# DISSERTATION

submitted to the
Combined Faculties for the Natural Sciences and Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of

## doctor rerum naturalium

presented by

## Dipl.-Phys. Andreas Jäger

born in Tashkent

Oral examination: April 15th, 2015

# Measurement of indirect CP asymmetries in $D^0 \to K^- K^+$ and $D^0 \to \pi^- \pi^+$ decays with the LHCb experiment

Referees:   Prof. Dr. Stephanie Hansmann-Menzemer

Prof. Dr. Hans-Christian Schultz-Coulon

## Abstract

This thesis presents a measurement of indirect CP asymmetries in the charm system. The indirect CP violation for the charm-meson system is expected to be small in the Standard Model and has not been observed experimentally so far. Effects of beyond Standard Model physics phenomena but as well perturbatively non-calculable processes can manifest themself in a potential non-zero CP violation. The knowledge of the size of CP violation makes a significant contribution for the understanding of the Standard Model and has a direct impact on the effective parameters in the charm system.

The LHCb experiment has collected the world largest data samples of charm mesons. The proton-proton collisions data used in this thesis have been recorded in the first run period in 2011-2012. They correspond to an integrated luminosity of $3.0\,\mathrm{fb}^{-1}$. The measurement uses the singly Cabibbo-suppressed $D^0 \to K^- K^+$ and $D^0 \to \pi^- \pi^+$ decays. The charm mesons exploited in this analysis are produced in semi-muonic B-meson decays. Thereby, the charge of the created muon determines the flavour of the charm meson at production time. To perform this measurement, a robust method is developed that reaches a sensitivity at sub-permille level.

The indirect CP asymmetry manifests itself in the asymmetries of effective lifetimes, $A_\Gamma$, of the $D^0$ and $\bar{D}^0$ decays and is measured for both decays

$$A_\Gamma(K^- K^+) = (-0.134 \pm 0.077\,{}^{+0.026}_{-0.034})\%\,,$$

$$A_\Gamma(\pi^- \pi^+) = (-0.092 \pm 0.145\,{}^{+0.025}_{-0.033})\%\,,$$

where the first uncertainty is statistical and the second systematic. The results are compatible with previous measurements and with the zero hypothesis of no CP violation. This analysis provides a significant contribution to the effective description of the charm meson system and to the world average. The results have been submitted for publication to the Journal of High Energy Physics (JHEP) [1].

## Kurzfassung

In dieser Arbeit wird die Messung der indirekter CP-Asymmetrie im Charm-System vorgestellt. Die indirekte CP-Verletzung im $D^0$-Mesonen System wird im Standardmodell als sehr klein vorhergesagt und konnte bislang experimentell nicht nachgewiesen werden. Physik Phänomene jenseits des Standardmodells aber auch störungstheoretisch nicht berechenbare Prozesse können dadurch potenziell nachgewiesen werden. Die Kenntnis der Größe der CP-Verletzung leistet einen wichtigen Beitrag für das Verständnis des Standardmodells und hat einen direkten Einfluss auf die Parameter der effektiven Beschreibung des Charm-Systems.

Das LHCb-Experiment verfügt über den weltweit größten Daten-Satz an Charm-Mesonen, die in den Proton-Proton-Kollisionen aufgenommen wurden. Der hier verwendete Daten-Satz wurde in der Periode von 2011-2012 aufgezeichnet und entspricht einer integrierten Luminosität von $3.0\,\mathrm{fb}^{-1}$. Diese Messung verwendet die einfach Cabibbo-unterdrückten $D^0 \to K^- K^+$ und $D^0 \to \pi^- \pi^+$ Zerfälle. Das besondere an dieser Analyse ist, dass die analysierten Charm-Mesonen in semi-myonischen B-Zerfällen erzeugt werden. Hierbei legt die Ladung des dabei entstehendes Myons den Charm-Meson-Flavour zum Produktionszeitpunkt fest. Um diese Messung durchführen zu können wurde eine Methode entwickelt, die eine Sensitivität von unter $\mathcal{O}(10^{-3})$ erreicht.

Die indirekte CP-Asymmetrie zeigt sich in der Asymmetrie der effektiven Lebensdauern zwischen $D^0$-und $\bar{D}^0$-Zerfällen, $A_\Gamma$. Folgende Werte wurden gemessen

$$A_\Gamma(K^- K^+) = (-0.134 \pm 0.077\,{}^{+0.026}_{-0.034})\%\,,$$

$$A_\Gamma(\pi^- \pi^+) = (-0.092 \pm 0.145\,{}^{+0.025}_{-0.033})\%\,.$$

Dabei ist die erste Unsicherheit statistisch und die zweite systematisch. Die Resultate sind kompatibel mit den vorhergegangenen Messungen und mit der Null-Hypothese, d. h. CP-Erhaltung. Die Messung leistet einen signifikanten Beitrag zur effektiven Beschreibung des Charm-Systems. Die Resultate sind bei Journal of High Energy Physics (JHEP) zur Veröffentlichung eingereicht, siehe Ref. [1].

# Contents

# Introduction

Symmetries have always been of great interest not only in physics, but in mathematics, philosophy and in many other fields. Symmetries form the basis of modern physics. Based on symmetry considerations, the Standard Model (SM) of particle physics was developed throughout the past decades. The SM has been tested in many experiments; many of its predictions have been discovered afterwards. Recently, the last missing and long searched for particle, the Higgs boson, was discovered at the Large Hadron Collider by the ATLAS and CMS experiments. The SM, as we know it today, is able to explain the known fundamental interactions, except gravitation.

However, it is known that the SM is not complete. For example, one of the important questions, "Why does our universe contain more matter than antimatter?", remains unsolved. An important process that is responsible for such an asymmetry is the violation of charge-parity (CP) symmetry. In the SM, CP asymmetry is introduced through the Cabibbo-Kobayashi-Maskawa matrix. However, the size introduced by this mechanism is not sufficient to explain the asymmetry observed in the universe. Cosmological observations show that the universe contains a small amount of baryonic matter, which is described by the SM, with a density of 4.99%, but it is dominated by the dark energy (68.5%) and cold dark matter (26.5%), Ref. [2]. Until now, the SM can not explain the amount of dark matter in the universe. Additionally, 28 parameters are required to describe all processes of electroweak and strong interactions. There is, however, no explanation for the number of free parameters nor of their values. All these shortcomings of the SM let physicists believe that there must be a more fundamental underlying theory, so-called New Physics. Therefore, it is important to search for the hints towards such a theory that would revolutionize our understanding of the universe.

One important system in the Stardard Model is the neutral $D^0$ meson system. The $D^0$ meson can mix with its anti-particle $\overline{D}^0$. The $D^0$ mixing, however, is rather small and was just recently established with a single experiment at LHCb, Ref. [3]. This small non-zero mixing opens up the possibility to search for the CP asymmetry that can appear through mixing and interference with the decay, which manifests in an indirect CP violation. CP violation in the charm system is predicted to be small in the SM and has not been observed so far. However, non-perturbative effects or New Physics can have an impact on the observed asymmetries. The objective of this thesis is to develop a method that can measure indirect CP asymmetry in the $D^0$-meson system with a precision below $\mathcal{O}(10^{-3})$.

The LHCb experiment has recorded the word largest samples of D mesons. The dataset used for the measurement presented in this thesis was recorded in 2011 and 2012 and corresponds to an integrated luminosity of $3\,\mathrm{fb}^{-1}$. The two main channels studied in this analysis are $D^0 \to K^- K^+$ and $D^0 \to \pi^- \pi^+$. Both decays are singly-Cabibbo suppressed (SCS) and correspond to CP-even final states. The $D^0$ mesons exploited in this thesis are produced in the semi-muonic B decays. The charge of the muons of the $B$ decays, $B \to \overline{D^0}\mu^+ X$ and $\overline{B} \to D^0\mu^- \overline{X}$, respectively, is used to identify the $D^0$ flavour at production time (flavour tagging). This is the first measurement of indirect $CP$ violation in $D^0 \to K^- K^+$ and $D^0 \to \pi^- \pi^+$ decays using semileptonic $B$ decays.

This measurement has an impact on the limits of CP observables in the charm sector. It therefore improves the understanding of the SM and has the potential for a first indirect indication of New Physics. The results have been submitted by the author of this thesis to JHEP [1], and are also documented in an internal LHCb Note [4].

The thesis is structured in the following way: In the first part, the basics of the Standard Model

and its symmetries are introduced. This is followed by Chapter 2 which describes the LHCb experiment with all the relevant details of the detector apparatus. Afterwards, in Chapter 3 the properties of charm meson mixing are explained, and the relevant observables for this thesis are introduced. Additionally, besides the detector induced asymmetries, the production asymmetries and flavour tagging induced asymmetries are introduced here. In the second part, starting with Chapter 6, the method for the extraction of the indirect CP asymmetry is presented. The selection requirements are explained together with the B candidate reconstruction in Chapter 5. The results are shown in Chapter 7 and are followed by a detailed determination of systematic uncertainties in Chapter 8. Finally, the results are summarised in Chapter 9 and the conclusions are drawn in Chapter 10.

# 1

# Theory

This chapter summarizes the Standard Model (SM) of particle physics. It gives an overview of all fundamental particles involved and explains the gauge group $SU(3) \times SU(2) \times U(1)$. Further, the interactions are discussed and also the Cabibbo-Kobayashi-Maskawa matrix and the mechanism which introduces CP violation in the Standard Model is described. This basic introduction is followed by a more detailed discussion of the theoretical concepts relevant for this thesis in Chap. 3.

## 1.1 The Standard Model of particle physics

The Standard Model of particle physics is a theoretical model that describes the electromagnetic, weak and strong interaction of the fundamental particles. The basis of the SM (see [5] and [6]), was developed during the second part of the last century. The SM is very successfully describing and predicting experimental observations and reflects our current best knowledge of particle physics.

The SM origins from a relativistic quantum field theory, thereby the system can be described through a Hamilton or Lagrange density. The quanta of the field can be identified as fundamental particles.They can be split in bosons (integer spin) and fermions (half-integer spin).

All the fundamental fermions of the SM are grouped in two species, namely leptons and quarks as shown in Tab. 1.1. The leptons appear in doublets of three generations. That means, for each lepton (electron, muon or tau), which has a negative charge, a corresponding neutrino, which has no electric charge, from the same generation exists. The quarks appear in doublets as well. The so called up-type quarks are u (up), c (charm) and t (top or sometimes called truth) and have the charge +2/3. For each generation the corresponding down-type quarks, d (down), s (strange) and b (bottom or sometimes called beauty) have the electric charge −1/3. Additionally, the quarks have a color (red, green or blue). For every fermion from Tab. 1.1 a corresponding anti fermion with the same mass exists. Anti fermions have an opposite electric charge and an anti color.

Furthermore, the interaction observed in nature can be described by gauge bosons. Table 1.2 shows all the fundamental spin-1 bosons. Every fundamental interaction in the SM is based on a symmetry and follows from the according gauge invariance. To find out the form of interaction one needs to find the corresponding symmetry.

**Table 1.1:** *Fundamental fermions of the SM. The right part of the table shows also the left and right-handed components.*

| Fermions | Generation | | | el. charge | color | weak isospin T | |
|---|---|---|---|---|---|---|---|
| | I | II | III | Q | | left-handed, $T$, $T_3$ | right-handed, T |
| Leptons (l) | $\nu_e$ | $\nu_\mu$ | $\nu_\tau$ | 0 | – | $\binom{\nu_l}{l}_L$, 1/2, $\begin{matrix}1/2\\-1/2\end{matrix}$ | - |
| | $e$ | $\mu$ | $\tau$ | -1 | | | $l_R$, 0 |
| Quarks (q) | $u$ | $c$ | $t$ | 2/3 | r,g,b | $\binom{q^u}{q^d}_L$, 1/2, $\begin{matrix}1/2\\-1/2\end{matrix}$ | $q^u{}_R$,0 |
| | $d$ | $s$ | $b$ | $-1/3$ | | | $q^d{}_R$,0 |

**Table 1.2:** *The gauge bosons of the SM.*

| Gauge bosons | Interaction | Mass[ GeV/$c^2$] | Symmetry group | |
|---|---|---|---|---|
| $\gamma$ | electromagnetic | 0 | $U(1)_{\text{em}}$ | $U(1)_Y \times SU(2)_{\text{L}}$ |
| $W^\pm, Z^0$ | weak | $M_{W^\pm} \approx 80.4$, $M_Z \approx 91.2$ | $SU(2)_{\text{L}}$ | |
| 8 Gluons ($g_i$) | strong | 0 | $SU(3)_C$ | |

**Electroweak unification**

In case of Quantum Electrodynamics (QED) every charge carrying field transforms under the unitary symmetry group U(1). Noether's Theorem predicts for such a symmetry of global transformation a corresponding conserved quantity. Going further to the gauge principle, a local gauge phase transformation can be used. Under local phase transformation a field can be deduced, that can be identified as the photon. Similary to a global phase invariance, that leads to the charge conservation, a local phase invariance here gives rise to a photon field.

Studying the nuclear interactions, additional symmetries and conservation laws have been proposed. Motivated by the observation of parity violation, a chiral gauge theory was developed. It splits the fields into chiral fields and introduces the left-handed and right-handed fermions. The symmetry group accordingly rotates the leptons and quarks within the generation, e.g. rotating the left-handed electron to a left-handed neutrino which is not possible with right-handed particles. The weak interaction requires the introduction of a $SU(2)$ symmetry group (special unitary or Lie group). One generation of left-handed quarks or leptons builds a doublet, with isospin $T = 1/2$, where the $SU(2)_{\text{L}}$ acts on the pair of left-handed fermions, exchanging the $W^\pm$-bosons (see Tab. 1.1). This gives rise to the $W^i$ field and the corresponding $W^\pm$ and $W^0$ bosons which build a triplet of the weak isospin T ($T = 1, T_3 = \pm 1, 0$). However, $W^0$ is not a physical observable field. For the electroweak unification a $U(1)_Y$ group is required, which is connected to the weak hypercharge

$$Y = 2(Q - T_3).$$

Thereby Q is the electric charge and $T_3$ the third component of the weak isospin T. The corresponding field that couples to this hypercharge is $B^0$. The $U(1)_Y$ does not mix up the weak isospin and the $SU(2)_{\text{L}}$ does not mix up the hypercharge. Therefore it can be written as product structure. Which results in the electroweak theory $SU(2)_{\text{L}} \times U(1)_Y$. Further, rotating the $W^0$ and $B^0$ fields using the weak mixing angel or Weinberg angle $\theta_W$, one can obtain the physical fields: $Z^0$, $A_\mu$ (or $\gamma$) and $W^\pm$.

**Symmetry breaking**

Beside the massless $\gamma$, Tab. 1.2 shows that the $W^\pm$ and the $Z^0$ bosons are pretty heavy particles. However, one can not just add a mass term to the Lagrange density for these particles, the theory

would be not gauge invariant. In the SM this is done through symmetry breaking, that requires the introduction of a Higgs field. The particle that is associated to this field is the Higgs boson with spin 0. The source of this Higgs boson is very different compared to the gauge bosons which originate from the gauge symmetries. Here, the Higgs field can be directly postulated. The field can change a left-handed particle to a right-handed one and the other way around giving a mass term to the fermions. The Higgs field together with the Yukawa coupling determines the mass of the fermions. Additionally, the Higgs field also gives the mass to the gauge bosons. All the gauge bosons have been measured with an incredible accuracy and recently also the Higgs boson (with mass near $125\,\mathrm{GeV}/c^2$) has been discovered at the LHC.

**Strong Interaction**

The Quantum Chromodynamics (QCD) describes the strong interactions and is based on a symmetry $SU(3)$ group, also called color group $SU(3)_C$ [1]. In the SM each quark carries a color charge, similarly to the electric charge in QED. The leptons and the $\gamma$, $Z^0$ and $W^\pm$ bosons are color neutral and thus do not participate in this process. The $SU(3)_C$ group has 8 generators that can transform the quark color. These generators are associated with the new massless gauge field and the existence of gluons (see Tab. 1.2). The 8 gluons themselves need to carry a color and anti color. This leads to the fact that the gluons can interact with each other. Stable particle in nature are color neutral, therefore a quark can not be observed as a free particle. The observed particles that are bound together by the strong interaction are hadrons. They consists of quarks and interacting gluons. They can be further separated in mesons consisting of a quark and an anti quark, and baryons consisting of three quarks. Recently, a candidate for a four-quark state has been discovered by the LHCb experiment.

**Cabibbo-Kobayashi-Maskawa (CKM) matrix and generation of CP violation**

So far, the $SU(2)_L$ symmetry, with interacting W bosons, is considered as a symmetry transformation within one generation of leptons or quarks. However, in case of quarks a transition between different generations is observed in reality. This is described in the SM by the mechanism of the Cabibbo-Kobayashi-Maskawa (CKM) matrix. There is a difference between the free quarks that are described by the mass eigenstates $(d,s,b)$ and the quarks that take part in the weak interaction the so-called flavour eigenstates $(d',s',b')$. The mass is given by the Yukawa coupling and the Higgs field. In the SM the change of the bases is described by the Cabibbo-Kobayashi-Maskawa (CKM) matrix

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \underbrace{\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}}_{V_{CKM}} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.1}$$

The nine free complex (18 real) elements are not independent. The unitarity requires:

$$\sum_{k=1}^{3} V_{ik} V_{jk}^* = \delta_{ij}, \quad i,j \in \{1,2,3\}. \tag{1.2}$$

This removes 9 degrees of freedom. Additionally, the global phase is not an observable therefore every quark field can absorb one additional parameter. Finally, four parameters remain. They can be parametrised as 3 rotation angles and leave one non trivial phase, which can lead to an observable interference effects and is the only source of CP violation in the SM.

---

[1]The $SU(3)$ group was postulated to explain the existence of the $\Delta^{++}$ baryon, since without additional quantum number the Pauli principle would be violated.

The matrix is almost diagonal, thus the absolute values of the off-diagonal elements decrease changing the generation. One possible parametrization of the CKM matrix is the Wolfenstein parametrization, which is described by the following four real parameters, taken from Ref. [2].

$$\lambda \approx 0.22537 \tag{1.3}$$
$$A \approx 0.814 \tag{1.4}$$
$$\bar{\rho} \approx 0.117 \tag{1.5}$$
$$\bar{\eta} \approx 0.353 \tag{1.6}$$
$$\tag{1.7}$$

with $\bar{\rho} = \rho(1 - \lambda/2 + \cdots)$ and $\bar{\eta} = \eta(1 - \lambda/2 + \cdots)$. Up to $\mathcal{O}(\lambda^5)$ this can be expressed, as in Ref. [7],

$$V_{\text{Wolf}}^{\text{(CKM)}} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} - \frac{\lambda^4}{8} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda + \frac{\lambda^5}{2}A^2(1 - 2\rho - 2i\eta) & 1 - \frac{\lambda^2}{2} - \frac{\lambda^4}{8}(1 + 4A^2) & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) + \frac{\lambda^5}{2}A(\rho + i\eta) & -A\lambda^2 + \frac{\lambda^4}{2}A(1 - 2\rho - 2i\eta) & 1 - \frac{\lambda^4}{2}A^2 \end{pmatrix} + \mathcal{O}(\lambda^6) \,. \tag{1.8}$$

The complex phase difference is contained in the term $(\eta + i\rho)$ [2] which is suppressed by $V_{cd} \sim \mathcal{O}(A^2\lambda^5)$ and $V_{td}, V_{ub} \sim \mathcal{O}(A\lambda^3)$ and $V_{ts} \sim \mathcal{O}(A\lambda^4)$ depending on the matrix element. Note that imaginary terms depending on the order can be transfered between the matrix elements satisfying the unitarity relations (original parametrizations can be found in Ref. [8] and [7]). This parametrization illustrates clearly the smallness of CP violation in the SM.

---

[2]Both parameters $\eta$ and $\rho$ are of the same order.

# 2

# LHCb experiment

## 2.1    Particle production at high-energy colliders

In the past decades several experiments have used different particle production mechanisms to study elementary particles in a wide range of energies. Among those are fixed target experiments and $e^+$-$e^-$, proton-electron, proton-meson, proton-anti-proton and proton-proton colliders.

Currently, the highest center of mass energies, $\sqrt{s}$, in collider experiments are achieved in proton-proton (pp) collisions, which is the main topic of this section. Going to higher $\sqrt{s}$ the production cross sections of especially heavy particles increase drastically. Understanding the mechanism of hadron production in these collisions requires a deep understanding of the proton structure and of the QCD processes involved.

**Proton Structure**
In the simplest model, the proton is built out of three valence quarks (up, up, down). The gluons are responsible for binding the proton. Beside the valence quarks, the proton contains pairs of quarks and anti-quarks (sea quarks) that are created and annihilate. The densities of the partons are described by the Parton Density Functions (PDFs). Proton PDFs can not be calculated peturbatively, However, they can be measured, e.g. at H1 and ZEUS experiments at HERA, as presented in Ref. [9]. The commonly-used combined results are shown in Fig. 2.1. Here, $Q$ is the transfered four-momentum and $x$ the Bjorken variable quantifying the fraction of the proton momentum carried by the struck quark. At small $Q^2$, the proton momentum is carried mainly by the valence quarks, where each of the partons carries the main momentum fraction x. By increasing $Q^2$, one is looking deeper into the proton and the densities of the gluons and sea quarks increase.

**Particle production**
The production in pp collisions can be separated into two different phases, as described in Ref. [10]. The first one is a perturbative phase, which happens in the hard scattering process of the protons and where the quarks and lepton are mainly produced. The second phase is non-perturbative phase and takes place at low energies where the hadronization happens. [1]

---

[1] Notice, that a correct approach requires the consideration of the factorization scale $\mu^2 = Q^2$ for the parton PDFs of the proton and not only the renormalization scale for the process between the partons.
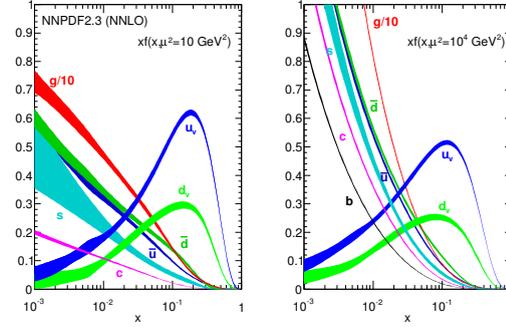
**Figure 2.1:** *The parton distribution function $f(x)$ multiplied by x, for $Q^2 = \mu^2$ values of $10\,\text{GeV}^2$ (left) and $10^4\,\text{GeV}^2$ (right), (Figures taken from Ref. [2]).*
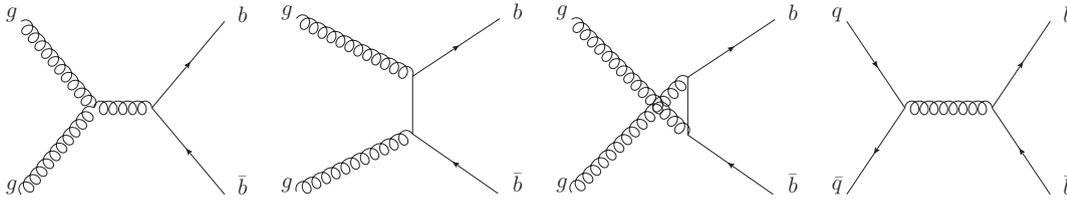


**Figure 2.2:** *Leading Order QCD diagrams for production of heavy quarks (The beauty quark can be replaced by a charm or top quark).*

The first one that takes place at short length scale, where quarks can be produced directly. The description of this process builds the basis for every reliable simulation (see e.g. PYTHIA [11] and HERWIG [12] event generators). In the simulation these processes are often approximated with the parton shower approach, where the processes are split into initial state radiation, hard cascade, and final state radiation. Finally, after production of this parton states, a non-perturbative hadronization or fragmentation of the remaining quarks and gluons takes place when producing final states particles. This can be modeled phenomenologically.

**Calculation of the proton-proton cross section**

The energy scale above which QCD is renormalizable is $\Lambda_{QCD} \approx 200\,\text{MeV}$. Below this scale the strong coupling becomes $\alpha_s(\Lambda_{QCD}) > 1$ and diverges. Therefore, light quarks, with masses $m_u, m_d, m_s \ll \Lambda_{QCD}$ can not be treated perturbatively. The charm, beauty and top quark masses are significantly above this scale, $m_c, m_b, m_t > \Lambda_{QCD}$ and non-perturabtive contributions can be controlled, e.g. in Heavy Quark Effective Theory (HQET). The lifetime of the top quark is too short to form hadrons. For beauty and charm quarks this is possible. The dominant mechanisms (of the leading order $\mathcal{O}(\alpha_s)$ QCD contributions) in pp collisions are visualized in Fig. 2.2, where additional diagrams with gluons in the final state are omitted. The consideration of the gluon emission in the final state requires next-to-leading order (NLO) calculations, see Ref. [13]. Mathematically, one is now able to calculate the cross sections for proton-proton scattering. This requires the knowledge of the structure of the proton, which is described by the PDFs. The short range interactions between the partons can be now calculated perturbatively. Combining these two in the factorization approach, the overall cross section can be calculated.

For charm the calculation can be done up to NLO and the results are shown in Fig. 2.3(b). One can also see that this calculation describes well the data in the shown energy range. Fig. 2.3(a)

**(a)** *NNLO production cross sections*                        **(b)** *NLO production cross section for charm*
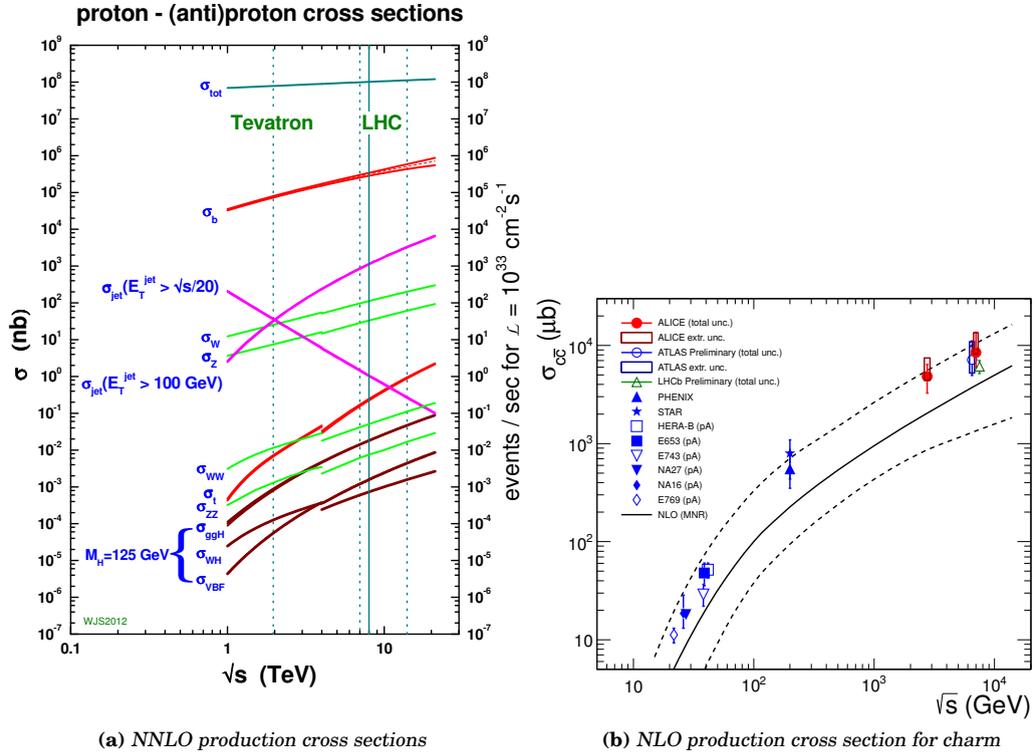
**Figure 2.3:** *(a) The total proton-proton cross section is shown in blue. Other contributions, starting with bottom cross section, are calculated in NLO ( figure taken from MSTW group from Ref. [14], W.J. Stirling, private communication). The small step of the cross section in the intermediate region between Tevatron and LHC correspond to the small difference between proton proton and proton anti-proton collisions.*
*(b) Charm pair production cross section (taken from Ref. [15]), the theoretical prediction has been calculated in NLO (see Ref. [13]).*

shows the next-to-next-to leading order (NNLO) calculations of the cross sections and their dependence on $\sqrt{s}$. For hard scattering the cross sections rise with increasing $\sqrt{s}$.

**Heavy quark production mechanism**

In simulation, often a simplified approach, including initial and final state radiation and the hard processes is used. This simulation approach is called parton shower. Parton shower is able to describe the leading log scale. For deeper understanding of the heavy quark production mechanism one can split the main production processes into three classes: pair creation, flavour excitation and gluon splitting. [2] The process type depends on how many heavy flavour quarks participate in the hard scattering. Fig. 2.4 shows these three main processes. At LHC energies $\sqrt{s} \approx 7 - 8\,\mathrm{TeV}$ the dominant process for beauty as well as for charm production is the flavour excitation (see Ref. [10]).

