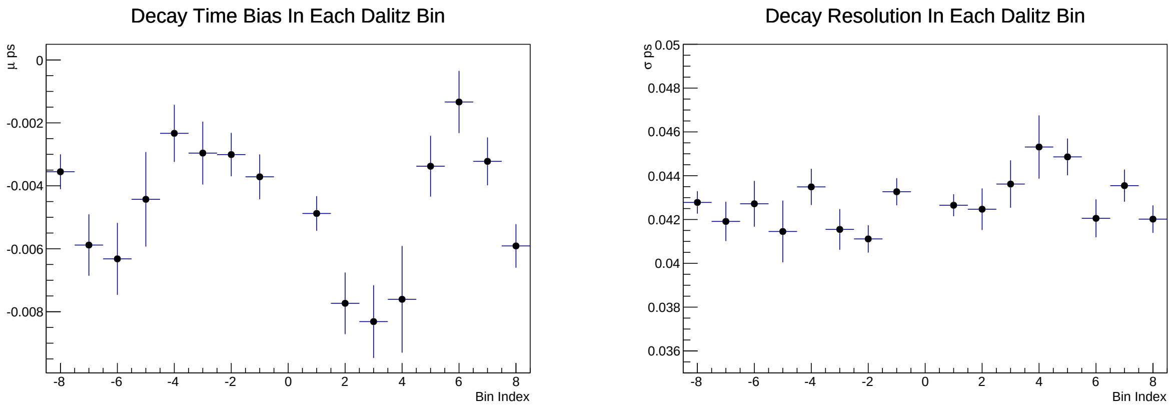
(a) (b)

Figure 6.5: (a) Gaussian to extract values of x_D for 1000 toys that are generated with decay time resolution effects. (b) Equivalent plot for y_D.

for each bin are also listed in Table 6.3. It is clear that the bias does vary from bin to bin, however, the resolution remains relatively constant. The difference in the bias is likely driven by the variation in the decay time acceptance across the Dalitz plot. This is simple to test using toys generated in the manner described earlier in this section. As the decay time biasing effects in the toys are generated uniquely for each Dalitz bin using the turning point distributions measured from data, if the variation in the decay time resolution across each bin is purely driven by the acceptance effects, then applying a global resolution will produce a similar effect in the toys. Figure 6.7 shows the variation in fitted decay time bias and resolution in each of the Dalitz plot bins taken from the same toy data used to produce the plots in Figure 6.4. It is clear that the spread of the resolution parameters across each bin is less for the toys compared with that seen in the signal MC. However, this is perhaps to be expected given the higher statistics involved. Despite this, given the variation in resolution parameters cross the Dalitz plot bins observed in signal MC is relatively small in comparison to the overall resolution, it is not felt that there will be any additional systematic effect in the data that is not accounted for in the study performed earlier in this section.

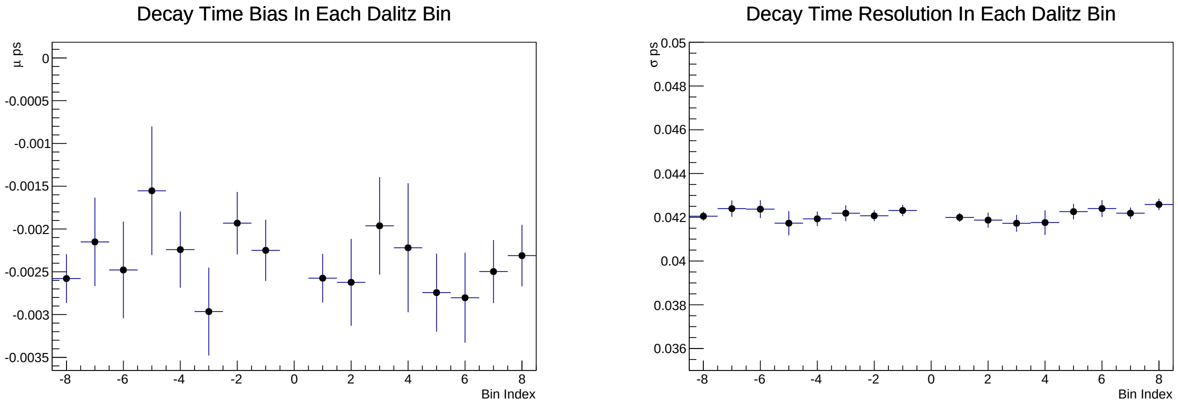
Bin	μ ps	σ ps
1	-0.00488 ± 0.00055	0.0427 ± 0.0005
2	-0.00773 ± 0.00098	0.0425 ± 0.0010
3	-0.00831 ± 0.00115	0.0436 ± 0.0011
4	-0.00761 ± 0.00170	0.0453 ± 0.0014
5	-0.00338 ± 0.00097	0.0449 ± 0.0008
6	-0.00134 ± 0.00099	0.0421 ± 0.0009
7	-0.00322 ± 0.00076	0.0435 ± 0.0007
8	-0.00591 ± 0.00069	0.0420 ± 0.0006
-1	-0.00355 ± 0.00055	0.0428 ± 0.0005
-2	-0.00588 ± 0.00098	0.0419 ± 0.0009
-3	-0.00632 ± 0.00114	0.0427 ± 0.0011
-4	-0.00443 ± 0.00150	0.0415 ± 0.0014
-5	-0.00233 ± 0.00091	0.0435 ± 0.0008
-6	-0.00296 ± 0.00100	0.0415 ± 0.0009
-7	-0.00301 ± 0.00069	0.0411 ± 0.0006
-8	-0.00371 ± 0.00071	0.0433 ± 0.0006

Table 6.3: Fitted decay time resolution parameters for individual Dalitz plot bins taken from signal MC.



(a) (b)

Figure 6.6: (a) Fitted decay time bias in individual Dalitz plot bins taken from signal MC. (b) Fitted decay time resolution in individual Dalitz plot bins taken from signal MC.



(a) (b)

Figure 6.7: (a) Fitted decay time bias in individual Dalitz plot bins taken from toy MC (b) Fitted decay time resolution in individual Dalitz plot bins taken from toy MC.

6.3 Turning Point Resolution

Another resolution effect that is not accounted for in the fit is the turning point resolution. Although the turning points are measured to a very high precision using the iterative procedure described in Section 4.6, the small, but non-zero resolution may have an adverse effect on the extracted mixing parameters. This is particularly important for the first turning point as any significant deviation from the true turning point position at low decay times could have a large effect on the size of the acceptance correction applied in the fit. The procedure to estimate the size of any potential effect is similar to what was done in the previous section. The turning point resolution is measured for the first turning point using data. This resolution is then used to smear the generated turning points in 1000 toy experiments produced using the same setup as for the mixing fit validation studies, but ignored in the fit. The same resolution is applied to both TP_1 and TP_{diff}, but with each smeared independently. As mentioned, it is unlikely that the smearing of TP_{diff} will have as much of an effect as TP_1, but smearing both provides an upper limit on the size of the effect likely to be seen in data resulting in a more than conservative estimate of any associated systematic.

The turning point resolution for TP_1 is measured by subtracting TP_1 from the reconstructed decay time and plotting the resulting distribution. In the limit of perfect turning point resolution this plot would have a sharp cut off at $t = 0$. In reality the resolution

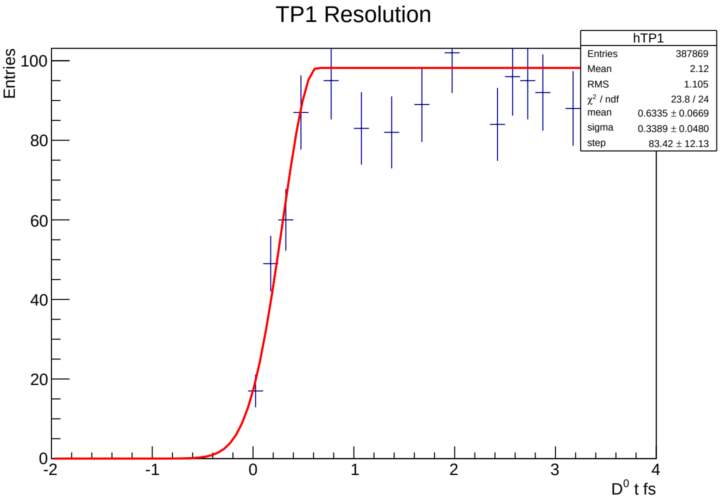


Figure 6.8: Fit to TP_1 resolution measured from real data.

smooths this boundary. By fitting a frequency function to the distribution in the region around zero one can take the fitted width as an estimate of the turning point resolution. Figure 6.8 shows the fitted resolution of TP_1 for $K^0_S\pi^+\pi^-$ data. The measured resolution is $\sigma = 0.338 \pm 0.048$ fs. In addition, a small bias in the measured turning points is observed with $\mu = 0.634 \pm 0.067$ fs. Both of these are much smaller than the parameters describing the overall decay time resolution.

1000 toy datasets are generated with both TP_1 and TP_{diff} smeared by a Gaussian with $\mu = 0.7$ fs and $\sigma = 0.4$ fs such that the resolution effect induced is slightly larger than that measured from real data. The smearing is applied to all fit components in the same way. The distribution of fitted values for x_D and y_D for the 1000 toys are shown in Figures 6.9 (a) and (b), respectively. A Gaussian fit is performed to each distribution and the results are summarised in Table 6.4. The shifts away from the generated values for each parameter are $\Delta x_D(fit - gen) = -0.0178 \pm 0.0139\,\%$ and $\Delta y_D(fit - gen) = -0.0124 \pm 0.0155\,\%$. However, in this case the differences between the mean fitted values and those expected from the fit validation study are smaller than for the decay time resolution giving $\Delta x_D(fit - exp) = -0.0020 \pm 0.0197\,\%$ and $\Delta y_D(fit - exp) = -0.0001 \pm 0.0221\,\%$. As done previously for the decay time resolution, the magnitude of the relative shifts between the fitted and expected values summed in quadrature with their uncertainties are assigned as systematic uncertainties on x_D and y_D such that $\sigma_{x_D}(\text{syst.}) = 0.0198\,\%$ and $\sigma_{y_D}(\text{syst.}) = 0.0221\,\%$. It is clear from these numbers that the turning point resolution has very little effect on the extracted mixing parameters and the systematic is dominated by the uncertainties from the toys.

Parameter	Generated Value	Mean Fitted Value	Expected Value
x_D (%)	-1.00	-1.0178 ± 0.0139	-1.0158 ± 0.0140
y_D (%)	-1.00	-1.0124 ± 0.0155	-1.0125 ± 0.0157

Table 6.4: Mean of Gaussian fit to distributions of fitted values for x_D and y_D from toys that have been generated with turning point resolution effects. The last column is the biased value expected from the fit validation toys performed in the Section 6.1.

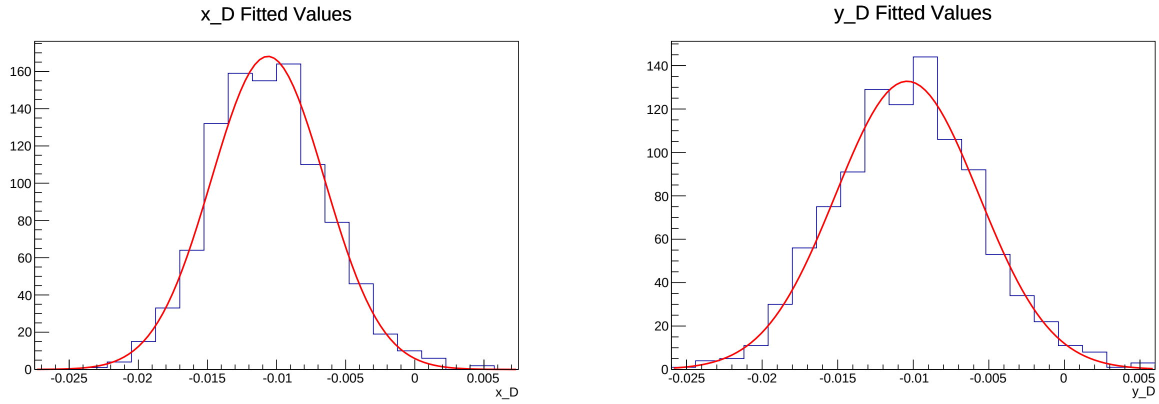
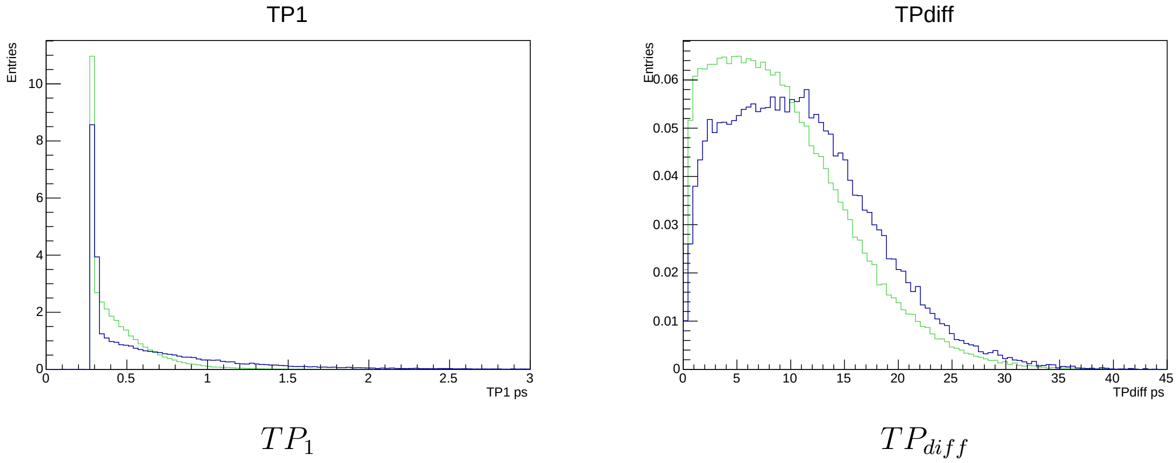


Figure 6.9: (a) Distribution of fitted values for x_D from 1000 toys that were generated with TP_1 and TP_{diff} smeared according to the turning point resolution measured from data. (b) Equivalent plot for y_D.

6.4 Turning Point Distributions

In the mixing fit, it is assumed that the turning point distributions for prompt and secondary components are the same. However, it is likely that this is not true of the data as explained in Section 5.2.4. To estimate the size of any systematic effect due to this assumption a series of toy experiments are conducted. The generation of each toy is conducted in exactly the same way as for the fit validation study except that the distributions used to generate the turning points which bias the decay time distribution for prompt and secondary are taken from data in different regions of $ln(\chi^2_{IP})$. Both the prompt right-π^+_{slow} and prompt wrong-π^+_{slow} components have turning points generated from distributions taken from data that



TP_1 TP_{diff}

Figure 6.10: Example of difference in turning point distributions used to generate prompt (green) and secondary (blue) events. The difference closely resembles that observed in the studies of signal MC performed in Section 5.2.4.

have $ln(\chi^2_{IP}) < 1.0$ within each Dalitz plot bin. In contrast both the secondary right-π^+_{slow} and secondary wrong-π^+_{slow} components have turning points generated from data that have $ln(\chi^2_{IP}) > 3.0$ in each Dalitz plot bin (see Figure 6.10). The turning point PDFs used to model these distributions in the fit are then fixed to the shape of the distribution for all generated data within a given Dalitz plot bin with all four of these components treated in the same way. The combinatoric component has turning points generated using distributions taken from the D^0 mass sidebands in each Dalitz plot bin and the PDFs used to describe these are fixed to those used in the generation. The idea is to isolate the effect of mis-modelling the turning point distributions for prompt and secondary events.

The generation and fit step is run 1000 times and the resulting distributions of fitted values for x_D and y_D are plotted in Figures 6.11 (a) and (b), respectively. The results of a Gaussian fit to each of the resulting distributions are summarised in Table 6.5. The differences between the mean fitted values and those expected from the fit validation study are $\Delta x_D(fit-exp) = -0.0473 \pm 0.0192\,\%$ and $\Delta y_D(fit-exp) = -0.0056 \pm 0.0220\,\%$. It is clear that mis-modelling of the turning point distributions has little effect on the extracted value of y_D. There is, however, a significant shift observed in x_D. As has been done previously, the systematic for each parameter is taken as the sum in quadrature of the magnitude of the bias and its uncertainty. This results in systematic uncertainties of $\sigma_{x_D}(\text{syst.}) = 0.0511\,\%$ and $\sigma_{y_D}(\text{syst.}) = 0.0227\,\%$.

Parameter	Generated Value	Mean Fitted Value	Expected Value
x_D (%)	-1.00	-1.0631 ± 0.0138	-1.0158 ± 0.0134
y_D (%)	-1.00	-1.0069 ± 0.0154	-1.0125 ± 0.0157

Table 6.5: Mean of Gaussian fit to distributions of fitted values for x_D and y_D from toys that have been generated with different turning point distributions for the prompt and secondary components. The last column is the biased value expected from the fit validation toys performed in the Section 6.1.

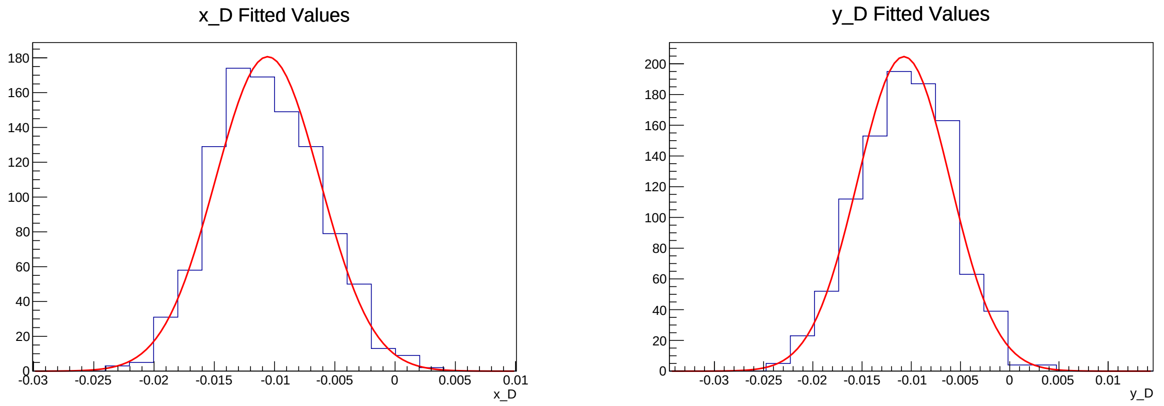
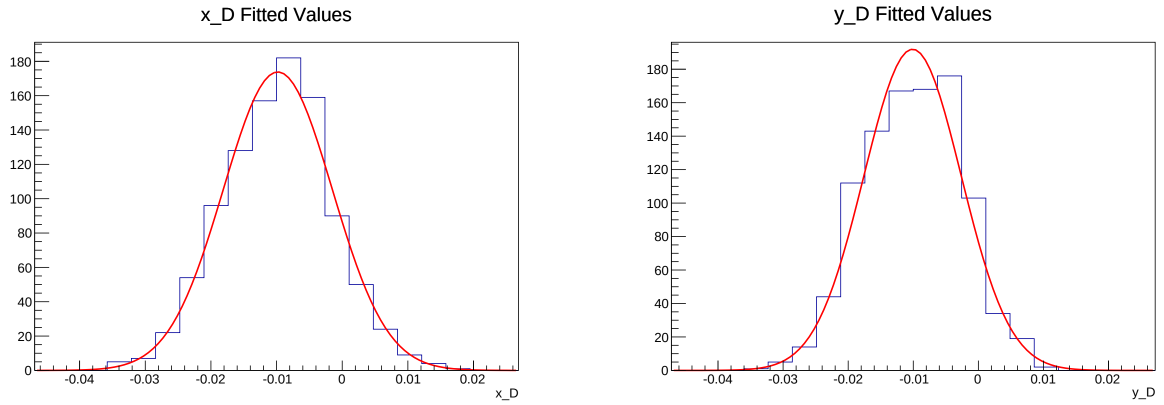


Figure 6.11: (a) Distribution of fitted values for x_D from 1000 toys that were generated with different turning point distributions for prompt and secondary components. (b) Equivalent plot for y_D.

6.5 Dalitz Plot Efficiency

No correction is applied to the data to account for the non-uniform efficiency as a function of position in the Dalitz plot. Instead, the effect this has on the extracted mixing parameters is included as a systematic uncertainty. This is done using toy Monte Carlo which has a Dalitz plot distribution generated according to the amplitude model described in [53] and provided by CLEO in bitmap form in [58]. The first step is to generate the invariant mass squared of each pair of daughters; m_{12}^2 and m_{13}^2, which determine in which Dalitz plot bin the event resides. The decay time, $ln(\chi^2_{IP})$ and turning points are then generated using the same generation process as for the other toys used in the fit validation studies. Note that only the right-π^+_{slow} signal component is generated for the purposes of this study.

Initially, a test sample of 1000 toys of 180k events each are generated assuming perfect Dalitz plot efficiency and with input mixing parameters $x_D = y_D = -1\,\%$. A fit is performed to each sample and the resulting fitted values of x_D and y_D are plotted in Figures 6.12 (a) and



(a) (b)

Figure 6.12: (a) Fitted values of x_D for toy MC generated assuming perfect Dalitz plot efficiency. (b) Equivalent plot for y_D.

(b), respectively. The values of c_i and s_i are allowed to float in the fit, but are constrained in the same way as for the fit to data. Since this is signal MC and not data it can have different values for c_i, s_i, T_i, and T_{-i}. Blindly plugging in the CLEO measured values could introduce a second systematic effect. Instead, the central values are extracted from the bitmap of the amplitude model used to generate the MC, ensuring that the input parameters accurately describe the strong phase variation in the generated data. The method used to extract these from the model is described in Appendix A. A Gaussian is fitted to each distribution with resulting means of $x_D = -0.9810 \pm 0.0267\,\%$ and $y_D = -1.0108 \pm 0.0244\,\%$. Both the means of the x_D and y_D distributions are consistent with their generated values within uncertainties. However, as we wish to understand the size of any effect relative to this ideal scenario the systematic uncertainties due to the addition of the Dalitz plot efficiency are measured with respect to the fitted values from these toys rather than from the true generated values.

A second set of toys is now generated, but with generated events filtered using the accept/reject technique according to the 2011 Dalitz plot efficiency model described in Section 4.8. This sculpts the Dalitz plot distribution of the toys in a way that mimics the behaviour in real data. The fits are then run using exactly the same configuration as for the previous set of toys with no attempt to correct for the Dalitz plot acceptance effects. The resulting distributions of fitted values for x_D and y_D are shown in Figures 6.13 (a) and (b). The fitted means of these distributions are $x_D = -0.9411 \pm 0.0298\,\%$ and

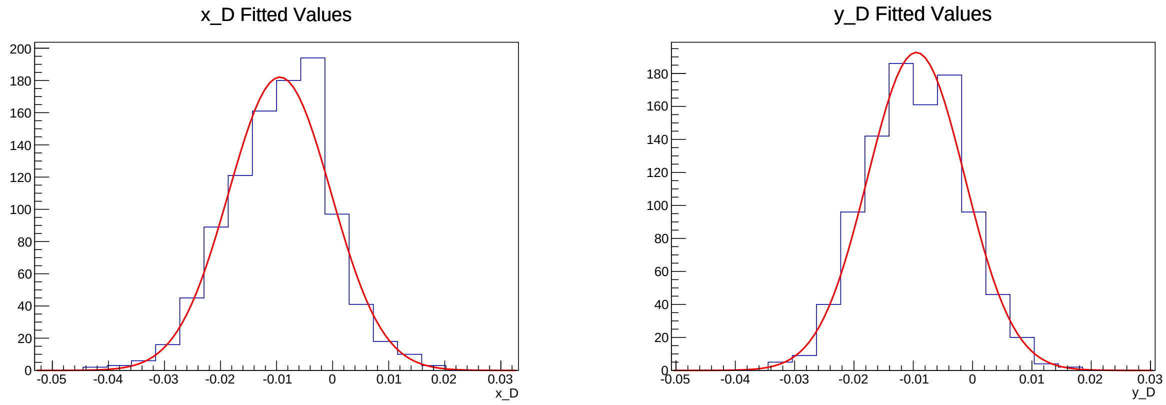


Figure 6.13: (a) Fitted values of x_D for data that has been filtered according to the 2011 Dalitz plot efficiency model. (b) Equivalent plot for y_D.

$y_D = -0.9495 \pm 0.0265\,\%$, respectively. Comparing these values with the central values found in the fits to MC with perfect Dalitz plot efficiency indicates a slight shift in the extracted mixing parameters of $\Delta x_D = +0.0399\,\%$ and $\Delta y_D = +0.0613\,\%$ as a result of the efficiency being applied. This gives an estimate of the size of the systematic uncertainty, but there are additional cross-checks that should be made.

The study conducted so far gives an estimate of the effect of the Dalitz plot efficiency assuming that the model used to describe the efficiency variation is accurate and that the efficiency is symmetric about the $m^2_{12} = m^2_{13}$ axis. To see what effect any potential mis-modelling of the efficiency can have, two approaches are taken. The first is to fit a more complicated efficiency model to the phase space MC (the 2011HO model described in Appendix B). No significant difference is observed between this and the preferred model. However, a further model (denoted the 2010 model), which was developed using a much older set of MC that had the trigger and stripping cuts applied by hand, is also used to filter events to see what effect this has on the toy data. Although it is known that this model is not an accurate description of the true efficiency seen in the data, it is useful to examine its effect on the mixing parameters to understand how sensitive they are to any mis-modelling of the Dalitz plot efficiency. The 2010 model is also briefly described in Appendix B. The resulting distributions of fitted mixing parameters after fitting to 1000 toys that were filtered by the older model are shown in Figures 6.14 (a) and (b). Fitting a Gaussian to each these distributions results in means of $x_D = -0.9310 \pm 0.0295\,\%$ and $y_D = -0.9543 \pm 0.0268\,\%$ which corresponds to

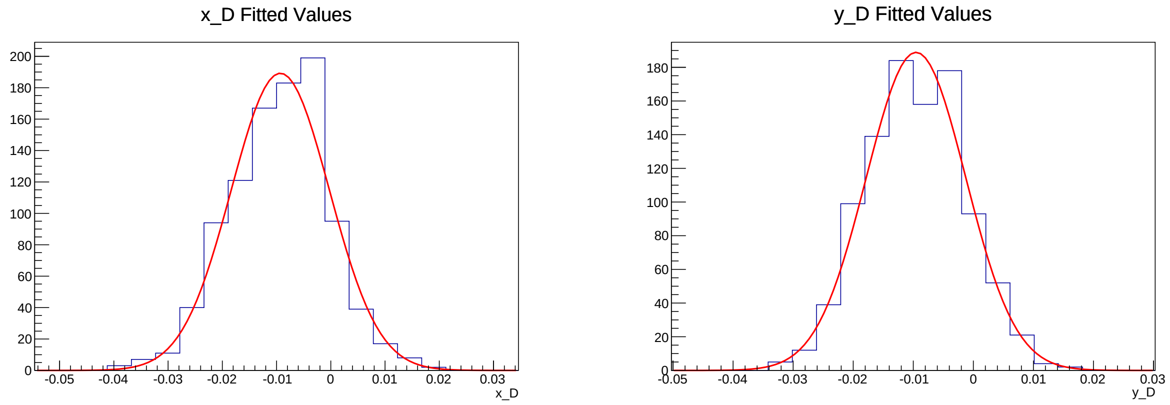


Figure 6.14: (a) Fitted values of x_D for data that has been filtered according to the older, alternative Dalitz plot efficiency model. (b) Equivalent plot for y_D.

shifts in the mixing parameters compared to the perfect efficiency case of $\Delta x_D = +0.0500\,\%$ and $\Delta y_D = +0.0574\,\%$.

Finally, the effect of any potential asymmetry in the Dalitz plot efficiency is also investigated. The more complicated version of the preferred model (2011HO) is fitted to the phase space MC with a subset of the parameters allowed to float independently either side of the lead diagonal. The resulting asymmetric model is denoted the 2011ASYM model (also described in Appendix B). This gives the model freedom to describe both halves of the Dalitz plot individually. Any significant asymmetry could affect the ratio T_i/T_{-i} and undermine one of the fundamental assumptions made in the fit. Using the 2011ASYM model and filtering the toy MC in the same way as for the other models an additional 1000 toy studies are conducted. The resulting distributions of x_D and y_D are shown in Figures 6.15 (a) and (b), respectively. The fitted means are $x_D = -0.9341 \pm 0.0284\,\%$ and $y_D = -0.9867 \pm 0.0267\,\%$ which corresponds to shifts in the mixing parameters compared to the perfect efficiency case of $\Delta x_D = +0.0469\,\%$ and $\Delta y_D = +0.0241\,\%$.

The results for each efficiency model are listed in Table 6.6. It is clear that the variation of the mixing parameters is relatively consistent regardless of the chosen efficiency parameterisation. This gives confidence that any mis-modelling of the efficiency will not affect the overall size of the systematic effect estimated using these toys. Therefore, the final systematic uncertainty is taken as the sum in quadrature of the size of the shift observed from the

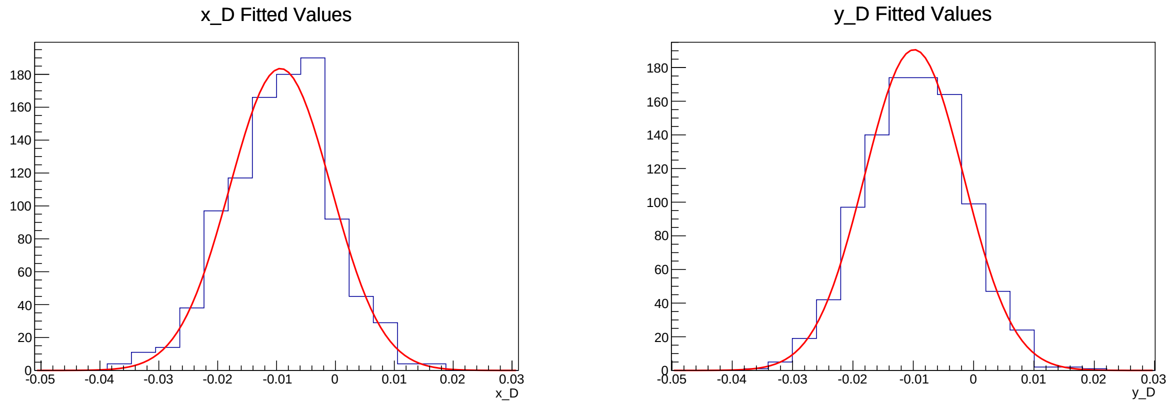


Figure 6.15: (a) Fitted values of x_D for data that has been filtered according to the asymmetric Dalitz plot efficiency model. (b) Equivalent plot for y_D.

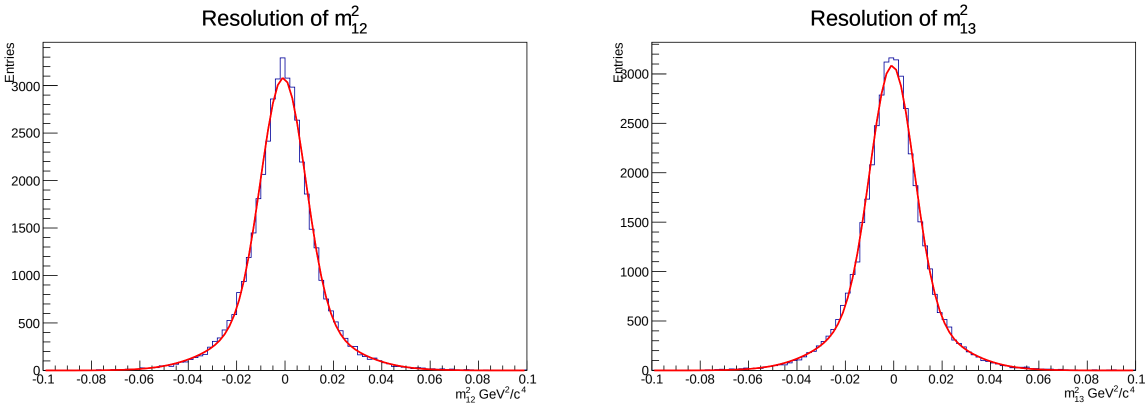
preferred model and its uncertainty as this model most accurately describes the efficiency variation in real data. This corresponds to systematic uncertainties of $\sigma_{x_D}(\text{syst.}) = 0.0565\,\%$ and $\sigma_{y_D}(\text{syst.}) = 0.0711\,\%$, respectively.

Value	Uniform Efficiency	2011	2010	2011ASYM
x_D (%)	-0.9810 ± 0.0267	-0.9411 ± 0.0298	-0.9310 ± 0.0295	-0.9341 ± 0.0284
y_D (%)	-1.0108 ± 0.0244	-0.9495 ± 0.0265	-0.9543 ± 0.0268	-0.9867 ± 0.0267
Δx_D (%)	N/A	$+0.0399$	$+0.0500$	$+0.0469$
Δy_D (%)	N/A	$+0.0613$	$+0.0574$	$+0.0241$

Table 6.6: Mean fitted parameters and their difference for each of the Dalitz plot efficiency toy studies.

6.6 Invariant Mass Resolution

The reconstructed invariant mass squared of both the $K^0_\mathrm{S}\,\pi^+$ and $K^0_\mathrm{S}\,\pi^-$ pairs are used to determine in which Dalitz plot bin an event belongs. If the invariant mass resolution is large or if any bias is present in the reconstructed values then events may migrate from one bin to another and bias the physics observables. In addition to an overall resolution there may be differences in the quality of the invariant mass reconstruction as a function of position in the Dalitz plot. This is likely to have greatest effect in bins that cover regions close to the kinematic boundary of the Dalitz plot where one or more particles have low momentum and are subsequently less well reconstructed. As `DecayTreeFitter` (DTF) is used to re-fit the



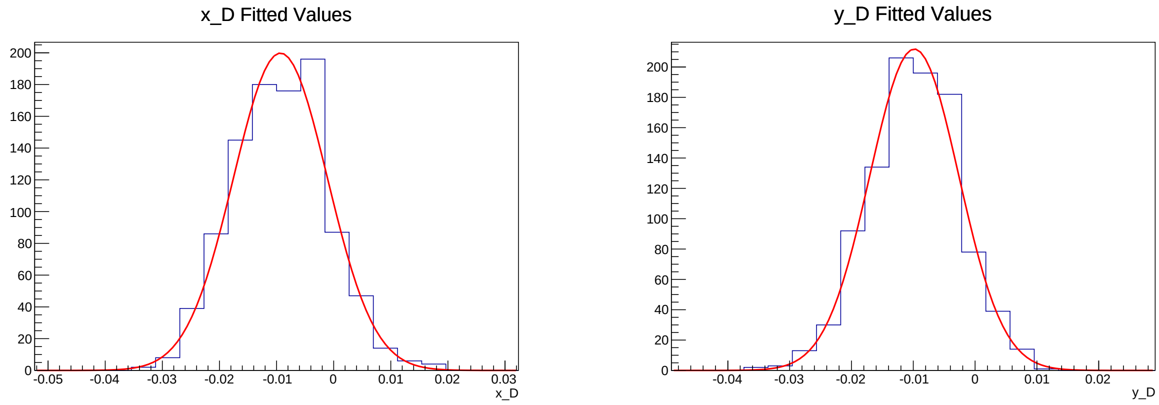
(a) (b)

Figure 6.16: Invariant mass squared resolution of (a) m^2_{12} and (b) m^2_{13} measured using signal MC for prompt signal events.

daughter momenta it is expected that the invariant mass resolution will be small. However, the use of DTF could potentially induce other biasing effects that may need accounting for.

The invariant mass resolution is examined for prompt signal using truth-matched signal MC. Plotting the difference between the reconstructed and true invariant mass squared gives an idea of the size of the resolution. Figure 6.16 shows the total resolution using all data for m^2_{12} and m^2_{13}. A double Gaussian is fitted to the resulting distribution. DTF has the effect of "squeezing" the resolution (as is clearly seen when applied in the calculation of Δm, see Figure 4.1) which is why the distribution is not perfectly Gaussian. The weighted means and sigmas of the two Gaussians for each of the Dalitz plot variables are $\mu_{m^2_{12}} = (-0.85 \pm 0.22) \times 10^{-3}$ GeV$^2/c^4$ and $\sigma_{m^2_{12}} = (-11.40 \pm 0.38) \times 10^{-3}$ GeV$^2/c^4$, and $\mu_{m^2_{13}} = (-0.81 \pm 0.21) \times 10^{-3}$ GeV$^2/c^4$ and $\sigma_{m^2_{13}} = (-11.40 \pm 0.36) \times 10^{-3}$ GeV$^2/c^4$.

To test whether there is any significant variation in the invariant mass resolution across the individual Dalitz plot bins a single Gaussian is fitted to the distribution of the difference between the true invariant mass squared and the reconstructed value in each Dalitz plot bin. A single Gaussian was adopted as the fit became unstable with the double Gaussian parameterisation due to low statistics in a handful of the Dalitz plot bins. The resulting fitted parameters for each Dalitz plot bin are given in Table 6.7 with the resulting plots for m^2_{12} and m^2_{13} given in Figures 6.18 and 6.19, respectively.



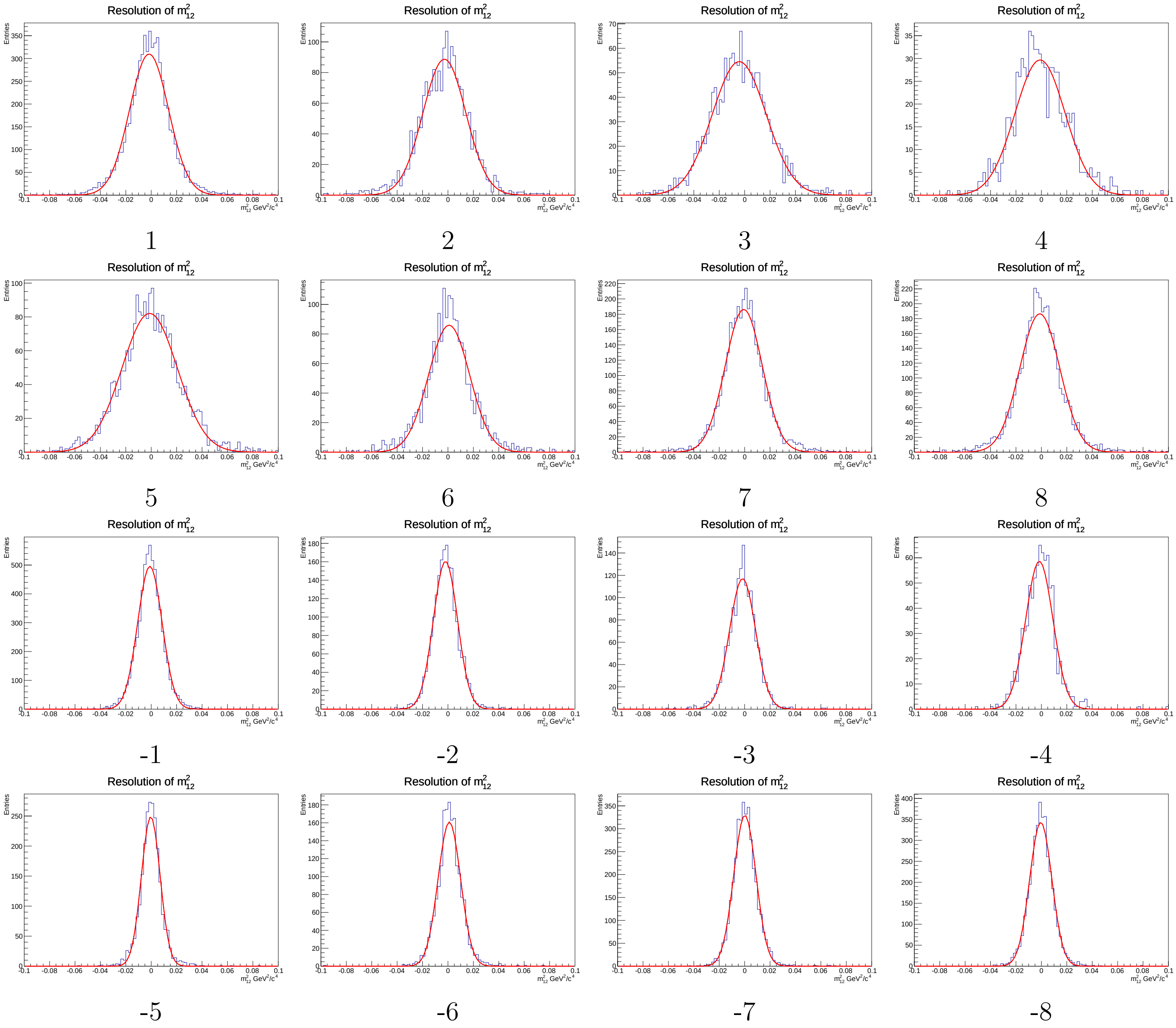
(a) (b)

Figure 6.17: (a) Distribution of fitted values of x_D from toys that were generated with smearing applied to the invariant mass parameters m_{12}^2 and m_{13}^2. (b) Equivalent plot for y_D.