So far, we have considered only the strong interaction, where $b\bar{b}$ quarks are produced in pairs. However, there is also a single $b$ production possible, e.g. via weak interaction, naively considering the $W$ or top production. For the $b$ quark this contribution is, using this naive calculation, more than $\mathcal{O}(10^{-5})$ suppressed at LHC energies compared to the $b\bar{b}$-pair production.

---

[2]This is also required to avoid double counting.

**(a)** *Pair creation*    **(b)** *Flavour excitation*    **(c)** *Gluon splitting*
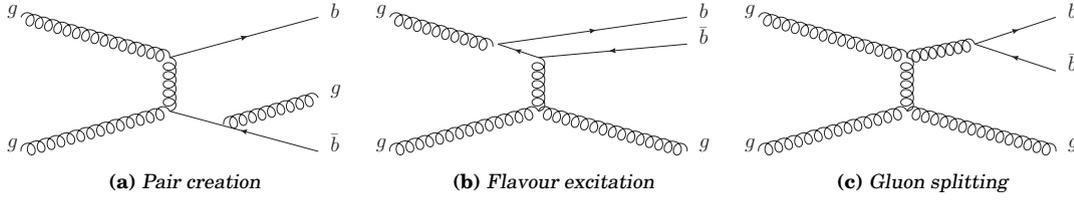
**Figure 2.4:** *NLO QCD feynman graphs. Classification in three groups (a) 2 b quarks, (b) 1 b quarks, (c) 0 b quarks participate in the hard scattering process.*
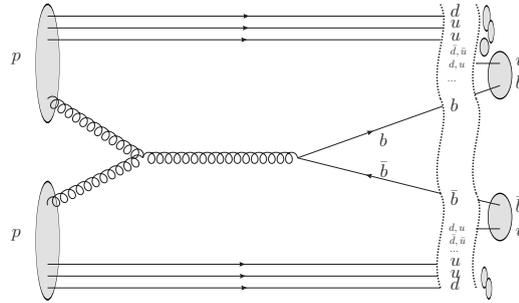


**Figure 2.5:** *Example of a hard scattering process in a pp collision. The heavy quark production is displayed in the center, the hadronization is indicated by the vertical curves on the right side.*

### Hadronization and fragmentation

The obtained partonic states need to hadronize to produce the final state of colorless hadrons. This step can not be calculated perturbatively. However, phenomenological models, e.g. the Lund String model Ref. [16], which treats every color singlet subsystem (e.g string), can handle this.

Fig. 2.5 shows the production mechanism, where $b$ quarks [3] are produced in the hard scattering. In the next step the hadronization is visualized. Thereby, a hadron with an anti-$b$ quark and one with a $b$ quark appears in the final state.

### Production asymmetry

In pp or p$\bar{\text{p}}$-collisions the quarks, as explained previously, are predominantly produced in pairs. In the following, the focus will lie on the pp-collisions. For visualization, Fig. 2.5 shows the production of a $b\bar{b}$ pair, that after fragmentation results here in a $\bar{B}^0(b\bar{u})$ and $B^0(\bar{b}u)$ meson-pair.

For the heavy quark pair production in the perturbative QCD no asymmetry ($A = 0$) is expected, Ref. [17]. However, many experiments have measured production asymmetries for different energies and in different kinematic phase-space region, Ref. [17], [18]. This requires further possible explanation. Different models with theoretical calculations are available that can handle this, e.g. Ref. [19]. In the following, some artefacts are explained.

For the final asymmetry in the production of hadrons one needs to consider the valence quarks of both protons. One general argument is the baryon number conservation law. Since the initial state two protons, the final state has to have the same baryon number, and therefore $\Lambda_b$ production is preferred relative to the $\overline{\Lambda}_b$ production. As the sum of the $b$ hadrons and anti-$b$ hadrons is the same the $b$-baryon asymmetry results as well in a $b$-meson asymmetry.

The second effect is, that the protons provide for the combination the valence $(uud)$ quarks but not their anti-quarks. In the final states more $B^0(\bar{b}u)$ and $B^+(\bar{b}d)$ mesons can be produced

---

[3] The $b$ quark can be replaced by other heavy quarks.
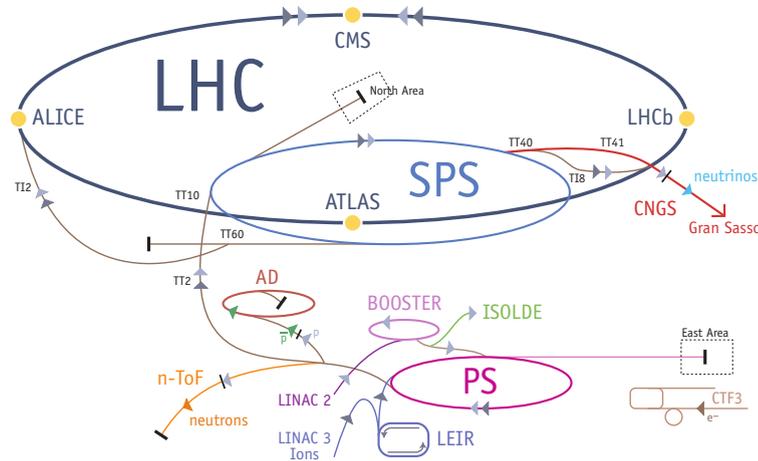
**Figure 2.6:** *Accelerator complex at CERN (taken from Ref. [23]).The four main collision points placed at the LHC ring, where the experiments are set up, are highlighted in yellow.*

with respect to $\overline{B^0}(b\bar{u})$ and $B^-(b\bar{d})$ and therefore create a meson asymmetry. This has then to be compensated by on opposite asymmetry of the remaining B-hadrons. In the fragmentation one needs also to consider different non-trivial color topologies of allowed final states including the beam remnants, that can also lead to an asymmetry in production, depending on the phase space. Current results show that the asymmetries at LHCb are at a level below $\mathcal{O}(1\%)$, Ref. [18].

Additional contribution can arise already before the fragmentation, if one allows some "recombination" in the proton. The proton can e.g. fluctuate in a baryon meson pair in a so called "Meson Cloud Model" Ref. [19], which would result in an observable asymmetry of e.g. charm mesons.

Studying and disentangling the different effects is one of the main topics of currently ongoing production-asymmetry analyses. This thesis will discuss later on how the presence of production asymmetries needs to be taken into account in the analysis of the studied CP asymmetries in the charm system.

**Charm and beauty production cross sections**

Theoretical predictions of the cross sections are well in agreement with the measurements from LHCb. For this analysis the beauty and charm cross sections are of interest. The beauty cross section is measured at LHCb to be $\sigma_{pp\to b\bar{b}X} = (284 \pm 4 \pm 48)\,\mu b^{-1}$, Ref. [20]. At $sqrts = 7\,\text{TeV}$, the charm hadrons cross section (see Refs. [15], [21] ) is about a factor 20 higher compared to the one of the beauty hadrons, shown in Figs. 2.3. This analysis uses so called secondary charm production, where the charms hadrons are originating from the $b$-quark decays, $b \to W^+ c$. Beauty quarks mainly hadronize in B mesons that have large lifetimes. In the boosted system at LHCb they are easy to identify and provide a powerful signal identification. This makes this analysis, despite the lower production cross section, competitive with the analysis which exploits charm mesons produced promptly in the fragmentation, Ref. [22].

## 2.2   LHC machine

The Large Hadron Collider (LHC) is a part of the accelerator complex at the European Organization for Nuclear Research laboratory (CERN) established in 1964 and build next to Geneva, as shown in Fig. 2.6.

The LHC is a superconducting accelerator that is able to control and collide different hadron beams: proton-proton (p-p), lead-proton (Pb-p and p-Pb) or lead-lead (Pb-Pb). It consists of superconducting magnets in a ring with a circumference of 27km with two vacuum pipes, where the beams can circulate.

The beam bunches are collided at four nominal interaction points, where the four main experiments are set up, which are LHCb, ATLAS, CMS and ALICE. Before the injection into the LHC ring, the protons have to pass several accelerator systems before they reach the injection energies of some hundreds of GeV. First of all they are accelerated at the Linear Accelerator (LINAC 2) to energies of 50 MeV and are injected into the Proton Synchrotron Booster (BOOSTER or PBS). There, the beams reach energies of 1.5 GeV and can be further accelerated in the Proton Synchrotron (PS) to 25 GeV. In the Super Proton Synchrotron (SPS) they reach the required energy of 450 GeV for the injection to the LHC ring. Finally, using the radiofrequency (RF) cavities in the main LHC ring the beams reach the required energy and together with the magnets system (different dipol-, quadro-, sextu- and octupol magnets) sort the protons into packets, so-called bunches. The current design of LHC allows to run at stable conditions and high luminosities at energies of 4 TeV per proton beam, this will be increased to 6.5 TeV in 2015. The operational parameters of LHC are listed in Tab. 2.1, more details in Ref. [24].

The event rate for any kind of analysis is dependent on two parameters, which are the cross section and the luminosity,

$$\frac{dN}{dt} = \mathscr{L} \cdot \sigma. \tag{2.1}$$

The cross section depends on the physical process we are interested in and the luminosity is one of the main figures of merit for a collider. For two colliding proton bunches this can be described through the following equation.

$$\mathscr{L} = \frac{f_{rev} \, \gamma \, N^2 \, k_b}{4\pi\epsilon_n \, \beta^*} F \tag{2.2}$$

Thereby, the bunches revolve with a frequency $f_{rev}$, $k_b$ is the number of occupied bunches circulating in LHC, N is the number of protons contained in each bunch and $\gamma$ is a relativistic correction factor. The size of the beam is characterized by two quantities, the amplitude function $\beta^*$ and transverse emittance $\epsilon_n$. Additionally, due to different beam crossing angles, a geometric factor $F$ between the beams needs to be considered. The achieved performance parameters of LHC are summarized in Tab. 2.1.

**Table 2.1:** *LHC operating parameters, Ref. [25]*

|  | Unit | 2011 | 2012 | Design |
|---|---|---|---|---|
| Energy per nucleon | [TeV ] | 3.5 | 4 | 7 |
| $\beta^*$(beta function) | [m ] | low(0.6)-high(3.0) | | 0.55 |
| Bunch spacing | [ns ] | 50 | 50 | 25 |
| Number of bunches |  | 1374 | 1374 | 2808 |
| Number of protons in a bunch |  | $1.6 - 1.7 \cdot 10^{11}$ | | $1.15 \cdot 10^{11}$ |
| Peak luminosity $\mathscr{L}$ | [cm$^{-2}$s$^{-1}$] | $7.7 \cdot 10^{33}$ | $7.7 \cdot 10^{33}$ | $1 \cdot 10^{34}$ |

LHCb can not operate at the full luminosity regime like ATLAS and CMS, since the particle flux and detector occupancies would be too high. Therefore, LHCb is operated at a higher beta function( $\beta^* = 3.0$), resulting in lower instantaneous luminosities. However, during a run when the number of particles per bunch decreases, the machine can adjust the beam parameters (e.g. $\beta^*$) such that the luminosity stays constant. In 2012, LHCb could operate at an instantaneous luminosity of up to $4 \cdot 10^{32}$ cm$^{-2}$s$^{-1}$, mainly taking the data with lower luminosity but stable conditions.
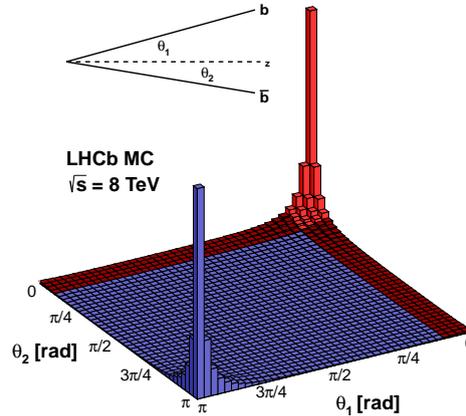
**Figure 2.7:** *Production angle of b-quarks, obtained from simulation (Figure taken from Ref. [26]).*

## 2.3  LHCb experiment

The LHCb experiment is designed to study the physics of beauty and charm quarks. As explained above, most of these heavy quarks are produced in pairs in pp collisions. Fig. 2.7 shows the angular distribution of the b-quarks with respect to the beam axis. The pairs are boosted in the forward and backward direction. The LHCb experiment is designed to cover about 25% of b-quarks produced in the pp collisions at LHC energies, as it is shown in Fig. 2.8. LHCb is a single arm dipole spectrometer with a pseudorapidity coverage of $2 < \eta < 5$.

LHCb has also a possibility of flipping the magnet polarity, which is important for asymmetry measurements and very useful for further checks. The main detector components are the silicon vertex detector surrounding the interaction point and the tracking system placed in layers before and after the magnet. For particle identification two ring imaging cherenkov detectors, placed before and after the magnet, are used. They are followed by the calorimeters and muon stations. The following sections give more details about the subdetector systems.
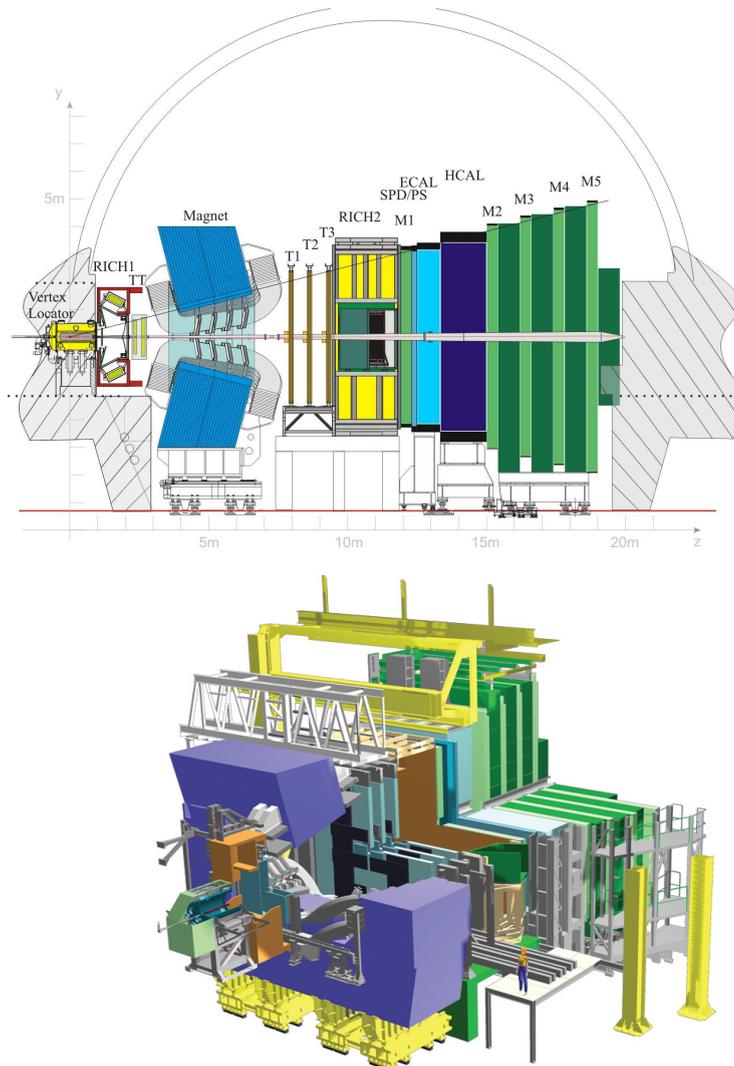
**Figure 2.8:** *Up: Side view of the LHCb detector, from Ref. [27].*
*Bottom: Rotated top-front view of the experiment, from Ref. [26].*
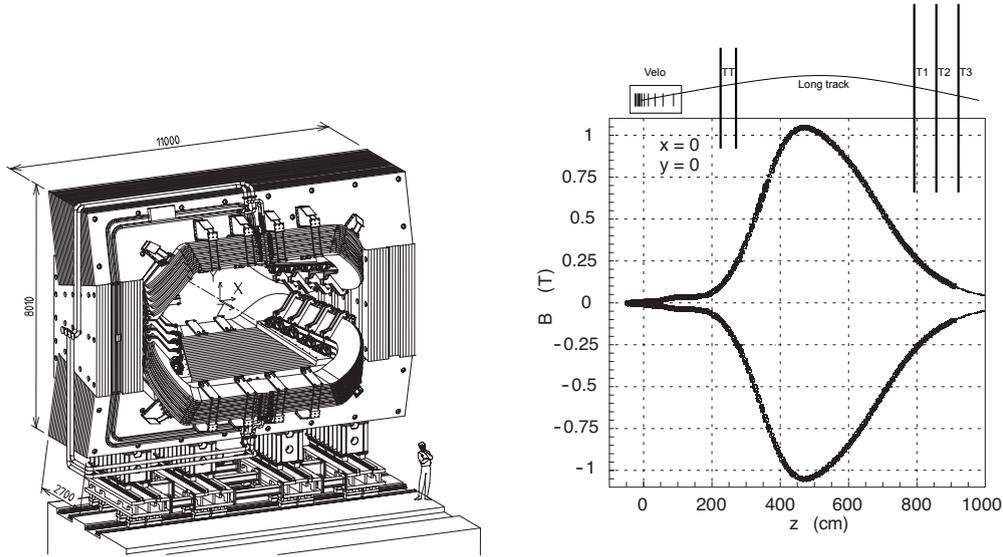
**Figure 2.9:** *Left: LHCb magnet design (taken from Ref. [28]). Right: The main component $B_y$ of magnetic field measured along the z-Axis (from Ref. [28]). Overlapped on top with the main components of the tracking system (Vertex Locator, Tracker Turiencies and Tracking Stations). The field for two different magnet polarities is shown.*

### 2.3.1 Magnet

To perform a measurement of charged particles momenta, a warm [4] dipole magnet is placed in the region between the Tracker Turiencies and the Tracking Stations. The tracking system around the magnet has a long lever arm of about 10m, to provide a good momentum resolution. [5] The heavy material of the magnet is placed outside of the acceptance region. The magnet consists of two coils (with a weight of 54 tons) placed in an iron-yoke window, with a weight of 1500 tons, shown in Fig. 2.9, with an inductance of $L = 1.3$H. It reaches a magnetic field strength above 1T and has an integrated magnetic field of around $\int B\mathrm{d}l = 4$Tm. The magnet fulfills the RICH operation requirements to have low magnetic fields inside the RICH detectors. The magnetic field has the main direction along the $y$-axis deflecting charged particles mainly in $\pm x$-direction and is not homogeneously distributed along the $x$-axis. A precise measurement of the magnetic field was performed using hall sensors. It is shown in Fig. 2.9.

To achieve a good momentum resolution, the magnetic field is measured with a relative precision of about $4 \times 10^{-4}$. The three dimensional field map covers most of the LHCb acceptance and can be reproduced for both polarities, see Ref. [28].

### 2.3.2 Vertex locator

The VErtex LOcator, or short VELO, Ref. [29], is a silicon detector placed closest to the collision point. Its main purpose is to measure tracks precisely and find displaced vertices that can be used directly in the high level trigger. It is a stereo silicon strip detector using a cylindrical geometry (r and $\phi$ strips). With this geometry design of the VELO the track reconstruction is able to be performed fast already on trigger level. The VELO consists of 42 sensors group in (R,$\phi$) sensors that are mounted on modules. The VELO is split into two halves which are A-and C-sides (positive and negative x). The modules on both sides have a slight offset and are placed in 21 parallel layers

---

[4] i.e. operated without superconducting technology

[5] Long tracks, used in this analysis, pass the full tracking system, as shown in Fig. 2.9.
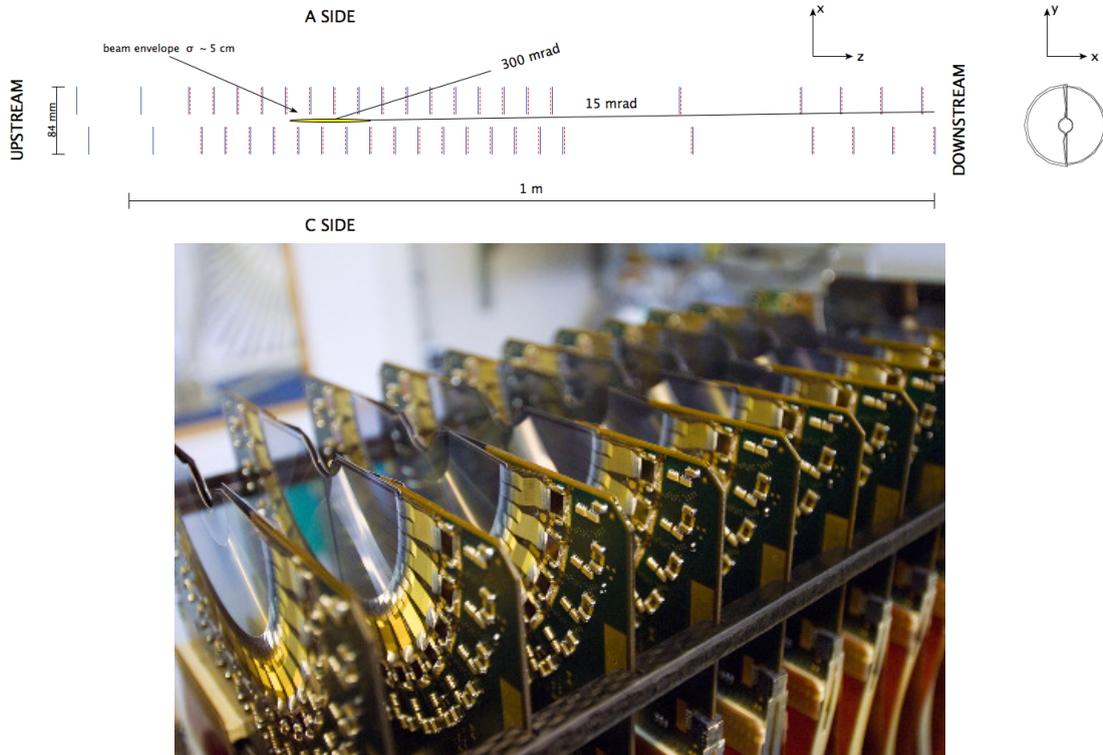
**Figure 2.10:** *(a) Position of the 42 + 4 VELO sensors, shown are the A and C sides.*
*(b) A couple of sensors already mounted on the modules, before the installation of the device*
*(Figures taken from [28], [29], [30]).*

along the beam axis, as shown in Fig. 2.10. Most of the modules are placed downstream of the nominal interaction point (positive z-position), beside some modules that cover the beam envelope and are placed upstream (negative z-position). The four modules placed at the most upstream position were planed to be used for vetoing pile-up and further studies, but are irrelevant for the datasets used in this analysis. VELO covers an acceptance region of $1.6 < |\eta| < 4.9$ [6] required by the design and the other sub detectors. Additionally, it has a high cluster finding efficiency above 99.9% (excluding bad strips, Ref. [30]) and track finding efficiency typically above 98%, see Ref. [31], [30]. A key feature, especially for this analysis, is that the VELO is able to precisely measure vertex position and therefore determine the decay times with high precision. The VELO is a moveable device. To guaranteer save operations it is moved out of the beam line during the beam injection and ramping phase. During the period of stable beam operation, the sensors are brought with an accuracy of $10\,\mu m$ close to the beam. The active silicon sensors are up to $8.2\,mm$ far away from the beam in this position.

The sensors are $n^+$-on n strips on oxygenated silicon [7], the n-type bulk has a backplane of $p^+$ type. The silicon is $300\,\mu m$ thin with a pitch size varying between 40 and 100 $\mu m$ (see Fig. 2.11) minimizing the occupancies. The average occupancies obtained on 2011 data (pp collisions with average number of visible interaction per bunch crossing $\mu = 1.7$) is about 1.1% for both $r$ and $\phi$ sensors. Both sensor types have 2048 readout channels each.

The modules provide the electrical readout of the sensors. Before the digitization, the readout

---

[6]For primary vertices produced within $z = \pm 10.6\,cm$.

[7] Oxygen enriched silicon makes the sensors radiation hard (see Ref. [32]).

Two $p^+$-on n sensors have been used, mainly for testing, in the upstream region

**(a)** *Sensors*                    **(b)** *Module*
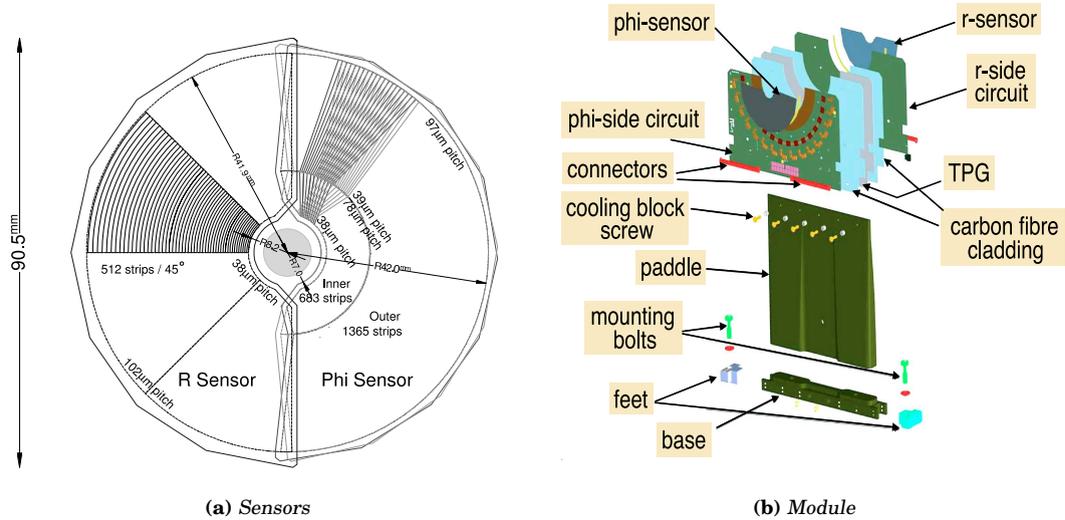
**Figure 2.11:** *(a) Geometry of the r and φ VELO sensors ( taken from Ref. [28]). Only parts of the strips are shown.*
*(b) The composition of a module.*

of each sensor is done by 16 Beetle Chips (Ref. [33]) with an analog front end. Each module holds one r and one $\phi$ sensor, see Fig. 2.11. The core of the module consist of 400 μm thermal pyrolytic graphite (TPG), that is able to transport the heat from the modules to the cooling blocks. The modules are further used by the cooling management system (details in Ref. [34]). The full electronic chain of the modules can be found in Ref. [28].

The modules are shielded by a radio-frequency foil (RF foil), made of AlMg3, from the beam vacuum and from induction through the beam currents. The 300 μm-foil, coated with insulator, fulfills this protection and keeps additional multiple scattering effects from the foil at a low level.

**(a)** *OT-Stations*        **(b)** *OT-Module*
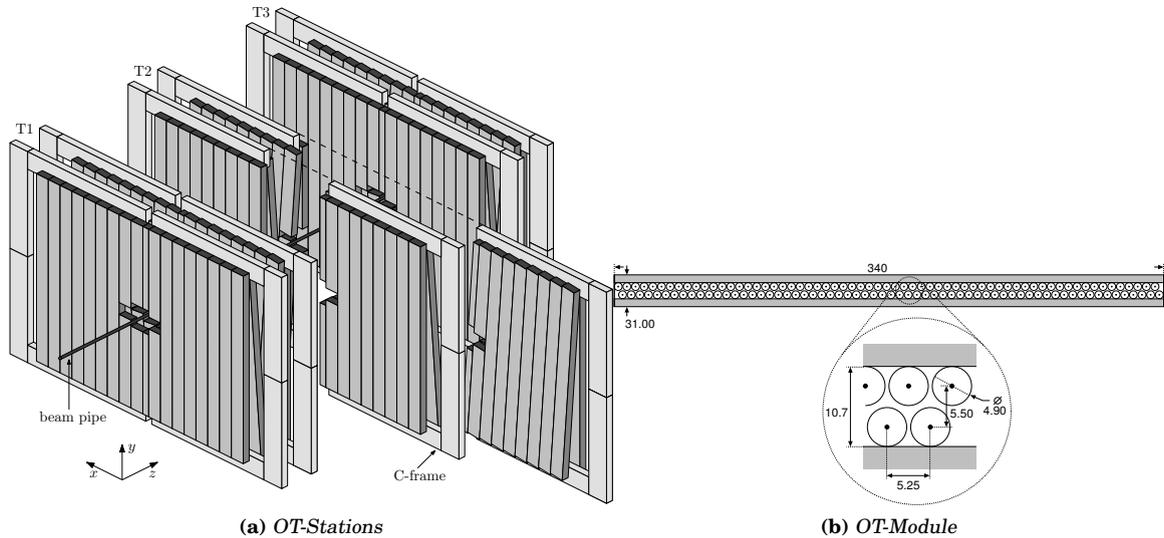
**Figure 2.12:** *(a) The three gaseous Outer Tracker stations, composed of four double layers each (b) Outer Tracker module design with a double layer, the values listed in* cm *(taken from Ref. [35]).*

### 2.3.3 Outer Tracker

The Outer Tracker (OT) is a gaseous ionization detector built from straw tubes that are operated as proportional counters. The design of OT is shown in Fig. 2.12a.