To evaluate whether the invariant mass resolution has any significant effect on the mixing parameters a toy study is conducted. 1000 toy samples containing 180k events each are generated in the same way as for the Dalitz plot efficiency study outlined in the previous section, but with the invariant mass smeared according to the resolution measured in each Dalitz plot bin. The “true”, generated invariant mass is used to determine in which Dalitz plot bin the event belongs. Then the appropriate parameterisation from Table 6.7 is used to smear the Dalitz plot co-ordinates and the Dalitz plot bin association is re-run. This means that events which are smeared outside of the acceptance will be recaptured by the bin finding algorithm described in Section 4.4. A Gaussian is fitted to the distribution of fitted values for x_D and y_D taken from the toys (see Figures 6.17 (a) and (b)). The means of the Gaussian fits are $x_D = -0.9323 \pm 0.0266\,\%$ and $y_D = -0.9838 \pm 0.0234\,\%$. A systematic uncertainty equal to the sum of the difference between the mean fitted and generated values plus its uncertainty taken in quadrature is assigned to both parameters resulting in $\sigma_{x_D}(\text{syst.}) = 0.0727\,\%$ and $\sigma_{y_D}(\text{syst.}) = 0.0284\,\%$.

m_{12}^2			m_{13}^2		
Bin	μ GeV$^2/c^4$	σ GeV$^2/c^4$	Bin	μ GeV$^2/c^4$	σ GeV$^2/c^4$
1	-0.00169 ± 0.00020	0.0154 ± 0.0002	1	-0.00118 ± 0.00013	0.0098 ± 0.0001
2	-0.00278 ± 0.00039	0.0169 ± 0.0003	2	-0.00192 ± 0.00022	0.0096 ± 0.0002
3	-0.00427 ± 0.00056	0.0212 ± 0.0005	3	-0.00230 ± 0.00028	0.0104 ± 0.0002
4	-0.00102 ± 0.00075	0.0197 ± 0.0007	4	-0.00083 ± 0.00040	0.0109 ± 0.0004
5	-0.00125 ± 0.00046	0.0218 ± 0.0004	5	-0.00029 ± 0.00016	0.0076 ± 0.0001
6	0.00091 ± 0.00039	0.0162 ± 0.0005	6	0.00077 ± 0.00020	0.0086 ± 0.0002
7	-0.00069 ± 0.00025	0.0146 ± 0.0002	7	0.00037 ± 0.00015	0.0087 ± 0.0001
8	-0.00125 ± 0.00026	0.0158 ± 0.0003	8	-0.00053 ± 0.00014	0.0087 ± 0.0001
-1	-0.00104 ± 0.00012	0.0096 ± 0.0001	-1	-0.00134 ± 0.00020	0.0152 ± 0.0002
-2	-0.00195 ± 0.00022	0.0095 ± 0.0002	-2	-0.00310 ± 0.00041	0.0173 ± 0.0004
-3	-0.00163 ± 0.00027	0.0102 ± 0.0002	-3	-0.00313 ± 0.00059	0.0221 ± 0.0004
-4	-0.00161 ± 0.00039	0.0107 ± 0.0003	-4	-0.00098 ± 0.00078	0.0209 ± 0.0007
-5	-0.00037 ± 0.00015	0.0073 ± 0.0002	-5	-0.00084 ± 0.00046	0.0221 ± 0.0004
-6	0.00110 ± 0.00021	0.0089 ± 0.0002	-6	0.00043 ± 0.00039	0.0164 ± 0.0004
-7	0.00018 ± 0.00015	0.0088 ± 0.0001	-7	-0.00031 ± 0.00024	0.0144 ± 0.0002
-8	-0.00047 ± 0.00014	0.0087 ± 0.0001	-8	-0.00135 ± 0.00027	0.0161 ± 0.0002

Table 6.7: Fitted parameters for invariant mass squared resolution in individual Dalitz plot bins.



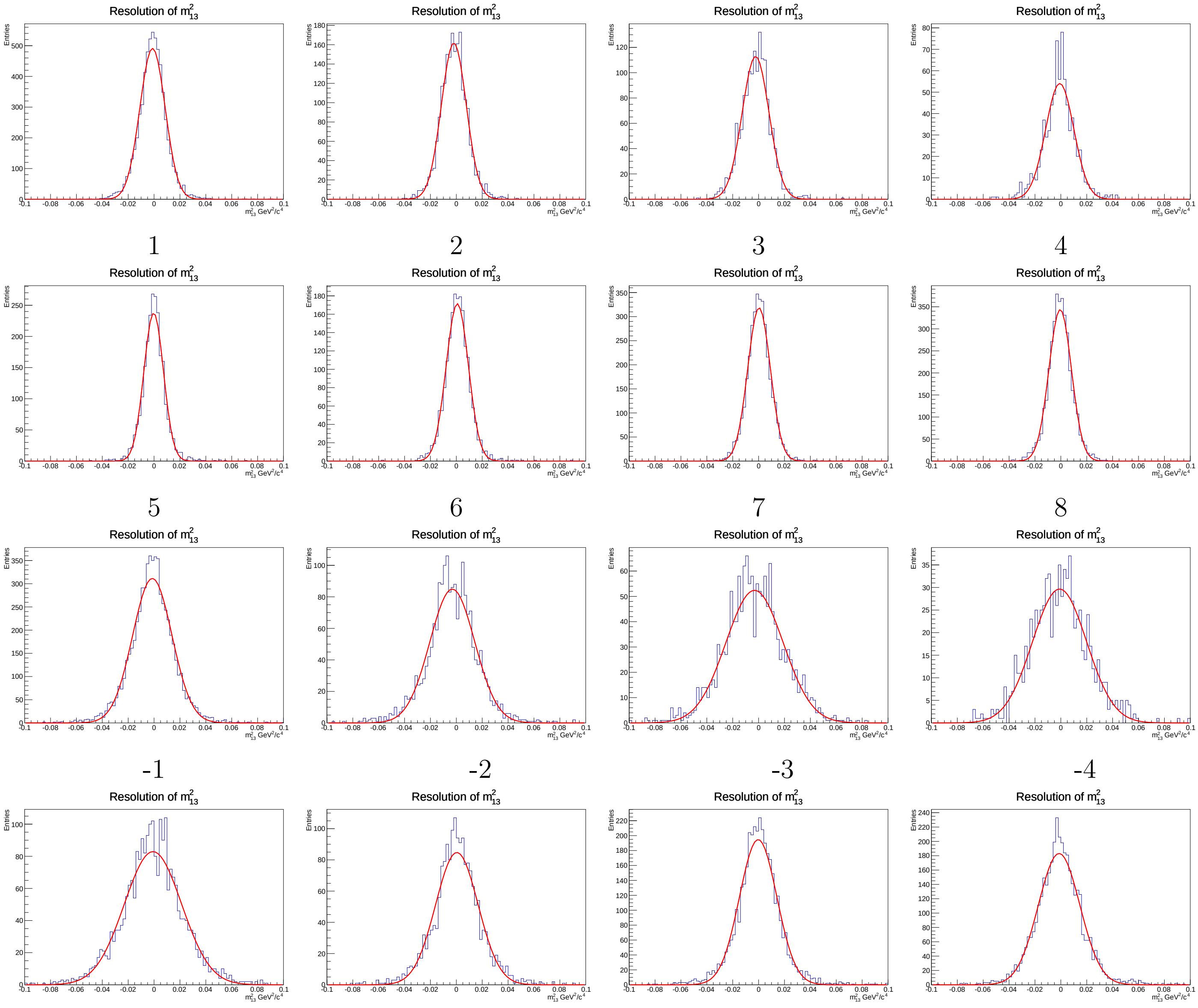
1 2 3 4

5 6 7 8

-1 -2 -3 -4

-5 -6 -7 -8

Figure 6.18: Fitted projections of invariant mass squared resolution for m^2_{12} in individual Dalitz plot bins.



1 2 3 4

5 6 7 8

-1 -2 -3 -4

-5 -6 -7 -8

Figure 6.19: Fitted projections of invariant mass squared resolution for m^2_{13} in individual Dalitz plot bins.

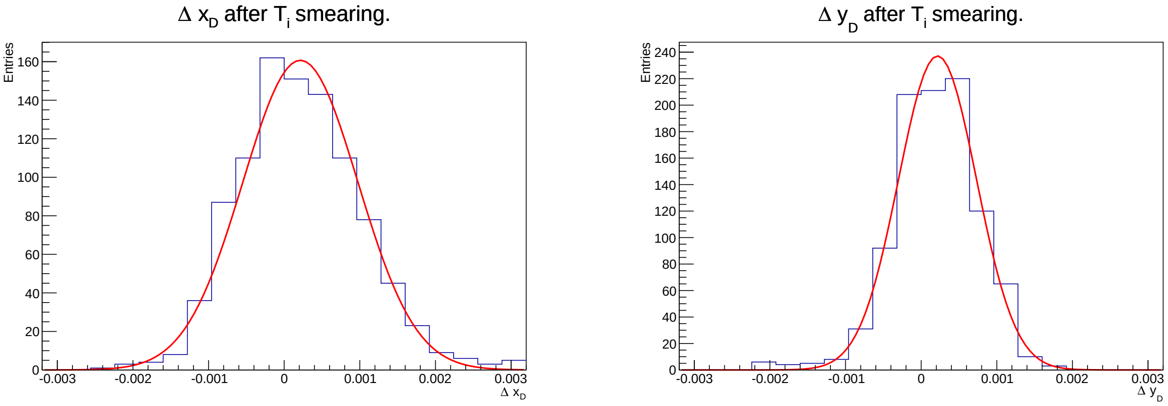
6.7 Uncertainty from T_i and T_{-i}

The uncertainty due to the CLEO measurements of T_i and T_{-i} is evaluated using real data. The complete fit to data is initially performed with T_i and T_{-i} set to the central values as measured by CLEO and the values of x_D and y_D are saved. The fit is then repeated 1000 times, but with the values of T_i and T_{-i} fixed to values that are Gaussian-smeared with a width equal the total error on each parameter as measured by CLEO. Although there is no correlation between the values of T_i, such that the smearing can be done independently, the smeared values are re-normalised such that the sum of all the T_i across all Dalitz plot bins is equal to 1. The values of x_D and y_D are saved for each fit and the difference between the smeared and un-smeared results plotted. The resulting plot should be centred around zero assuming that the smearing of T_i and T_{-i} does not bias the extracted mixing parameters.

The differences between the smeared and un-smeared values for x_D and y_D are shown in Figures 6.20 (a) and (b), respectively. The Gaussian fitted parameters—NB, not pulls—are $\mu = (0.0214 \pm 0.0026)\,\%$ and $\sigma = (0.0761 \pm 0.0019)\,\%$ for x_D and $\mu = (0.0216 \pm 0.0018)\,\%$ and $\sigma = (0.0516 \pm 0.0012)\,\%$ for y_D. The uncertainty on x_D and y_D from T_i and T_{-i} is taken as the fitted width of the distributions for each parameter. There is a slight bias on the extracted values of x_D and y_D. Although the biases are smaller than the total width of the distribution, they are significant and are accounted for in the final systematic. The systematic uncertainties assigned to x_D and y_D are then taken as the size of the width and the magnitude of the bias of each distribution summed in quadrature. This gives $\sigma_{x_D}(\text{syst.}) = 0.0791\,\%$ and $\sigma_{y_D}(\text{syst.}) = 0.0559\,\%$, respectively.

6.8 Per-track Efficiency Parameterisation

A linear parameterisation is chosen for the tracking efficiency as a function of DOCAZ^2 as it adequately describes the distributions shown in Figures 4.5 and is what has been adopted in previous analyses [126]. However, the fit quality with a linear function is not totally satisfactory, particularly for the K^0_S daughter tracks. Therefore, the fits to the efficiency



(a) (b)

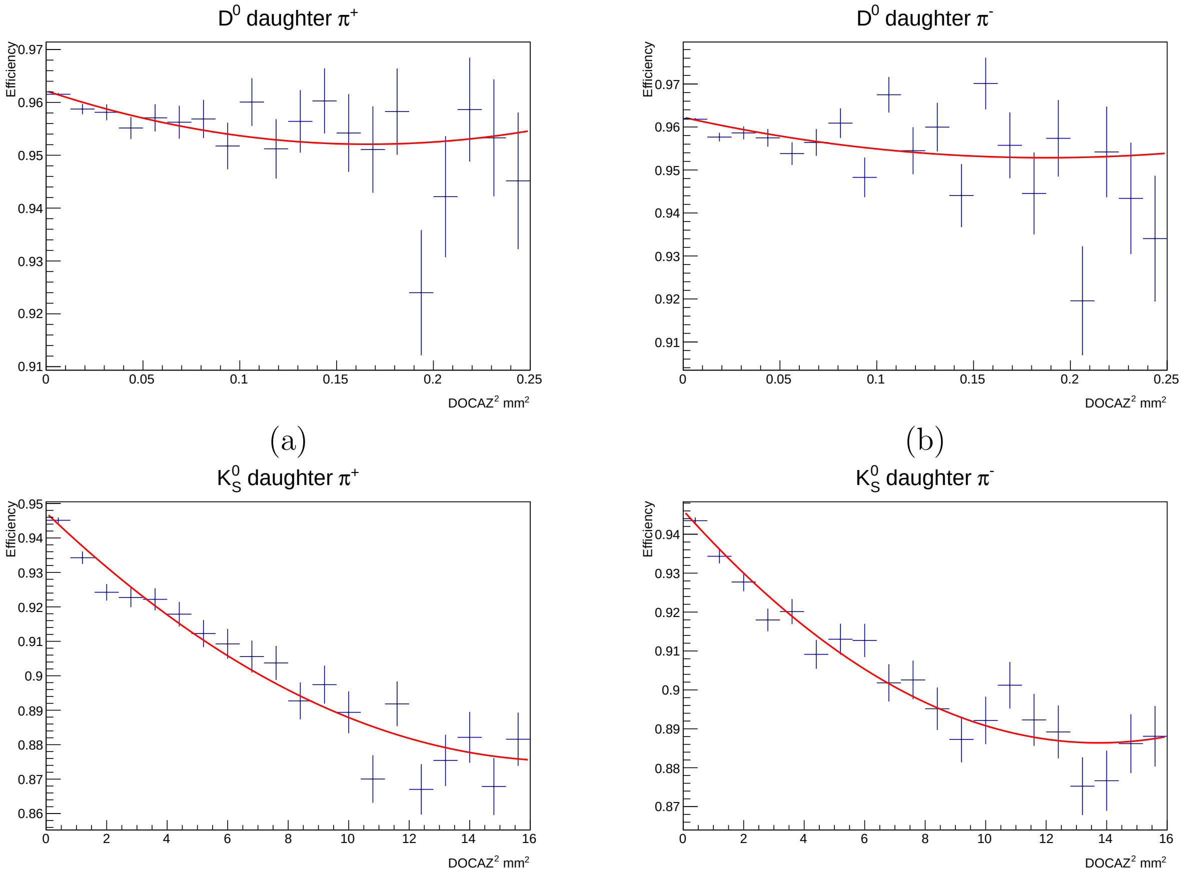
Figure 6.20: Distribution of the difference between fitted results for x_D (a) and y_D (b) when smearing the value of T_i and T_{-i}. The blind central value for x_D and y_D extracted from a fit with T_i and T_{-i} fixed to the CLEO measured values is subtracted from each fit result ensuring that this systematic check can still be performed blind.

Track	α	β mm^{-2}	γ mm^{-4}
D^0 daughter π^+	0.9622 ± 0.0003	-0.1213 ± 0.0298	0.3645 ± 0.1874
D^0 daughter π^-	0.9623 ± 0.0003	-0.1001 ± 0.0301	0.2669 ± 0.1940
K^0_S daughter π^+	0.9473 ± 0.0009	-0.0084 ± 0.0006	$0.0002 \pm 4.6\times10^{-5}$
K^0_S daughter π^-	0.9461 ± 0.0009	-0.0087 ± 0.0006	$0.0003 \pm 4.6\times10^{-5}$

Table 6.8: Alternative fitted parameters for efficiency as a function of DOCAZ^2 for each daughter track.

as a function of DOCAZ^2 are again performed for each track using a 2$^{\text{nd}}$ order polynomial. Figure 6.21 shows the fit results. It is clear from Figures 6.21 (a) and (b) that the deviation from a straight line for the D^0 daughter tracks is small and a linear approximation is sufficient. However, Figures 6.21 (c) and (d) show that the efficiencies for the K^0_S daughter tracks do indeed have a more quadratic like dependence on DOCAZ^2. The fitted parameters for each Track are listed in Table 6.8 where the polynomials are described as $\alpha + \beta t + \gamma t^2$.

To see if this alternative description affects the resulting lifetime and mixing parameters, both the lifetime and mixing fits are re-run with the quadratic efficiency parameterisation for both the K^0_S daughter tracks (the D^0 daughter track efficiencies remain linear). The fitted values of each parameter and the difference between the fitted values of each from the linear and quadratic efficiency parameterisations are summarised in Table 6.9. The shift in central value for all parameters is smaller than their respective statistical uncertainties. The



(a) (b)

(c) (d)

Figure 6.21: Alternative fits to efficiency as a function of DOCAZ^2 for each daughter track. In each case a 2$^{\text{nd}}$ order polynomial is used.

variation in the mixing parameters is also small and will not have a large effect on the total systematic uncertainty. However, for consistency, an uncertainty equal to the size of the shift is assigned to both parameters giving $\sigma_{x_D}(\text{syst.}) = 0.0145\,\%$ and $\sigma_{y_D}(\text{syst.}) = 0.0245\,\%$.

Parameter	Linear K^0_S Efficiency	Quadratic K^0_S Efficiency	Difference
τ (fs)	410.87 ± 1.14	410.00 ± 1.11	-0.87
Γ (MeV)	2.4351 ± 0.0069	2.4359 ± 0.0068	-0.0008
x_D (%)	-0.8540 ± 0.5170	-0.8395 ± 0.5199	-0.0145
y_D (%)	-0.0535 ± 0.4880	-0.0780 ± 0.4867	$+0.0245$

Table 6.9: Difference between the fitted values of x_D and y_D when using linear and quadratic parameterisations for the per-track efficiency of the K^0_S daughters.

6.9 Dalitz Plot Binning

The Dalitz plot binning scheme is provided by CLEO in [58]. It is defined using a grid of micro-bins which are 0.0054×0.0054 GeV$^2/c^4$ in area. Although constraining the D^0 mass brings most of the data points within the kinematic boundary of the binning scheme, due to resolution effects there are some events which still remain outside of the kinematically allowed region. To maximise the number of events in the final sample a scan is performed for each event that sits outside the Dalitz plot boundary and the index of the closest micro-bin is assigned to that event (discussed in more detail in Section 4.4). This procedure is only applied to a small fraction of events and it is unlikely that the use of the bin-finding will have a significant effect on the extracted mixing parameters. However, to err on the side of caution, the fit to data is run over a sample where events that sit outside of the Dalitz plot acceptance are rejected and the values of x_D and y_D are compared with the regular fit. The distribution of rejected candidates can be seen in Figure 6.22. There is an observed shift in x_D of $-0.0120\,\%$ and a shift in y_D of $+0.0061\,\%$. The magnitude of the shift is assigned as a systematic uncertainty on each parameter.

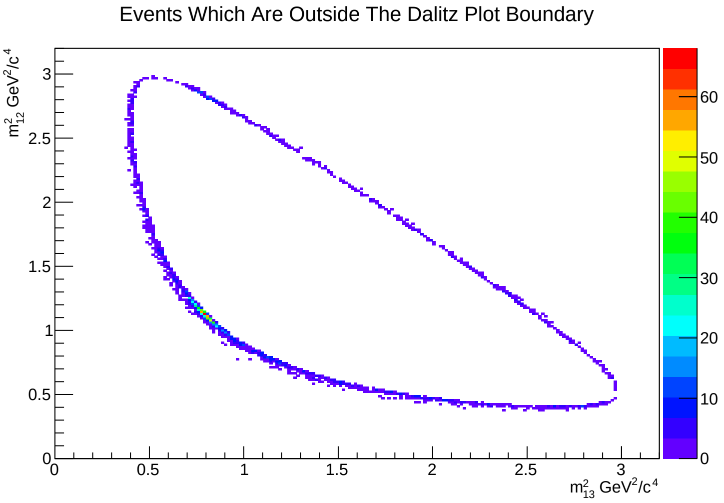


Figure 6.22: Events which remain outside of the Dalitz plot boundary despite having their Dalitz plot coordinates calculated from the re-fitted variables after constraining the D^0 mass. The bin finding algorithm captures all of these events for the fit. The slightly larger number of events outside of the lower edge are likely radiative D^0 decays.

6.10 Treatment of Individual Fit Components

6.10.1 Uncertainties from the m vs. Δm Fits

The mass vs. Δm fits are used to get a handle on the relative contribution from each class of event in each Dalitz plot bin i.e. the class probabilities $P(signal)$, $P(wrong\text{-}\pi^+_{slow})$ and $P(comb)$ (see Table 5.3). These values are fixed in the fit, but have uncertainties (although small) that need to be accounted for. To do this, the full mixing fit to data is run 500 times with the contribution from each class in each bin Gaussian-smeared within its uncertainty for every fit. The class probabilities are then re-normalised such that the sum of the smeared values is equal to 1. The difference between the values of the mixing parameters found in each smeared fit is subtracted from the un-smeared value and the resulting distribution plotted for each of the mixing parameters. A Gaussian is fitted to each distribution and the sum of the resulting width and its uncertainty taken in quadrature yields systematic uncertainties of $\sim 0.001\,\%$ for both x_D and y_D. As expected, the uncertainties associated with the class probabilities have very little influence on the mixing results and no additional systematic uncertainty is assigned.

6.10.2 Uncertainties from the Lifetime Fit

A large proportion of the parameters used to describe the $ln(\chi^2_{IP})$ and decay time distributions for the right-π^+_{slow} and wrong-π^+_{slow} components are fixed from the lifetime fit (described in Section 5.4) across all Dalitz plot bins. These parameters have uncertainties that need to be accounted for[1]. To assess this contribution to the overall uncertainty, the full mixing fit is run 750 times with the input parameters smeared within their uncertainties. The smearing is done such that it accounts for the covariance between the parameters (the full covariance matrix is given in Appendix C). The difference between the values of the mixing parameters found in each smeared fit is subtracted from the un-smeared value and the resulting distribution plotted for each of the mixing parameters. A Gaussian is fitted to each distribution giving $\mu = 0.0095 \pm 0.0007\,\%$ and $\sigma = 0.0176 \pm 0.0007\,\%$ for x_D and $\mu = 0.0152 \pm 0.0018\,\%$ and $\sigma = 0.0398 \pm 0.0012\,\%$ for y_D. Given there are significant shifts in the central values for both x_D and y_D that are sizeable with respect to the width, these are accounted for in the final systematic. Taking the sum in quadrature of the resulting width and the magnitude of the shift as the size of the systematic effect yields uncertainties of $\sigma_{x_D}(\text{syst.}) = 0.0200\,\%$ and $\sigma_{y_D}(\text{syst.}) = 0.0426\,\%$.

6.10.3 Combinatoric Background

To test how susceptible the extracted mixing parameters are to differences in the parameterisation of the combinatoric background, the fit is performed with the combinatoric shapes taken from data in the upper and lower mass sidebands separately and compared with the results from using both together. This also gives an indication as to whether the small contamination from $K^0_S K^{\mp}\pi^{\pm}$ and $K^0_S\pi^+\pi^-\pi^0$ decays in the lower mass sideband, investigated in Section 4.9.1, has an effect on the measured rate of mixing. This test is done with the mixing parameters blind. The results are summarised in the upper half of Table 6.10. Taking the average of the unsigned differences for each parameter as a systematic uncertainty gives

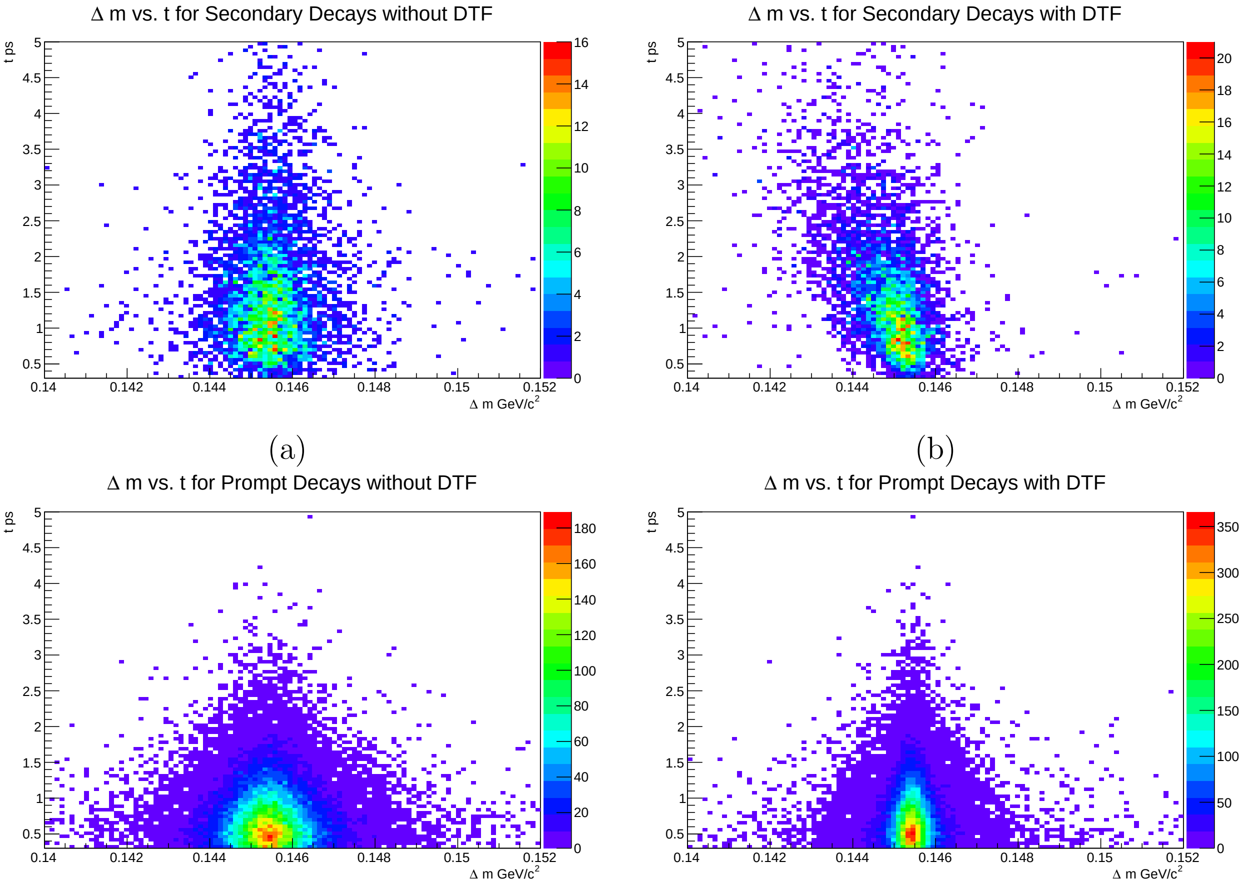
---

[1] Additionally, there can be effects due to the variation of these parameters across the Dalitz plot. These are addressed in Section 6.11.

$\sigma_{x_D}(\text{syst.}) = 0.0884\,\%$ and $\sigma_{y_D}(\text{syst.}) = 0.0504\,\%$. The difference between the extracted parameters in each case is much smaller than the statistical uncertainty, but larger than expected given the low level of combinatorics. However, a lack of combinatoric data may be the reason why such a large deviation is seen in the data. The kernel PDFs used to model the combinatoric data are very good at modelling statistical fluctuations. In the tests conducted the number of events in the sidebands for each Dalitz plot bin is reduced by a factor of 2 making the total sideband yield roughly half the total combinatoric yield in the signal window. In addition, the Dalitz plot distributions for the upper and lower sidebands are different. This means that the number of events in the sidebands of each Dalitz plot bin will also differ in each fit. The result is a potential inflation of any statistical fluctuations that the kernel PDFs model in the fits using individual sidebands and may go some way to explaining the large systematic effect. To test this, the fits are re-run with the sideband sample split by Dalitz plot bin, but not D^0 flavour. This has the effect of roughly doubling the statistics of the sideband samples used to model the combinatoric distributions. The results are listed in the lower half of Table 6.10. The average, unsigned deviation for x_D is $0.0337\,\%$ which is smaller than in the previous case. The equivalent value for y_D is $0.0534\,\%$ which is comparable to the previous case. These results indicate that there is a surprisingly strong dependence of the extracted mixing parameters on the choice of parameterisation for the combinatoric component in each Dalitz plot bin. Taking the average unsigned shift from all four fits for each parameter results in systematic uncertainties of $\sigma_{x_D}(\text{syst.}) = 0.0610\%$ and $\sigma_{y_D}(\text{syst.}) = 0.0519\,\%$. Although these values are a bit large for comfort (and it is not totally clear why given the small contribution from combinatoric background), they are still much smaller than the statistical uncertainty on both parameters and can be accommodated as systematic uncertainties.

6.10.4 Secondary D Decays

There are several sources of systematic uncertainty due to the presence of secondary charm decays. Firstly, it is assumed that the level of secondary contamination does not change sig-



(a) (b) (c) (d)

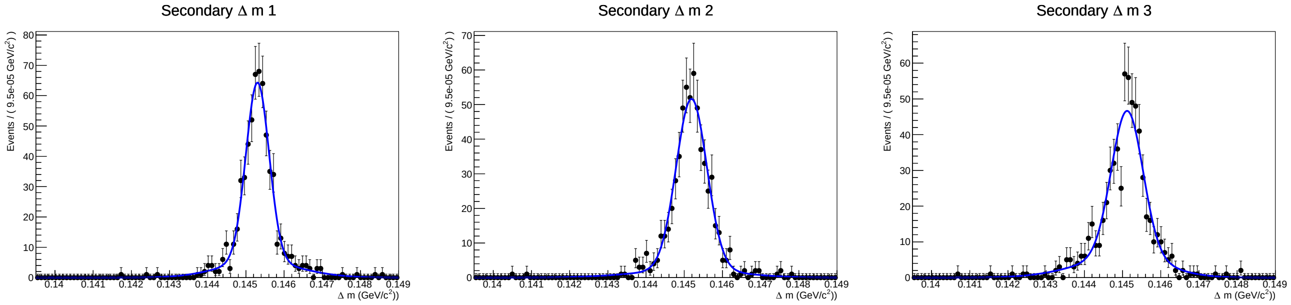
Figure 6.23: (a) Δm vs. decay time for matched, secondary signal MC without using re-fitted variables to calculate Δm. (b) Same as (a), but using re-fitted variables to calculate Δm. (c) and (d) are the same plots, but for prompt decays.

Parameter	Result Both	Difference Lower	Difference Upper
Split D^0 and $\overline{D}^0$ sideband data			
x_D (%)	-0.8450 ± 0.5170	-0.0764	$+0.1003$
y_D (%)	-0.0535 ± 0.4880	-0.0738	$+0.0269$
Combined D^0 and $\overline{D}^0$ sideband data			
x_D (%)	-1.0138 ± 0.5067	$+0.0075$	-0.0598
y_D (%)	$+0.0577 \pm 0.4823$	-0.0125	$+0.0943$

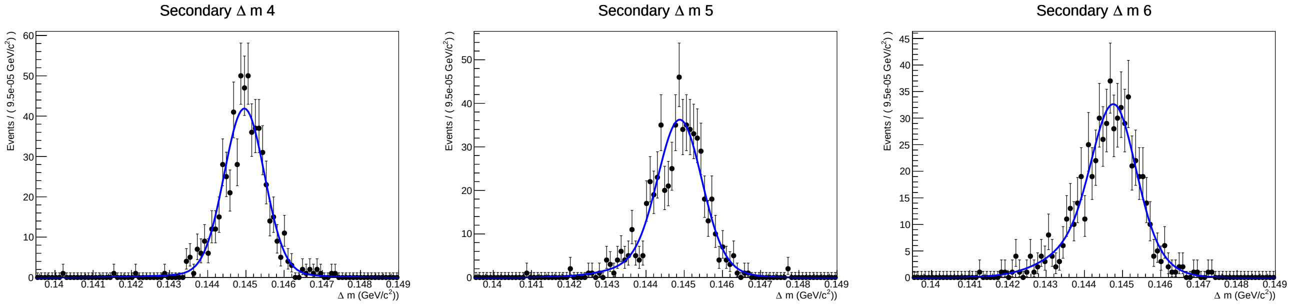
Table 6.10: Variation in extracted parameters for fits where the combinatoric parameterisation is modelled on data from different regions in D^0 mass.

nificantly across the Dalitz plot bins. The effect of fixing this is investigated in Section 6.11. In addition, the parameters describing the secondary $ln(\chi^2_{IP})$ and decay time distributions are fixed from the lifetime fit. These parameters carry an uncertainty which is included in Section 6.10.2. There is one other systematic effect which is more subtle and stems from the use of `DecayTreeFitter` to constrain the soft pion to the primary vertex and re-calculate the value of Δm using the re-fitted kinematic variables. As Δm is not well defined for secondary decays, this can induce a time-dependent bias in the Δm distribution for secondary decays. Figure 6.23 (a) shows the distribution in signal MC of Δm vs. t for matched secondary events where Δm is calculated without any re-fitted variables. Figure 6.23 (b) is the same plot, but where Δm has been calculated with the re-fitted variables. It is clear that the distribution of Δm calculated using the re-fitted variables has a time-dependence for secondary decays. This could lead to the fit under or over estimating the size of the secondary background and may bias the extracted mixing parameters. Note that this does not affect prompt decays as can be clearly seen from Figures 6.23 (c) and (d). To test this a series of toy studies is conducted where the Δm distribution for the right-π^+_{slow} secondary component is generated with a time dependence which is subsequently ignored in the fit. Note that this may have a small effect on secondary decays which have a non-peaking structure in Δm, but as the contamination due to wrong-π^+_{slow} secondary decays is less than 1 % it can safely be ignored.

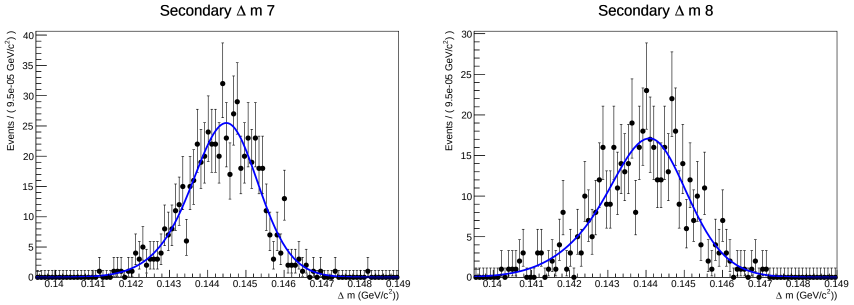
Both the mean and the width of the right-π^+_{slow} secondary Δm distribution evolve as a function of decay time. The parameterisation of the time evolution is taken from the secondary matched signal MC shown in Figure 6.23 by fitting to the Δm distribution in bins



(a) 0 - 0.71 ps (b) 0.71 - 0.92 ps (c) 0.92 - 1.14 ps



(d) 1.14 - 1.38 ps (e) 1.38 - 1.67 ps (f) 1.67 - 2.07 ps

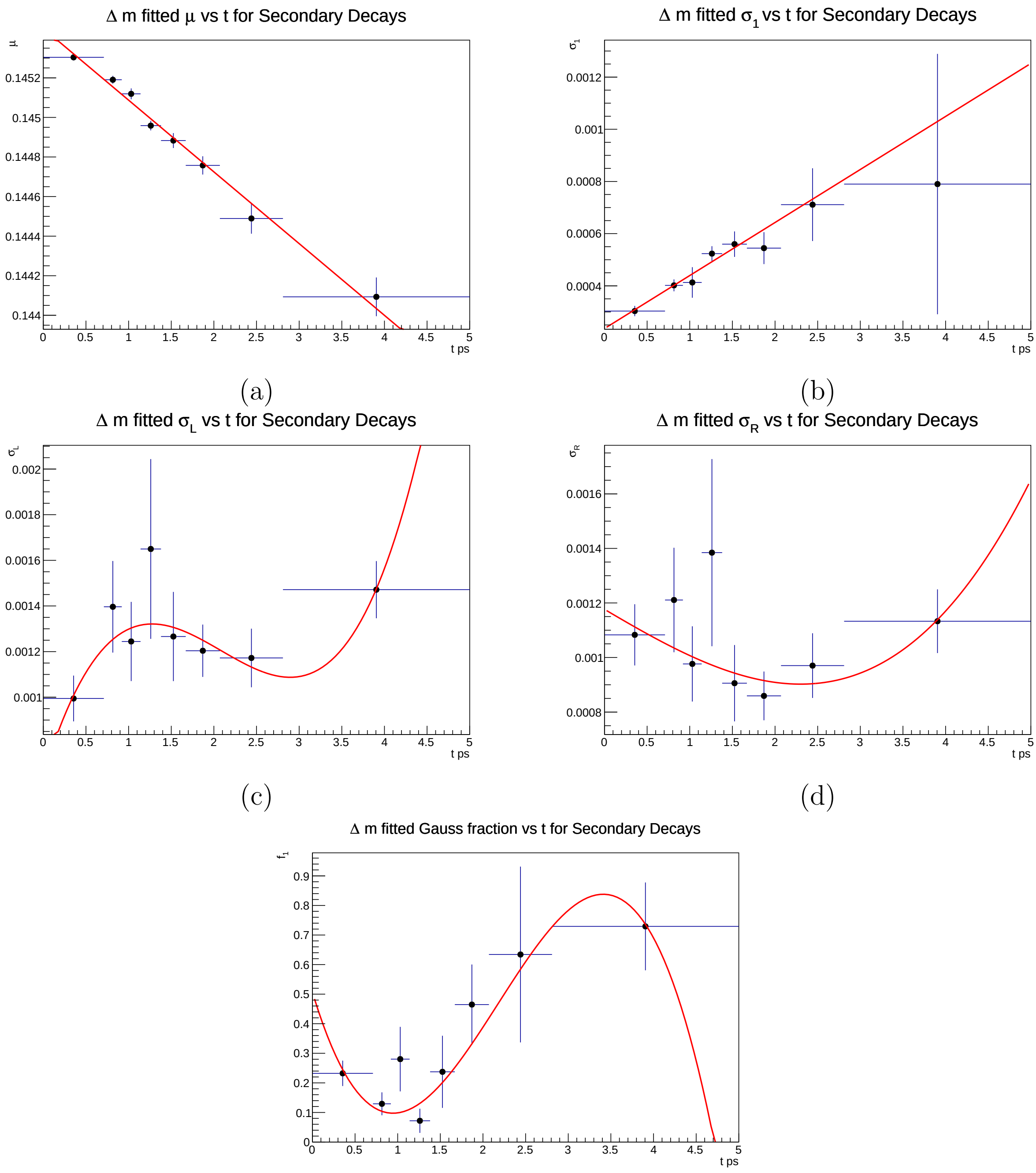


(g) 2.07 - 2.81 (h) 2.81 - 5.00 ps

Figure 6.24: Fits to Δm in slices of reconstructed decay time from secondary signal MC.

of time. It is assumed that this time-dependence does not vary across the Dalitz plot given that the kinematics of the soft pion and the D are largely uncorrelated. It is also assumed that the size of any systematic due to some global time dependence on Δm for secondaries will be much larger than any additional effect due to the time dependence varying across the Dalitz plot. Figures 6.24 (a)–(h) show the Δm projections in slices of reconstructed decay time. The PDF used to describe the peaking Δm component for secondaries in data is given in Equation 5.40 and takes the form of a double Gaussian plus a bifurcated Gaussian. Due to the low statistics of the secondary MC a simpler model consisting of a single Gaussian plus a bifurcated Gaussian is used to prevent instabilities in the fit. The values of the parameters in each time bin are plotted and a fit is performed to each of the resulting distributions (see Figures 6.25 (a)–(e)). The time evolution of the mean, μ and width of the single Gaussian, σ_1 are fitted with linear functions of the form $a + bt$. The evolution for the left and right hand widths of the bifurcated Gaussian, σ_L and σ_R, respectively, and the fit fraction of each Gaussian, f_1 are fitted with 3$^{\text{rd}}$ order polynomial functions of the form $a+bt+ct^2+dt^3$. The fitted parameters are listed in Table 6.11. Figures 6.25 (c) and (e), and to some extent (d), are likely not accurately modelling the time dependence of these parameters in the region $t > 4$. However, this is not too much of a concern given that there is little data out at high decay times and these models are used to estimate the size of an effect, not correct for it in data. It should also be noted that all σ_i and f_1 are forced to be positive definite. In addition to performing the fit in slices of time, a time integrated fit is also performed. The fitted projection and parameters are given in Figures 6.26 (a) and Table 6.12, respectively. It is noted that the fit quality in Figure 6.26 (a) is relatively poor. Adding an additional Gaussian as is used in the fits to data improves things significantly as can be seen in Figure 6.26 (b). However, the time integrated Δm parameterisation is simply used to generate a sample with no time-dependent Δm for secondaries which is used to verify the fit procedure. Therefore, a totally accurate description of the distribution from MC is not required, and the simpler parameterisation is used for consistency.