It consists of three stations ($T_1$, $T_2$ and $T_3$) with overall 24 single layers of tubes and 53,760 readout channels. Each station is split into four double layers. The two innermost double layers are tilted with respect to the vertically oriented outer double layers by $\pm 5°$ to provide a real stereo measurement and avoid ambiguities. Each double layer consists of several straw-tube modules, shown in Fig. 2.12b. Each module has 64 drift tubes that are ordered in two monolayers. The drift tubes have an inner diameter of 4.9 mm and are about 2.4 m long. The inner tungsten wire anode of each tube is coated with gold and has a diameter of 25 μm is operated at $+1550V$ with respect to the electrically conducting plated cathode tube. Additionally, the straw tubes are filled with a gas mixture of $Ar/CO_2/O_2$ (70/28.5/1.5%) to keep the drift times below 50ns for reasonable operation at the LHC bunch crossing rate. Measuring the drift times of this gas mixture a spatial resolution of below 200 μm can be obtained. Together with the magnetic field it keeps the momentum resolution at a precise level of $\delta p/p = 0.5\%$ for particles with momentum below 20 GeV/$c$ (low momentum dominated by multiple scattering, Ref. [28]) and reaches $\delta p/p = 0.8\%$ for particles at 100 GeV/$c$, which has a direct influence on every mass resolution measurement at LHCb. The hit efficiency for tracks passing close to the center of a tube is above 99%. Typical occupancies are at the order of 10% and increase with higher bunch crossing, so that the drift times can overlap between the bunch crossings for operations above the nominal bunch crossing of the design (as described in Ref. [35]). Finally, after the readout of the modules the information is digitized and after the $L$0 trigger decision it can be further processed at 1 MHz by the front end electronics (detail can be found in Ref. [35]).

### 2.3.4 Silicon Tracker

There are two silicon micro-strip subdetectors at LHCb. These are the Tracking Turiencies (TT) placed upstream of the magnet and the Inner Tracker (IT) placed downstream of the magnet.

**(a)** *TT-Stations* **(b)** *Module*
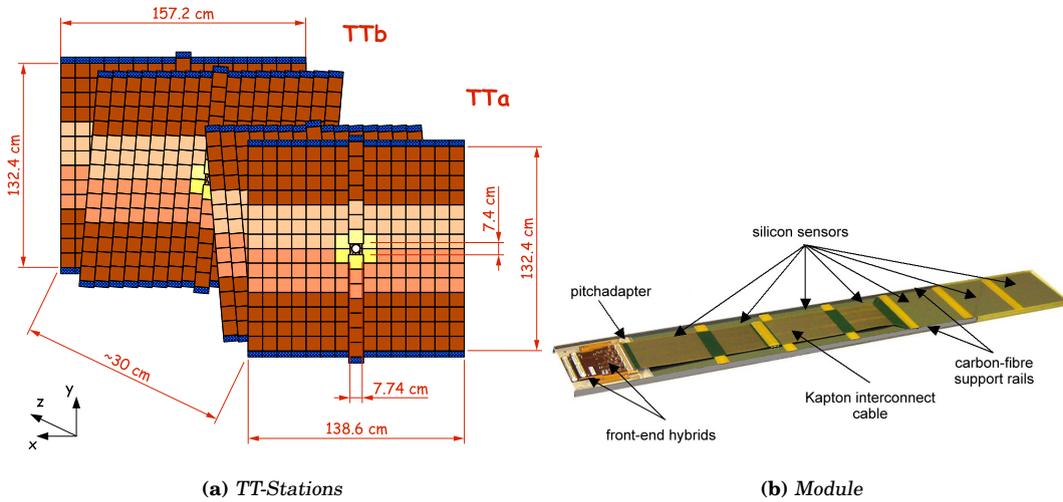
**Figure 2.13:** *(a) Tracker Turiencies with the 4 stereo layers (taken from Ref. [36])*
*(b) a half module of one of the TT-stations, with the sensors and frond end hybrid electronics*
*(taken from Ref. [36]).*
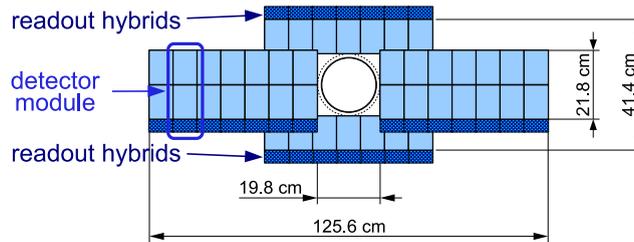


**Figure 2.14:** *One Inner Tracker station and its composition of the modules and the readout hybrids (taken*
*from Ref. [28]).*

The IT covers the high pseudorapidity region in the inner region of the Tracking Stations. The TT improves the momentum resolution. The aim of the IT is to provide precise momentum measurements in the high rapidity region of LHCb.

**Tracking Turiencies(TT)**

The TT is a micro-stip silicon detector that consists of four silicon sensor layers which are grouped in two stations (TTa and TTb, see Fig. 2.13). The stations are about 30 cm separated along the beam axis. The strips of the first and last layer are aligned vertically in y-direction. The layers in the middle are rotated with a stereo-angle of $+5°$ and $-5°$ to avoid ambiguities and provide the necessary resolution in y-direction. Each layer is built out of rows of half modules. Each module has 7 silicon sensors and the readout is done with hybrid front end electronics, see Fig. 2.13. The core part of the readout is the Beetle Chip, Ref. [33]. that reads out the strips and sends an analog signal to the digitization boards. The sensors are $p^+$-on-n silicon with n-doped Si bulk and $p^+$ strips, with a $n^+$ layer on the backside. Unlike the VELO sensors, which are $n^+$-on-n sensors, the $p^+$-on-n sensors can not be operated at voltages below the full charge collection, see Ref. [37]. However, due to their position the expected dose for this sensors is some orders lower.

The bulk is a $500\,\mu$m thick silicon with a pitch size of $183\,\mu$m. Overall, there are 143,360 readout strips with a hit efficiency above 99% (Ref. [38]). Additionally, the number of readout channels is reduced through simultaneous readout, considering also the occupancies, that stays in order of a few percent.

**Inner Tracker(IT)**

The IT is a micro strip silicon detector, consisting of three stations. They cover the high pseudorapidity region and are mounted in the three tracking stations ($T_1$,$T_2$,$T_2$). Together with the OT-stations they buildup the T-stations. Each station, as shown in Fig. 2.14, is built of four boxes placed around the beam pipe. Every box contains 7 modules and has four layers with two of them mounted with a stereo angle $\pm 5°$. The top and bottom boxes have sensors with a bulk of $320\,\mu$m thickness. The left and the right boxes, with two-sensor modules, have a thickness of $410\,\mu$m. The sensors are the same as in the TT built out of $p^+$-on-n silicon with a strip pitch size of $198\,\mu$m. The readout electronics is part of the module and, similarly to the TT, it is integrated into the front-end hybrid readout. Overall, the IT contains 129,024 strips, with a hit efficiency of above 99% and the occupancies are in order of few percent (for details, see Ref. [38]).

### 2.3.5 Particle identification

In the following, the main detector components used for Particle IDentification (PID) are introduced. This are the Cherenkov detectors combined with the calorimetry and muon system. Based on the output of the PID system a combined probability variable is provided and different particle hypothesis can be tested. In such a way, combined Log-Likelihood ($\log \mathcal{L}$) is calculated on a track basis. E.g. to separate kaons and pions the difference between the Likelihoods is considered $\Delta \log \mathcal{L}_{K\pi} = \log \mathcal{L}(K) - \log \mathcal{L}(\pi)$. This quantity is important for many analysis at LHCb with the kaons and pions in the final state, that are otherwise dominated by combinatorial background.

**Cherenkov detectors**

LHCb makes use of two Ring Imaging CHerenkov detectors (RICH). A charged particle crossing a medium with a velocity, $\beta c$, faster than the local phase velocity of light in a medium, with refractive index $n > 1$, emits Cherenkov light. The photons are emitted in the medium under Cherenkov angle, $\Theta_C$, with respect to the particle momentum direction

$$\Theta_C = \frac{1}{n\beta}.$$

Therefore, the RICH provides the information about the velocity of a particle. Knowing the measured momentum and the trajectory of a particle one can test different mass hypotheses, if light has been emitted under the corresponding Cherenkov angle. LHCb has two RICH detectors. The first one, RICH1, placed closed to the VELO, upstream of the magnet, covers an angular acceptance from 25mrad to 300mrad. It is responsible for particle identification and covers the low momentum particles in a range $1 - 60\,$GeV/$c$. Two media are used to achieve this: aerogel and $C_4F_{10}$, which produce different Cherenkov angles. The full design of RICH1 is shown in Fig. 2.15a.

An optical system, split in two symmetrical parts, containing a spherical mirror array and a carbon fibre mirror, is used to focus the Cherenkov photons. This photons are detected with pixel-Hybrid Photon Detectors (HPDs). These are vacuum photo detector tubes. The photons are converted on the surface of the photo cathode into electrons and, finally, after acceleration the electrons are detected by a silicon pixel sensor. RICH1 contains 196 of those tubes, of which the final sensors are segmented into 1024 pixels each.
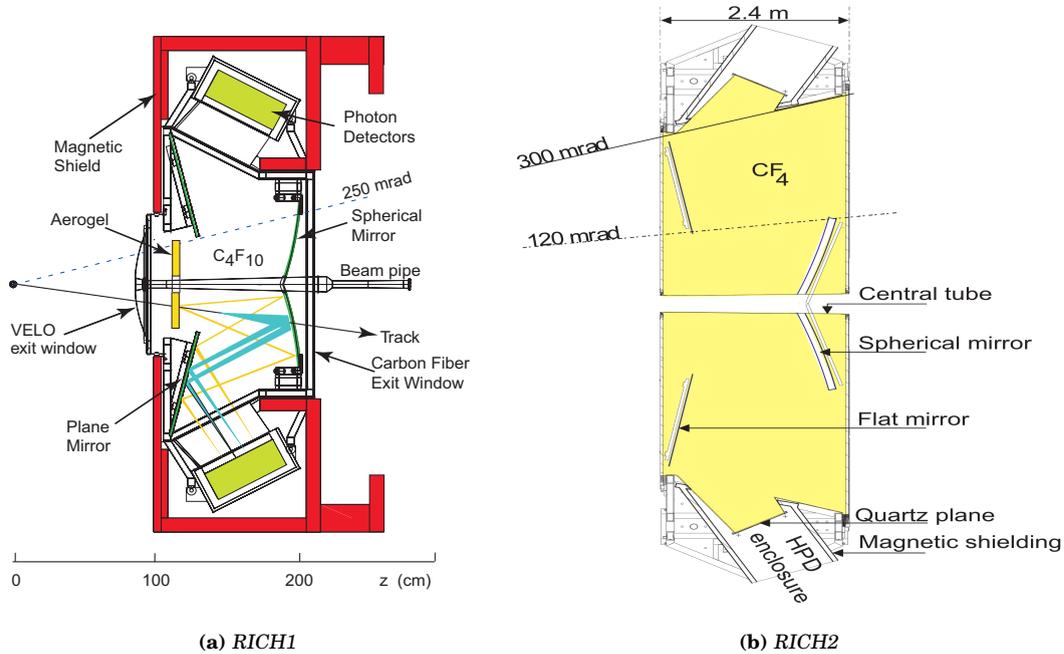
**(a)** *RICH1*

**(b)** *RICH2*

**Figure 2.15:** *(a) RICH1 and (b) RICH2 showing and design and the arrangement of the spherical mirrors that focus the Cherencov light on the Photon Detectors (HPDs). (both figures taken from Ref. [28])*

RICH2 has a very similar design, shown in Fig. 2.15b. The detector is placed after the OT. It contains 288 HPDs and covers a horizontal angel acceptance from 15 mrad to 120 mrad. The gas medium used is $CF_4$. It covers the high momentum range of $15 - 100\,\text{GeV}/c$.

**Calorimeters**

The main propose of the LHCb calorimetry system is to provide already on hardware trigger level the information about identification of hadrons, electrons and photons with large focus on the energy. Additionally, the system aims to provide energy position and particle identification information for the offline analysis, Ref. [39]. The calorimetry system consists of four main components. These are the Scintilating Pad Detectors (SPDs) followed by a wall of lead and a PreShower detector (PS), the electromagnetic and the hadronic calorimeter, see Fig. 2.16.

The SPD allows to separate charged and uncharged particles which, after scattering in the electromagnetic shower, can be detected in the PS. The unit of the PSD/PS is therefore a core component for identification and background rejection for electrons, hadrons and photons. In the ECAL, the deposited energy of electrons and photons can be measured. Finally, the deposited energy of hadrons is measured in the HCAL. All four subsystems use scintillating light that is produced and transmitted to the Photomultipliers (PMTs) through wave shifting fibers. After the readout in the Front End electronics the signal is send via optical links directly to the trigger and Data AcQuisition (DAQ) system. Fig. 2.17 shows one quadrant in the $x - y$ plane for the SPD, PS and ECAL. Each of them has 6016 readout channels and a very similar cell sizes, scaling their size down going to higher pseudorapidity region. The SPD/PS detectors have a small thickness of 0.1 nuclear interaction length ($\lambda_{\text{int}}$) and 2.0 electromagnetic interaction lengths($X_0$). The ECAL is built in a scintillator-lead layer structure with a thickness of 1.1 $\lambda_{\text{int}}$ and $25 X_0$ to collect the full
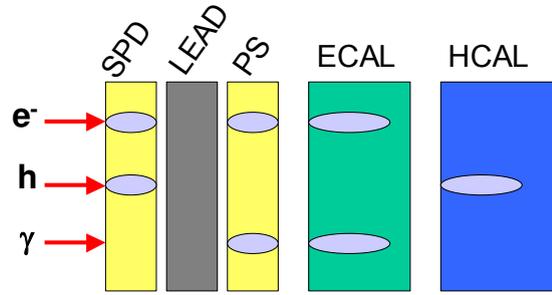
**Figure 2.16:** *Calorimetry system. For visualisation of detector functions a possible amount of scintillating light is roughly visualized by ellipses (taken from Ref. [40]).*



| | |
|---|---|
| **(a)** *SPD/PS/ECAL* | **(b)** *HCAL* |
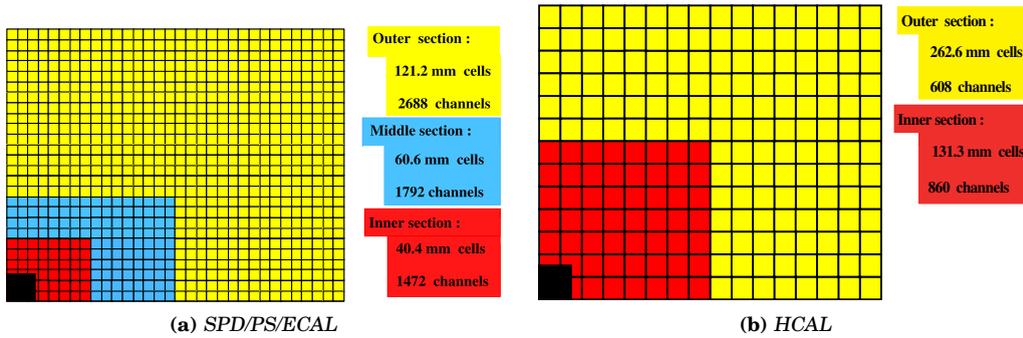
**Figure 2.17:** *(a) Front view of one quadrant of (a) SPD/PS and ECAL and (b) HCAL (both figures taken from Ref. [28]).*

electromagnetic shower. Its energy resolution is, as summarized in Ref. [28],

$$\delta E/E = 10\%/\sqrt{E/\mathrm{GeV}} \oplus 1\%.$$

The HCAL has larger cells with 1488 readout channels and has alternating scintillator-iron layers with a thickness of 5.6 $\lambda_{\mathrm{int}}$. In Ref. [39] its energy resolution was obtained to be

$$\delta E/E = 69\%/\sqrt{E/\mathrm{GeV}} \oplus 9\%.$$

This and further performance details of the calorimeter system are given in Refs. [39] and [28].

**Muon System**

The Muon system is designed to efficiently detect the muons in the LHCb acceptance. It is a key ingredient in many analyses at LHCb that investigate decays with a muon in the final state, especially this present analysis. Therefore, the muon system provides the trigger information and identification for tracks in the offline analysis. The Muon system consists of five stations (M1-M5), see Fig. 2.18, with an inner and outer acceptance in the bending (and non-bending) plane of $20 - 306$mrad ($16 - 258$mrad). Each station can be subdivided into C-side (negative-x) and A-side (positive-x) and further divided into four regions (R1-R4). To account for different occupancies, the cell size decreases, going to higher pseudorapidities. All the station are placed downstream of the magnet. The first one (M1), before the calorimeter, requires to be able to cope with high radiation
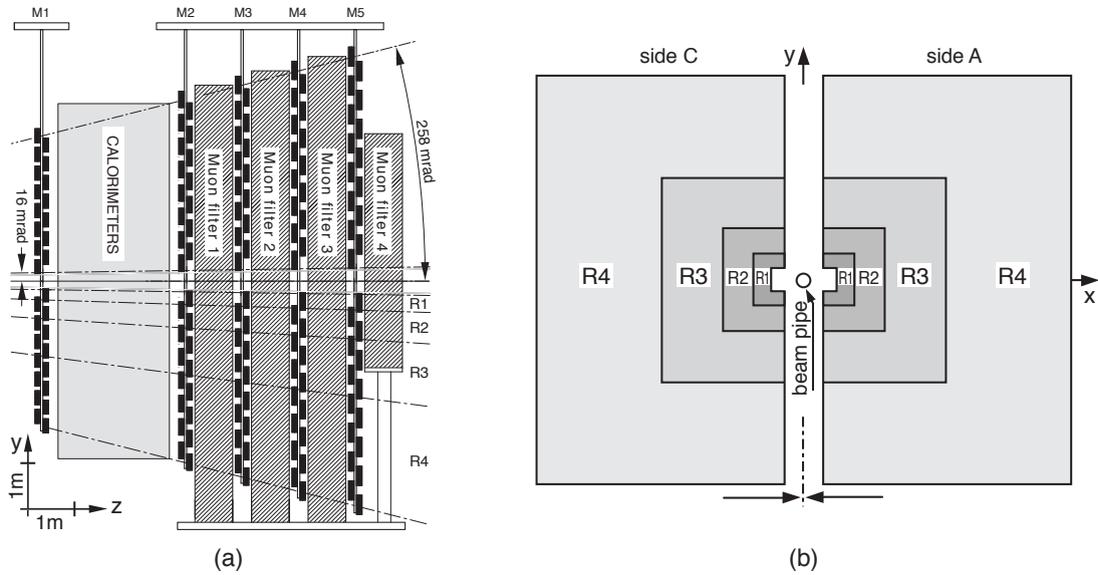
**Figure 2.18:** *Muon system: shown are the five stations with a side view (a) and a front view (b). Each station is subdivided into A and C side and four regions. (figures taken from Ref. [41]).*

dose and occupancies. M1 is used to improve the momentum measurement in the trigger. The calorimeters with a thickness of approximately 7 nuclear interaction lengths serve as shielding, letting only muons with momenta above some GeV/$c$ through to the other four stations.

All the stations are using Multi-Wire Proportional Chambers (MWPC) filled with a fast gas mixture $Ar/CO_2/CF_4 (40:50:5)$. Except for M1 the innermost R1 region of M1, where radiation hard triple Gas Electron Multiplier detectors (triple GEM), with slightly different fast gas mixture $Ar/CO_2/CF_4 (40:15:40)$ (see Ref. [41]), is used. The full muon system consists of 1368 MWPCs and 12 GEM chambers.

After the readout in the Front End electronics the signal is send via optical links to the trigger and Data AcQuisition (DAQ) system. Using all the five stations, muon detection efficiencies above 99% are reached, Ref. [41]. Together with the calorimeters the muon system builds the core part of the hardware trigger.

## 2.3.6 Trigger description

LHCb Trigger System is designed to separate pp collisions with beauty and charm production. The requirement is to analyze events at 40 MHz bunch crossing and to reduce the frequency to 5 kHz output rate, which is required to be able to record from minimum bias events and to reduce the rate to 5 kHz output rate and use them later in the offline analyses This is done in several steps.

**Level-0 Trigger (L0)**
The L0 is a low level trigger and is implemented in hardware electronics synchronized with the LHC bunch crossing. The main goal is to reduce the rate down to 1 MHz. Beside the filtering of interesting physics data, it also provides events for calibration, luminosity measurements and other purposes. The L0 aims to find high $E_T$ hadrons, electrons and photons in the calorimeter and high $p_T$ muons in the muon chambers. The thresholds for transverse momentum and energy are chosen at a few GeV range (details see Ref. [42]). Global variables, like total transverse energy

(a) *Muon track finding*



(b) *Muon towers*
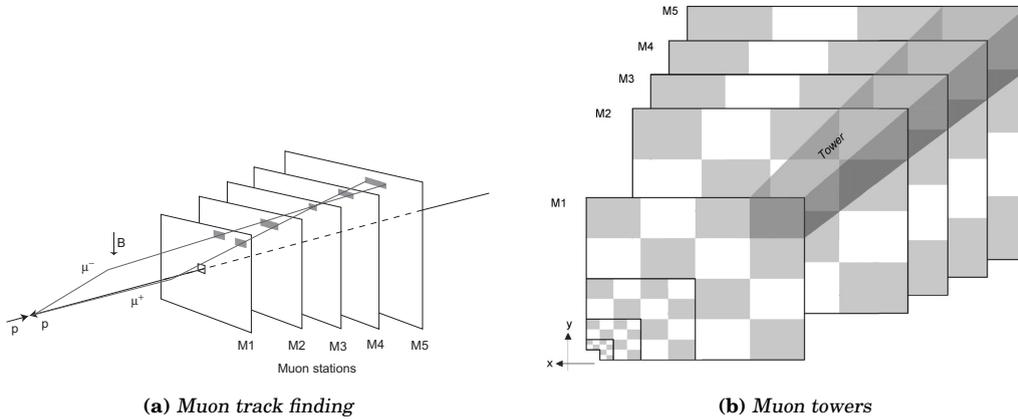
**Figure 2.19:** *(a) Illustration of L0 track finding in the Muon stations, starting with M3 as a seed possible candidates in M2,M4 and M5 are considered, and finally using the hit in M1 the momentum is estimated. (b) Region is split in symmetrical towers pointing to the collision point. The same logical operators are used in each tower. Figures taken from Ref. [43].*

deposit and total number of SPD hits are also used to select interesting events.

The relevant L0 trigger used in this analysis is the L0 trigger for muon finding (L0Muon, Ref. [43]). It performs a track search, using the active muon pads by combining them logically. Fig. 2.19a shows in gray the active search regions used in muon track finding, details can be found in Ref. [43]. The core idea is to start with a seed in M3 and obtain a certain region under straight track assumption for hits in M2, M4 and M5 stations. In combination with the hits from the first station M1 one can get a first estimate on the muon momentum. To perform this fast on hardware level, the muon stations are subdivided in $48 \times 4 = 198$ towers, that point to the interaction region as shown in Fig. 2.19b. Using the information of all five muon station with certain search windows [8] high momentum muons can be efficiently reconstructed. Typically, the transverse momentum threshold for a muon is $1 - 2\,\text{GeV}/c$.

**High Level Trigger (HLT)**

After passing the L0, the data can be further processed by the HLT system that operates on a computer farm asynchronously to bunch crossing. This is further subdivided in two steps, Level-1 (HLT1) and Level-2 (HLT2). Both steps are realized in the software. HLT1 reduces the rate to 80 kHz. It partially reconstructs the event starting with seeds in the VELO and searching for corresponding patterns in the T-stations. [9] Furthermore, track impact parameters (IPs) are calculated and a track quality is used to confirm or reject events. HLT1 makes its decision on a track basis. The analysis presented in this thesis uses a commonly used HLT1 trigger for hadrons (TrackAllL0). The tracks are required to be separated from the primary vertex, to have a good track quality and a high transverse momentum. Additionally, the HLT1-trigger for muons (TrackMuon) is used. Thereby, hits in the muon stations are used for tracking and particle identification. To suppress the background, a displaced muon with high momentum is required.

In the last HLT2 trigger step, the trigger rate is reduced to 5 kHz. Here, a full event reconstruction with very precise track information is performed. Only tracks with a transverse momentum above $0.3\,\text{GeV}/c$ are considered. At this stage, a more complex selection can be considered. This analysis uses so-called topological triggers which exploit the full decay topology. It is optimized to

---

[8] Details on muon track finding can be found in Ref. [43].
[9] See Ref. [44] for details.

identify B-meson decays with several daughter traces ($\pi$, $K$, $\mu$) in the final state. The typical event size recorded in $pp$ collision ($2011 - 2012$) is 35 kbytes and can be written on permanent storage at this rate. This and further details about the LHCb trigger system can be found in references [42] and [28].

### 2.3.7 Data processing

After being accepted by the trigger system, the data is recorded in streamed files (so called RAW files). Thereby, the data is managed by the LHCbDirac framework [45].

**Reconstruction**

The recorded data is sent offline to different TIER [10] computing centers. There it is replicated, reconstructed and after successful verification, it is stored. Furthermore, the raw and reconstructed data can be used to visualize the performance for every subdetector system to report and fix possible problems. After passing the data quality requirements, the data can be flagged for further usage.

**Core software frameworks**

Following LHCb core software frameworks are used in this analysis and are described briefly together with their tasks. First of all, Brunel [46] is a framework that is able to do a full event reconstruction. Thereby the main tasks are track finding and also adding particle identification information using the RICHs, calorimetry and the muon systems. As input, raw data but also digitized simulation samples are used. Its output is further utilized in the analysis framework DaVinci [47]. DaVinci is one of the commonly used frameworks at LHCb. This framework builds the main physics analysis software in LHCb. Thereby, different particles hypothesis are applied and further reconstruction and selection of the signal decays are performed. Gauss [48] is responsible for event generation and detector simulation. It provides the simulated particles, vertices and hit information for further usage. Finally, the digitization for each detector component is performed with the Boole [49] project. The output can be finally used by Brunel and DaVinci frameworks.

**Stripping**

Most of the physics analyses require a very particular selection of reconstructed events. Since the application of these selections is very CPU-time and storage consuming, the selection and reconstruction step is centralized. Events are grouped in stripping selections, that are preselections for different analyses. The Stripping is done with the previously explained DaVinci framework [47].

The stripping step can be repeated, using e.g. upgraded reconstruction versions or improved alignment. The data files are compressed, reducing them to the relevant objects for the analysis, e.g. MDST files, Ref. [45]. On this preselected samples, that are also replicated and distributed at different TIERs, the users can perform their analyses fast.

The full datasets available for analysis and recorded at LHCb in pp collisions correspond to an integrated luminosity of $1.1\,\mathrm{fb}^{-1}$ (in 2011) and $2.1\,\mathrm{fb}^{-1}$ (in 2012). The recorded and delivered luminosity at LHCb for pp collisions during the first run period are shown in Fig. 2.20 and show that LHCb is very efficiently taking data.

---

[10]TIER: Telecomunications Infrastructure Standard for Data Center
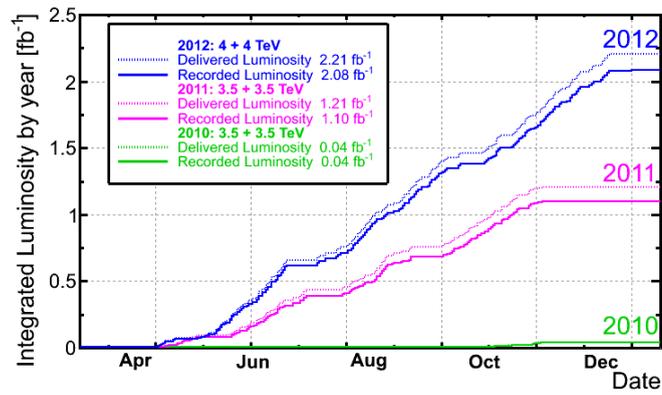
**Figure 2.20:** *Delivered and recorded integrated luminosities for pp collision at LHCb in the years 2010-2012 (taken form Ref. [26]).*

# 3

# From theory to the measurement

This chapter introduces the quantities needed for direct and indirect *CP* violation measurements in the charm system. It introduces the theoretical predictions of the observables to be measured in this analysis and gives the connection to commonly used observables from other measurements.

## 3.1 Neutral-meson mixing and *CP* violation

Neutral mesons are produced as flavour eigenstates which are different from mass eigenstates. In the Standard Model (SM) quark flavour transitions are described by the Cabibbo-Kobayashi-Maskawa (CKM) matrix. This matrix introduces a weak phase in $D^0$-decay amplitudes. The interference of several amplitudes with different weak and strong phases introduces direct *CP* violation. The interference due to mixing and decay diagrams introduces indirect *CP* violation.