Two sets of 1000 toys each containing 280k events are generated. The first contains secondary events that have been generated with no Δm time-dependence whereas the sec-



(a)

(b)

(c)

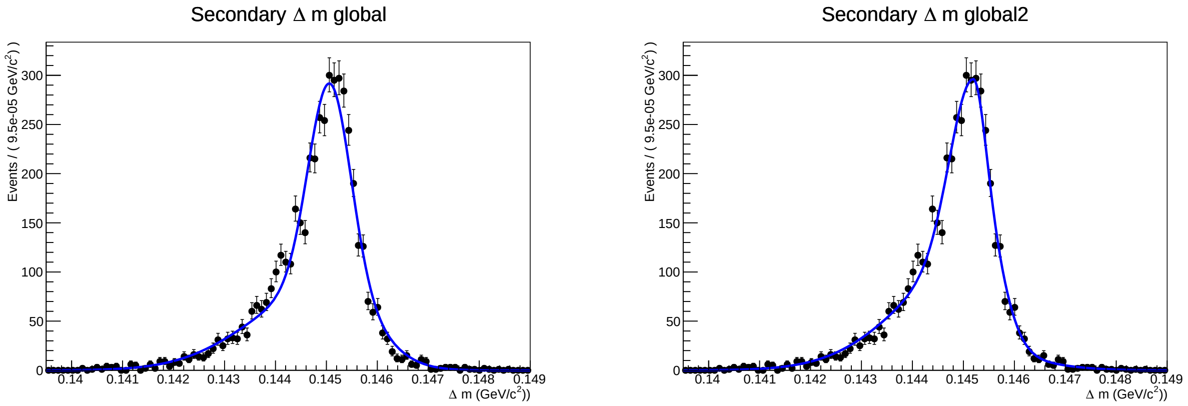
(d)

(e)

Figure 6.25: Time evolution of fitted parameters for secondary Δm distributions given in figures 6.24 (a)–(e)

	a	b	c	d
μ	0.14545 ± 0.00002	-0.00036 ± 0.00002	N/A	N/A
σ_1	0.00024 ± 0.00003	0.00020 ± 0.00003	N/A	N/A
σ_L	0.00066 ± 0.00024	0.00122 ± 0.00059	-0.00069 ± 0.00036	0.00011 ± 0.00006
σ_R	0.00118 ± 0.00024	-0.00019 ± 0.00054	0.00001 ± 0.00032	0.00001 ± 0.00005
f_1	0.507 ± 0.118	-0.955 ± 0.326	0.646 ± 0.241	-0.099 ± 0.043

Table 6.11: Fitted parameters describing the time evolution of the secondary Δm parameterisation.


Secondary Δ m global

Secondary Δ m global2

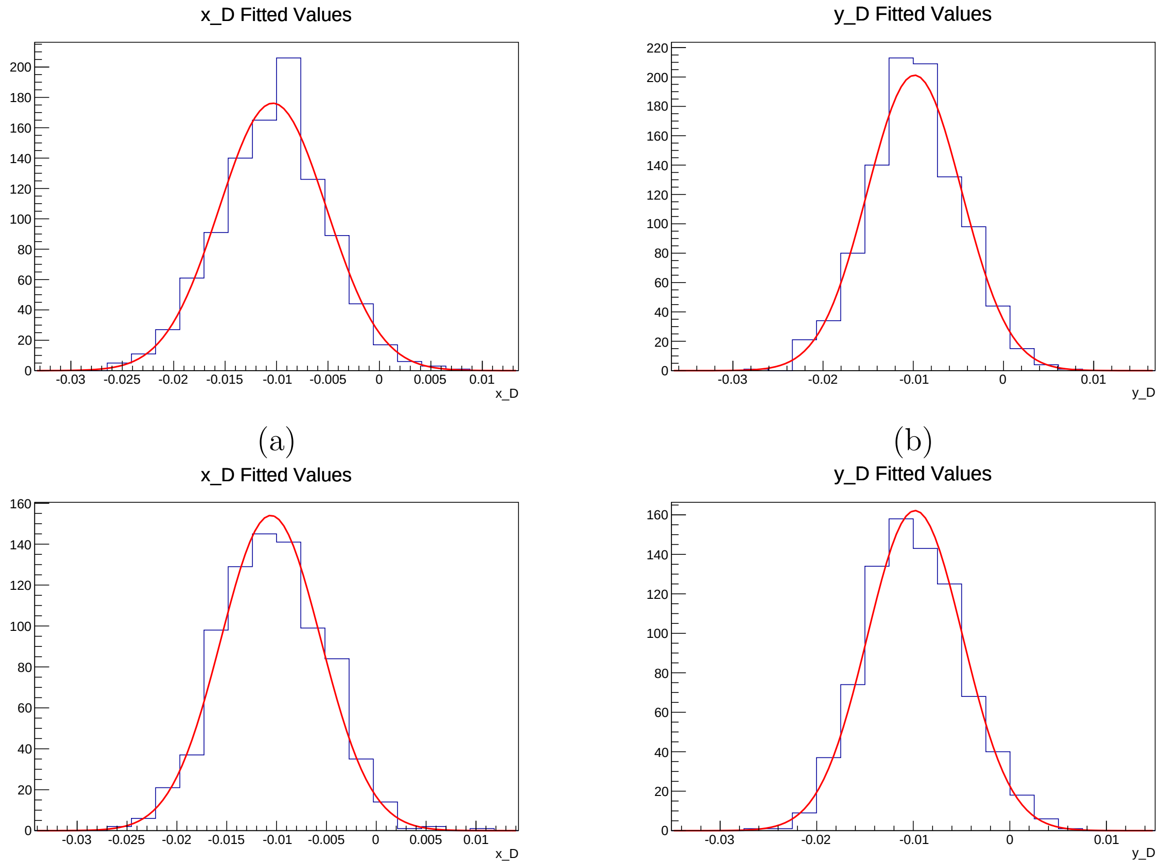
Figure 6.26: Fit projection of time-integrated Δm fit to secondary signal MC. (a) Single Gaussian plus a bifurcated Gaussian. (b) Double Gaussian plus a bifurcated Gaussian.

Parameter	Fitted Value
μ (GeV)	0.14506 ± 0.00002
f_1	0.536 ± 0.029
σ_1	0.00041 ± 0.00002
σ_L	0.00145 ± 0.00004
σ_R	0.00081 ± 0.00002

Table 6.12: Fitted parameters for global Δm fitted to secondary signal MC.

ond does. In both cases the fit is performed without accounting for the time dependence. The generation is initially performed in exactly the same way as for the toys used in the fit validation studies with the D^0 flavour, Dalitz plot bin, decay time, $ln(\chi^2_{IP})$, TP_1 and TP_{diff} being generated first. This step also decides which class of event is being generated. The values of m and Δm are then generated from the appropriate PDF for the class of event. Events which are prompt, right-π^+_{slow} have their Δm value generated using the time-integrated Δm fit conducted earlier in this section. Events which are secondary, right-π^+_{slow} have their Δm values generated using the time-integrated fit for the first set of toys and from the time-dependent fits for the second set. The remaining components are generated using parameters taken from previous fits to data. The generated data is then fed through exactly the same fit procedure as is used for fits to data. Note that the larger sample size is required as events are generated across a broad range in mass and Δm such that a large proportion are lost after the signal window cuts are applied. The number generated is chosen such that the dataset that is fed to the fit has approximately 180k events after applying the mass window cuts. The relative contribution from each class generated prior to the mass fits is also chosen such that, after the mass cuts, the contribution from each class of event closely emulates that seen in data with approximately 93 % right-π^+_{slow}, 4 % wrong-π^+_{slow} and 3 % combinatoric.

The fits are performed with c_i and s_i allowed to float within the covariant constraint described in Section 5.5.2. The input mixing parameters are $y_D = x_D = -1\,\%$. The resulting distribution of fitted values for x_D and y_D are shown for each set of toys in Figures 6.27 (a)–(d). Figures 6.27 (a) and (b) are the fitted values for the toys with no time-dependent Δm component and (c) and (d) are the results where the time dependence is generated, but not accounted for in the fit. A Gaussian is fitted to each of the distributions and the results summarised in Table 6.13. It is clear that there is a significant shift in the value of x_D when introducing a time-dependence on Δm for secondary decays. There is no significant shift observed in the value of y_D. However, for consistency, the size of the systematic for both x_D and y_D is taken as the difference between the central values from the two sets of toys and their uncertainties summed in quadrature. This yields $\sigma_{x_D}(\text{syst.}) = 0.0459\,\%$ and



(a) (b) (c) (d)

Figure 6.27: (a) Distribution of fitted values of x_D for toys that have been generated without a time-dependent Δm component. (b) Equivalent plot for y_D. (c) Distribution of fitted values of x_D for toys that have been generated with a time-dependent Δm component. (d) Equivalent plot for y_D.

$\sigma_{y_D}(\text{syst.}) = 0.0254\,\%$.

Parameter	Regular Δm	t-dependent Δm
x_D (%)	-1.0159 ± 0.0170	-1.0549 ± 0.0171
y_D (%)	-0.9856 ± 0.0170	-0.9780 ± 0.0173

Table 6.13: Summary of fitted parameters for toy MC that has been generated with and without a time-dependent Δm component for secondary D decays.

6.10.5 Wrong-π^+_{slow} Decays

It is assumed that the wrong-π^+_{slow} component is an even mix of D^0 and $\overline{D}^0$ events i.e. p_{D^0} is fixed to 0.5. It is also assumed that the ratio of $\overline{D}^0$ to D^0 events in a given bin is equal to T_{-i}/T_i measured by CLEO. The latter assumption is investigated in Section 6.7

and a systematic uncertainty assigned. The former assumption is investigated here. As the kinematics of the π^+_{slow} are relatively uncorrelated with that of the D^0 it can be assumed that the slow pion reconstruction efficiency should not vary across the Dalitz plot. However, there may be small asymmetries in the production and detection of D^0 and $\overline{D}^0$ candidates which may render this assumption void. There will likely be very little asymmetry in the detection efficiencies of D^0 and $\overline{D}^0$. Therefore, the measured D^0 production asymmetry taken from [127] of $A_p(D^0) = -1.08\pm0.32(\text{stat.})\pm0.12(\text{syst.})\,\%$ is likely a good approximation. The fraction of D^0 events expected in the prompt wrong-π^+_{slow} background given this asymmetry is $p_{D^0} = 0.4946$. Some simple cross-checks to verify that the $D^0/\overline{D}^0$ asymmetry in our data is comparable to the measured value are performed. Firstly, the yield asymmetry of D^{*+}/D^{*-} is measured from data within the mass and Δm signal windows. This provides an estimate of $p_{D^0} = 0.4988$. It is also possible to estimate the $D^0/\overline{D}^0$ asymmetry by making the following approximations:

$$p_{D^0} \cdot N(D^{*+})_{fake} = A \cdot N_{real}(D^0) \cdot N_{\pi^+} \tag{6.1}$$

$$(1 - p_{D^0}) \cdot N(D^{*+})_{fake} = A \cdot N_{real}(\overline{D}^0) \cdot N_{\pi^+} \tag{6.2}$$

$$p_{D^0} \cdot N(D^{*-})_{fake} = A \cdot N_{real}(D^0) \cdot N_{\pi^-} \tag{6.3}$$

$$(1 - p_{D^0}) \cdot N(D^{*-})_{fake} = A \cdot N_{real}(\overline{D}^0) \cdot N_{\pi^-} \tag{6.4}$$

where, A is assumed a constant, $N(D^{*\pm})_{fake}$ is the total number of fake $D^{*\pm}$ candidates in the signal windows, $N_{real}(D^0/\overline{D}^0)$ is the number of real $D^0/\overline{D}^0$ in the signal windows, and $N_{\pi^\pm}$ is the number of $\pi^\pm$ reconstructed. Taking the ratio $N(D^{*+})_{fake}/N(D^{*-})_{fake}$ from real data provides an estimate of N_{π^+}/N_{π^-} and the equations can be solved for p_{D^0}. This gives $p_{D^0} = 0.4981$. Therefore, all of these estimates (although not perfect) give similar values for p_{D^0} and the size of the $D^0/\overline{D}^0$ asymmetry in all cases is small. The mixing fit is run with $p_{D^0} = 0.4946$ as it is the largest deviation from 0.5 and the resulting variation in the mixing parameters is summarised in Table 6.14. Taking the magnitude of the difference as the systematic uncertainty on each parameter results in $\sigma_{x_D}(\text{syst.}) = 0.0013\,\%$ and $\sigma_{y_D}(\text{syst.}) = 0.0059\,\%$.

Parameter	$p_{D^0} = 0.5$	$p_{D^0} = 0.4946$	Difference
x_D (%)	-0.8540 ± 0.5170	-0.8527 ± 0.5222	-0.0013
y_D (%)	-0.0535 ± 0.4880	-0.0476 ± 0.4570	-0.0059

Table 6.14: Variation in fitted mixing parameters for different estimates of p_{D^0}.

6.11 Variation of Signal Components Across the Dalitz Plot

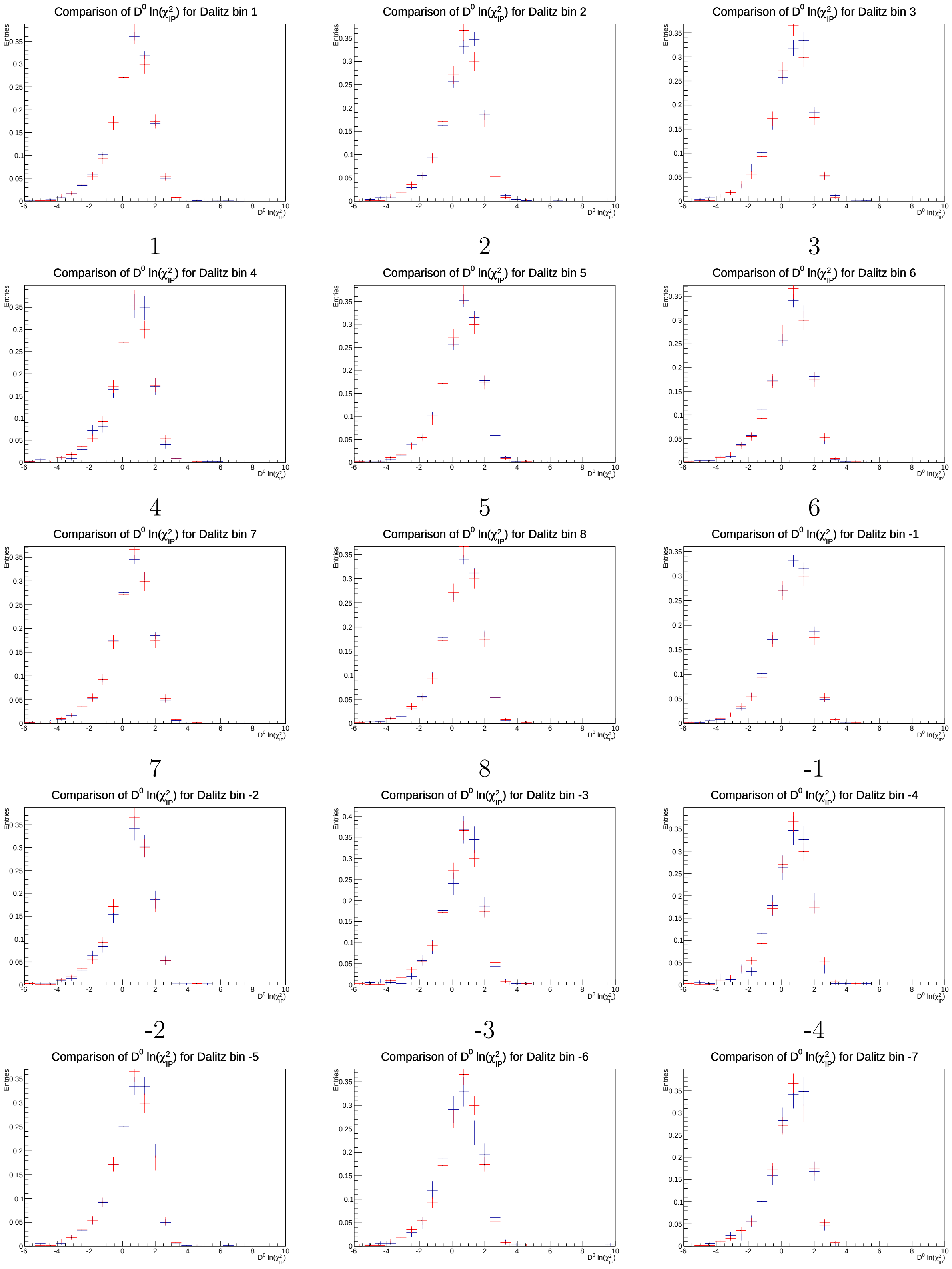
A potential source of systematic uncertainty stems from the assumption that some of the signal component distributions are the same across all bins of the Dalitz plot (signal components refers to: prompt/sec right-π^+_{slow} and prompt/sec wrong-π^+_{slow}). In particular, the shapes of the $ln(\chi^2_{IP})$ distributions for the prompt and secondary components of both the right-π^+_{slow} and wrong-π^+_{slow} are fixed to that obtained in the lifetime fit. This relies on the assumption that there is little variation in the shapes of these distributions as a function of position in the Dalitz plot. The strong assumption is there is no variation at all. However, as both $ln(\chi^2_{IP})$ and the Dalitz plot variables are momentum-dependent, the selection cuts may well cause a sculpting of the $ln(\chi^2_{IP})$ distribution in different regions of the Dalitz plot. The weaker assumption is that the cumulative distribution within individual Dalitz plot bins does not vary significantly bin-to-bin. In addition, any differences between the $ln(\chi^2_{IP})$ distributions in bins of the Dalitz plot may not be the same for prompt and secondary events given the missing momentum in the latter. To examine this, the $ln(\chi^2_{IP})$ distributions for each Dalitz plot bin are compared with each other. This is done using the *signalmixfiltered* MC sample with separate plots produced for truth-matched prompt and secondary signal events, respectively. The matched distributions for prompt and secondary events in bins $-7 \rightarrow 8$ are normalised and plotted against that from bin -8 (see Figures 6.28 and 6.29). A summary of the χ^2 and p-value between each of the distributions for each bin is also listed in Table 6.16. Examining the plots and the table of χ^2 values it is clear that there is some variation between the cumulative $ln(\chi^2_{IP})$ distributions of prompt and secondary events across each of the Dalitz plot bins. In particular, many of the χ^2 values listed in Table 6.16 are significantly less than 1. In addition, there are likely more subtle effects that the limited statistics of the signal MC may not be sensitive to (i.e. differences in the time evolution of the $ln(\chi^2_{IP})$ distributions for secondary decays). Therefore, it is desirable to investigate any potential differences in real data. To this end, the full lifetime fit procedure (described in Section 5.4) is run separately in each bin of the Dalitz plot. Each fit is performed on all data in a given bin with no separation of D^0 and $\overline{D}^0$ and the value of the lifetime in each bin is initially blinded. All parameters describing the prompt and secondary signal components are allowed to float and the combinatoric component is taken from the mass sidebands for a given Dalitz plot bin. Running the fit in this manner and comparing the fitted parameters gives an indiction of the variation in the decay time and $ln(\chi^2_{IP})$ distributions in each Dalitz plot bin from data. The fitted parameters for all bins are listed in Table 6.17. It is clear that many of the parameters do not vary significantly from bin to bin. However, there are some exceptions. One of particular concern is the difference in the secondary decay time constants found. To see if these differences have an effect on the measured mixing parameters, the mixing fit is run on real data with the parameterisations for each bin taken from Table 6.17. This is done such that the same parameterisation for the signal components is used for D^0 and $\overline{D}^0$ in bin i. The c_i and s_i parameters are allowed to float. The difference between the mixing parameters extracted from this fit and the regular fit are listed in Table 6.15. The shift observed is small for both the mixing parameters. Assigning the magnitude of the difference as a systematic uncertainty gives $\sigma_{x_D}(\text{syst.}) = 0.0128\,\%$ and $\sigma_{y_D}(\text{syst.}) = 0.0172\,\%$.

Parameter	Regular Fit	Per-bin Fit	Difference
x_D (%)	-0.8540 ± 0.5170	-0.8322 ± 0.5167	-0.0128
y_D (%)	-0.0535 ± 0.4880	-0.0363 ± 0.4865	-0.0172

Table 6.15: Difference between fit using fixed parameterisation for the signal components from all Dalitz plot bins and using a unique parameterisation for each bin.

Prompt Decays			Secondary Decays		
Bin	χ^2/ndf	p-value	Bin	χ^2/ndf	p-value
1	1.0121	0.4449	1	1.1833	0.2760
2	1.0588	0.3836	2	1.0512	0.3975
3	0.7542	0.7980	3	0.7779	0.7038
4	0.5984	0.9380	4	0.5488	0.9140
5	0.5226	0.9730	5	0.9156	0.5457
6	0.7949	0.7477	6	0.9406	0.5172
7	0.8121	0.7252	7	1.7042	0.0428
8	0.7290	0.8266	8	1.3268	0.1756
−1	0.8298	0.7014	−1	0.4632	0.9590
−2	0.4302	0.9930	−2	0.6963	0.7908
−3	0.7843	0.7613	−3	0.5662	0.9024
−4	0.7326	0.8227	−4	0.5283	0.9267
−5	0.5072	0.9778	−5	1.0674	0.3812
−6	0.7151	0.8415	−6	0.9107	0.5514
−7	0.5511	0.9622	−7	1.2168	0.2496

Table 6.16: Chi-square tests for $ln(\chi^2_{IP})$ distributions in each Dalitz plot bin. Comparison is done w.r.t reference bin −8.



1 2 3

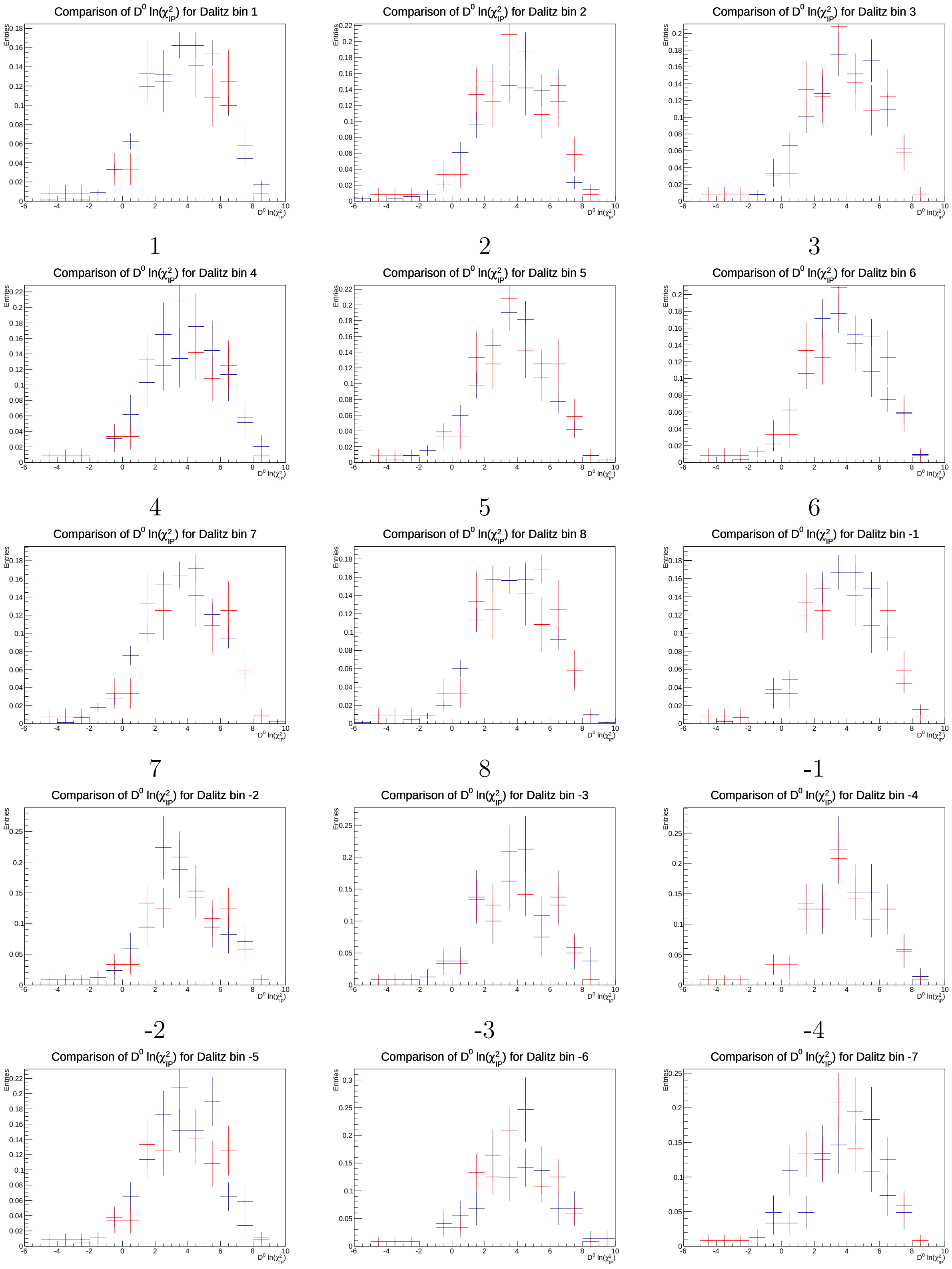
4 5 6

7 8 -1

-2 -3 -4

-5 -6 -7

Figure 6.28: Comparison of $ln(\chi^2_{IP})$ distributions in individual Dalitz plot bins for truth-matched prompt events. All distributions (blue) are plotted against the distribution taken from bin -8 (red) which is taken as a reference.



Comparison of D^0 ln(χ^2_{IP}) for Dalitz bin 1

1

Comparison of D^0 ln(χ^2_{IP}) for Dalitz bin 2

2

Comparison of D^0 ln(χ^2_{IP}) for Dalitz bin 3

3

Comparison of D^0 ln(χ^2_{IP}) for Dalitz bin 4

4

Comparison of D^0 ln(χ^2_{IP}) for Dalitz bin 5

5

Comparison of D^0 ln(χ^2_{IP}) for Dalitz bin 6

6

Comparison of D^0 ln(χ^2_{IP}) for Dalitz bin 7

7

Comparison of D^0 ln(χ^2_{IP}) for Dalitz bin 8

8

Comparison of D^0 ln(χ^2_{IP}) for Dalitz bin -1

-1

Comparison of D^0 ln(χ^2_{IP}) for Dalitz bin -2

-2

Comparison of D^0 ln(χ^2_{IP}) for Dalitz bin -3

-3

Comparison of D^0 ln(χ^2_{IP}) for Dalitz bin -4

-4

Comparison of D^0 ln(χ^2_{IP}) for Dalitz bin -5

-5

Comparison of D^0 ln(χ^2_{IP}) for Dalitz bin -6

-6

Comparison of D^0 ln(χ^2_{IP}) for Dalitz bin -7

-7

Figure 6.29: Comparison of $ln(\chi^2_{IP})$ distributions in individual Dalitz plot bins for truth-matched secondary events. All distributions (blue) are plotted against the distribution taken from bin -8 (red) which is taken as a reference.

Bin	f_{sec} (%)	A	f_1	μ	σ_1
1	9.426 ± 0.451	0.066 ± 0.020	0.415 ± 0.029	1.237 ± 0.033	0.929 ± 0.021
2	13.752 ± 1.121	0.038 ± 0.032	0.412 ± 0.037	1.205 ± 0.028	0.864 ± 0.035
3	9.965 ± 0.588	0.003 ± 0.037	0.322 ± 0.049	1.353 ± 0.054	0.918 ± 0.054
4	10.403 ± 1.436	0.030 ± 0.050	0.498 ± 0.058	1.177 ± 0.042	0.924 ± 0.046
5	10.002 ± 0.648	0.045 ± 0.032	0.488 ± 0.040	1.199 ± 0.045	0.919 ± 0.028
6	11.962 ± 0.774	0.031 ± 0.032	0.484 ± 0.021	1.201 ± 0.002	0.943 ± 0.028
7	8.989 ± 0.483	0.036 ± 0.023	0.471 ± 0.028	1.207 ± 0.009	0.946 ± 0.023
8	11.183 ± 0.607	0.028 ± 0.027	0.471 ± 0.032	1.201 ± 0.037	0.913 ± 0.023
-1	9.555 ± 0.385	0.078 ± 0.019	0.479 ± 0.022	1.172 ± 0.019	0.911 ± 0.016
-2	10.370 ± 0.738	0.072 ± 0.038	0.481 ± 0.050	1.133 ± 0.052	0.914 ± 0.033
-3	10.110 ± 0.844	0.046 ± 0.045	0.380 ± 0.058	1.273 ± 0.067	0.904 ± 0.048
-4	13.929 ± 1.304	0.082 ± 0.054	0.515 ± 0.065	1.097 ± 0.067	0.887 ± 0.054
-5	10.289 ± 0.617	-0.003 ± 0.033	0.442 ± 0.038	1.264 ± 0.044	0.883 ± 0.030
-6	10.166 ± 0.586	0.019 ± 0.031	0.439 ± 0.042	1.203 ± 0.003	0.903 ± 0.032
-7	10.854 ± 0.428	-0.007 ± 0.026	0.481 ± 0.032	1.202 ± 0.037	0.921 ± 0.024
-8	9.720 ± 0.452	0.055 ± 0.019	0.461 ± 0.025	1.202 ± 0.003	0.920 ± 0.020

Bin	σ_L	σ_R	τ (ps)	B	D
1	1.957 ± 0.024	0.716 ± 0.038	`blind` ± 0.003	6.014 ± 0.272	-0.079 ± 0.082
2	1.952 ± 0.040	0.777 ± 0.054	`blind` ± 0.005	5.981 ± 0.347	-0.128 ± 0.103
3	1.918 ± 0.042	0.740 ± 0.073	`blind` ± 0.006	4.877 ± 0.249	0.054 ± 0.102
4	2.043 ± 0.073	0.794 ± 0.097	`blind` ± 0.017	6.625 ± 0.662	-0.093 ± 0.291
5	2.006 ± 0.042	0.772 ± 0.056	`blind` ± 0.005	5.773 ± 0.207	0.101 ± 0.206
6	1.995 ± 0.013	0.728 ± 0.048	`blind` ± 0.005	6.744 ± 0.435	-0.435 ± 0.026
7	1.993 ± 0.031	0.757 ± 0.044	`blind` ± 0.004	5.288 ± 0.271	-0.074 ± 0.107
8	2.026 ± 0.034	0.782 ± 0.044	`blind` ± 0.004	5.798 ± 0.313	-0.194 ± 0.111
-1	2.011 ± 0.025	0.764 ± 0.033	`blind` ± 0.003	5.325 ± 0.156	-0.323 ± 0.054
-2	1.928 ± 0.047	0.784 ± 0.064	`blind` ± 0.005	5.197 ± 0.371	-0.138 ± 0.129
-3	1.955 ± 0.048	0.805 ± 0.066	`blind` ± 0.006	5.613 ± 0.192	-0.172 ± 0.120
-4	2.028 ± 0.076	0.819 ± 0.095	`blind` ± 0.007	5.743 ± 0.513	-0.387 ± 0.085
-5	1.981 ± 0.040	0.781 ± 0.057	`blind` ± 0.005	5.063 ± 0.307	-0.037 ± 0.104
-6	1.924 ± 0.041	0.766 ± 0.047	`blind` ± 0.005	5.287 ± 0.298	-0.034 ± 0.133
-7	2.009 ± 0.034	0.794 ± 0.048	`blind` ± 0.004	5.519 ± 0.088	-0.124 ± 0.085
-8	2.010 ± 0.028	0.755 ± 0.018	`blind` ± 0.004	5.410 ± 0.282	-0.185 ± 0.090

Bin	C	μ_{sec}	α	τ_1 (ps)	τ_2 (ps)
1	-0.517 ± 0.075	2.098 ± 0.235	1.045 ± 0.046	0.472 ± 0.109	0.680 ± 0.092
2	-0.647 ± 0.109	1.364 ± 0.409	1.236 ± 0.084	0.348 ± 0.083	0.726 ± 0.059
3	-0.709 ± 0.165	2.330 ± 0.426	1.000 ± 0.319	0.645 ± 0.051	0.639 ± 0.043
4	-0.670 ± 0.199	0.996 ± 0.835	1.057 ± 0.119	0.381 ± 0.189	0.736 ± 0.132
5	-1.039 ± 0.180	0.777 ± 0.411	1.005 ± 0.052	0.526 ± 0.031	0.526 ± 0.036
6	-1.114 ± 0.075	-0.341 ± 0.408	1.267 ± 0.067	0.560 ± 0.039	0.586 ± 0.028
7	-0.624 ± 0.101	2.090 ± 0.285	1.056 ± 0.057	0.604 ± 0.029	0.605 ± 0.028
8	-0.548 ± 0.091	1.815 ± 0.288	1.159 ± 0.059	0.569 ± 0.026	0.570 ± 0.026
-1	-0.784 ± 0.104	1.554 ± 0.279	1.143 ± 0.045	0.570 ± 0.022	0.569 ± 0.024
-2	-0.782 ± 0.177	1.665 ± 0.523	1.122 ± 0.077	0.614 ± 0.043	0.615 ± 0.036
-3	-0.441 ± 0.142	2.368 ± 0.509	1.141 ± 0.104	0.649 ± 0.052	0.632 ± 0.043
-4	-0.669 ± 0.219	1.175 ± 0.546	1.258 ± 0.101	0.300 ± 0.115	0.720 ± 0.047
-5	-0.685 ± 0.040	2.046 ± 0.127	1.086 ± 0.065	0.595 ± 0.036	0.602 ± 0.034
-6	-0.768 ± 0.151	1.613 ± 0.395	1.000 ± 0.045	0.576 ± 0.040	0.577 ± 0.026
-7	-0.732 ± 0.095	1.479 ± 0.270	1.091 ± 0.048	0.590 ± 0.026	0.591 ± 0.025
-8	-0.535 ± 0.079	2.243 ± 0.232	1.026 ± 0.048	0.605 ± 0.045	0.610 ± 0.038

Table 6.17: Fitted parameters for lifetime fit performed in each Dalitz plot bin.

6.12 Summary

Table 6.18 lists each source of systematic uncertainty investigated and its contribution to the mixing parameters. The sum in quadrature of each contribution is then given at the bottom for each parameter.

Systematic	x_D (%)	y_D (%)
Fit bias	0.0211	0.0201
Decay time resolution	0.0654	0.0391
Turning point resolution	0.0198	0.0221
Turning point distributions	0.0511	0.0227
Dalitz plot efficiency	0.0565	0.0711
Invariant mass resolution	0.0727	0.0284
Uncertainty from T_i and T_{-i}	0.0791	0.0559
Tracking efficiency parameterisation	0.0145	0.0245
Dalitz plot binning	0.0120	0.0061
Uncertainties from mass vs. Δm fits	Neg.	Neg.
Uncertainties from lifetime fit	0.0200	0.0426
Combinatoric background	0.0610	0.0519
Treatment of secondary D decays	0.0459	0.0254
Wrong-π^+_{slow} background	0.0013	0.0059
Variation of signal components across the Dalitz plot	0.0128	0.0172
Total systematic uncertainty	0.1709	0.1343

Table 6.18: Summary of all contributions to the systematic uncertainties on x_D and y_D.

The final results including both the statistical and systematic uncertainties are

$$x_D = -\,(0.863 \pm 0.527\,(\text{stat.}) \pm 0.171\,(\text{syst.}))\,\%,$$

$$y_D = -\,(0.026 \pm 0.463\,(\text{stat.}) \pm 0.134\,(\text{syst.}))\,\%.$$

The most recent world averages [48] taken from published analyses across all D decay modes for the no-CPV hypothesis are $x(\text{HFAG}) = 0.490 \pm 0.150\,\%$ and $y(\text{HFAG}) = 0.617 \pm 0.079\,\%$. Note that the CP formalism used by HFAG differs from that used in this analysis such that $x_D = -x$ and $y_D = -y$. The correct values for comparison with the results presented here are $x_D(\text{HFAG}) = -0.490 \pm 0.150\,\%$ and $y_D(\text{HFAG}) = -0.617 \pm 0.079\,\%$. Both mixing parameters are consistent with their world average values with x_D 0.63σ and y_D 1.21σ apart, respectively. Both mixing parameters are also consistent with the current world averages from analyses of $D^0 \to K^0_\mathrm{S} h^+ h^-$ decays, $x_D(\text{HFAG-}K^0_\mathrm{S} h^+ h^-) = 0.419 \pm 0.211\,\%$ and $y_D(\text{HFAG-}K^0_\mathrm{S} h^+ h^-) = 0.456 \pm 0.186\,\%$, respectively. Taking a naive combination of the world average from all charm decays and the measurement presented here gives $x_D = -0.570 \pm 0.145\,\%$ and $y_D = -0.534 \pm 0.078\,\%$, where the correlation between x_D and y_D has been ignored. It is clear that the statistical precision of the dataset limits the impact this measurement has on the world averages. However, the work presented in this thesis is a vital step forward in the development of new tools for measuring mixing in $D^0 \to K^0_\mathrm{S} \pi^+ \pi^-$ decays at LHCb and other experiments in the future where much larger yields are anticipated.

Chapter 7

Conclusions and Outlook

Presented in this thesis is a measurement of charm mixing parameters x_D and y_D via a time-dependent Dalitz plot analysis of prompt $D^0 \to K^0_S\pi^+\pi^-$ decays. A model-independent description of the Dalitz plot was utilised for the first time in the context of charm mixing with measurements of the strong phase difference between D^0 and $\overline{D}^0$ provided by the CLEO collaboration used as input, thus removing any systematic uncertainty associated with modelling the resonate structure of the Dalitz decay. In addition, a data-driven technique for correcting decay time acceptance effects known as "swimming" is utilised for the first time in multi-body charm decays at LHCb. The resulting mixing parameters were measured to be

$$x_D = -\,(0.863 \pm 0.527\,(\text{stat.}) \pm 0.171\,(\text{syst.}))\,\%,$$

$$y_D = -\,(0.026 \pm 0.463\,(\text{stat.}) \pm 0.134\,(\text{syst.}))\,\%.$$

Both parameters are consistent with the current world averages, $x_D(\text{HFAG}) = -0.490 \pm 0.150\,\%$ and $y_D(\text{HFAG}) = -0.617 \pm 0.079\,\%$. The results presented here have larger uncertainties than the world averages due to the limited statistical precision of the dataset. The analysis was performed on 1 fb^{-1} of data collected by LHCb during 2011 corresponding to a signal yield of 164k D^0 decays. At the time of writing, LHCb has amassed a total of 3 fb^{-1} of data. The additional 2 fb^{-1} was recorded during 2012 with the LHC running at a centre-of-mass energy of 8 TeV and LHCb operating at a higher luminosity increasing the rate of charm production. Improvements to the trigger and the addition of an inclusive trigger line for charm means that the projected yield per fb^{-1} of prompt $K_{\mathrm{S}}^{0}\pi^{+}\pi^{-}$ decays is around a factor of 6 larger in 2012 compared to 2011. This data can also be combined with $D^0 \to K_{\mathrm{S}}^{0}\pi^{+}\pi^{-}$ decays coming from semi-leptonic B decays of which LHCb has collected several hundred thousand to date. Therefore, LHCb has roughly a factor of 15 more $D^0 \to K_{\mathrm{S}}^{0}\pi^{+}\pi^{-}$ decays already on tape. Assuming naively[1] that the statistical uncertainty scales as $\sqrt{N}$ then a statistical precision of less than 0.2 % for both x_D and y_D is potentially achievable. This level of precision is comparable to the uncertainty on the current world average for x_D and would contribute significantly to our understanding of mixing in charm. In addition to the data already on tape, LHCb is due to start collecting more data during Run II of the LHC. It is projected that LHC will take around 5–6 fb^{-1} of data initially at 13 TeV and then at 14 TeV. The higher energy will increase the charm production cross-section by a factor 1.5–2 and, coupled with further improvements to the trigger and data acquisition, may result in 5–10M $K_{\mathrm{S}}^{0}\pi^{+}\pi^{-}$ events being collected over the coming years. This will reduce the statistical uncertainty on x_D and y_D by an additional factor of 2. Looking further into the future, the LHCb upgrade is projected to collect tens of millions of $K_{\mathrm{S}}^{0}\pi^{+}\pi^{-}$ events at which point the statistical uncertainties on the mixing parameters will be much less than 0.1 %. It should be noted that the uncertainties from the current strong-phase measurements provided by CLEO will begin to dominate the mixing measurements so additional input from on-resonance running at BES-II would be required to improve things further. However, LHCb will be able to make precision measurements of the CP-violation parameters α_{CP} and r_{CP}. In addition to the future exploits of LHCb, the Belle II experiment is currently being commissioned with the aim to start collecting data in 2016. Belle II is projected to collect a total of 50 ab^{-1} during its lifetime which will likely correspond to tens of millions of $K_{\mathrm{S}}^{0}\pi^{+}\pi^{-}$ events and reach a statistical precision that is comparable to

---

[1] This is not true in this case as the statistical uncertainties quoted have a contribution due to the uncertainties on c_i and s_i from the CLEO measurements. The total uncertainty is around 0.5 % with a contribution from the CLEO measurements of around 0.13 %. The difference in quadrature yields an LHCb uncertainty of around 0.48 % and scaling this with the projected yields for the full 3 fb^{-1} of LHCb data results in a statistical uncertainty comparable to the contribution from the CLEO measurements.

what is expected at LHCb over a similar time period. At this very high level of precision, control of the systematic uncertainties will become absolutely critical. The techniques used in this analysis remove completely the systematic uncertainty associated with modelling the resonant structure of the Dalitz plot and significantly reduce those associated with correcting for decay time acceptance effects with simulated data. The former technique introduces another source of uncertainty due to its dependence on external measurements which, given the current CLEO measurements, is determined to be approximately 0.13 %. Therefore, this additional uncertainty will not dominate until the end of Run II of the LHC. However, work is already underway from BES III to improve these measurements [71] meaning that this uncertainty will likely come down in the future.