Using a naive picture to estimate the effect of *CP* violation in the SM, first of all, short-distance penguins with $b$-quarks in the loop as well as simple box diagrams (Fig. 3.1) can be calculated. However,this can explain only a very small amount of *CP* violation [50]. For the full





(a)  (b)
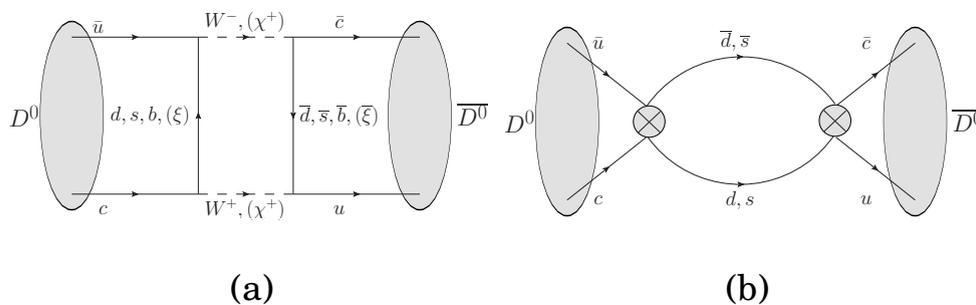
**Figure 3.1:** *(a) One possible contributions to CP violation from potential new physics in the $D^0$ mixing box diagram. In the GIM mechanism, the contributions of the three down-type quarks mainly cancel. However, new physics (here e.g. from a weak-isosinglet quark (ξ) or a SUSY scalar χ) would have direct impact on indirect CP violation.*
*(b) One possible long range process, with a non-perturbative QCD contribution.*
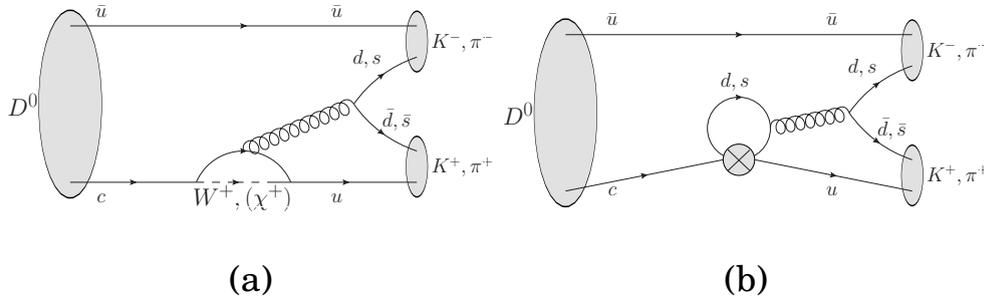
**Figure 3.2:** *Higher order Feynman graphs for $D^0 \to hh$ decays are shown: (a) a penguin diagram(with possible new physics contribution $\chi$) and (b) a penguin contraction for non-perturbative QCD effects. The interference of the tree level diagram with these penguins is the soures of direct CP violation in $D^0 \to hh$ decays.*

calculation one has to consider also non-perturbative contributions. This is typically done within the operator product expansion (OPE) framework [51]. Naive calculations in the SM result in *CP* asymmetries below $\mathcal{O}(10^{-3})$ [50]. However, they have large hadronic uncertainties. Depending on their treatment one can explain a possible sizeable direct *CP* violation in the charm sector of up to $\mathcal{O}(10^{-2})$. On the one hand, the SM contribution to indirect *CP* violation remains small due to the small mixing in the $D^0$ system. Direct *CP* violation measurements are sensitive to non-perturbative QCD effects. An example process with non-perturbative effects is shown for SM gluonic penguin in Fig. 3.2 and for a long distance in Fig. 3.1. On the other hand, one should not forget the possibility that new physics can lead to a significant contribution to *CP* violation (see Fig. 3.1). Such a contribution could result in an enhancement of indirect *CP* violation. From the *B*-system it is known that heavy quark expansion and factorisation work well and could be able to work also in the charm system [51]. Therefore, new physics could potentially indirectly be measured in the charm system.

### 3.1.1 Effective Hamiltonian and mixing formalism

For any essential numerical calculation of the mixing phenomena an effective theory is used, see Ref. [51], [52]. The neutral meson system can be represented by a $(2 \times 2)$ non-hermitian Hamiltonian[1]. This is a typical two state system commonly used in many quantum physics problems. The time evolution of neutral particles can be described by the Schroedinger equation using this effective Hamiltonian $\mathcal{H}$.

The solutions of this two state system can be written in the following way, where the flavour eigenstates can be connected to the mass eigenstates

$$\begin{pmatrix} |D^0(t)> \\ |\overline{D^0}(t)> \end{pmatrix} = Q \begin{pmatrix} e^{-iM_L t - \Gamma_L t/2} & 0 \\ 0 & e^{-iM_H t - \Gamma_H t/2} \end{pmatrix} Q^{-1} \begin{pmatrix} |D^0> \\ |\overline{D^0}> \end{pmatrix} \tag{3.1}$$

with

$$Q = \begin{pmatrix} p & p \\ q & -q \end{pmatrix}. \tag{3.2}$$

The details of the assumptions that were used to determine the solution can be found in [53]. Notice that the parameters p and q define also a phase shift. The mass eigenstates have a clearly defined mass and decay width. Due to the splitting in mass $\Delta m = M_H - M_L$ and decay width

---

[1]In the flavor basis with $CP|D^0> = -|\overline{D^0}>$.

$\Delta\Gamma = \Gamma_H - \Gamma_L$ [2] can be used to define dimensionless parameters,

$$x = \frac{\Delta m}{\Gamma}$$
$$y = \frac{\Delta\Gamma}{2\Gamma}. \tag{3.3}$$

Additionally, decay amplitudes to a final state f can be introduced:[3]

$$\overset{(-)}{A}_f = <f|\mathcal{H}|\overset{(-)}{D^0}> . \tag{3.4}$$

Thus the time-dependent decay rates for $D^0$ mixing can be described in the following way [54], [55]

$$\Gamma(\overset{(-)}{D^0}(t) \to f) = \mathcal{N}_f \frac{1}{2} e^{-\tau} \left|\overset{(-)}{A}_f\right|^2 \{(1 + \left|\lambda_f^{(-1)}\right|^2)\cosh(y\tau) + (1 - \left|\lambda_f^{(-1)}\right|^2)\cos(x\tau)$$
$$+ 2\Re(\lambda_f^{(-1)})\sinh(y\tau) - 2\Im(\lambda_f^{(-1)})\sin(x\tau)\}, \tag{3.5}$$

where $\Gamma = \frac{1}{N}\frac{dN}{dt}$ is the decay rate, with $dN$ number of transitions in the decay-time range from $t$ to $t + dt$, and N the total number of $D^0$s produced at decay time $t = 0$[4]. $\mathcal{N}_f$ is a normalization factor. The unitless parameter $\tau$ is defined as $\tau = \Gamma t$. The variable describing $CP$ violation in here is the imaginary part of

$$\lambda_f = \frac{q}{p}\frac{\bar{A}_f}{A_f}. \tag{3.6}$$

A useful parametrization for $\lambda_f$ is found to be [55]

$$\lambda_f = -\eta_{CP}\left|\frac{q}{p}\right|\left|\frac{\bar{A}_f}{A_f}\right|e^{i\phi} \tag{3.7}$$

$$= -|R_m||R_f|e^{i\phi}.$$

Here, $\eta_{CP} = \pm 1$ is the eigenvalue of the $CP$ eigenstates $f$. Kaons and pions are pseudoscalars as well as the $D^0$ meson($J^P = 0^-$). Thus our final states $K^+K^-$ and $\pi^+\pi^-$ are $CP$ even and have a $CP$ eigenvalue of $\eta_{CP} = 1$.

Several scenarios in which $CP$ violation can appear and to which the observables in this parametrisation are sensitive are discussed in the following. The ratio $R_m = \frac{q}{p}$ is determined by mixing parameters only. If $|R_m| \neq 1$ there is $CP$ violation in mixing, namely the transition probability of a $D^0$ to a $\bar{D}^0$ is not the same as the transition probability of a $\bar{D}^0$ to a $D^0$ ($P(D^0 \to \bar{D}^0) \neq P(\bar{D}^0 \to D^0)$).

In the similar way, the ratio $R_f = \frac{\bar{A}_f}{A_f}$ is characterized only by decay amplitudes into the final state $f$. In case of $|R_f| \neq 1$ there is $CP$ violation in decay (direct $CP$ violation). However, $CP$ violation can be also caused by the relative phase difference $\phi$ between the decay amplitudes ($R_f$) and the mixing parameters ($R_m$) [54]. The following asymmetries are introduced:

$$A_m = (|R_m|^2 - 1) \tag{3.8}$$

for contribution from $CP$ violation in mixing and

$$A_d = (|R_f|^2 - 1) \tag{3.9}$$

from $CP$ violation in decay.

Exploiting our current experimental limits, only $A_d$ can be considered as small, Ref. [56] [57]. The three ways of CP violation and according parameters are depicted in Figure 3.3.

---

[2]The indices $H$, $L$ stand for heavy and light, so that $\Delta m > 0$ by definition. Following this definition $\Delta\Gamma > 0$ in the $D$ system contrary to the B system

[3]In this analysis $f = \bar{f}$ and thus $\bar{A}_f = \bar{A}_{\bar{f}}$.

[4]$\hbar$ is set to 1 in this notation
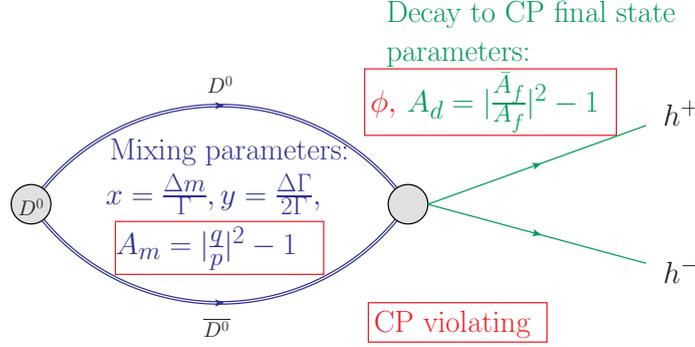
**Figure 3.3:** *Total contribution to CP violation (here in $\lambda_f$ parametrisation, Eq. 3.8) is coming from 1st) mixing, 2nd) decay and 3rd) interference between mixing and decay.*

### 3.1.2   Direct and indirect *CP* violation

The time dependent *CP* asymmetry is defined in the following way

$$A_{CP}(t) = \frac{\Gamma_{D^0 \to f}(t) - \Gamma_{\overline{D^0} \to f}(t)}{\Gamma_{D^0 \to f}(t) + \Gamma_{\overline{D^0} \to f}(t)} \tag{3.10}$$

$$= \frac{\begin{matrix} (|q/p|^2 - 1)[(1 + |\lambda_f|^2)\cosh(y\tau) & +2\,\Re(\lambda_f)\sinh(y\tau) & ] \\ +(|q/p|^2 + 1)[(1 - |\lambda_f|^2)\cos(x\tau) & -2\,\Im(\lambda_f)\sin(x\tau) & ] \end{matrix}}{\begin{matrix} (|q/p|^2 + 1)[(1 + |\lambda_f|^2)\cosh(y\tau) & +2\,\Re(\lambda_f)\sinh(y\tau) & ] \\ +(|q/p|^2 - 1)[(1 - |\lambda_f|^2)\cos(x\tau) & -2\,\Im(\lambda_f)\sin(x\tau) & ] \end{matrix}}.$$

We use a linear approximation of this expression in the following. The correction of the higher order term is beyond our current measurement precision of $\mathcal{O}(10^{-4})$. Due to the small $x$, and $y$ $\lesssim \mathcal{O}(10^{-2})$ and also considering the time window of this analysis $\tau = \Gamma t \lesssim \mathcal{O}(1)$ one can expand $A_{CP}(t)$ using $x\tau$ and $y\tau \lesssim \mathcal{O}(10^{-1})$. This expansion leads us to the following equation

$$A_{CP}(t) = A_{CP}^{dir} + A_{CP}^{indir} \Gamma\, t. \tag{3.11}$$

With the following CP violating asymmetries

$$A_{CP}^{dir} = 1 - \frac{2|\lambda_f^2|}{|\lambda_f^2| + |\frac{q}{p}|^2} = \frac{|\frac{q}{p}|^2 - |\lambda_f^2|}{|\lambda_f^2| + |\frac{q}{p}|^2} \tag{3.12}$$

$$= \frac{1 - |R_f|^2}{1 + |R_f|^2} = \frac{-A_d}{2 + A_d}.$$

$$A_{CP}^{indir} = -2|R_m|^2 \frac{\left[ (1 + |\lambda_f|^2)\, x\, \Im(\lambda_f) + (1 - |\lambda_f|^2)\, y\, \Re(\lambda_f) \right]}{\left( |R_m|^2 + |\lambda_f|^2 \right)^2} \tag{3.13}$$

$$= -2(1 + A_m) \frac{\left[ (1 + (1 + A_m)(1 + A_d))\, x\, \Im(\lambda_f) + (1 - (1 + A_m)(1 + A_d))\, y\, \Re(\lambda_f) \right]}{((1 + A_m) + (1 + A_d)(1 + A_m))^2}.$$

The full calculation can also be found in the Appendix A.8. Assuming $A_m$ and $A_d$ to be small this can be further expanded to a commonly used expression

$$A_{CP}^{indir} \approx x\, \sin(\phi) - \frac{1}{2}(A_m + A_d)\, y\, \cos(\phi). \tag{3.14}$$

At the same time the equations above define the direct and indirect *CP* violation contribution. Eq. 3.12 and A.25 illustrate that direct *CP* violation contains only the parameters of the decay. However, the indirect *CP* violation is also sensitive to the mixing part. $A_m$ does not depend on the final state and is therefore a universal observable. This is contrary to the decay amplitudes $A_f/\overline{A}_f$, which depend on the final state. Accordingly, $A_d$ and the phase difference $\phi$ can be different for both final states studied in this analysis [57].

### 3.1.3 $A_\Gamma$ observable

The difference between effective lifetimes of $D^0$ and $\overline{D^0}$ is parametrized by $A_\Gamma$:

$$A_\Gamma = \frac{\hat{\Gamma}_{D^0} - \hat{\Gamma}_{\overline{D^0}}}{\hat{\Gamma}_{D^0} + \hat{\Gamma}_{\overline{D^0}}}. \tag{3.15}$$

where $\hat{\Gamma}$ is defined as

$$1/\hat{\Gamma} = \frac{\int t\,\Gamma(t)dt}{\int \Gamma(t)dt}. \tag{3.16}$$

After expansion in $x\tau$ and $y\tau$ up to $\mathcal{O}((x\tau)^2) + \mathcal{O}((y\tau)^2) + \mathcal{O}(yx(\tau)^2)$ this simplifies to

$$A_\Gamma = \frac{1}{2}|\lambda_f^{-1}|^2[(1 + |\lambda_f|^2)\,x\,\Im(\lambda_f) + (1 - |\lambda_f|^2)\,y\,\Re(\lambda_f)]. \tag{3.17}$$

Comparing this with Eq. A.25 one obtains

$$A_\Gamma = -A_{CP}^{indir}\frac{(|R_m|^2 + |\lambda_f|^2)^2}{4\,|R_m|^2|\lambda_f|^2} \tag{3.18}$$

$$= -A_{CP}^{indir}\frac{1}{4}(|R_f| + |R_f^{-1}|)^2$$

$$= -A_{CP}^{indir}\frac{1}{4}(\sqrt{1 + A_d} + \frac{1}{\sqrt{1 + A_d}})^2.$$

For small violation $A_d$, the correction factor ($\sqrt{1 + A_d} + \frac{1}{\sqrt{1+A_d}})^2 \approx 4 + A_d^2 + \mathcal{O}(A_d^3)$) can be neglected. Thus to the precision level of $\mathcal{O}(10^{-4})$, [5] $A_\Gamma$ can be directly expressed as:

$$A_\Gamma \approx -A_{CP}^{indir} \approx \frac{1}{2}(A_m + A_d)y\cos\phi - x\sin\phi. \tag{3.19}$$

And using our previous notation it can be written in a more handable form

$$A_\Gamma \approx (A_m/2 - A_{CP}^{dir})y\cos\phi - x\sin\phi. \tag{3.20}$$

### 3.1.4 Definition of indirect *CP* violation

In theoretical papers one can often find different definitions of indirect *CP* violation.

I) All the phases that can not be eliminated in the amplitudes, describing a $\Delta F = 2$ process [58] (we call it $a_{CP}^{ind}$).

II) Indirect CP violation is the time-dependent part of the asymmetry: so $A_{CP}^{indir}$ or $-A_\Gamma$ [57].

III) The terms containing the *CP* violation in mixing + the terms containing interference between decays with and without mixing [59].

---

[5] works even below $\mathcal{O}(10^{-6})$ for $A_d \lesssim \mathcal{O}(10^{-2})$
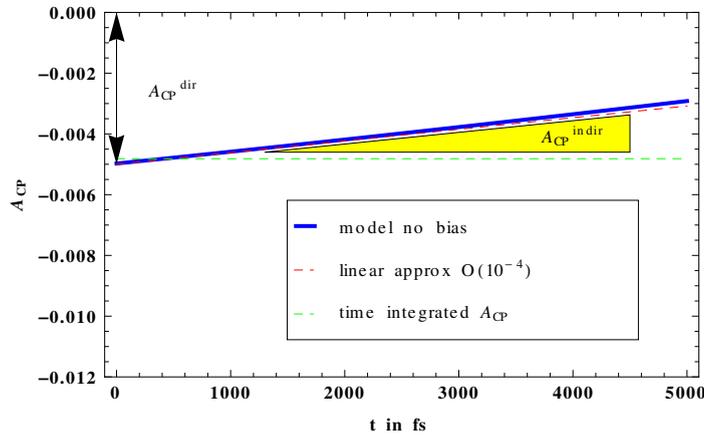
**Figure 3.4:** *$A_{CP}$ time dependence, shown is in blue the full model according to Eq. 3.10. In red the linear approximation Eq. A.25 and in green the time integrated total CP asymmetry $< A_{CP} >$. Also the offset that can be approximated as $A_{CP}^{dir}$ and the slope that is $A_{CP}^{indir} \approx -A_\Gamma$ are shown.*

In the absence of direct *CP* violation all definitions are equivalent. The first definition is the most fundamental one. However, in case of the presence of direct *CP* violation there are some small corrections with additional terms. E.g. the second (II) case contains also the direct *CP* violation $\frac{1}{2} A_d \ y \ \cos\phi$. So, the quantity in case (II) itself can be non zero in absence of indirect *CP* violation. If one uses the first (I) definition one can rewrite:

$$A_{CP}^{indir} = a_{CP}^{ind} - \frac{1}{2} A_d \ y \ \cos\phi \qquad (3.21)$$

What is measured in this analysis is the second (II) quantity. So, according to the first (I) definition we measure $a_{CP}^{ind}$ plus an additional term (it is suppressed by $y \sim 1\%$ and $A_d < 1\%$) containing a tiny part of direct *CP* violation.

## 3.2   Measurement of indirect *CP* violation

Following the guidelines in previous paragraphs indirect *CP* violation can be experimentally measured in two different ways. Either, one can measure the effective lifetimes of the $D^0$ and $\overline{D^0}$ to *CP* eigenstates separately for both channels and measure $A_\Gamma$ as described in Eq. 3.15. Or it can be obtained through directly measuring the time dependence of $A_{CP}(t)$ (Eq. 3.11) at our current experimental sensitivity level. To illustrate the size and correlation of the involved parameters, the following scenarios have been studied.

|  | x | y | $\frac{1}{\Gamma}$ [fs] | $\phi$[rad] | $\eta_{CP}$ | $A_d$ | $A_m$ | $A_\Gamma$ | $A_{CP}^{dir}$ |
|---|---|---|---|---|---|---|---|---|---|
| scenario I | 0.002 | 0.005 | 410 | 0.1 | 1 | 0.01 | 0.007 | -0.00016 | -0.005 |
| scenario II | 0.001 | 0.01 | 410 | 1.5 | 1 | 0.01 | 0.01 | -0.00099 | -0.005 |

The time dependence of $A_{CP}(t)$ for scenario II is shown in Fig. 3.4.

In previous measurements of LHCb [60], Belle [61], Babar [62] additionally the time dependent ratio $r(t)$ of decay rates was studied as cross-check. It is defined as

$$r(t) = \frac{\Gamma_{D^0 \to f}(t)}{\Gamma_{\overline{D^0} \to f}(t)}. \qquad (3.22)$$
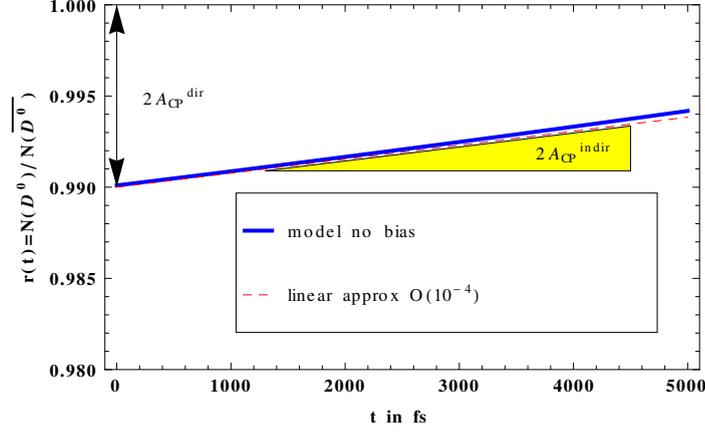
**Figure 3.5:** *Ratio time dependence, similar to Fig. 3.4. The offset and slope are increased by factor two. At the order $O(10^{-4})$, there is no difference in the slope extraction between r(t) and $A_{CP}(t)$.*

Expending terms up to $O(10^{-4})$ order this can be simplified to

$$r(t) \approx 1 + 2A_{CP}^{dir} - 2A_\Gamma \tau = 1 + 2A_{CP}^{dir} - 2A_\Gamma \Gamma t,$$

using the previous notation $\tau = \Gamma t$. The time dependence of the ratio and the quantities, which can be extracted are shown in Fig. 3.5. $A_{CP}(t)$ can also be directly connected to the ratio defined above: $A_{CP}(t) \approx \frac{1}{2}(r(t)-1)$. Using $A_{CP}(t)$ or $r(t)$ should be equivalent as long as you look at $\mathcal{O}(10^{-4})$ precision level (for details see Appendix B.6). In this note $A^{CP}(t)$ will be considered as default method due to statistical advantages. The ratio $r(t)$, which has asymmetric uncertainties, is however also calculated as a cross check.

## 3.3   Raw asymmetry measurement

A unique feature of this analysis is that it exploits $D^0$ and $\bar{D}^0$ mesons produced in B decays, $B \to \overline{D^0}\mu^+ X$ and $\bar{B} \to D^0\mu^-\overline{X}$. Thereby, the muon charge is used as a tag for the initial flavour determination of the charm meson, which introduces additional asymmetries that are discussed below.

The time dependent raw asymmetry $A_{CP}^{RAW}(t)$ is defined as

$$A_{CP}^{RAW}(t) = \frac{\Gamma_{D_{tag}^0}(t) - \Gamma_{\bar{D}_{tag}^0}(t)}{\Gamma_{D_{tag}^0}(t) + \Gamma_{\bar{D}_{tag}^0}(t)} \tag{3.23}$$

$$= \frac{dN(D_{tag}^0, t_{bin}) - dN(\bar{D}_{tag}^0, t_{bin})}{dN(D_{tag}^0, t_{bin}) + dN(\bar{D}_{tag}^0, t_{bin})}, \tag{3.24}$$

where $D_{tag}^0$ and $\bar{D}_{tag}^0$ are $D$ mesons tagged as $D^0$ and $\bar{D}^0$ respectively. Thereby $dN(D^0{}_{tag}, t_{bin})$ is the number of observed events in bin $t_{bin}$ with $\Gamma(t) = \frac{1}{N}\frac{dN(t)}{dt}$. The raw asymmetry includes initial production asymmetry and muon detection asymmetries. The $D^0$ meson detection asymmetry cancels since it is a $CP$ final state. The muon detection asymmetry $A_\mu$ is caused by reconstruction, selection and detection of the muon. Since we are using $D^0$ from $B$ decays, a production asymmetry
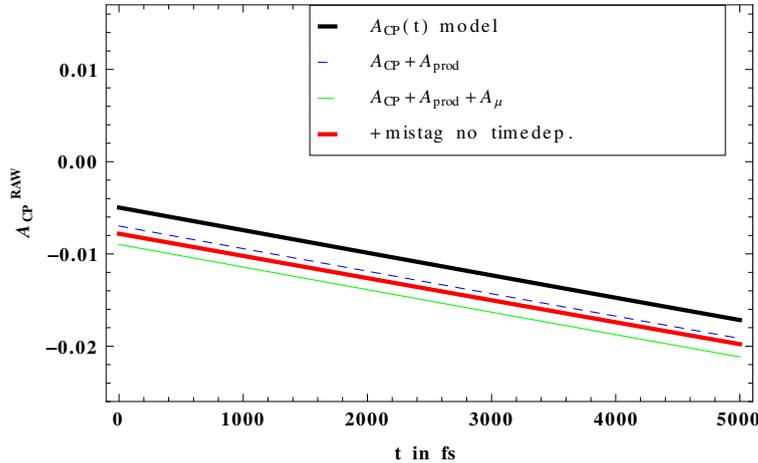
**Figure 3.6:** *Impact of different effects is shown. In black is the $A^{CP}(t)$, blue adding the production asymmetry, green adding the muon asymmetries and red introducing mistag probabilities. All effects have tiny impact on the extracted slope. As input the physics parameters of scenario II have been used. To simulate the mistag $\Delta\omega=0.001$ and $\omega=0.009$ have been chosen.*

of the $B$ meson has a direct impact on measured raw asymmetry. For small asymmetries and neglecting the effect of mistag this can be rewritten as (see Appendix B.10)

$$A_{CP}^{RAW}(t) = A^{CP}(t) + A_\mu + A_{prod} + \mathcal{O}(A^3). \tag{3.25}$$

In general, muon and production asymmetries are independent of $D^0$ decay time. However, due to second-order effects, a small $D^0$ decay time dependence can appear. This effect is very small, compared to current sensitivity level, as will be shown in the systematics Chapter 8 To illustrate effect of a (time-independent) detection and production asymmetry, a model with the expected size of these asymmetries is shown in Fig. 3.6.

### Mistag probability
Additionally, one also has to consider the probability that the muon not always tags the correct $D^0$ flavour. For this the following mistag probabilities are defined:

$$\omega^{\bar{D}^0} = P(\bar{D}^0_{tag}|D^0 \wedge B_{rec})$$
$$\omega^{D^0} = P(D^0_{tag}|\bar{D}^0 \wedge B_{rec}),$$

with the average mistag probability and mistag difference

$$\bar{\omega} = \frac{\omega^{\bar{D}^0} + \omega^{D^0}}{2}$$
$$\Delta\omega = \omega^{\bar{D}^0} - \omega^{D^0}.$$

where $D^0 \wedge B_{rec}$ means a true $D^0$ at production time which has been reconstructed in a semi-muonic $B$ decay and the index $tag$ indicates the tagging decision. Thus $P(\bar{D}^0_{tag}|D^0 \wedge B_{rec})$ describes the probability that a $D^0$ at production time is combined with a muon to a reconstructed $B$ candidate, however tagged as $\bar{D}^0$.

Taking the mis-tag probability into account the relation of physics asymmetry $A^{CP}$ and measured raw asymmetry is modified to be (see Appendix B.10):

$$A^{RAW}_{CP} = (1 - 2\overline{\omega})(A^{CP} + A_{mu} + A_{prod}) - \Delta\omega. \tag{3.26}$$

As can be seen in this equation, $\Delta\omega$ introduces a shift and $\overline{\omega}$ a damping factor for the measured asymmetry. The effect of time independent mistag rates is again visualized in Fig. 3.6. The impact on the measurement of the time dependence of the mistag probabilities and mistag difference is considered in the systematic studies described in Chapter 8.

# 4

# Analysis overview

This chapter presents the general strategy of the analysis and gives the reader an outlook how the measurement is performed. The tools and methods will be introduced in more detail in the upcoming chapters.

## 4.1 Introduction

The aim of the analysis is to measure the indirect CP violation in $D^0 \to K^- K^+$ and $D^0 \to \pi^- \pi^+$ decays with a sensitivity below $\mathcal{O}(10^{-3})$. This work exploits $D^0$ mesons originating from semileptonic b-hadron (B) decays, $B \to D^0 \mu^- \bar{\nu}_\mu X$. Thereby, the charge of the accompanying muon is exploited as a tag for the initial charm-meson flavour. Additionally, the $D^0$-meson decay time is determined using the distance given by the $B$ and $D^0$ decay vertex and the $D^0$ momentum.