In summary, the future of mixing measurements in charm is promising. The amount of available data will increase dramatically over the coming years offering unrivalled precision, but with this increased precision will come the challenging task of reducing systematic and experimental uncertainties. While the measurement performed here will have limited impact on the current world averages, the pioneering techniques utilised for the first time in the context of mixing in $K^0_S\pi^+\pi^-$ decays will play a critical role in subsequent measurements with much higher yields at LHCb and Belle-II.

Appendix A

Extraction of c_i and s_i from BaBar 2010 Amplitude Model

For toy MC data that has been generated using the BaBar 2010 amplitude model described in [47] the values of c_i and s_i measured by CLEO may not agree with that generated. As this MC is used to evaluate the size of the systematic uncertainty due to the Dalitz plot efficiency, it is crucial that the input values of c_i and s_i are correct for the data otherwise it could bias any measurement of x_D and y_D, rendering any conclusions drawn from the data unreliable. Therefore, c_i and s_i were measured for the BaBar model using a bitmap of the overall amplitude and phase for a fine grid of data points across the Dalitz plot. In addition, the expected values of T_i and T_{-i} are also extracted. For each bin of the Dalitz plot c_i and s_i can be expressed as

$$c_i = \frac{1}{N}\sum_j a_{12,13}a_{13,12}\cos(\delta_{12,13} - \delta_{13,12}) \times A_j \tag{A.1}$$

$$s_i = \frac{1}{N}\sum_j a_{12,13}a_{13,12}\sin(\delta_{12,13} - \delta_{13,12}) \times A_j \tag{A.2}$$

$$\tag{A.3}$$

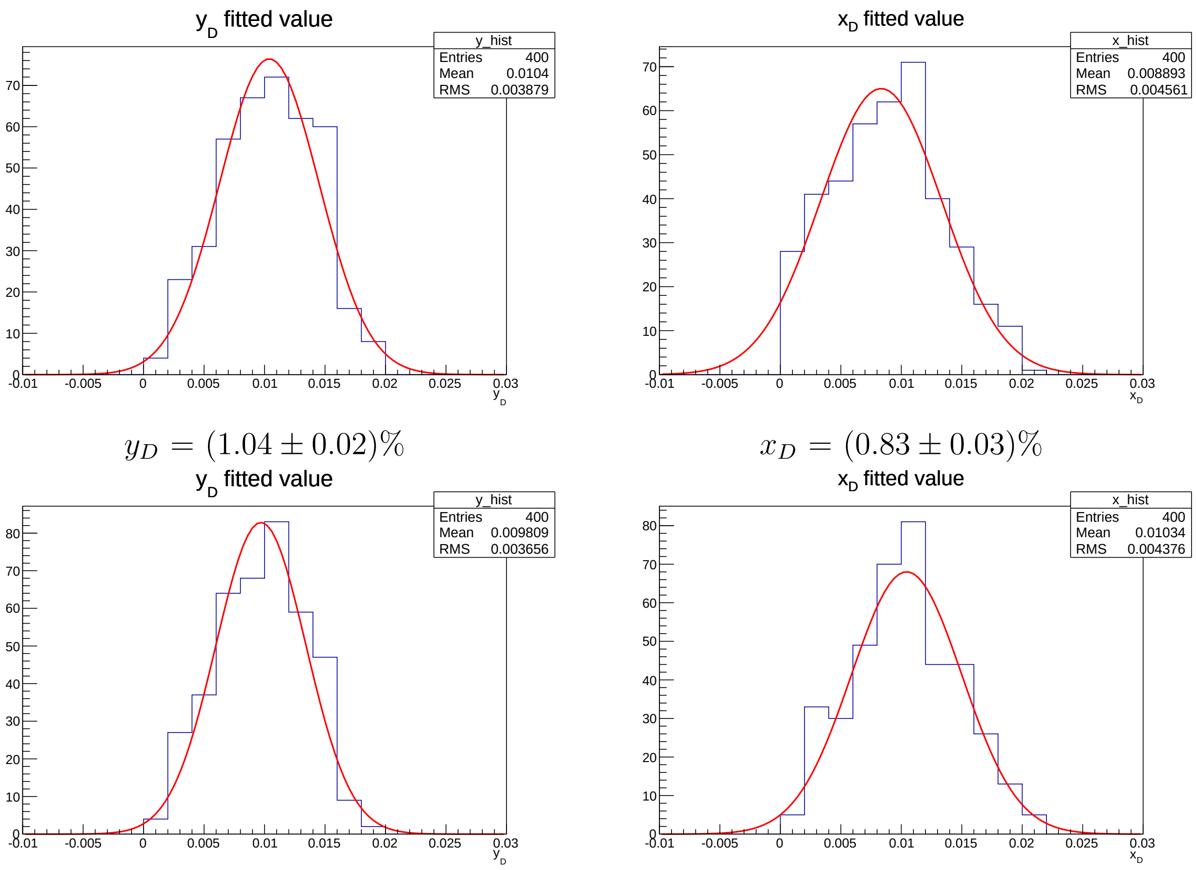
Bin	c_i	s_i	T_i (%)	T_{-i} (%)
1	0.690	-0.016	16.48	8.00
2	0.625	0.419	8.15	1.94
3	0.070	0.858	6.81	1.90
4	-0.497	0.777	2.56	1.71
5	-0.946	0.096	8.78	5.24
6	-0.634	-0.490	6.22	1.46
7	-0.064	-0.811	13.42	1.29
8	0.422	-0.450	13.45	2.60

Table A.1: Values of c_i, s_i, T_i and T_{-i} extracted from the BaBar 2010 amplitude model described in [47].

where j runs over all sub-bins within Dalitz plot bin i and A_j is the area of the Dalitz plot covered by the sub-bin. The equivalent expression for T_i is

$$T_i = \frac{1}{N}\sum_j |a^2_{12,13}| \times A_j. \tag{A.4}$$

Replacing $a_{12,13}$ in Equation A.4 with $a_{13,12}$ gives the value for T_{-i}. The values of c_i, s_i, T_i and T_{-i} as extracted from the model are listed in Table A.1. For all cases where fits are performed to toy MC that has been generated using this model the parameters in Table A.1 are used. When c_i and s_i are floated in these fits, the uncertainties from the CLEO measurements are used along with the published covariance matrix in [58]. The effect of using the published CLEO values, rather than those extracted from the model, on the extracted values of x_D and y_D can be clearly seen in Figure A.1. The distributions of fitted values of x_D and y_D for 400 fits to roughly 180k events are shown. The distributions in the top row are the fitted values using the fixed c_i and s_i parameters as measured by CLEO and those in the bottom row use the parameters fixed to the values listed in Table A.1.



$y_D = (1.04 \pm 0.02)\%$ $x_D = (0.83 \pm 0.03)\%$

$y_D = (0.97 \pm 0.02)\%$ $x_D = (1.04 \pm 0.03)\%$

Figure A.1: Comparison of fitted values of x_D and y_D from toy MC data generated using the BaBar 2010 amplitude model. The top row shows fits with c_i and s_i fixed to the values measured by CLEO. The bottom row shows fits with c_i and s_i set to the parameters listed in Table A.1.

Appendix B

Alternative Dalitz Plot Efficiency Models

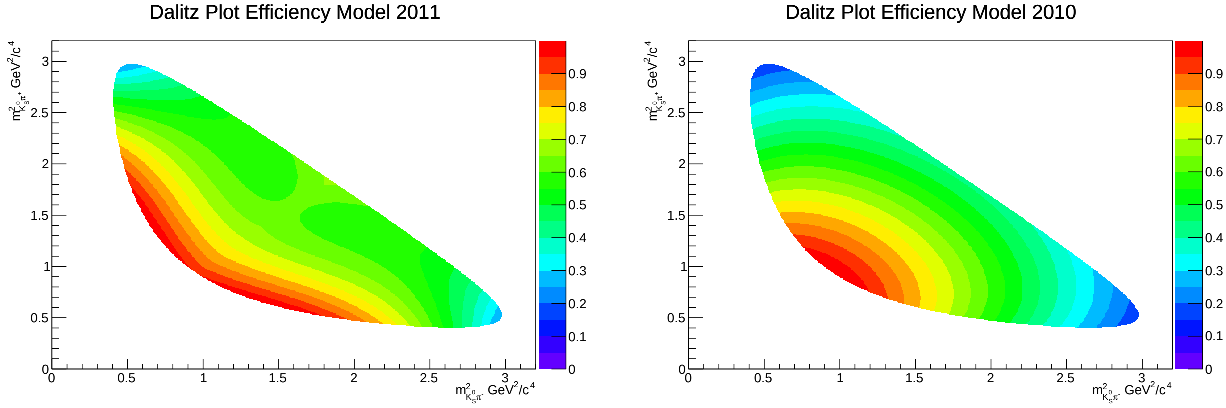
An additional model of the Dalitz plot efficiency was developed prior to that described in Section 4.8. The model is based on a fit to a modest sample of Monte Carlo events that was generated using an older version of the LHCb generation, reconstruction and stripping code. However, the majority of the trigger, stripping and offline selection cuts documented in Chapter 4 were applied to the data by hand. This model is therefore not a totally accurate description of the true Dalitz plot efficiency observed in data, but is a reasonable approximation. It is therefore useful to compare the effect of the Dalitz plot efficiency from this model and that described in Section 4.8 on the physics observables to understand how sensitive the fit is to any discrepancies in the modelling of the Dalitz plot efficiency. The model is parameterised using an empirical probability density function of the form

$$\epsilon(m_{12}^2, m_{13}^2) = a + b(m_{12}^2 + m_{13}^2) + c\frac{(m_{12}^2 - m_{13}^2)^2}{(m_{12}^2 + m_{13}^2)^2}. \tag{B.1}$$

The fitted parameters are listed in Table B.1. For comparison, both the model described in Section 4.8 (hereby referred to as the 2011 model) and this model (hereby referred to as the

Parameter	Value
a	21.23 ± 2.55
b	-4.20 ± 0.84 GeV$^2/c^4$
c	-8.99 ± 2.01

Table B.1: Fitted parameters for 2010 Dalitz plot efficiency model.



(a) (b)

Figure B.1: (a) 2011 Dalitz plot efficiency model. (b) 2010 Dalitz plot efficiency model.

2010[1] model) are plotted in Figures B.1 (a) and (b), respectively.

In addition to the 2010 and 2011 efficiency models a more complex parameterisation is fitted to the same MC data as used to estimate the 2011 model. A polynomial one order higher than that used in the 2011 model is used. Two fits are performed; the first is done assuming the model is symmetric around the axis $m^2_{12} = m^2_{13}$, the second allows for an asymmetry by letting certain parameters float independently in each half of the Dalitz plot. The first fit is used to check if there is any additional structure that the 2011 model is not sensitive to. This model is referred to as 2011HO. The latter fit is of particular importance given the large extent to which the analysis relies on the efficiency being symmetric. This model is referred to as the 2011ASYM model. The higher order polynomial used in both the 2011HO and 2011ASYM fits can be written as

$$\begin{aligned}\epsilon(m^2_{12}, m^2_{13})_{HO} = 1 + au + bu^2 + c|v| + dv^2 + eu|v| + fu^3 + gv^2|v| + hu^2|v| \\ + iuv^2 + ju^2v^2 + ku^3|v| + lu^4 + muv^2|v| + nv^4\end{aligned} \tag{B.2}$$

---

[1] This is because the model is based on MC data generated according to the 2010 running conditions of LHCb. However, the trigger, stripping, reconstruction and selection is re-run such that is similar to that applied to the MC used to develop the 2011 model.

where $u = m_{12}^2 + m_{13}^2$ and $v = m_{12}^2 - m_{13}^2$. However, in the case of the 2011ASYM model the parameters, c, d, e, g, h, i, k and m are allowed to float independently for $v < 0$ and $v > 0$ i.e. the upper and lower halves of the Dalitz plot. The fitted parameters for the 2011ASYM model are listed in Table B.2 and plotted in Figure B.2. No significant difference is found between the 2011 and the 2011HO models. However, there is some subtle variation in each half of the Dalitz plot picked out in the 2011ASYM model that is not accounted for in the 2011 model (although the uncertainties on the parameters are still large in comparison to the size of the variation). The 2011ASYM model is used in addition to the 2011 and 2010 models in the systematic cross checks described in Section 6.5.

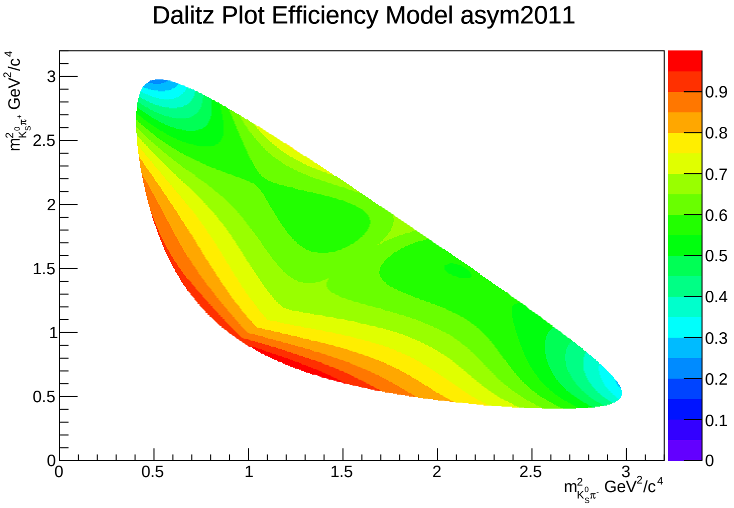


Figure B.2: 2011ASYM Dalitz plot efficiency model.

Parameter	Value
a	-0.165 ± 0.191
b	0.281 ± 0.342
c ($v >= 0$)	0.018 ± 0.434
c ($v < 0$)	-0.124 ± 0.415
d ($v >= 0$)	0.473 ± 0.671
d ($v < 0$)	0.532 ± 0.649
e ($v >= 0$)	-0.771 ± 0.552
e ($v < 0$)	-0.795 ± 0.578
f	-0.313 ± 0.459
g ($v >= 0$)	-0.456 ± 0.400
g ($v < 0$)	-0.413 ± 0.394
h ($v >= 0$)	-0.123 ± 0.661
h ($v < 0$)	-0.145 ± 0.628
i ($v >= 0$)	0.739 ± 0.617
i ($v < 0$)	0.798 ± 0.462
j	0.016 ± 0.475
k ($v >= 0$)	0.057 ± 0.585
k ($v < 0$)	0.379 ± 0.595
l	0.202 ± 0.429
m ($v >= 0$)	-0.252 ± 0.281
m ($v < 0$)	-0.372 ± 0.285
n	0.123 ± 0.090

Table B.2: Fitted parameters for the asymmetric Dalitz plot efficiency model.

Appendix C

Covariance Matrix

Covariance matrix for parameters that are fixed from the lifetime fit in the mixing fit.

	f_{sec}	A	f_1	μ	σ_1	σ_L	σ_R	B	D	C
f_{sec}	9.122×10^{-7}	1.834×10^{-7}	1.229×10^{-7}	-4.748×10^{-11}	-7.339×10^{-7}	4.554×10^{-9}	-1.365×10^{-8}	-1.677×10^{-6}	-1.336×10^{-6}	9.811×10^{-7}
A		2.054×10^{-6}	8.59×10^{-7}	1.198×10^{-9}	-9.308×10^{-7}	2.035×10^{-8}	-4.551×10^{-8}	-4.935×10^{-6}	-7.543×10^{-6}	1.602×10^{-6}
f_1			1.153×10^{-6}	7.421×10^{-10}	-6.323×10^{-7}	1.331×10^{-8}	-2.994×10^{-8}	-3.224×10^{-6}	-4.868×10^{-6}	1.064×10^{-6}
μ				6.345×10^{-9}	-5.659×10^{-10}	2.59×10^{-11}	-5.535×10^{-11}	-4.236×10^{-9}	-5.857×10^{-9}	1.251×10^{-9}
σ_1					1.846×10^{-5}	-1.579×10^{-8}	3.93×10^{-8}	3.551×10^{-6}	7.698×10^{-6}	-6.811×10^{-7}
σ_L						1.151×10^{-7}	-5.695×10^{-10}	-7.586×10^{-8}	-1.223×10^{-7}	2.515×10^{-8}
σ_R							1.695×10^{-7}	1.699×10^{-7}	2.875×10^{-7}	-5.314×10^{-8}
B								9.104×10^{-5}	3.182×10^{-5}	-3.725×10^{-6}
D									0.0001603	-1.204×10^{-5}
C										1.481×10^{-5}
μ_{sec}										
α										
τ_1										
τ_2										

	μ_{sec}	α	τ_1	τ_2
f_{sec}	-5.303×10^{-7}	1.141×10^{-6}	-1.809×10^{-6}	-3.736×10^{-7}
A	-8.283×10^{-8}	9.53×10^{-6}	2.146×10^{-7}	-1.207×10^{-6}
f_1	-7.65×10^{-8}	6.125×10^{-6}	5.065×10^{-8}	-7.868×10^{-7}
μ	8.944×10^{-11}	9.277×10^{-9}	2.994×10^{-9}	-8.596×10^{-10}
σ_1	-2.122×10^{-7}	-6.247×10^{-6}	4.465×10^{-6}	1.345×10^{-6}
σ_L	-2.09×10^{-9}	1.392×10^{-7}	-1.932×10^{-8}	-2.015×10^{-8}
σ_R	2.684×10^{-9}	-3.098×10^{-7}	6.838×10^{-8}	4.79×10^{-8}
B	-1.137×10^{-6}	-3.479×10^{-5}	5.22×10^{-6}	5.102×10^{-6}
D	2.511×10^{-6}	-5.315×10^{-5}	6.056×10^{-6}	7.638×10^{-6}
C	1.611×10^{-6}	1.137×10^{-5}	-2.768×10^{-6}	-1.78×10^{-6}
μ_{sec}	2.739×10^{-5}	-5.632×10^{-7}	2.091×10^{-6}	3.096×10^{-7}
α		7.461×10^{-5}	2.872×10^{-6}	-8.407×10^{-6}
τ_1			7.356×10^{-5}	2.388×10^{-6}
τ_2				6.787×10^{-6}

(C.1)

Bibliography

[1] M. Thompson et al. Proceedings of the 2011 RAL School for Experimental High Energy Physics Students. RAL-TR-2011-014, 2011.