In the charm system, the expected CP-asymmetry can be approximated as

$$A_{CP}(t) \approx A_{CP}^{dir} - A_\Gamma \frac{t}{\tau},$$

(4.1)

with $A_\Gamma$, the observable of our interest, see Eq. 3.11 and 3.19. Therefore, the CP asymmetry is measured as a function of decay time and the according slope is extract.

However, in reality the raw asymmetries are measured. These can be simplified, see Eq. 3.25

$$A_{CP}^{RAW}(t) \approx A_{CP}(t) + A_\mu + A_{prod}.$$

(4.2)

Thereby, $A_\mu$ is the muon production asymmetry introduced through the detection and selection of the muons. $A_{prod}$, is the production and selection asymmetry of the $B$-hadron decays. In the first order the additional asymmetries are time independent. This approach makes it possible to extract the indirect CP violation from the time dependence of the raw asymmetries. The validity of these assumptions is studied in details in the Chap. 8. This measurement is however dominated by the statistical uncertainty as will be shown later.

To verify the proper analysis procedure, the channel $D^0 \to K^- \pi^+$ provides a very high statistics. This control channel is a Cabibbo-favoured mode in which the indirect CP violation is negligible, since the contribution from $D^0 \to K^+ \pi^-$ decays is highly suppressed in this decay mode.

## 4.2   Analysis steps

The analysis is divided into four main steps.

In the first stage, a robust selection is developed and optimized, which is described in Chap. 5. Thereby, already at this stage many possible backgrounds and sources of systematic uncertainties are identified and removed. One important feature of the analysis is, that it covers a wide range of decay-times. Due to the large flight distance of the mother $B$ decays as well $D$ mesons with small decay times are possible to identify and to select. The final selection is validated and does not introduce additional systematic effects, thereby the significance of the measurement is maximized.

In the second part, a fitter framework is set up that can handle the model parameter extraction. It relies as default on maximum likelihood method. The method determines the signal yields and asymmetries. This extraction is done thought the invariant mass spectrum of the $D^0$ and $\bar{D}^0$ meson decays where the background can be removed from all the candidates. Moreover, different parametrizations and extraction methods are developed and tested, as described in Chap. 6. The asymmetries are extracted in bins of decay time. This method also considers shape variations between the decay time bins and between the $D^0$ and $\bar{D}^0$ candidates. The method is optimized and is validated on simulation and data.

As described in Chap. 7, in the third step, the indirect CP-violation parameter, $A_\Gamma$, is extracted. This is done by using a $\chi^2$ fit to the time dependence of the asymmetry. An individual, robust binning scheme is developed without losing in sensitivity of the analysis. Thereby, the full fit procedure is validated.

In the last step, the systematic uncertainties are validated in Chap. 8. Furthermore, the time-dependent effects of the remaining asymmetries are evaluated. Additionally, the mistag probabilities that can dilute the observed asymmetry are studied. Several cross checks are performed, e.g. splitting the datasets in data taking periods, magnet polarity and other subsamples and confirm the uncertainties which are referenced in this Chap. 8. Finally, the full set of systematic uncertainties is evaluated.

# Reconstruction and selection of $D^0$ mesons from semi-muonic $B$ decays

This analysis exploits $D^0$ ($c\bar{u}$) and $\overline{D^0}$ ($\bar{c}u$) mesons produced in semi-muonic $B$ decays. [1] These $D^0$ mesons have their origin in the two primary channels $B^+ \to \overline{D}^0 \mu^+ \nu_\mu X$ and $B^0 \to \overline{D}^0 \mu^+ \nu_\mu X$. [2] Thereby, X denotes other possible particles that can be produced in the decays. Contrary to a fully hadronic decay, the $D^0$ meson is accompanied through a muon. The charge of the muon can be used to determine at production time the flavour of the $D^0$ mesons, which are produced as flavour eigenstates . The muon, originating from the first vertex, is used for the flavour determination (tagging) of the $D^0$ mesons. The second vertex can be reconstructed using the $D^0$ decay products. For the measurement CP final states $D^0 \to \pi^- \pi^+$ and $D^0 \to K^- K^+$ are used and $D^0 \to K^- \pi^+$ is used as a control channel.

The decay chain reconstruction requires the determination of two vertices, where the charm quark is produced and the vertex where the $D^0$ meson decays. For precise decay time determination, the information of both vertices can improve the resolution. Therefore the full decay chain $B \to \overline{D}^0 \mu^+$ is fitted with Kalman Fitter. The fitter simultaneously extracts the momenta, positions and decay time and together with their correlations and uncertainties for particles in the decay chain, the details of the method are described in Ref. [63]. Thereby only the muon and the $D^0$ meson build the B vertex and missing particles, e.g. neutrino, are not reconstructed. The formed B-meson candidate is therefore only partly reconstructed. However, the $D^0$ decay is fully reconstructed and $D^0$-meson mass and decay time are accurately determined.

One of these reconstructed events recorded with LHCb detector is shown in Fig. 5.1. In the following sections the datasets used and the reconstruction and selection of these decays is described. For further studies, additionally to the datasets taken by LHCb, large simulation samples are generated and used here.

## 5.1 Datasets

We are using two disjoint pp-collision datasets recorded with LHCb at a center of mass energies of $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV. They correspond to $1.0\,\text{fb}^{-1}$ of data collected in 2011 and $2.0\,\text{fb}^{-1}$ of

---

[1]Semi-muonic $B$ decays: $B \to \overline{D}^0 \mu^+ X$

[2]Charge conjugated decays are also implied, unless it is explicit stated.
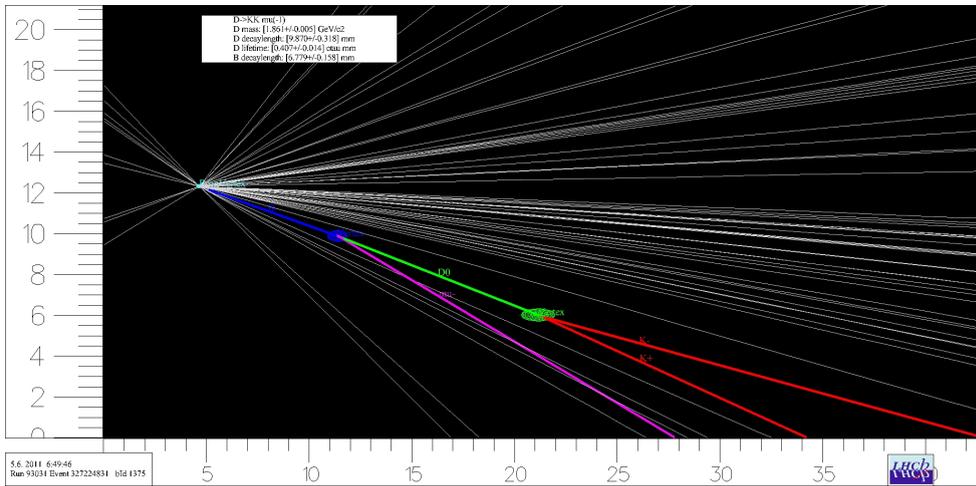
**Figure 5.1:** $\overline{B} \to D^0 \mu^- X$, $D^0 \to K^- K^+$ *candidate event.*
*Recorded at LHCb in the year 2011 (taken from Ref. [64]). Shown are the tracks in the x-z plane in units of* mm *of the innermost VELO region. The primary vertex is highlighted in light blue, the path of the B-meson and the decay vertex are visualized in blue. The reconstructed muon is shown in pink and the $D^0$ meson in green. Finally, the decay vertex of the $D^0$-meson is visualized in green and the reconstructed kaons in red.*

data recorded in 2012 that pass all data quality criteria [3]. The data is reconstructed with Brunnel framework (Ref. [46]) and for the analysis DaVinci framework (Ref. [47]) is used, as described in Chapter 2.

To study efficiencies in several selection steps and also their dependencies versus true decay time a Monte Carlo simulation sample (MC) using the Gauss framework (Ref. [48]) is generated . In the MC production, the events are generated with Pythia8 and go through the full detector simulation, listed in Appendix A.7. Thereby, the events are reconstructed with the same software as the data and a trigger with the same configuration as on data [4] . For the trigger the Moore framework (Ref. [65]) is used .

## 5.2   Preselection and Trigger

This section describes the preselection and the trigger selection. The huge datasets require a centrally managed reconstruction, that is necessary to reduce the datasets to the interesting events that are used in the analysis. This preselection step is called "Stripping" (see Chapter. 2.3.7) and this analysis uses a configuration called Stripping20 for 2012 data and Stripping20r1 for 2011 data which are explained in the following. Finally, the trigger selection together with the stripping step is described.

### Preselection (Stripping)
The signal yields are determined in four different channels, namely $D^0 \to \pi^+\pi^-$ and $D^0 \to K^+K^-$ decays for the $A_\Gamma$ measurement and $D^0 \to K^\pm\pi^\mp$ decays for systematic studies. They are extracted from a fit to the $D^0$ mass distribution.

Examples of $D^0$ candidate mass distributions for the $D^0 \to \pi^+\pi^-$ and $D^0 \to K^+K^-$ channels are shown in Figure 5.2. Besides combinatorial background, in both channels a physics background

---

[3] All detector subsystems were on and the detector performance was verified for the used data samples.

[4]Trigger Configuration Key Number `Trig0x409f0045`

coming from $D^0 \to K^+ \pi^-$ reflection is seen in the low mass region of the $D^0 \to \pi^+ \pi^-$ distribution and in the high mass region of the $D^0 \to K^+ K^-$ distribution. Additionally, there can be $D^0$ and $D^+$ multi-body decays with one or more missing tracks, appearing at lower masses. Further, decays of other hadrons that match our $D^0$ decay chain but where the particle ID is wrong or some products are missing would contribute to the background. The different backgrounds are studied and are suppressed by the selection choice.

The following section concentrates on describing the selection used in this analysis. Furthermore a possible gain in significance ($S/\sqrt{S+B}$) of the offline selection is studied. However not only the total signal yield is essential but also keeping events at large $D^0$ decay times as they have more weights to the sensitivity of the $A_\Gamma$ analysis. Both effects are followed up in this section.

The requirements of the preselection are described in the following. The applied selection cuts are summarized in Table 5.1.

First of all, track-based cuts on the stable particles are applied to reduce efficiently the background. This is done by requiring minimum momenta ($p$) and transverse momenta ($p_t$) for the $D^0$ daughters and the muon. Additionally, to reduce the misidentified events, particle identification (PID) cuts for the $B$ and $D^0$ daughters are applied. To remove clone tracks, a minimum Kullback-Liebler (KL) distance is required for all tracks [66]. To remove ghosts and obtain better resolution, fit quality criteria on the track ($\chi^2_{\text{track}}/$ndof) are applied. This analysis is interested in $D^0$ mesons produced in $B$ decays. To efficiently suppress the prompt $D^0$ mesons the tracks are required to have a large impact parameter (IP) $\chi$ value with respect to the primary vertex (PV). The $D^0$ daughters tracks are required to have a small $\chi^2$ of the distance of closest approach ($\chi^2$ DOCA), to be able to form a proper vertex. Before explicitly reconstructing the $D^0$ and $B$ mesons, the mass from the four-vector sum of the three particles (two hadrons from the $D^0$ decay and the muon) has to be below $6.2\,\text{GeV}/c^2$ and the sum of transverse momenta of the $D^0$ daughters has to be above $1.4\,\text{GeV}/c$. Once the $D^0$ and $B$ mesons are reconstructed, a loose cut on the signal masses is applied. Also, cuts on the quality of the $B$ and $D^0$ decay vertices are applied ($\chi^2_{\text{vertex}}/$ndof). Additionally, to reduce background, for $B$ and $D^0$ mesons the cosine of the angle $\cos\alpha$ (DIRA) between the momentum of the particle and the direction of flight from the best PV to the decay vertex is computed and minimum requirements on their values are applied.

The efficiency of the stripping selection is evaluated on MC samples (details, see Appendix A.7). Different selection steps before the stripping are studied to understand the contribution to the total efficiency. In the used MC samples, already on generator level the signal particles are required to be in the detector acceptance. Therefore, the following efficiencies are quoted relative to events with the muon and the $D^0$ in the detector acceptance. Different effects that have an impact on the efficiency are listed in increasing order in Tab. 5.2. The main effect is that the
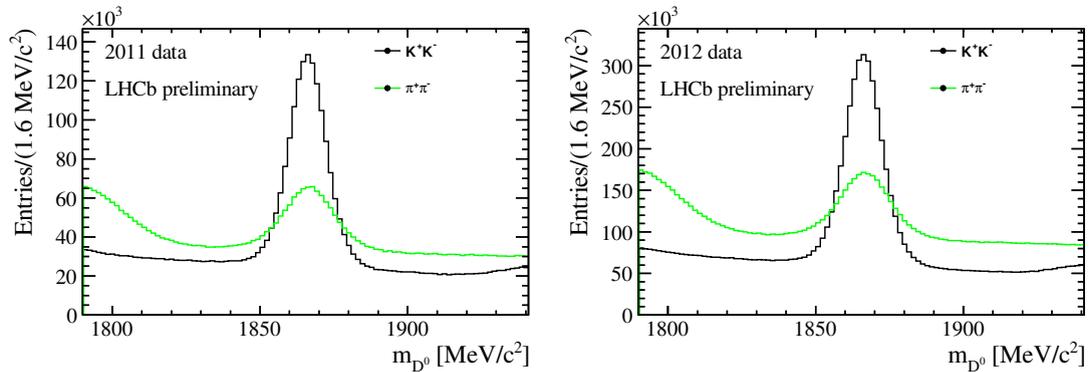


**Figure 5.2:** *$D^0 \to K^+K^-$ and $D^0 \to \pi^+\pi^-$ invariant mass distributions, after the stripping selection is applied; on the left for 2011 (`stripping20r1`) and on the right for 2012 data (`stripping20`).*

**Table 5.1:** *Stripping selection requirements.*

|  | $K$ | $\pi$ | $\mu$ |
|---|---|---|---|
| $p$ | $> 2\,\text{GeV}/c$ | $> 2\,\text{GeV}/c$ | $> 3\,\text{GeV}/c$ |
| $p_\text{T}$ | $> 300\,\text{MeV}/c$ | $> 300\,\text{MeV}/c$ | $> 800\,\text{MeV}/c$ |
| $D_{KL}$ distance | $< 5000$ | $< 5000$ | $< 5000$ |
| $\chi^2_\text{track}/\text{ndof}$ | $< 3$ | $< 3$ | $< 3$ |
| $\chi^2(\text{IP})$ | $> 9$ | $> 9$ | $> 4$ |
| $\log \mathscr{L}(\text{PID or DLL})$ | $\log \frac{\mathscr{L}(K)}{\mathscr{L}(\pi)} > 4$ | $\log \frac{\mathscr{L}(\pi)}{\mathscr{L}(K)} < 10$ | $\log \frac{\mathscr{L}(\mu)}{\mathscr{L}(\pi)} > 0$ |

|  | $D^0$ | $\mu D^0$ |
|---|---|---|
| invariant mass | $\in [1785, 1945]\,\text{MeV}/c^2$ | $\in [2.5, 6.0]\,\text{GeV}/c^2$ |
| mass from four-vector sum | - | $< 6.2\,\text{GeV}$ |
| sum $p_\text{T}$ $D^0$ daughters | $> 1400\,\text{MeV}/c$ | - |
| $\chi^2$ (DOCA) | $< 20$ | - |
| $\chi^2_\text{vertex}/\text{ndof}$ | $< 6$ | $< 6$ |
| $\chi^2$-distance $D^0$ vertex–PV | $> 100$ | - |
| $\cos \alpha (DIRA)$ | $> 0.99$ | $> 0.999$ |

**Table 5.2:** *The definitions of different selection steps.*

| cut name | requirements |
|---|---|
| generator level | signal tracks in LHCb acceptance |
| fiducial | $p_t(\mu) > 0.8\,\text{GeV}/c$ and $p(\mu) > 3\,\text{GeV}/c$ and $D^0$ daughters $p_t > 0.3\,\text{GeV}/c$ |
| reconstructible | 3 R-and $\phi$ hits in Velo and 1 x-and stereo hit in each of the T-stations |
| loose reco. | loose reconstruction using `StdAllLooseParticles`, see Tab.5.4 |
| stripping | full stripping selection and `StdLooseParticle`, see Tab.5.1 |

tracks are not reconstructible. This can be either due to their rather low momentum, due to decays in flight of the particle or due to interaction with the material ( [67], [68]). Requiring the tracks to pass fiducial cuts at generator level on the momentum of the final state particles involved in the decay and asking them to be reconstructible already lowers the efficiency to about 40% (Tab.5.3).

After this step, a loose reconstruction is performed which is described in Table 5.4. The loose reconstruction reconstructs about 25% of the signal events where the $D^0$ and $\mu$ are generated in the LHCb acceptance. Requiring the additional cuts of the full stripping selection the yield is further reduced to $10-15\%$ depending on the decay channel (Tab. 5.3). The MC samples, split in $D^0$ produced from $B^+$ and $B^0$ (or c.c.), show very similar efficiencies. Additionally, the decay time dependence of the efficiency is studied. Not only absolute efficiencies are important for this measurement, but also to have efficient selection for high decay times as these events have a higher weight in the determination of $A_\Gamma$. On generator level no acceptance bias on the $D^0$ decay time is visible. Furthermore, requiring fiducial cuts, reconstructible daughter tracks and loose reconstruction keeps the acceptance constant. Applying the full stripping selection, however, introduces a small drop in the decay time acceptance at low decay times (Fig.5.3). The reason for this fall are the $\chi^2(IP)$ cuts for the kaon and pion. The small time dependence does not change the strategy for the measurement of $A_\Gamma$, since the $\chi^2(IP)$ cuts are independent of the muon charge. Moreover, these cuts are applied at the charge symmetric final state ($K^+K^-$ or $\pi^+\pi^-$).

**Table 5.3:** *The table shows the selection efficiencies calculated on simulated events using the same selection as in* `stripping20`. *Every cut is applied on top of the previous set of cuts. The studied efficiencies are $\epsilon$(fiducial cuts), $\epsilon$(fid. cuts $\wedge$ reconstructible ), $\epsilon$(fid. cuts $\wedge$ reconstructible $\wedge$ loose reco), $\epsilon$(fiducial cuts), $\epsilon$(fid. cuts $\wedge$ reconstructible ), $\epsilon$(fid. cuts $\wedge$ reconstructible $\wedge$ loose reco $\wedge$ stripping). They are quoted with the respect to the number of signal $D^0$ which are in the LHCb acceptance (100%).*

| | efficiencies in % | | | |
|---|---|---|---|---|
| decay | fiducial cuts | $\wedge$ reconstructible | $\wedge$ loose reco | $\wedge$ **stripping** |
| $B^+(D^0 \to KK)$ | 66.2 | 38.3 | 22.3 | **10.8** |
| $B^0(D^0 \to KK)$ | 69.0 | 39.8 | 23.1 | **10.9** |
| $B^+(D^0 \to K\pi)$ | 67.9 | 39.7 | 23.8 | **10.0** |
| $B^0(D^0 \to K\pi)$ | 71.0 | 41.3 | 24.7 | **10.3** |
| $B^+(D^0 \to \pi\pi)$ | 69.5 | 42.0 | 25.3 | **14.4** |
| $B^0(D^0 \to \pi\pi)$ | 72.8 | 44.0 | 26.4 | **14.6** |

**Table 5.4:** *Loose reconstruction requirements. $K,\pi$ and $\mu$ are reconstructed using* `StdAllLooseParticles`, *which have no $p_t > 250\,\mathrm{MeV}/c$ and $\chi^2(IP) > 4$ compared to* `StdLooseParticles`.

| selection in loose reconstruction | $D^0$ (from `StdAllLoose`) | $\mu\,D^0$ (from `StdAllLoose`) |
|---|---|---|
| invariant mass | $\in [1785, 1945]\,\mathrm{MeV}/c^2$ | $\in [2.5, 6.0]\,\mathrm{GeV}/c^2$ |
| mass from four-vector sum | - | $< 6.2\,\mathrm{GeV}$ |

| | additional detector requirement |
|---|---|
| `StdAllLoosePions` | CALO, RICH |
| `StdAllLooseKaons` | CALO, RICH, $\log \frac{\mathscr{L}(K)}{\mathscr{L}(\pi)} > -5$ |
| `StdAllLooseMuons` | Muon stations, IsMuon=true [5] |
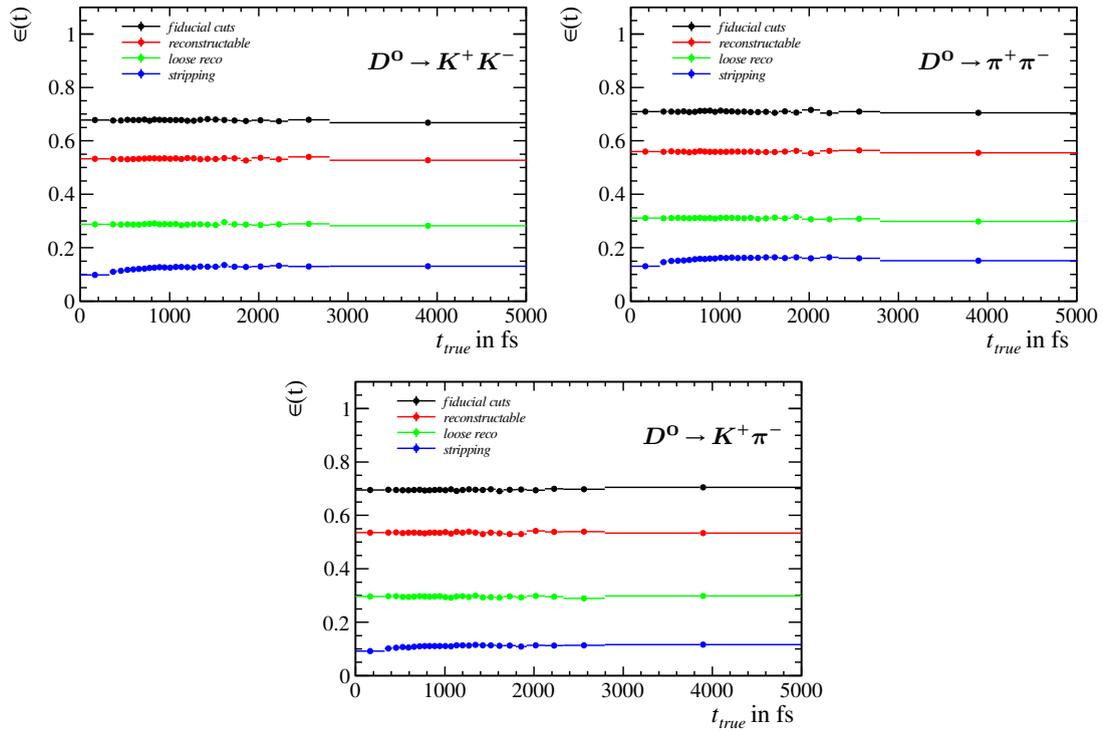
**Figure 5.3:** *Time dependent efficiencies for different selection steps are shown for $D^0 \to K^- K^+$, $D^0 \to \pi^- \pi^+$ (up) and $D^0 \to K^- \pi^+$ (down). Shown is $\epsilon$(fid. cuts), $\epsilon$(reco), $\epsilon$(loose reco) and full trigger selection $\epsilon$(stripping) versus true decay time. All efficiencies are determined with respect to generator level cuts, where the $D^0$ and the $\mu$ are required to be in the LHCb acceptance.*

**Table 5.5:** *Different trigger efficiencies $\epsilon(trigger|stripping)$ determined on MC. The efficiencies are calculated in the sequential order: $\epsilon(L0|stripping), \epsilon(Hlt1|L0 \wedge stripping)$ and $\epsilon(Hlt2|Hlt1 \wedge L0 \wedge stripping)$.*

| MC after | | | |
|---|---|---|---|
| **Trigger line** | $D^0 \to K^- K^+$ | $D^0 \to \pi^- \pi^+$ | $D^0 \to K^- \pi^+$ |
| **L0Muon on $\mu$** | 68.1% | 67.5% | 68% |
| TrackAllL0 on $B$ | 77.2% | 77.5% | 76.8% |
| TrackMuon on $\mu$ | 79.2% | 78.8% | 78.7% |
| **HLT1 combined** | 88.5% | 88.4% | 88.1% |
| TopoMu2BodyBBDT on $B$ | 62.7% | 64% | 62.2% |
| TopoMu3BodyBBDT on $B$ | 58.6% | 59% | 56.9% |
| SingleMuon on $\mu$ | 33.7% | 33% | 32.8% |
| **HLT2 combined** | 81% | 82.4% | 80.7% |
| **combined trigger** | 48.8% | 49.2% | 48.3% |

**Trigger**

Our decay topology contains a muon in the final state, which are very efficiently triggered at LHCb, see Chapter 2 for more details. The muon candidates can be reconstructed already at the hardware level with the L0 muon trigger with a momentum precision of 20 %. A trigger configuration with the requirement on the transverse momentum of the muon $p_T > 1.76\,\text{GeV}/c$ [6] is used. After the full reconstruction of B-decay candidates, it is checked that the according B-daughter muon has fired the L0Muon trigger. In the software trigger (HLT) a trigger line is used that selects tracks with good track quality and removes the ones with a very low transverse momentum ($p_T > 1.6\,\text{GeV}$).

The reconstructed events have to pass a well-defined trigger setup. An efficient signal selection is chosen with the triggers described below. All the selected trigger lines have to fulfill the trigger requirement comparing the detector hits used with the selected signal particles. Thereby, only candidates from the signal reconstructed decay $B \to \overline{D^0}\mu^+$ are used.

- For L0: The trigger requirement is directly on the muon (L0Muon on $\mu$).

- For HLT1: One of the reconstructed B-daughter tracks has to pass the TrackAllL0 requirements (TrackAllL0 on one of the $B$ daughters) or the muon has to fire the trigger (TrackMuon on $\mu$).

- Finally, for HLT2: The Topological trigger lines [69] are used, that takes into account the $B$ decay chain (TopoMu2BodyBBDT on $B$ or TopoMu3BodyBBDT on $B$). Both lines are based on a single displaced track and a significantly displaced vertex containing this track and additional 1-2 other tracks. The multivariate BBDT [7] algorithm is very efficient in selecting the signal b-hadron events and works very robust, as described in Ref. [70]. Furthermore, a single muon trigger is used (SingleMuon on $\mu$), without requirements on the $D^0$ daughters.

This trigger selection keeps about 49% of the signal events selected after the stripping (Tab. 5.5). However, looking at the data the efficiency of our trigger selection with respect to the stripping and all the triggers currently available at LHCb, efficiency is about 75%. These numbers are consistent for different channels. This actually means that a fraction ($1 - \frac{49\%}{75\%} \approx 0.35$) of events that

---

[6] For 2011 trigger configuration, the threshold was $p_T > 1.48\,\text{GeV}/c$ but the momentum was not correctly estimated. The solution to correct for this effect required an emulation of the trigger and further momentum selection of $p_T > 1.64\,\text{GeV}/c$, reducing the yield by 7%.

[7] Bonsai Boosted Decision Tree algorithm

**Table 5.6:** *Different trigger efficiencies determined on data with respect to events passing the stripping selection and any available triggers. The efficiencies are calculated in the sequential order: $\epsilon(L0|\ stripping \wedge\ any\ trigger)$, $\epsilon(\ Hlt1\ |\ L0 \wedge\ stripping \wedge\ any\ trigger)$ and $\epsilon(Hlt2\ |\ Hlt1 \wedge\ L0 \wedge\ stripping \wedge\ any\ trigger)$. The HLT2SingleMuon line has a prescale of 0.5.*

| **2011 data after** | | | |
|---|---|---|---|
| **Trigger line** | $D^0 \to K^- K^+$ | $D^0 \to \pi^- \pi^+$ | $D^0 \to K^- \pi^+$ |
| **L0Muon on $\mu$** | 81.5% | 80.7% | 81.0% |
| TrackAllL0 on $B$ | 88.6% | 89.4% | 88.8% |
| TrackMuon on $\mu$ | 81.5% | 80.9% | 81.4% |
| **HLT1 combined** | 97.5% | 97.6% | 97.6% |
| TopoMu2BodyBBDT on $B$ | 66.6% | 66.9% | 66.8% |
| TopoMu3BodyBBDT on $B$ | 56.1% | 57.9% | 56.3% |
| SingleMuon on $\mu$ | 19.3% | 17.9% | 18.8% |
| **HLT2 combined** | 85.9% | 87.1% | 86.6% |
| **combined trigger** | 68.8% | 69.1% | 68.9% |
| **2012 data** | | | |
| **Trigger line** | $D^0 \to K^- K^+$ | $D^0 \to \pi^- \pi^+$ | $D^0 \to K^- \pi^+$ |
| **L0Muon on $\mu$** | 85.4% | 84.5% | 84.9% |
| TrackAllL0 on $B$ | 86% | 86.7% | 86.2% |
| TrackMuon on $\mu$ | 87.5% | 86.8% | 87.3% |
| **HLT1 combined** | 98.5% | 98.6% | 98.4% |
| TopoMu2BodyBBDT on $B$ | 68.9% | 69.1% | 69.4% |
| TopoMu3BodyBBDT on $B$ | 62.2% | 64.1% | 62.4% |
| SingleMuon on $\mu$ | 20.4% | 19.1% | 20% |
| **HLT2 combined** | 87.6% | 88.9% | 88.4% |
| **combined trigger** | 73.7% | 74.1% | 73.9% |

could pass the stripping are not triggered by any of the triggers (note that this is MC). For data, the contributions of every trigger line can be found in Tab. 5.6. The triggers are chosen not only to maximize the efficiency but also to keep enough events in the upper decay time region, since the covering of a large decay time region is necessary for the precision on $A_\Gamma$. They are validated and commonly used in other LHCb analyses [71], [72]. Fig. 5.4 shows the time dependence of the different trigger efficiencies. HLT1 and L0 trigger selection have no big time dependence. For Hlt2 triggers there is a decrease in efficiency for larger decay times visible. This also propagates into the final selection. The main reason for this are the TopoBody triggers, which preferably group muons and the D decay products to one vertex. This leads to a decrease in acceptance at higher decay times. However, this effect is studied and has no influence on the measurement of $A_\Gamma$. This assumption is based on the fact that acceptance is caused by effects in the trigger that do not dependent on the muon charge. It is validated at our precision level on simulation and in Chapter 7 by measuring $A_\Gamma$ in the $D^0 \to K^- \pi^+$ control sample.
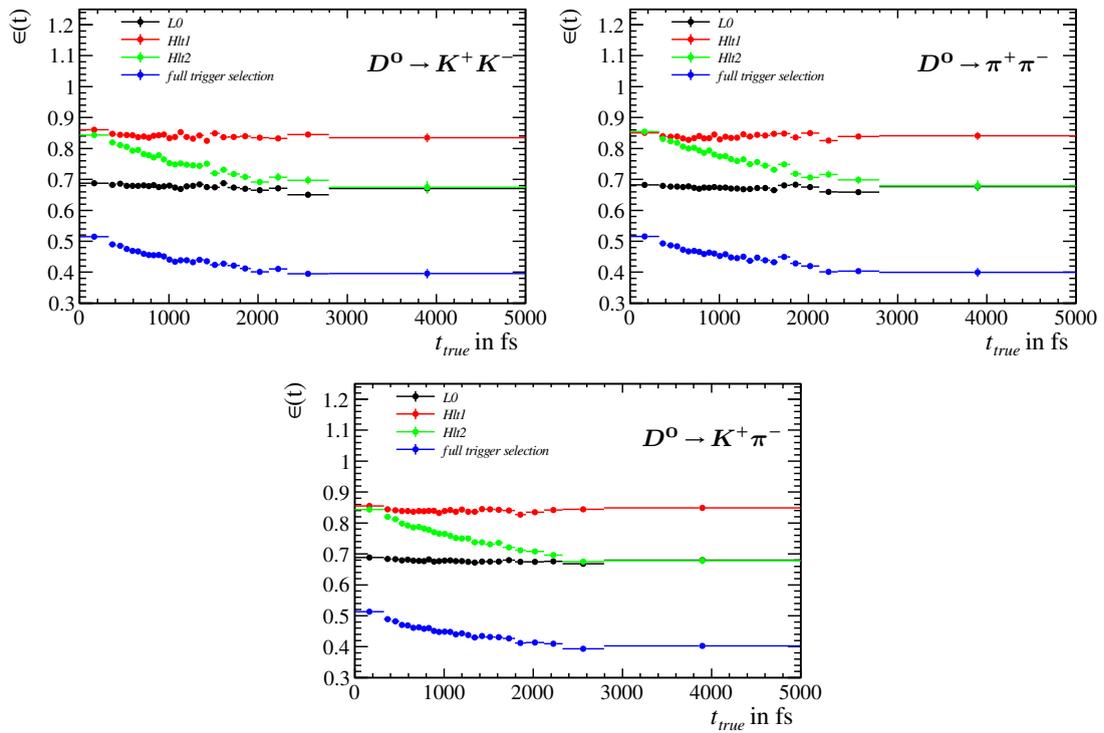
**Figure 5.4:** *Time dependence of relative trigger efficiencies in simulated events, shown for $D^0 \rightarrow K^- K^+$, $D^0 \rightarrow \pi^- \pi^+$ (up) and $D^0 \rightarrow K^- \pi^+$ (down). Shown is $\epsilon(L0|stripping)$, $\epsilon(Hlt1|L0 \wedge stripping)$, $\epsilon(Hlt1|Hlt2 \wedge L0 \wedge stripping)$ and full trigger efficiency $\epsilon(L0 \wedge Hlt1 \wedge Hlt2|stripping)$ versus true decay time.*

## 5.3 Offline selection

The datasets used in this analysis are required to have a proper data quality flag. Especially, all the sub-systems should be flagged as good. For 2012 data all the run numbers below 114000 are excluded, due to possible asymmetry biases in this dataset.

Finally, after stripping and trigger cuts the significance of the signal yields are optimized. All the variables that have an impact on reducing the background have to be considered in offline selection. However, it is avoided here to use variables strongly correlated with $D^0$ decay time not to decrease the sensitivity on $A_\Gamma$. More details on the selection quantities and their separation power can be found in Appendix B.11. The default selection that we introduce here is a cut-based method. It is applied to the variables with the highest separation power. The separation power of the variables is shown in Tab. B.1.

To remove kaon and pion ($K \leftrightarrow \pi$) misidentification, PID cuts on the DLL values are applied. The optimal value is chosen maximizing the significance $S/\sqrt{S+B}$. For the $D^0 \to \pi^-\pi^+$ channel this resulted in a slight increase in the DLL cut with respect to the stripping. One small background is coming from the $J/\psi(1S)$ and $\psi(2S)$ resonances decaying into two muons, where one of the muons can be reconstructed as a pion or a kaon. To remove it, one combines the muon from the $B$ decay with the $D^0$ daughter with the opposite charge under the muon hypothesis. Requiring the kaon or pion to pass the muon stations with additional `isMuon` flag, and the invariant mass m($\mu^+\mu^-$) to be inside the mass windows $\in [3040, 3155]\,\mathrm{MeV}/c^2$ or $[3630, 3745]\,\mathrm{MeV}/c^2$ gives us an efficient $J/\psi$ and $\Psi(2S)$ veto. Additionally, for the same charge combination a peak at $2\times$m($\mu$) is visible. These are cloned tracks which have the same momentum. They are vetoed requiring invariant mass of $M(\mu\mu) > 240\,\mathrm{MeV}/c^2$. The impact of these two vetos is shown in Fig. 5.5. All 3-body $B$ decays are suppressed with a cut on the maximum invariant $D^0$-$\mu$ mass of the $B$ candidates of $5\,\mathrm{GeV}/c^2$.

For the final selection, events with negative reconstructed decay time are removed, since they do not contribute significantly to the measurement of $A_\Gamma$, however, have a high background, especially in the $D^0 \to \pi^+\pi^-$ channel. The full offline selection is shown in Tab. 5.7.

The final efficiency and yield numbers of the offline selection are determined with respect to stripping and trigger selection (Tab. 5.8). The efficiency numbers are similar for $D^0 \to K^-\pi^+$ and $D^0 \to K^-K^+$ channel but are lower in case of $D^0 \to \pi^-\pi^+$, since we apply a slightly harder DLL cut there. The time acceptance of the offline selection shows no further time dependence, which would have an impact on the $A_\Gamma$ significance. Removing events with negative decay times leads to a small drop in efficiency at very low decay times. The relative offline selection efficiency stays flat for higher decay times (Fig. 5.6).

Finally, to visualize the improvement of the offline selection Tab. 5.9 shows the final significance gain that is directly proportional to the significance gain on $A_\Gamma$. In $D^0 \to K^-K^+$ we obtain a

**Table 5.7:** *Offline selection on top of the stripping and the trigger requirements. For the final result the decay time cut is tightened further.*

| Variable | $D^0 \to \pi^-\pi^+$ | $D^0 \to K^-K^+$ |
|---|---|---|
| $p_\mathrm{T}(D^0)$ | $> 0.5\,\mathrm{GeV}/c$ | $> 0.5\,\mathrm{GeV}/c$ |
| $p_\mathrm{T}(\mu)$ | $> 1.2\,\mathrm{GeV}/c$ | $> 1.2\,\mathrm{GeV}/c$ |
| $\chi^2(\mathrm{IP})$ of muon | $> 9$ | $> 9$ |
| $\mathrm{DLL}_{K\pi}$ of $D^0$ daughters | $< -2$ | $> 7$ |
| $M(B)$ | $< 5\,\mathrm{GeV}/c^2$ | $< 5\,\mathrm{GeV}/c^2$ |
| $\tau(D^0)$ (final) | $> 0.0\,\mathrm{ps}$ | $> 0.0\,\mathrm{ps}$ |
| $J/\psi$ veto(see text) | yes | yes |
| clones veto(see text) | yes | yes |
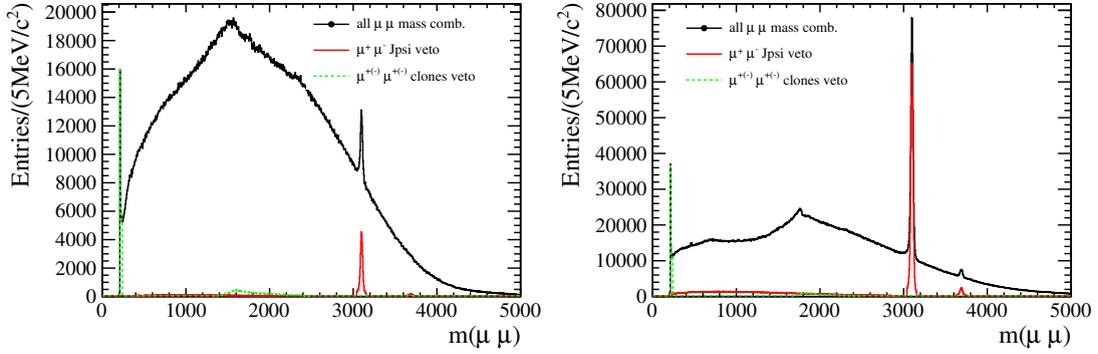
**Figure 5.5:** *Mass combinations under the muon hypothesis right after stripping. Combination of same and opposite charges are shown. In green the veto on the clones and in red the removed $J/\psi(1S)$ and $\psi(2S)$ resonances for $D^0 \to K^- K^+$ (left) and $D^0 \to \pi^- \pi^+$ (right candidates in data.)*

**Table 5.8:** *Offline selection efficiencies with respect to the stripping and trigger selection. In brackets for the datasets also the signal yields of $D^0$ mesons are shown.*

|  | offline selection efficiency and signal yields | | |
|---|---|---|---|
| **Sample** | $D^0 \to K^- K^+$ | $D^0 \to \pi^- \pi^+$ | $D^0 \to K^- \pi^+$ |
| **2012** | 79.7% (1.57 M) | 72.9% (0.56 M) | 77.8% (7.77 M) |
| **2011** | 80.4% (0.67 M) | 71.5% (0.23 M) | 77.8% (3.54 M) |
| **MC** | 83.1% | 73.6% | 80.6% |

gain of: 2-3% and in $D^0 \to \pi^- \pi^+$: 24-27% in significance.

**Table 5.9:** *Improvement of the offline selection. Shown are significances defined as $\frac{S}{\sqrt{S+B}}$ in the $M_{D^0} \in \{1815, 1915\}$ MeV/c² mass window. The value in brackets shows the significance value of the stripping and trigger selection.*

|  | Significance of the offline selection | |
|---|---|---|
| **Sample** | $D^0 \to K^- K^+$ | $D^0 \to \pi^- \pi^+$ |
| **2011** | 662(643) | 356(286) |
| **2012** | 1028(1005) | 549(432) |

Further possible gain in selection are studied using the TMVA package (Appendix B.11). Compared to the full offline cut-based analysis, a gain of 4% in significance is possible in $D^0 \to K^- K^+$ and about 3% in $D^0 \to \pi^- \pi^+$ exploiting a multivariate training. For simplicity reasons, however, the cut-based selection is taken for this analysis. Further improvements would require an update of the stripping and trigger selections.
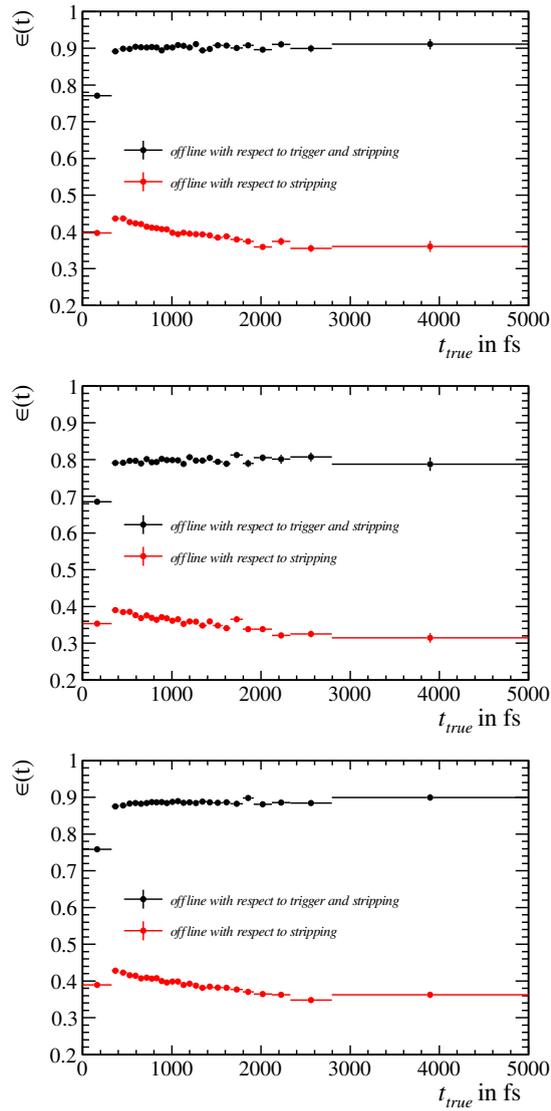
**Figure 5.6:** *Efficiencies from time dependence of offline selection obtained from MC for $D^0 \to K^- K^+$, $D^0 \to \pi^- \pi^+$ (up) and $D^0 \to K^- \pi^+$ (down) decays. Shown is $\epsilon$(offline selection|stripping) in red, $\epsilon$(offline selection|trigger $\wedge$ stripping) versus true decay time.*
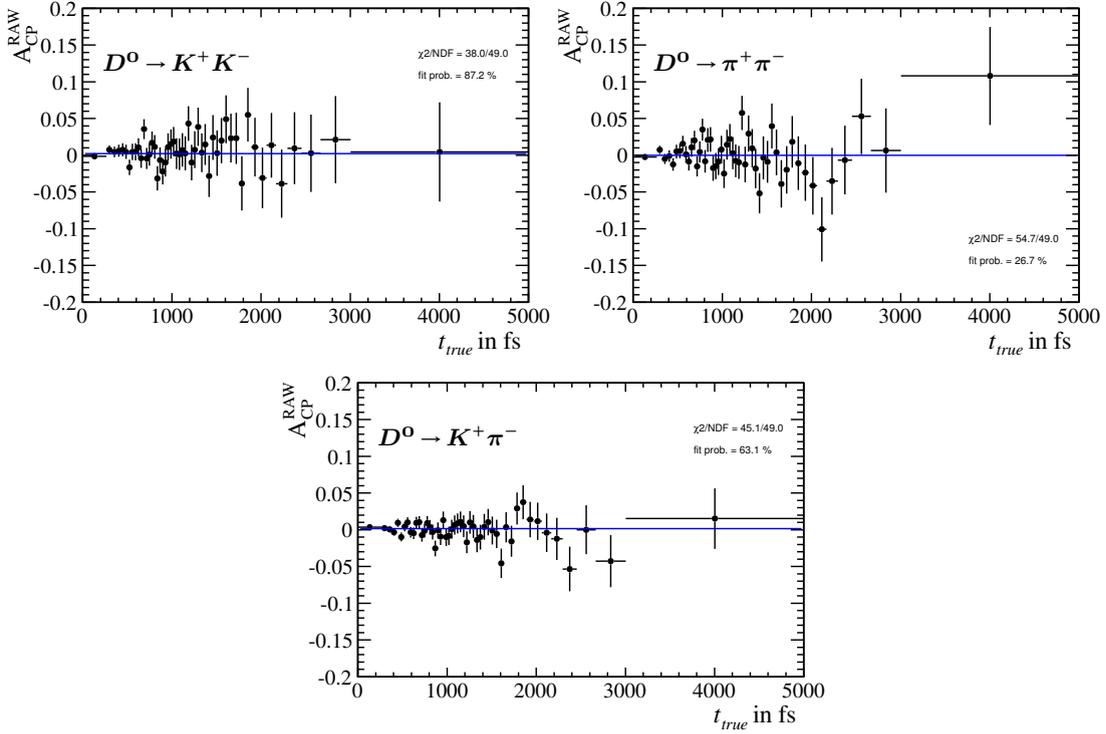
**Figure 5.7:** *Raw asymmetries determined on MC sample after the final offline selection. Shown for $D^0 \to K^- K^+$, $D^0 \to \pi^- \pi^+$ (up) and $D^0 \to K^- \pi^+$ (down). Here the samples $B^0$ and $B^+$ are merged. Additionally, $A_\Gamma$ is extracted and the uncertainty on the value is shown. The result are consistent with having no time dependence.*

### 5.3.1 MC raw asymmetries

Also, charge dependent effects are studied on MC. Figure 5.7 shows $A_{CP}^{RAW}$ versus true decay time determined on MC after the full offline selection. A constant is fitted to the asymmetry time dependence and probability determined from $\chi^2$ and number of degrees of freedom. This shows no significant variations from the hypotheses that the asymmetry has no time dependence. Also the $B^+$ and $B^0$ samples show no significant deviations. All the 3 channels are compatible with $A_\Gamma = 0$. Combining all the 3 channels on MC gives

$$A_\Gamma = (0.012 \pm 0.082)\%$$

and is also fully compatible with zero. In data, this dependence can be studied with even higher precision in the $D^0 \to K^- \pi^+$ channel and confirms this MC charge asymmetry study, for details see Chapter 7.

# 6

# Methods

This section describes in detail the analysis strategy and methods used for the determination of the signal yields or directly of the raw asymmetries ($A_{CP}^{RAW}$) and their time dependence. Finally, a method for extraction of the indirect CP violation is explained.

## 6.1 Methods for parameter estimation

Two main techniques are used for the parameter extraction; this is the maximum Likelihood and the minimum $\chi^2$ methods. Both are briefly derived here.

**Maximum Likelihood**

Maximum likelihood is an efficient method, which satisfies the consistency and with increasing statistics can estimate the true parameter in an with arbitrary precision. One considers $n$ measurements of a quantity set $\vec{x}$. Under the assumption that the underlying probability density function $\mathscr{PDF}(\vec{x}|\vec{\lambda})$ is known, where $\vec{\lambda}$ parametrizes the several unknown parameters, one can calculate the probability for such a result. The maximum likelihood method uses the following Likelihood function

$$\mathscr{L}(\vec{\lambda}) = \prod_{i=1}^{n} \mathscr{PDF}(\vec{x}_i|\vec{\lambda}). \tag{6.1}$$

It gives the probability to obtain from the $\mathscr{PDF}$ the data $\vec{x}_i$ ($i = 1, ..., n$), if parameters $\vec{\lambda}$ are chosen. The maximum Likelihood fit obtains as the best estimator, the parameters that maximize the Likelihood. Usually, this problem is transformed into a minimization problem, using the negative logarithm. The function one needs to minimize to find the best estimator is

$$\mathscr{F}(\vec{\lambda}) = -\log \mathscr{L}(\vec{\lambda}) = -\sum_{i=1}^{n} \mathscr{PDF}(\vec{x}_i|\vec{\lambda}).$$

**Minimum $\chi^2$ estimator**

In very typical case of a Gaussian $\mathscr{PDF}(x_i|\mu) = \text{const} \cdot \exp(-\frac{1}{2}(\frac{x_i-\mu}{\sigma_i})^2)$, with the parameter $\mu$ as the mean of the distribution, the Log-Likelihood can be further simplified to a sum of $\chi_i^2 = (\frac{x_i-\mu}{\sigma_i})^2$, which is squared of the difference between expectation and measured value divided by its uncertainty. Moreover, it gives the following $\chi^2$ to be minimized:

$$\mathscr{F}(\vec{\lambda}) = -\log\mathscr{L}(\vec{\lambda}) + const = \sum_i \frac{\chi_i^2}{2}$$

The method can be even further generalized using a general function $f(x)$, as long the uncertainties can be approximated by a Gaussian distribution. This works for a function $f(x_i|\vec{\lambda})$, with a measured value of $f_{\text{meas}}(x_i)$ and its uncertainty $\sigma_{f_{\text{meas}}(x_i)}$, depending on the data point $x_i$. For the extraction of the estimator of $\vec{\lambda}$, one needs to consider the following quantity

$$\mathscr{F}_{min}^{\chi^2} = 2\mathscr{F}(\vec{\lambda}) = \sum_i (\frac{f_{\text{meas}}(x_i) - f(x_i|\vec{\lambda})}{\sigma_{f_{\text{meas}}(x_i)}})^2,$$

that needs to be minimized.

**Minimization**

The minimization of such a quantity (in case of both previously described methods) requires

$$\frac{\partial\mathscr{F}}{\partial\lambda_k} = 0 \qquad\qquad \text{for} \qquad\qquad k = 1, 2, \cdots$$

for the estimator $\hat{\lambda} = \lambda$. However, one has to take care, that the obtained parameters are not in a local minimum, but reach the global minimum (details Ref. [73]). Especially, with a growing number of parameters this requires a check that a real minimum is obtained. In most of the cases, the Likelihood function is normally distributed around the estimator $\hat{\lambda}$. In a large sample limit $n \to \infty$ the Likelihood function is becoming more Gaussian and the variance $V[\hat{\lambda}] \to 0$. This means that $\mathscr{F}$ is hyper-parabolic and symmetric minimization function. Therefore, it can be estimated by using

$$\mathscr{F}(\lambda) = \mathscr{F}(\hat{\lambda}) + \frac{1}{2}\sum_{i,j} \frac{\partial^2\mathscr{F}}{\partial\lambda_i\partial\lambda_j}|_{\hat{\lambda}}(\lambda_i - \hat{\lambda}_i)(\lambda_j - \hat{\lambda}_j) + \cdots.$$

The second derivative $G_{ij}$ can be calculated

$$G_{ij} = \frac{\partial^2\mathscr{F}}{\partial\lambda_i\partial\lambda_j}|_{\hat{\lambda}}.$$

Further, also the covariances ($V_{ij} = cov[\lambda_i, \lambda_j]$) can be estimated by inverting the second derivative matrix $G$

$$V_{ij}(\hat{\lambda}) = (G^{-1})_{ij}.$$

Using this, one can obtain the parameter correlations and their uncertainties. For an asymmetric likelihood function, the parameters need to be further transformed or the phase space needs to be scanned for the asymmetric uncertainty determination. Different minimization programs can handle this, e.g. HESSE for symmetric uncertainties and MINOS for asymmetric ones. All of them are part of the MINUIT package, Ref. [73]. Additionally, Mathematica, Ref. [74], provides similar minimization methods.

## 6.2   Mass fits

The $D^0$ mass resolution is dominated by the reconstruction and detector effects, e.g. momentum resolution, Ref. [67]. The probability density function ($\mathscr{PDF}$) used to model the signal mass distribution consists of the sum of three Gaussians. The first two Gaussians have the same mean $\mu_1$ and different widths $\sigma_{1,2}$. The third one has a larger width $\sigma_3$ and a shifted mean $\mu_3$ which is able to describe a small asymmetric tail (e.g. coming from photon radiation).

$$\mathscr{PDF}_{Signal}(m|\vec{\lambda_{sig}}) \equiv \sum_{i=1}^{3} \frac{c_i}{\sqrt{2\pi}\sigma_i} \exp\left\{-\frac{1}{2}\left(\frac{m-\mu_i}{\sigma_i}\right)^2\right\},$$

$$\mu_2 = \mu_1, \mu_3 = \delta_\mu + \mu_1, \sigma_2 = s_{\sigma_2}\cdot\sigma_1, \sigma_3 = s_{\sigma_3}\cdot\sigma_1,$$

$$c_1 = f_1, c_2 = (1-f_1)\cdot f_2, c_3 = (1-f_1)\cdot(1-f_2),$$

$$\vec{\lambda_{sig}} = \{\mu_1, \delta_\mu, \sigma_1, s_{\sigma_2}, s_{\sigma_3}, f_1, f_2\}.$$

Together with the fractions $f_i$ of the Gaussians there are seven free parameters for the signal shape. Different other approaches that have less and more free parameters (e.g. Crystal Ball function) are studied and show only a tiny impact on the measurement which is further described in Sec. 8.8. This Gaussian signal shape is used for all three analyzed modes $D^0 \to K^- K^+$, $D^0 \to \pi^- \pi^+$ and $D^0 \to K^- \pi^+$, however with an individual set of parameters for each mode.

Different possible backgrounds are studied that could peak in the mass region exploited in the fit, e.g. multi-body $D^0$ and $B$ decays with mis-identified particles or missing particles in the final state. No peaking structure is found inside the considered mass region. To model the combinatorial background an exponential function is used

$$\mathscr{PDF}_{comb.bkg}(m|\vec{\lambda_{bkg}}) \equiv N\cdot\exp(m\,\alpha), \quad \vec{\lambda_{bkg}} = \{\alpha\}, \tag{6.2}$$

with decay constant $\alpha$ and normalization $N$ in the mass window of the fit. In Sec. 8.8 different background models with more and less parameters are studied and no significant impact on the measurement of $A_\Gamma$ is found. For the $D^0 \to \pi^+\pi^-$ channel there is a visible contribution from $K\pi$ reflection in the $D^0$ invariant mass below $1820\,\mathrm{MeV}/c^2$. In the $D^0 \to K^+K^-$ channel, the $K\pi$ contribution is negligible. Therefore, for the $D^0 \to \pi^+\pi^-$ channel this component is modelled with a Gaussian function:

$$\mathscr{PDF}_{K\pi refl.bkg}(m|\vec{\lambda_{refl.bkg}}), \ \vec{\lambda_{refl.bkg}} = \{\mu_{refl.bkg}, \sigma_{refl.bkg}\}. \tag{6.3}$$

The total $\mathscr{PDF}$ has the following form:

$$\mathscr{PDF}_{tot K^+K^- or K\pi}(m|\vec{\lambda}) \equiv N_{sig}\cdot\mathscr{PDF}_{Signal}(m|\vec{\lambda_{sig}}) + N_{bkg}\cdot(\mathscr{PDF}_{comb.bkg}(m|\vec{\lambda_{bkg}})) \tag{6.4}$$

$$\mathscr{PDF}_{tot \pi^+\pi^-}(m|\vec{\lambda}) \equiv N_{sig}\cdot\mathscr{PDF}_{Signal}(m|\vec{\lambda_{sig}}) + N_{bkg}\cdot((1-f_{refl.bkg})\mathscr{PDF}_{comb.bkg}(m|\vec{\lambda_{bkg}})$$
$$+ (f_{refl.bkg})\mathscr{PDF}_{refl.bkg}(m|\vec{\lambda_{refl.bkg}}))$$

Finally, to extract the signal yields $N_{D^0 tag}$ and $N_{\bar{D}^0 tag}$ the data sample is split into $\mu^-(D^0)$ and $\mu^+(\bar{D}^0)$ tagged events. The yields can then be parametrized as $N_{\mu^+} = \frac{1-A_{CP}^{RAW}}{1+A_{CP}^{RAW}}N_{\mu^-}$. Assuming different mass shapes for $D^0$ and $\bar{D}^0$ events the number of parameters would double: $\vec{\lambda} = \{\vec{\lambda}^+, \vec{\lambda}^-\}$. However, several parameters that do not depend on the tag are chosen to be the same for both samples, which improves the sensitivity of the fit. The simultaneous parameters are all the signal shape parameters ($\sigma_1, s_{\sigma_2}, s_{\sigma_3}, \mu_1, \delta_\mu, f_1, f_2$), except the overall normalisation. These parameters for the $D^0$ mass shape are the same since since the final states $KK$ and $\pi\pi$ are charge symmetric. With this parametrisation a simultaneous unbinned extended maximum likelihood fit is performed
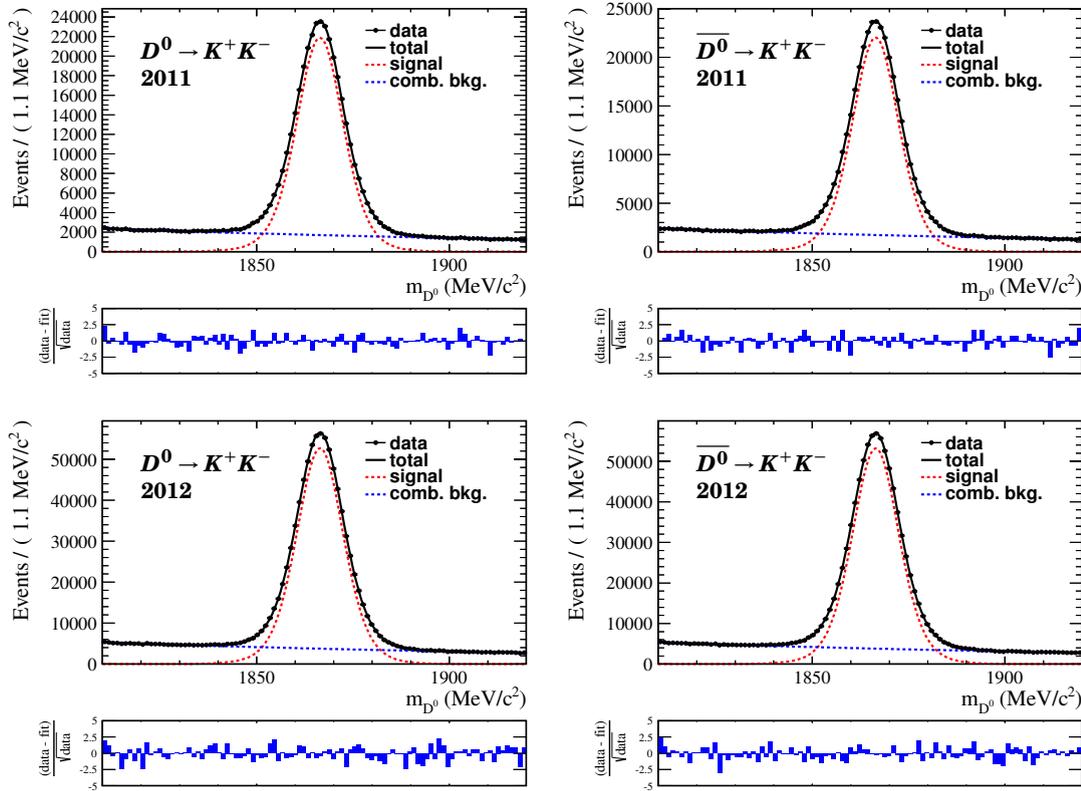
**Figure 6.1:** *Full data sets after the final selection, 2011 data (on top) and 2012 data (on bottom). The invariant mass distribution is fitted for $\mu^-(D^0)$-tagged (left) and for $\mu^+(\overline{D}^0)$-tagged candidates (right). The pulls are shown below.*

to extract the final global time integrated parameters for the $D^0 \to K^- K^+$ and $D^0 \to K^- \pi^+$ decay modes:

$$\vec{\lambda}_{global} = \{A_{CP}^{RAW}, \sigma_1, s_{\sigma_2}, s_{\sigma_3}, \mu_1, \delta_\mu, f_1, f_2, \vec{\lambda}_{bkg}^+, \vec{\lambda}_{bkg}^-\},  \tag{6.5}$$

and for $D^0 \to \pi^- \pi^+$ respectively:

$$\vec{\lambda}_{global} = \{A_{CP}^{RAW}, \sigma_1, s_{\sigma_2}, s_{\sigma_3}, \mu_1, \delta_\mu, f_1, f_2, \vec{\lambda}_{bkg}^+, \vec{\lambda}_{bkg}^-, f_{refl.bkg}^+, \vec{\lambda}_{refl.bkg}^+, f_{refl.bkg}^-, \vec{\lambda}_{refl.bkg}^-\}.  \tag{6.6}$$

The extracted parameters are listed in Appendix A.11. For the $D^0 \to K^- K^+$ channel the fit results are shown in Fig. 6.1.