[2] C. Burgess and G. Moore. *The Standard Model A Primer.* Cambridge University Press, 2007.

[3] J. Dalton. *A New System of Chemical Philosophy.* S. Russell, 1808.

[4] E. Rutherford. The Scattering of Alpha and Beta Particles by Matter and the Structure of the Atom. *Philosophical Magazine*, 21:669–688, 1911.

[5] J. J. Thomson. On the Structure of the Atom: An Investigation of the Stability and Periods of Oscillation of a Number of Corpuscles Arranged at Equal Intervals Around the Circumference of a Circle; With Application of the Results to the Theory of Atomic Structure. *Philosophical Magazine*, 7:237–265, 1904.

[6] P. Dirac. *The Principles of Quantum Mechanics (International Series of Monographs on Physics).* Oxford University Press, USA, 4^{th} edition, 1982.

[7] C. Anderson. The Positive Electron. *Physical Review*, 43:491–494, 1933.

[8] E. Bloom et al. High-Energy Inelastic e-p Scattering at 6° and 10°. *Physical Review Letters*, 23:930–934, 1969.

[9] M. Breidenbach et al. Observed Behavior of Highly Inelastic Electron-Proton Scattering. *Physical Review Letters*, 23:935–939, 1969.

[10] B. J. Bjorken and S. L. Glashow. Elementary Particles and SU(4). *Physics Letters*, 11:255–257, 1964.

[11] F. Abe et al. Observation of Top Quark Production in $p\overline{p}$ Collisions with the Collider Detector at Fermilab. *Physical Review Letters*, 74:2626–2631, 1995.

[12] M. L. Perl et al. Evidence for Anomalous Lepton Production in e^+-e^- Annihilation. *Physical Review Letters*, 35:1489–1492, 1975.

[13] C. Andreopoulos. Observation of Tau Neutrino Interactions. *Physics Letters B*, 504:218–224, 2001.

[14] D. I. Kazakov and G Smadja. Particle Physics and Cosmology: The Interface. Proceedings of the NATO Advanced Study Institute on Particle Physics and Cosmology: The Interface Cargèse, France, 2005.

[15] K. Freeman and G. McNamara. *In Search of Dark Matter.* Springer Praxis Books / Space Exploration. Praxis Publishing Limited, 2006.

[16] F. P. An et al. Observation of Electron-Antineutrino Disappearance at Daya Bay. *Physical Review Letters*, 108:171–803, 2012.

[17] Y. Fukuda et al. Evidence for Oscillation of Atmospheric Neutrinos. *Physical Review Letters*, 81:1562–1567, 1998.

[18] W. Pauli. The Connection Between Spin and Statistics. *Physical Review*, 58:716–722, 1940.

[19] S. N. Bose. Planck's Law and Light Quantum Hypothesis. *Zeitschrift für Physik*, 1924.

[20] A. Einstein. Quantum Theory of the Monatomic Ideal Gas. *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, pages 261–267, 1924.

[21] G. Aad et al. Observation of a New Particle in the Search for the Standard Model Higgs boson with the ATLAS Detector at the LHC. *Physics Letters B*, 716:1–29, 2012.

[22] S. Chatrchyan et al. Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC. *Physics Letters B*, 716:30–61, 2012.

[23] J. Beringer et al. 2013 Review of Particle Physics. *Physical Review D*, 86, 2012.

[24] S. P. Martin. A Supersymmetry Primer. *Perspectives On Supersymmetry*, 18:1–98, 1998.

[25] R. P. Feynman. *Q E D.* Penguin Books. Penguin Books, Limited, 1990.

[26] D. Hanneke, S. Fogwell Hoogerheide, and G. Gabrielse. Cavity control of a single-electron quantum cyclotron: Measuring the electron magnetic moment. *Physical Review A*, 83:052122, 2011.

[27] V. I. Borodulin, R. N. Rogalyov, and S. R. Slabospitsky. CORE 2.1 (COmpendium of RElations, Version 2.1). *ArXiv e-prints*, (hep-ph/9507456), 1995.

[28] B. V. Struminsky. Magnetic Moments of Baryons in the Quark Model. *JINR Publications*, (P-1939), 1965.

[29] M. Y. Han and Y. Nambu. Three-Triplet Model with Double SU(3) Symmetry. *Physical Review*, 139:1006–1010, 1965.

[30] O. W. Greenberg. Spin and Unitary-Spin Independence in a Paraquark Model of Baryons and Mesons. *Physical Review Letters*, 13:598–602, 1964.

[31] M. Perl, E. Lee, and D. Loomba. Searches for Fractionally Charged Particles. *Annual Review of Nuclear and Particle Science*, 59:47–65, 2009.

[32] P. W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Physical Review Letters*, 13:508–509, 1964.

[33] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Physical Review Letters*, 13:321–323, 1964.

[34] S. Weinberg. Physical Processes in a Convergent Theory of the Weak and Electromagnetic Interactions. *Physical Review Letters*, 27:1688–1691, 1971.

[35] M. Kobayashi and T. Maskawa. CP-Violation in the Renormalizable Theory of Weak Interaction. *Progress of Theoretical Physics*, 49:652–657, 1973.

[36] Z. Maki, M. Nakagawa, and S. Sakata. Remarks on the Unified Model of Elementary Particles. *Progress of Theoretical Physics*, 28:870–880, 1962.

[37] L. Wolfenstein. Parametrization of the Kobayashi-Maskawa Matrix. *Physical Review Letters*, 51:1945–1947, 1983.

[38] R. Aaij et al. A Model-Independent Dalitz Plot Analysis of $B^{\pm} \to DK^{\pm}$ with $D \to K^0_{\mathrm{S}} h^+ h^-$ $(h = \pi, K)$ Decays and Constraints on the CKM Angle γ. *Physics Letters B*, 718:43–55, 2012.

[39] UTFit. `http://www.utfit.org/UTfit/`.

[40] CKMFitter. `http://ckmfitter.in2p3.fr/www/html/ckm_main.html`.

[41] E. Noether. Invariante Variationsprobleme. *Nachrichten von der Gesellschaft der Wissenschaften zu Gttingen, Mathematisch-Physikalische Klasse*, pages 235–257, 1918.

[42] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2π Decay of the K^0_2 Meson. *Physical Review Letters*, 13:138–140, 1964.

[43] R. Aaij et al. Observation of $D^0 - \overline{D}^0$ Oscillations. *Physical Review Letters*, 110:101802, 2013.

[44] R. H. Good et al. Regeneration of Neutral K Mesons and Their Mass Difference. *Physical Review*, 124:1223–1239, 1961.

[45] A. Abulencia et al. Observation of $B^0_s - \bar{B}^0_s$ Oscillations. *Physical Review Letters*, 97:242003, 2006.

[46] H. Albrecht et al. Observation of $B^0 - \bar{B}^0$ Mixing. *Physics Letters B*, 192:245, 1987.

[47] P. del Amo Sanchez et al. Measurement of D^0-$\overline{D}^0$ Mixing Parameters Using $D^0 \to K^0_{\mathrm{S}} \pi^+ \pi^-$ and $D^0 \to K^0_{\mathrm{S}} K^+ K^-$ Decays. *Physical Review Letters*, 105:081803, 2010.

[48] Y. Amhis et al. Averages of B-Hadron, C-Hadron, and τ-Lepton Properties as of Early 2012. *ArXiv e-prints*, (arXiv:1207.1158v2), 2012.

[49] R. H. Dalitz. Decay of τ Mesons of Known Charge. *Physical Review*, 94:1046–1051, 1954.

[50] M. Blatt and V. Weisskopf. *Theoretical Nuclear Physics.* Springer-Verlag, 1979.

[51] F. von Hippel and C. Quigg. Centrifugal-barrier Effects in Resonance Partial Decay Widths, Shapes, and Production Amplitudes. *Physical Review D*, 5:624, 1972.

[52] G. Breit and E. Wigner. Capture of Slow Neutrons. *Physical Review*, 49:519–531, 1936.

[53] 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 $K^0_SK^+K^-$. *Physical Review D*, 78:034023, 2008.

[54] D. Aston et al. A study of $K\pi^+$ Scattering in the Reaction $K^-p \to K^-\pi^+n$ at 11 GeV/c. *Nuclear Physics B*, 296:493–526, 1988.

[55] K. Abe et al. Measurement of D^0-$\overline{D}^0$ Mixing Parameters in $D^0 \to K^0_{\mathrm{S}}\pi^+\pi^-$ decays. *Physical Review Letters*, 99:131803, 2007.

[56] D. M. Asner et al. Search for D^0-$\bar{D}^0$ Mixing in the Dalitz Plot Analysis of $D^0 \to K^0_S\pi^+\pi^-$. *Physical Review D*, 72:012001, 2005.

[57] A. Giri et al. Determining γ Using $B^{\pm} \to DK^{\pm}$ with Multibody D Decays. *Physical Review D*, 68:054018, 2003.

[58] J. Libby et al. Model-Independent Determination of the Strong-Phase Difference Between D^0 and $\bar{D}^0 \to K_{S,L}{}^0h^+h^-$ (h=π, K) and its Impact on the Measurement of the CKM Angle γ/ϕ_3. *Physical Review D*, 82:112006, 2010.

[59] A. Bondar and A. Poluektov. Feasibility Study of Model-Independent Approach to ϕ_3 Measurement Using Dalitz Plot Analysis. *The European Physical Journal C*, 47:347–353, 2006.

[60] C. Thomas and G. Wilkinson. Model-independent $D^0 - \bar{D}^0$ Mixing and CP Violation Studies with $D^0 \to K^0_{\mathrm{S}}\pi^+\pi^-$ and $D^0 \to K^0_{\mathrm{S}}K^+K^-$. *Journal of High Energy Physics*, 1210:185, 2012.

[61] TeVI Group. *Design Report Tevatron 1 Project.* Fermilab, 1984.

[62] P. Lukens. *The CDF IIb detector Technical design report.* Fermilab, 2003.

[63] A. Abashain et al. The Belle Detector. *Nuclear Instruments and Methods in Physics Research A*, 79:117–232, 2002.

[64] T. A. Aaltonen et al. Observation of $D^0 - \overline{D}^0$ Mixing Using the CDF II Detector. *Physical Review Letters*, 111:231802.

[65] B. R. Ko et al. Observation of $D^0 - \bar{D}^0$ Mixing in e^+e^- Collisions. *Physical Review Letters*, 112:139903, 2014.

[66] R. Aaij et al. Measurement of D^0-$\overline{D}^0$ Mixing Parameters and Search for CP Violation Using $D^0 \to K^+\pi^-$ Decays. *Physical Review Lettters*, 111:251801, 2013.

[67] R. Aaij et al. Measurement of mixing and CP violation parameters in two-body charm decays. *Journal of High Energy Physics*, 1204:129, 2012.

[68] G. Casarosa. Measurement of D^0–$\overline{D}^0$ Mixing and CP Violation at BaBar. Proceedings of the 5th International Workshop on Charm Physics, 2012.

[69] T. Abe et al. *Belle II Technical Design Report.* KEK Research Orgnisation, 2010.

[70] M. Ablikim et al. Design and Construction of the BESIII Detector. *Nuclear Instruments and Methods in Physics Research A*, 614:345–399, 2010.

[71] D. Ambrose and others. Measurement of the Relative Strong-Phase Difference Between D^0 and $\overline{D}^0 \to K^0_S\pi^+\pi^-$. Presented at the APR14 Meeting of the American Physical Society, Savannah, Georgia, USA, 2014.

[72] O. Brüning et al. *LHC Design Report.* CERN.

[73] G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3:S08003, 2008.

[74] S. Chatrchyan et al. The CMS Experiment at the CERN LHC. *Journal of Instrumentation*, 3:S08004, 2008.

[75] H. Bachacou. Searches for Phenomena Beyond the Standard Model at the Large Hadron Collider with the ATLAS and CMS Detectors. *Pramana*, 79:691–701, 2012.

[76] N. Rompotis. Beyond Standard Model Higgs boson Physics with the ATLAS Experiment at the LHC. Proceedings of the LHCP Conference, Columbia University, New York, 2014.

[77] M. Flechl. BSM Higgs Results from ATLAS and CMS. *European Physics Journal: Web Conference Series*, 60:02005, 2013.

[78] J. Kretzschmar. Electroweak Measurements with the ATLAS and CMS Experiments. Proceedings of the 47th Recontres de Moriond, EW Session, 2012.

[79] P. Starovoitov. Impact of ATLAS Data on Parton Density Functions. Proceedings of the Large Hadron Collider Physics (LHCP) Conference, New York, Jun 2014.

[80] I. V. Gorelov. Heavy Flavor Physics at ATLAS and CMS. *Nuclear Physics B Procedings and Supplements*, 245:168–172, 2013.

[81] S. N. White. Heavy Ion Physics with the ATLAS Detector. *Acta Physica Hungaria A*, 25:531–537, 2006.

[82] P. Cortese et al. *ALICE Physics Performance: Technical Design Report.* CERN, 2005.

[83] The ALICE Collaboration. *ALICE: Technical proposal for a Large Ion collider Experiment at the CERN LHC.* CERN, 1995.

[84] P. Foka. Overview of Results from ALICE at the CERN LHC. *Journal of Physics: Conference Series*, 455:012004, 2013.

[85] R. Preghenella. ALICE Physics Summary. *European Physical Journal: Web Conference Series*, 49:02003, 2013.

[86] S. Shingo. Heavy-Flavour Production in ALICE. *Nuclear Physics A*, 926:41–48, 2014.

[87] R. Aaij et al. *Framework TDR for the LHCb Upgrade: Technical Design Report.* CERN, 2012.

[88] R. Antunes-Nobrega et al. *LHCb Reoptimized Detector Design and Performance: Technical Design Report.* Technical Design Report LHCb. CERN, 2003.

[89] A. Augusto Alves et al. The LHCb Detector at the LHC. *Journal of Instrumentation*, 3:S08005, 2008.

[90] R. Aaij et al. Measurement of B Meson Production Cross-Sections in Proton-Proton Collisions at $\sqrt{s} = 7$ TeV. *Journal of High Energy Physics*, 1308:117, 2013.

[91] R. Aaij et al. Prompt Charm Production in pp Collisions at $\sqrt{s} = 7$ TeV. *Nuclear Physics B*, 871:1–20, 2013.

[92] A. J. Buras et al. On the Standard Model Prediction for $B_{\{s,d\}} to \mu^+\mu^-$. *European Physical Journal C*, 72:2172, 2012.

[93] F. Mahmoudi, S. Neshatpour, and J. Orloff. Supersymmetric Constraints from $B_s \to \mu^+\mu^-$ and $B \to K * \mu^+\mu^-$ Observables. *Journal of High Energy Physics*, 1208:92, 2012.

[94] R. Aaij et al. Measurement of the $B_s^0 \to \mu^+\mu^-$ Branching Fraction and Search for $B^0 \to \mu^+\mu^-$ Decays at the LHCb Experiment. *Physical Review Letters*, 111:101805, 2013.

[95] S. Chatrchyan et al. Measurement of the $B_{(s)} \to \mu^+ \ \mu^-$ Branching Fraction and Search for $B^0 \to \mu^+ \ \mu^-$ with the CMS Experiment. *Physical Review Letters*, 111:101804, 2013.

[96] Sheldon Stone. New Physics from Flavour. Proceedings of the 36th International Conference on High Energy Physics, Melbourne, Australia, 2013.

[97] A. Affolder et al. Radiation Damage in the LHCb Vertex Locator. *Journal of Instrumentation*, 8:P08002. 23 p, 2013.

[98] P. Rodríguez Pérez. The LHCb VERTEX LOCATOR performance and VERTEX LOCATOR upgrade. *Journal of Instrumentation*, 7, 2012.

[99] The LHCb VELO Group. Performance of the LHCb Vertex Locator. *ArXiv e-prints*, (arXiv:1405.7808v2), 2014.

[100] M. Adinolfi et al. Performance of the LHCb RICH Detector at the LHC. *European Physical Journal C*, 73:2431, 2013.

[101] R. Aaij et al. Opposite-side Flavour Tagging of B Mesons at the LHCb Experiment. *European Physics Journal C*, 72:2022, 2012.

[102] T. Bellunato et al. Performance of Aerogel as a Cherenkov Radiator. *Nuclear Instruments and Methods in Physics Research A*, 519:493–507, 2004.

[103] S. Amato et al. *LHCb Magnet: Technical Design Report.* CERN (LHCb-TDR-1), 2000.

[104] F. Archilli et al. Performance of the Muon Identification at LHCb. *Journal of Instrumentation*, 8:P10020. 17 p, 2013.

[105] R. Aaij et al. The LHCb Trigger and its Performance in 2011. *Journal of Instrumentation*, 8:P04022, 2013.

[106] R. Aaij et al. Measurement of Mixing and CP Violation Parameters in Two-Body Charm Decays. *Journal of High Energy Physics*, 1204:129, 2012.

[107] T. Sjöstrand, S. Mrenna, and P. Skands. PYTHIA 6.4 Physics and Manual. *Journal of High Energy Physics*, 5:26, 2006.

[108] D. J. Lange. The EvtGen particle decay simulation package. *Nuclear Instruments and Methods in Physics Research A*, 462:152–155, 2001.

[109] P. Golonka and Z. Was. PHOTOS Monte Carlo: A Precision Tool for QED Corrections in Z and W Decays. *European Physical Journal C*, 45:97–107, 2006.

[110] O. Shekhovtsova, I. M. Nugent, T. Przedzinski, P. Roig, and Z. Was. MC Generator TAUOLA: Implementation of Resonance Chiral Theory for Two and Three Meson Modes. Comparison with Experiment. American Institute of Physics Conference Series, 2012.

[111] V. Gligorov. A Single Track HLT1 Trigger. Technical Report LHCb-PUB-2011-003, CERN, Jan 2011.

[112] Wouter D. Hulsbergen. Decay Chain Fitting with a Kalman Filter. *Nuclear Instruments and Methods in Physics Research A*, 552:566–575, 2005.

[113] R. Bailey et al. Measurement of the Lifetime of Charged and Neutral D Mesons with High Resolution Silicon Strip Detectors. *Zeitschrift für Physik C Particles and Fields*, 28, 1985.

[114] T. Aaltonen et al. Measurement of the B^- Lifetime Using a Simulation Free Approach for Trigger Bias Correction. *Physical Review D*, 83, 2011.

[115] V. Gligorov. *Measurement of the CKM Angle Gamma and B Meson Lifetimes at the LHCb Detector.* University of Oxford, 2008.

[116] M. Gersabeck. *Alignment of the LHCb Vertex Locator and Lifetime Measurements of Two-Body Hadronic Final States.* University of Glasgow, 2009.

[117] M. Alexander. *Constraints on Mixing and CP-Violation in the Neutral Charmed Meson System at LHCb.* University of Glasgow, 2012.

[118] N. Torr V. Glivorov et al. Swimming : a Data Driven Acceptance Correction Algorithm. *Journal of Physics: Conference Series*, 396:022016, 2012.

[119] R. Aaij and other. Precision Measurement of the ${\Lambda_b}^0$ Baryon Lifetime. *Physical Review Letters*, 111:102003, 2013.

[120] W. Verkerke and D. Kirkby. The RooFit Toolkit for Data Modeling. *ArXiv e-prints*, (physics/0306116), 2003.

[121] F. James and M. Roos. MINUIT: A System for Function Minimization and Analysis of the Parameter Errors and Corrections. *Computer Physics Communications*, 10:343–367, 1975.

[122] M. Oreglia. *A Study of the Reactions* $\psi' \to \gamma\gamma\psi$. SLAC-0236, 1980.

[123] `http://en.wikipedia.org/wiki/Kernel_density_estimation`.

[124] B. W. Silverman. *Density Estimation for Statistics and Data Analysis.* Chapman & Hall, 1986.

[125] T. Bayes and R. Price. An Essay towards Solving a Problem in the Doctrine of Chances. *Philosophical Transactions*, 53, 1763.

[126] R. Aaij et al. Measurements of the B^+, B^0, B_s^0 Meson and Λ_b^0 Baryon Lifetimes. *ArXiv e-prints*, (1402.2554), 2014.

[127] The LHCb Collaboration. A search for Time-integrated CP Violation in $D^0 \to h^+h^-$ Decays and a Measurement of the D^0 Production Asymmetry. 13th International Conference on B-Physics at Hadron Machines, Amsterdam, Netherlands, 2011.