The fit results for the $D^0 \to \pi^- \pi^+$ channel are shown in Fig. 6.2. For both tags the pulls under the mass peaks illustrate an appropriate signal model description and also the background sidebands region is described well by the background $\mathscr{PDF}$.

The fit results for the control $D^0 \to K^- \pi^+$ channel are shown in Fig. 6.3. Small structure in the pulls is visible due to larger statistics. However, this has no impact on the extracted asymmetry (see Chapter 8).

The same procedure can be extended to extract the parameters also in bins of the $D^0$ decay time. However, many of the parameters remain constant and have no impact on $A_{CP}$ and are thus fixed to the values obtained from the global fit.
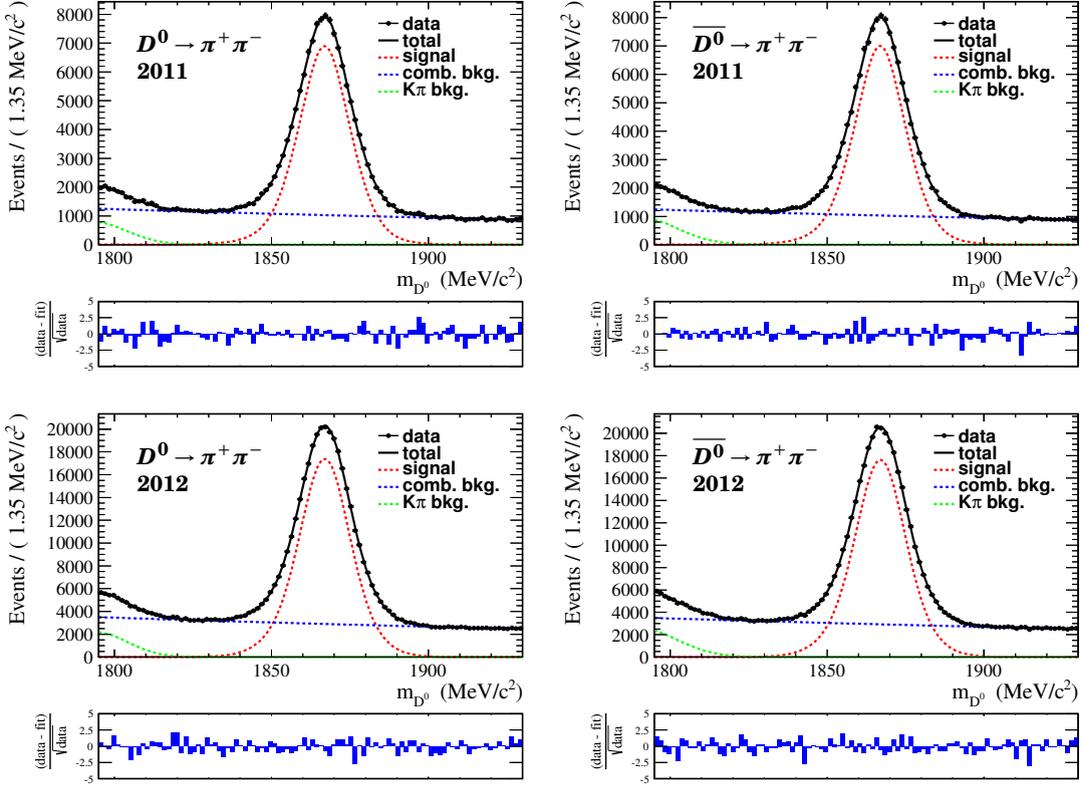
**Figure 6.2:** *Full data sets after the final selection, 2011 data (on top) and 2012 data (on bottom). Invariant $D^0$ mass is fitted for $D^0 \to \pi^- \pi^+$, left for $\mu^-(D^0)$-tag and the right for $\mu^+(\overline{D}^0)$-tag. The pulls are shown below.*

## 6.3   Raw-asymmetry fit in bins of time

To extract the asymmetry $A_{CP}^{RAW}(t_i)$ and their correct uncertainties, an extended unbinned fit is performed for every $t_i$ bin. Thereby, $A_{CP}^{RAW}(t_i)$ is used as a parameter which comes out as a result of a simultaneous fit of the $\mu^-(D^0)$ and $\mu^+(\overline{D}^0)$ tagged sample. Since the amount of background and its composition can change in time, the combinatorial backgrounds parameters used in the fit are allowed to float individually for each decay time bin (e.g. in $D^0 \to K^- K^+$ there can be a background from $D^+ \to K^+ \pi^- \pi^+$ where we miss a $\pi^+$ that can change the background slightly in time). Additionally, the mean $\mu_1$ and the widths $\sigma_1$ and $s_{\sigma_2}$ of the signal shape are allowed to float separately during the fit for every decay time bin. We expect small changes due to resolution effects (e.g. introduced through multiple scattering) that can slightly change the mean and width of events closer to the $B$ vertex and therefore change the weight and the mean for small decay times. Letting $\mu_1$, $\sigma_1$ and $s_{\sigma_2}$ float in the fit covers also these effects. The full parametrization can be then written in the following way:

$$\vec{\lambda}(t_i) = \{A_{CP}^{RAW}(t_i), \sigma_1(t_i), s_{\sigma_2}(t_i), \mu_1(t_i), \vec{\lambda_{bkg.}^+}(t_i), \vec{\lambda_{bkg.}^-}(t_i), \vec{\lambda'_{global}}\}. \tag{6.7}$$

where $\vec{\lambda'_{global}}$ is the subset of the global parameters of $\vec{\lambda_{global}}$ which is not varied individually for each time bin.

For the $D^0 \to \pi^- \pi^+$ channel additionally the fraction of reflection background $f_{refl.bkg.}^+(t_i)$ and $f_{refl.bkg.}^-(t_i)$ vary for each decay time bin. The time dependence of all non-global parameters is
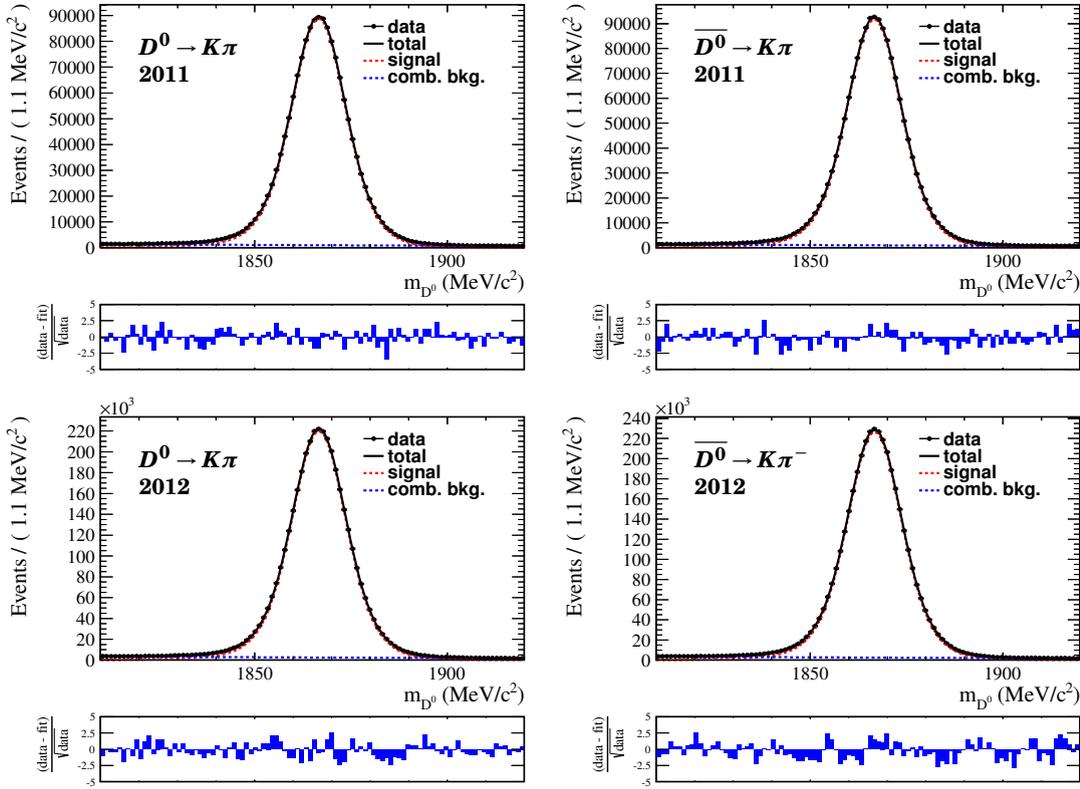
**Figure 6.3:** *Full data sets after the final selection, 2011 data (on top) and 2012 data (on bottom). Invariant $D^0$ mass is fitted for $D^0 \to K^-\pi^+$, left for $\mu^-(D^0)$-tag and the right for $\mu^+(\overline{D}^0)$-tag. The pulls are shown below.*

studied in Sec. 7.

## 6.4 Extraction of $A_\Gamma$

The time-dependent raw asymmetry can be written in the following way and as described in Chapter 3, it can be approximated by (see Appendix B.6 )

$$A_{CP}^{RAW}(t) \approx A_{CP}(t) + A_\mu + A_{prod} \tag{6.8}$$

$$\approx A_{CP}^{dir} + A_\mu + A_{prod} - A_\Gamma \frac{t}{\tau}. \tag{6.9}$$

The time dependence of the asymmetries and mistags as well as the damping factor from mistag probability is considered in Chapter 8. Therefore, the following linear function with two free parameters is fitted to the dataset,

$$A_{CP}^{RAW}(t|A_{const}^{RAW}, A_\Gamma) \approx A_{const}^{RAW} - A_\Gamma \frac{t}{\tau}. \tag{6.10}$$

The $D^0$ lifetime is fixed to $\tau = 410.1 \pm 1.5\,\text{fs}$ [54] and has a negligible effect on the determination of $A_\Gamma$, see Chapter 8.

The binning scheme is chosen to be fine enough not to loose in sensitivity. The bins at large decay times have larger significance on $A_\Gamma$ compared to the low decay-time bins. This is balanced
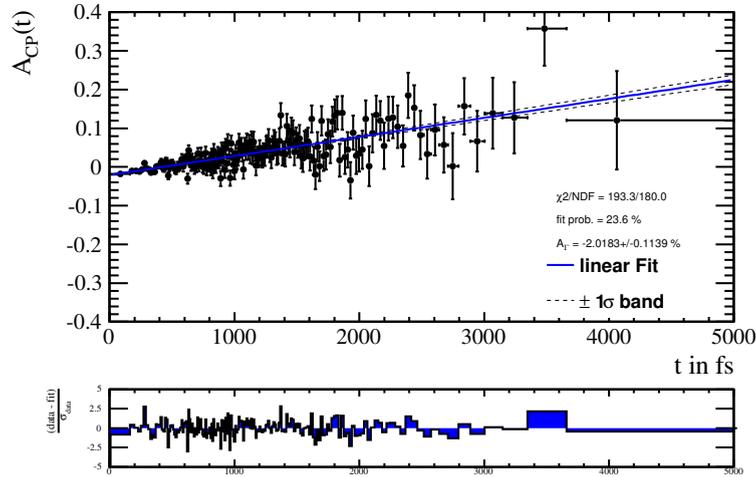
**Figure 6.4:** *Full method using raw asymmetry fits in bins of decay time applied on a toy MC sample (generated with $A_\Gamma = -2\%$, see Appendix B.6 for details).*

through larger statistics at low decay times. Therefore, the scheme is determined in a way, that every decay-time bin has roughly the same sensitivity, see Appendix B.12.

Due to the exponential decay time distribution the mean of the decay-time bins $< t_{bin_i} >$ are closer to the lower decay time bound. A fit to the combined $D^0$ and $\overline{D}^0$ samples is performed in each decay time bin and the mean determined using the sWeights procedure [75].

$$< t_{bin_i} > = \frac{\sum\limits_{t \in bin_i} w_k \cdot t_k}{\sum\limits_{t \in bin_i} w_k} \tag{6.11}$$

Finally, after obtaining the average decay time of the bins, a $\chi^2$ fit of the raw asymmetry (see Eq. 3.10) is performed and $A_\Gamma$ is extracted[1].

This method is robust and sensitive to $A_\Gamma$. For clarification of the presented method, Fig. 6.4 shows the results of the full method chain applied on a toy MC (generated $A_\Gamma$ =-2%). The validation of the method is documented in Chapter 8 and the results are discussed in the next Chapter 7.

---

[1]Fitting a linear function directly rather than using a binned approach has negligible effect on the determination of $A_\Gamma$ to the level of precision we want to achieve (see calculation in Appendix B.6)

<div align="right">

# 7

</div>

# Determination of indirect CP asymmetry

This chapter presents the fit results determined on the 2011 and 2012 data samples. The analysis is done blinded, see Ref. [76] for general blinding technique . Up to the final analysis validation, the sign and the extracted value of $A_\Gamma$ were hidden to the author, see Appendix A.1 for details about the blinding procedure. After the approval, the following unblinded results are obtained and presented in this thesis.

## 7.1   Fit of the $D^0/\overline{D}^0$ mass distribution in bins of decay time

The fits to the mass distribution are used for the signal yields extraction. They are performed simultaneously to $D^0$ and $\overline{D}^0$ mass distributions in 50 decay time bins. [1] The used model, its parametrization and further details are described in previous Chapter 6. Each result of the minimization is validated and presented in this section.

Looking into all time bins, channels and datasets ($50 \cdot 3 \cdot 2 = 300$ bins) the chosen model describes the data reasonably well. Figs. 7.1- 7.4 show the projections of the fits for five decay time bins for $D^0$ and $\overline{D}^0$ mass distributions separately for $D^0 \to K^- K^+$ and $D^0 \to \pi^- \pi^+$ signal candidates in 2011 and 2012 data. All the other mass fits show similar good description of the data.

Further, the time dependence of the determined parameterization is studied. The mean value of the first two Gaussian distributions ($\mu_{1,2}$), the average mass resolution $< \sigma_{M_{D^0}} >$, the average of the combinatorial background parameter for $D^0$ and $\overline{D}^0$ decays $\lambda = \frac{\alpha^+ + \alpha^-}{2}$ and the ratio of signal to background candidates $N_{sig}/N_{bkg}$ together with other parameters are studied. Additionaly, this results are also verified with simulation. The used method can determine the expected decay time dependencies fit parameters, which e.g. can not be handled by the sPlot technique [75]. In Appendix A.5 the evolution of several parameters determined on data is vizualized. Further, the bin-to-bin fluctuation is rather small, which confirms a rather stable fit model.

The $\chi^2$ of the fit to each mass bin is calculated and shown in Fig. 7.5 and 7.6. The number of $D^0$ and $\overline{D}^0$ mass bins is 50. Therefore, the total number degrees of freedom is $ndf = 2 \cdot 50 - ndf_{model}$. The $\chi^2/ndf$ values are fluctuating around 1. The p-values, $Prob(\chi^2, ndf)$, of the mass fits have a flat distribution for both the $KK$ and $\pi\pi$ channels. This confirms that in all decay time bins the model works well. The $D^0 \to K^- \pi^+$ channel has a more than 5 times larger statistics, therefore as expected small differences between the data distribution and the fit function become

---

[1]For the choice of the bins see Appendix B.12.

visible. Looking at the $\chi^2/ndf$ time dependence it increases for the lower decay time bins, which have higher statistics. It however stays below a $\chi^2/ndf$ of two. The corresponding probability $Prob(\chi^2, ndf)$ distribution shows a slightly larger fraction of fits with lower probability (Fig. 7.7). This has, however, a negligible impact on the asymmetry as it affects the extraction of the $D^0$ and $\overline{D}^0$ yields in the same way.

## 7.2   Raw asymmetry fit results and statistical significance

Simultaneous fits to the $D^0$ and $\overline{D}^0$ mass distributions as described in section 6 are used to measure the yields in each decay time bin. A $\chi^2$ fit is performed fitting a linear function to the decay time dependence of the raw asymmetry $A_{CP}^{RAW}(t)$. The outcome of these fits is shown for the $D^0 \rightarrow K^- K^+$ and $D^0 \rightarrow \pi^- \pi^+$ channel in Fig. 7.8 and Fig. 7.9. The fitted function describes the time dependence very well in both channels and in both data sets ($\chi^2/ndf \approx 1$). A cross-check on the $D^0 \rightarrow K^- \pi^+$ sample is additionally performed. Fig. 7.10 shows the result of this fit to the raw asymmetry. Here, $A_{CP}^{RAW}$ is expected to be time independent. The effect of different momentum distributions of the $D^0$ mesons and muons, that can cause a change of the detection asymmetries in time, is negligible for our current sensitivity (see section 8.2). One can see that the data are well described by the fitted function. The result is in good agreement with the expected flat distribution and the $A_{CP}$ offset is consistent with previous measurements.

Beside the $\chi^2/ndf$ and its fit probability $Prob_{\chi^2}(\chi^2, ndf)$ also the Kolmogorov-Smirnov probability $Prob_{K-S}$, which is more sensitive to the overall agreement of the shape of the data points and the fitted function is determined (see Appendix B.1). The results of all fits are summarized in Tab. 7.1. The model used for the extraction of $A_\Gamma$ describes well the data in both channels and

**Table 7.1:** *The measured $A_\Gamma$ values (Unblinded), the offset of the raw asymmetries and the goodness of the fit values for the 2011 and 2012 data sets.*

| 2011 data | $A_\Gamma[\%]$ | $A_{CP}^{RAW}(0)[\%]$ | $\chi^2/NDF$ | $Prob_{\chi^2}$ | $Prob_{K-S}$ |
|---|---|---|---|---|---|
| $D^0 \rightarrow K^- K^+$ | **-0.069±0.145** | -0.443±0.213 | 35.8/48 | 0.904 | 0.687 |
| $D^0 \rightarrow \pi^- \pi^+$ | **-0.346±0.273** | -1.117±0.399 | 41.8/48 | 0.724 | 0.877 |
| $D^0 \rightarrow K^- \pi^+$(check) | **-0.052±0.060** | -1.650±0.089 | 70.6/48 | 0.018 | 0.687 |
| 2012 data | | | | | |
| $D^0 \rightarrow K^- K^+$ | **-0.160±0.091** | -0.586±0.136 | 45.8/48 | 0.565 | 0.501 |
| $D^0 \rightarrow \pi^- \pi^+$ | **0.008±0.172** | -0.613±0.254 | 44.5/48 | 0.615 | 0.673 |
| $D^0 \rightarrow K^- \pi^+$(check) | **0.033±0.038** | -1.448±0.056 | 45.7/48 | 0.566 | 0.951 |

data sets. Combining the data sets, the results obtained in the three decay channels are

$$A_\Gamma(K^- K^+) = (-0.134 \pm 0.077)\% \,,$$
$$A_\Gamma(\pi^- \pi^+) = (-0.092 \pm 0.145)\% \,,$$
$$A_\Gamma(K^- \pi^+) = (\ \ 0.009 \pm 0.032)\% \,.$$

The final results are presented in Chapter 9.
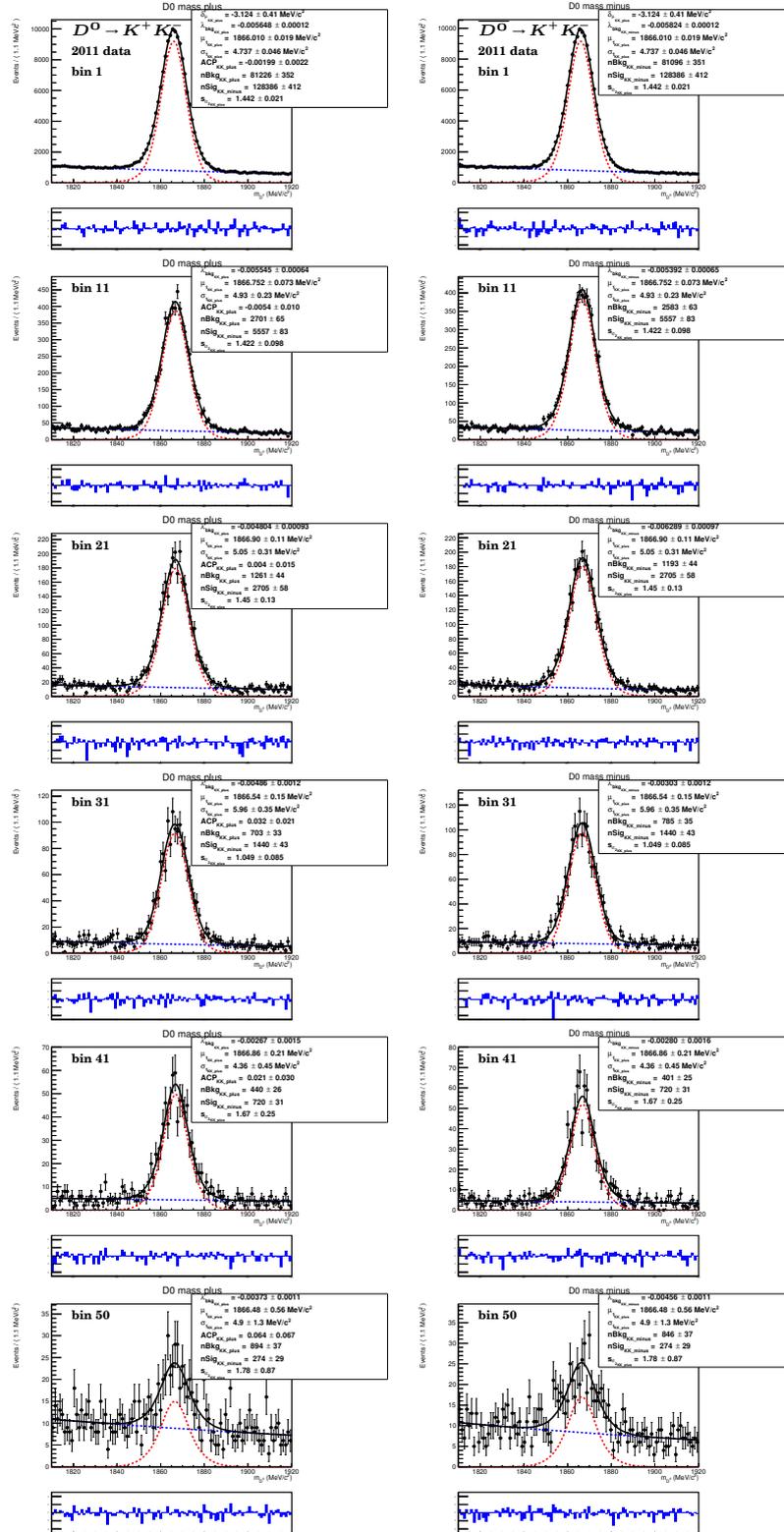
**Figure 7.1:** *Examples of mass fit projections for $D^0$ (left) and $\overline{D}^0$ (right) signal candidates for the decay $D^0 \to KK$ for 2011 data.*
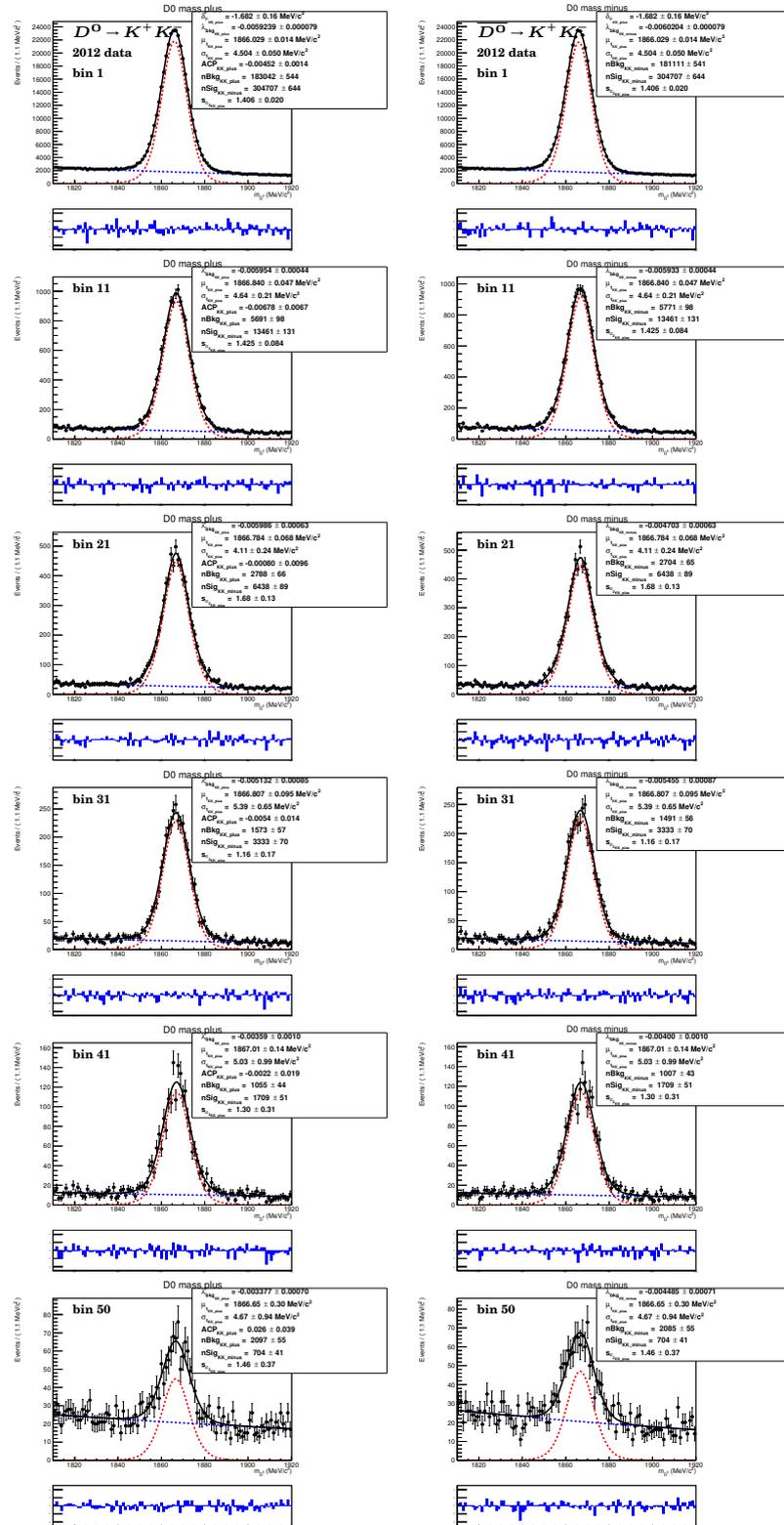
**Figure 7.2:** *Examples of mass fit projections for $D^0$ (left) and $\overline{D}^0$ (right) signal candidates for the decay $D^0 \to KK$ for 2012 data.*
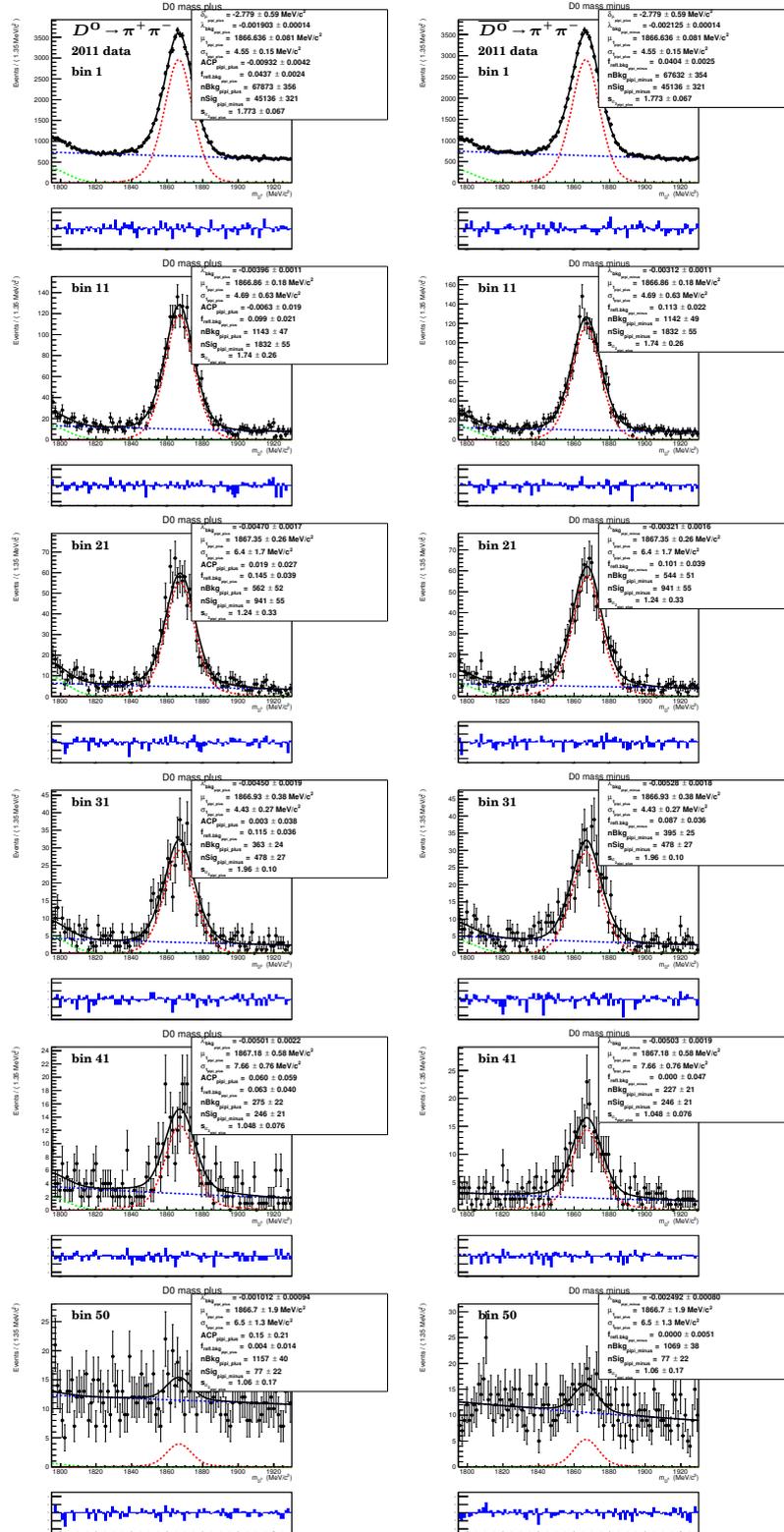
**Figure 7.3:** *Examples of mass fit projections for $D^0$ (left) and $\overline{D}^0$ (right) signal candidates for the decay $D^0 \to \pi\pi$ for 2011 data.*
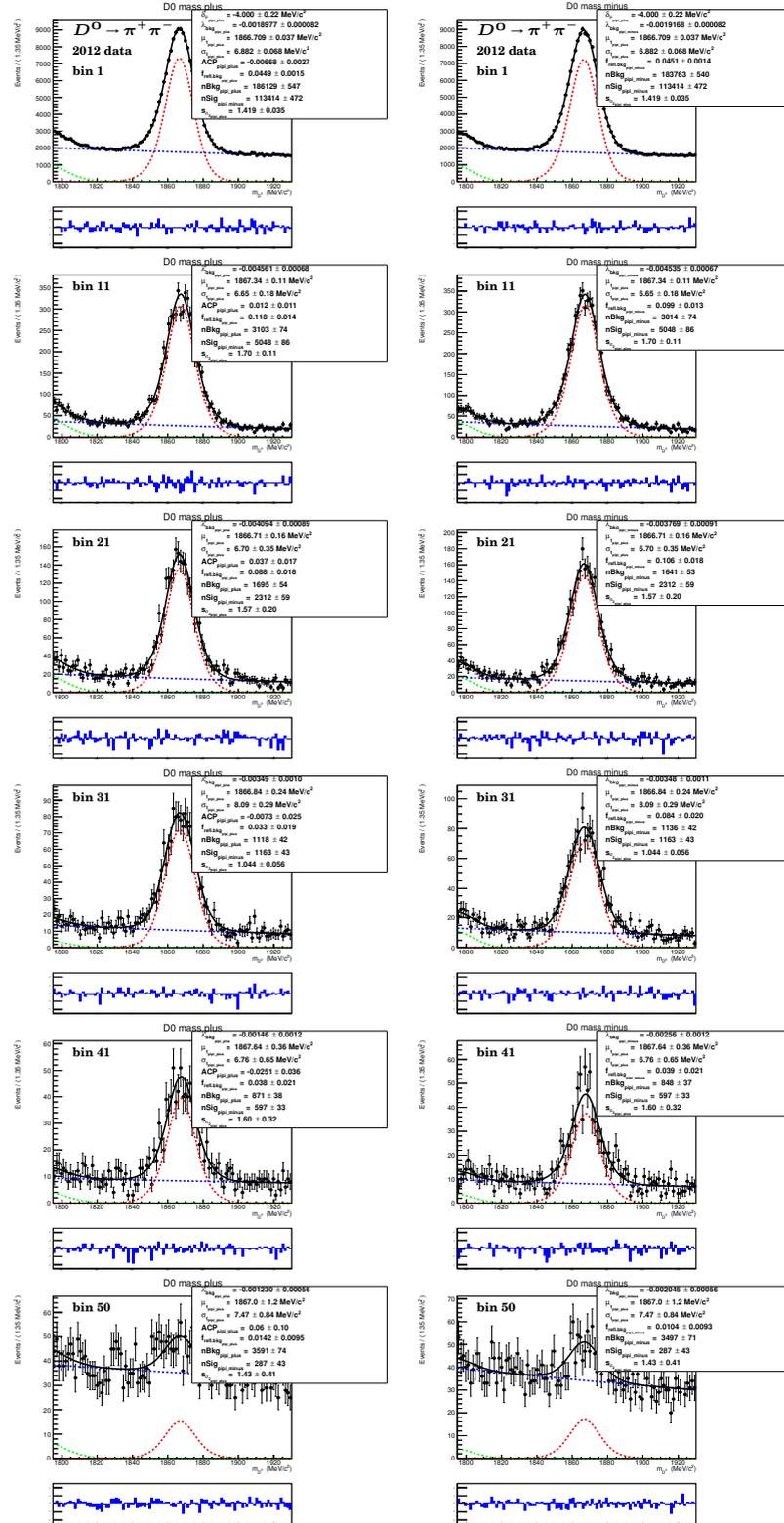
**Figure 7.4:** *Examples of mass fit projections for $D^0$ (left) and $\overline{D}^0$ (right) signal candidates for the decay $D^0 \to \pi\pi$ for 2012 data.*
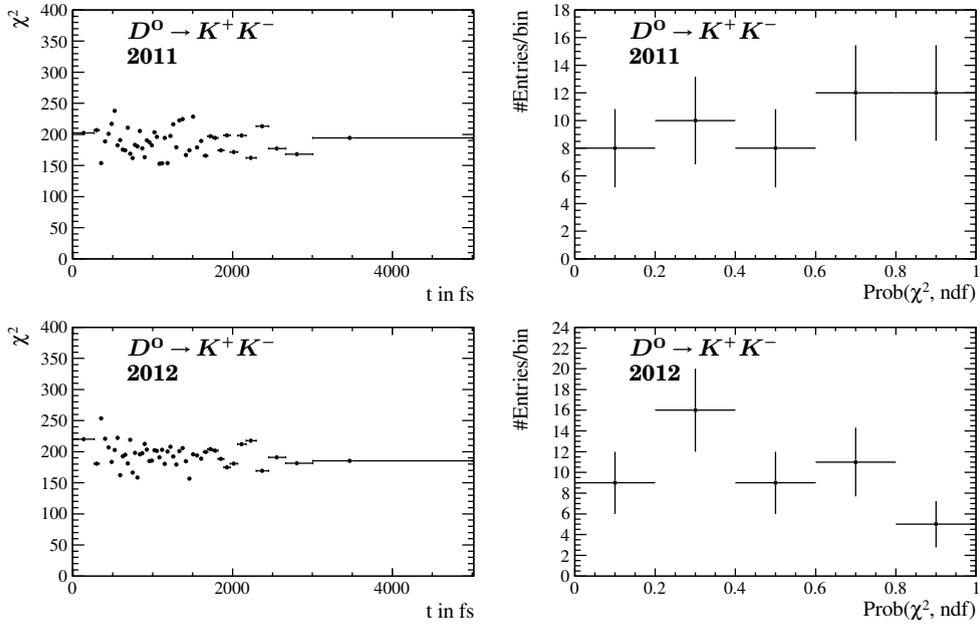
**Figure 7.5:** $\chi^2$ *of the* $D^0$ *and* $\bar{D}^0$ *mass fits vs. decay time (left) and their probability distribution (right) for 2011 (top) and 2012 (bottom) data in the* $D^0 \to KK$ *channel.*
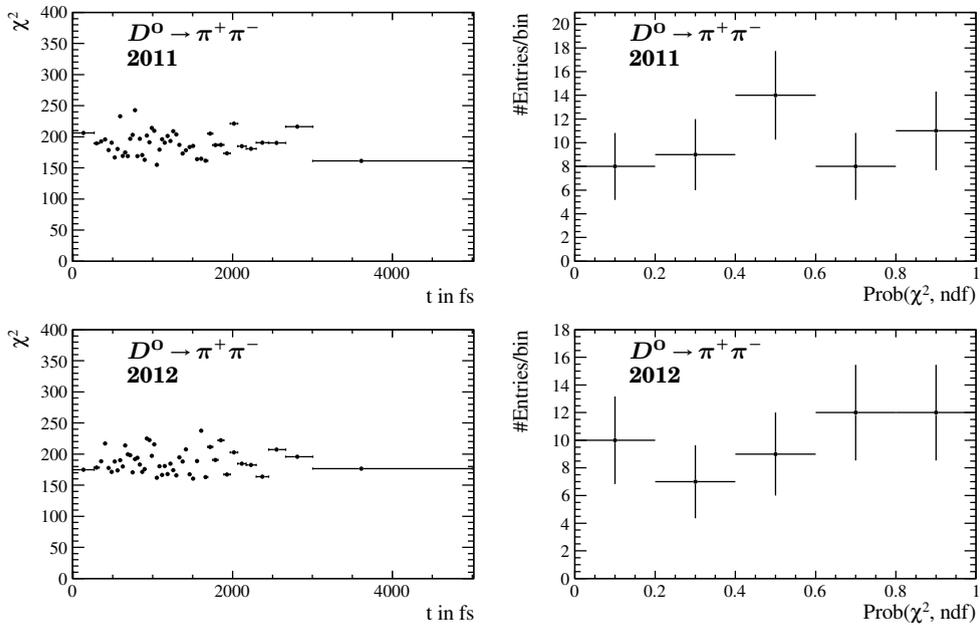


**Figure 7.6:** *The* $\chi^2$ *value of the* $D^0$ *and* $\bar{D}^0$ *mass fits vs. decay time (left) and their probability distribution (right) for 2011 (top) and 2012 (bottom) data in the* $D^0 \to \pi\pi$ *channel.*
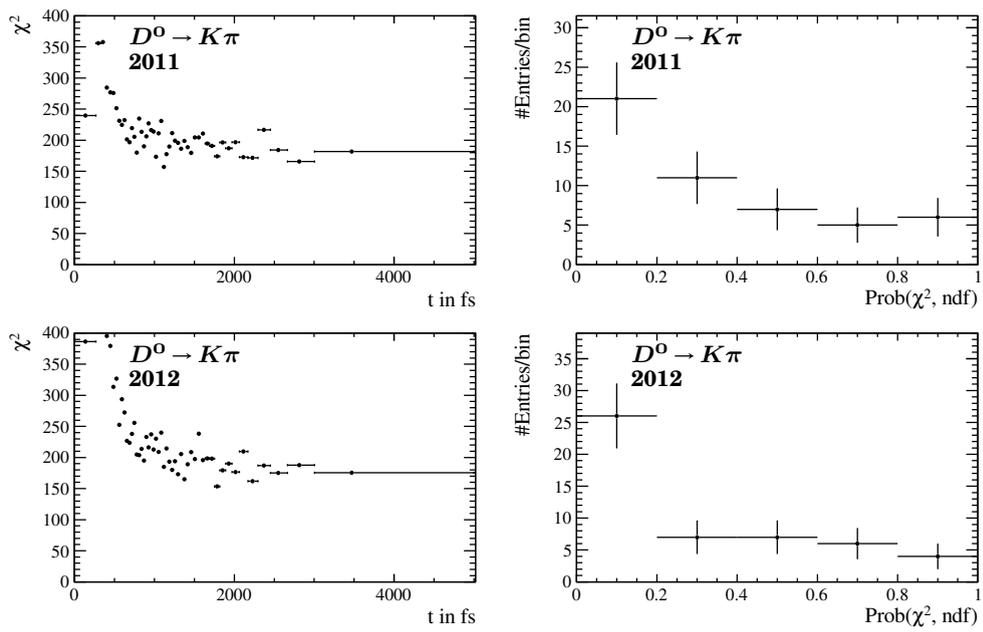
**Figure 7.7:** *The $\chi^2$ value of the $D^0$ and $\overline{D}^0$ mass fits as function of decay time (left) and their probability distribution (right) for 2011 (top) and 2012 (bottom) data in the $D^0 \to K\pi$ channel.*
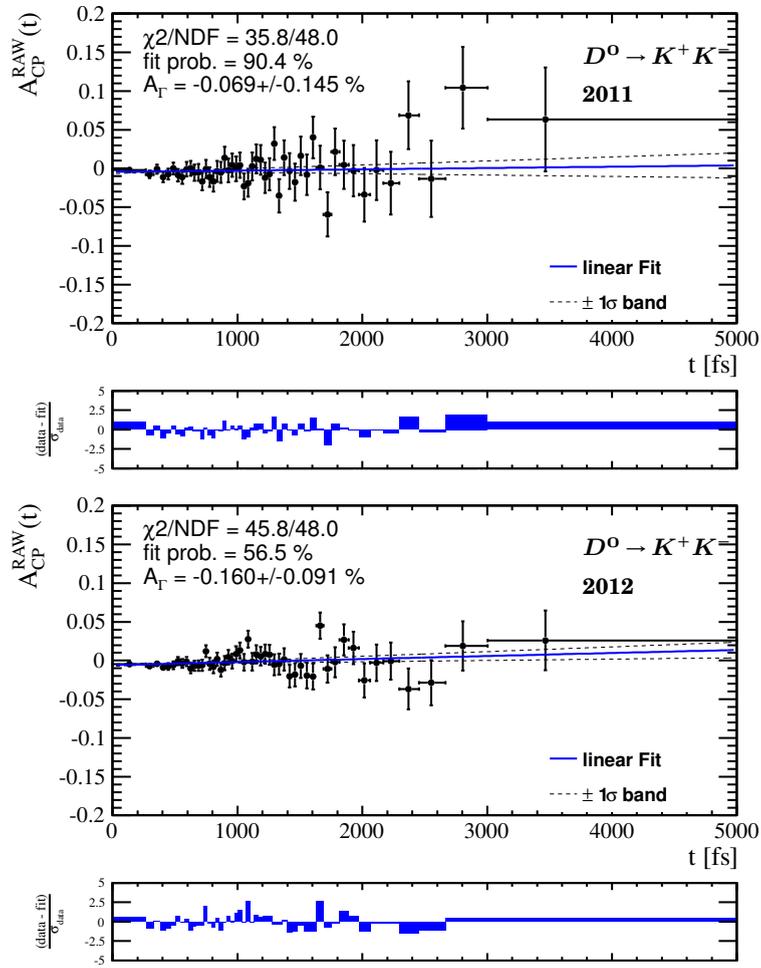
**Figure 7.8:** *Fit to the asymmetry distribution of the $D^0 \to K^- K^+$ signal candidates in 2011 (top) and 2012 (bottom) data.*
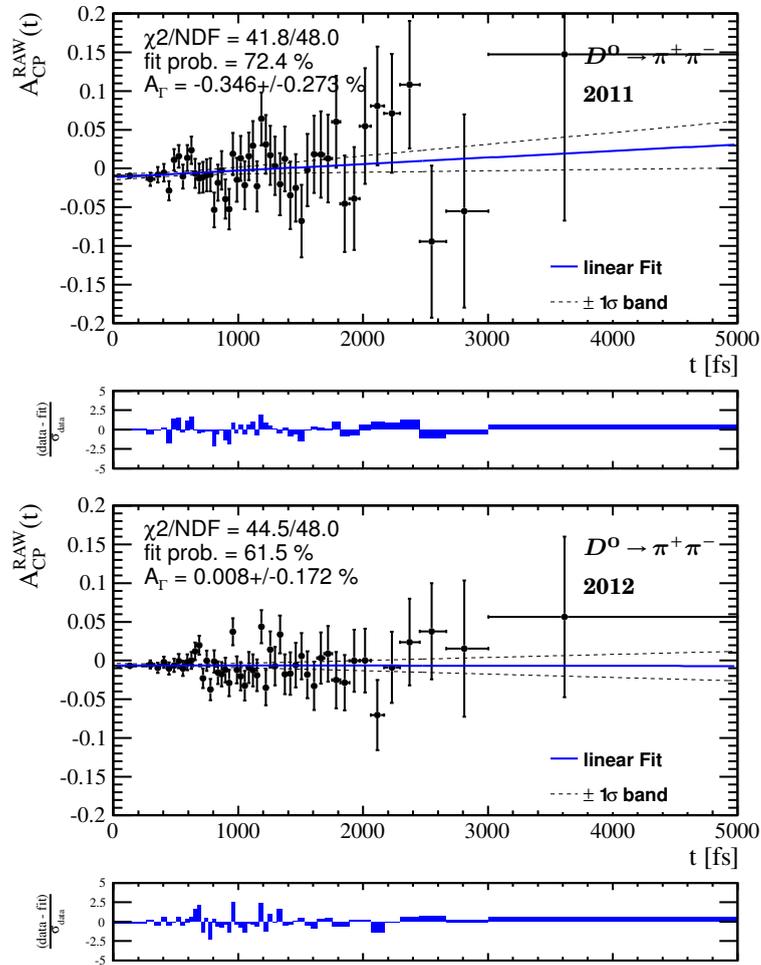
**Figure 7.9:** *Fit to the asymmetry distribution of the $D^0 \to \pi^-\pi^+$ signal candidates in 2011 (top) and 2012 (bottom) data.*
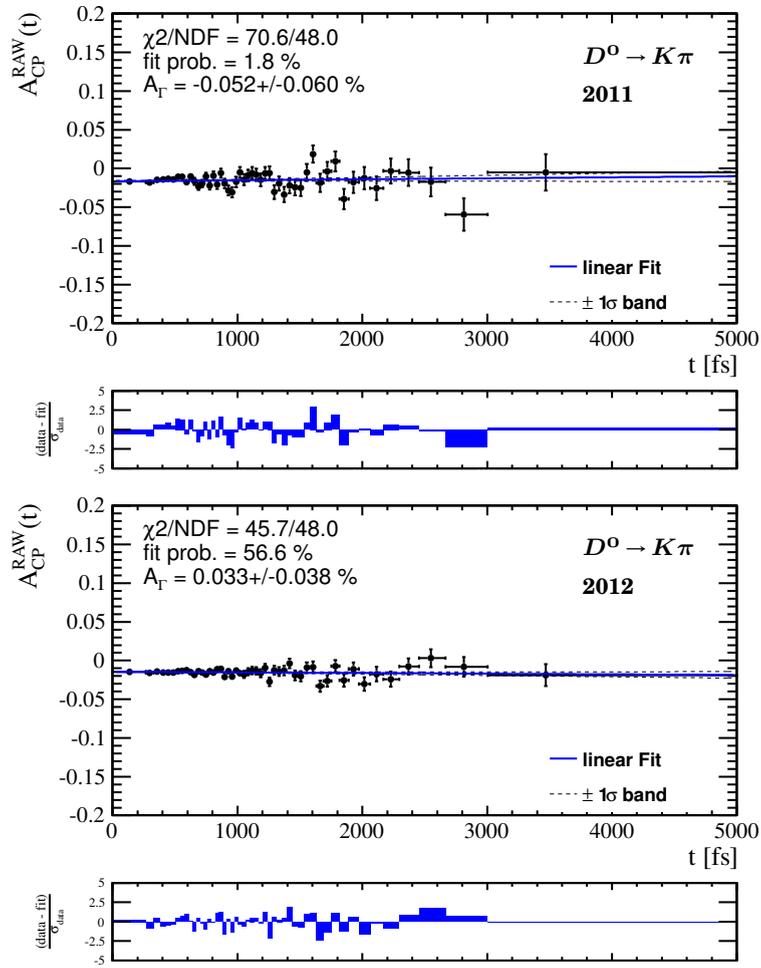
**Figure 7.10:** *Fit to the asymmetry distribution of $D^0 \to K^- \pi^+$ signal candidates in 2011 (top) and 2012 (bottom) data.*

# 8

# Systematic uncertainties

This chapter concentrates on a detailed description of the systematic effects that can influence this measurement. LHCb data sets, full MC simulation and toy MC simulation are used to quantify and visualize the possible biases on $A_\Gamma$. Furthermore, several cross checks are performed to confirm the stability and robustness of this analysis. The current analysis significance is dominated by the statistical uncertainty as will be shown in this chapter. A summary of the systematic uncertainties is given in Section 8.13.

## 8.1 Fitter validation

For the verification of the full fitting procedure with simultaneous fits of the $D^0$ and $\bar{D}^0$ yields in different time bins and the final $\chi^2$ fit of the slope, a toy model was introduced. Modeling the $D^0$ and $\bar{D}^0$ decay time distributions as a function of $A_\Gamma$ requires the use of the average $D^0$ lifetime ($<\tau>= 410.1 \pm 1.5\,\mathrm{fs}$ [54]). Additionally, acceptance functions and background distributions as found in data are used for this simulation, see Chapter. 5. The following studies are based on 1000 toys with 4M events each. Different binning in decay time have been investigated. Figure 8.1 shows the extracted $A_\Gamma$ and the pulls using the default fitting procedure with a binning scheme of 50 bins. The bins are chosen to have roughly the same statistical impact on $A_\Gamma$. This procedure is further described in Appendix B.12. The fitted $A_\Gamma$ value is found to be in agreement with the generated value of $A_\Gamma = -0.3\%$ and also the mean of the pulls is consistent with zero and their width with one. The default method using the asymmetry shows no bias and is independent from the binning. Also alternative methods have been studied in Appendix A.2. Choosing coarser binning, however, results in less precision. Choosing finer binning on the contrary results in less stable fits due the lower statistics in each decay time bin. As a compromise a binning scheme of 50 bins is chosen for this analysis. The obtained resolution is within statistical uncertainties comparable to the finer binning with 180 bins.

To give an upper limit of any potential bias on $A_\Gamma$ from the fit method a toy model with $A_\Gamma = 0$ is generated. Thereby, the full decay-time resolution is taken into account and used in the generation. Details are described in following Sec. 8.3, where the impact of decay-time resolution is studied. This value of $A_\Gamma$ is chosen to separate the bias from the fit method from a resolution-induced bias, which is proportional to $A_\Gamma$. The fitter can extract again the $A_\Gamma$ value without a bias at a 0.002% sensitivity level, as shown in Fig. 8.2. Thus, an uncertainty of 0.002% is assigned for the fitting method.
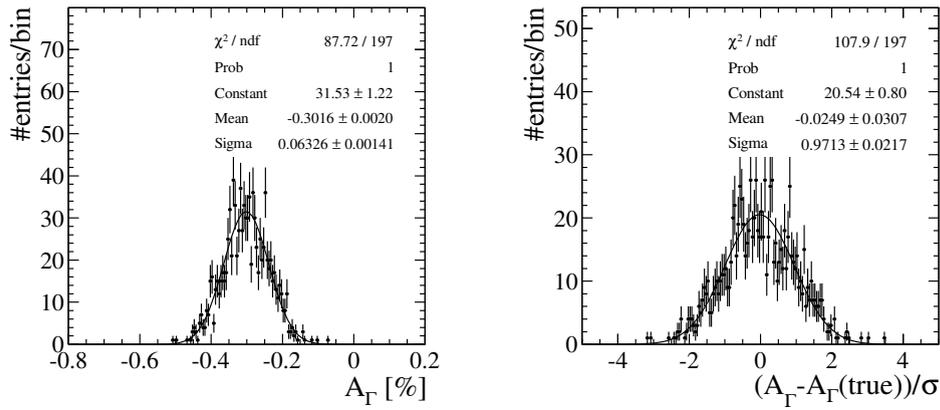
**Figure 8.1:** *Measured $A_\Gamma$ distribution and residuals. An input value of $A_\Gamma = -0.3\%$ is used in the generation.*
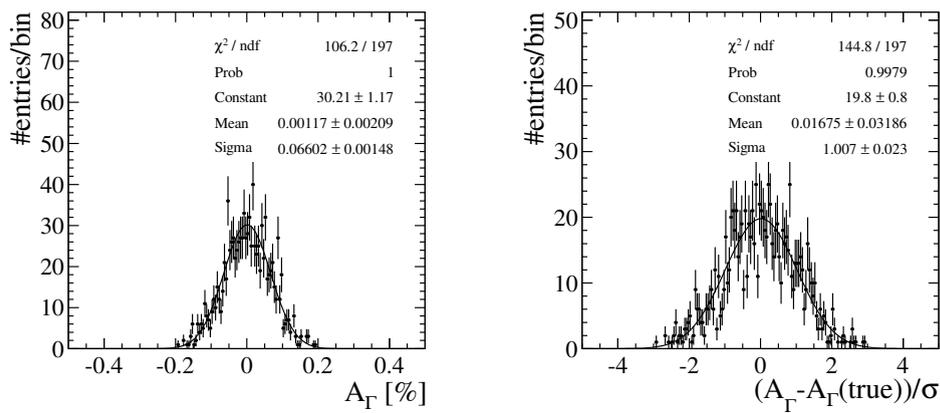


**Figure 8.2:** *Toy study for a fit with 50 bins with a generated input value of $A_\Gamma = 0$. Measured $A_\Gamma$ distribution (left) and residuals (right) are in good agreement with the generated value.*

## 8.2   Detection and production asymmetries and kinematics

Under the assumption that $A_\mu$ and $A_{prod}$ are time independent we derived the following relation of the raw asymmetry and the physics asymmetry (Eq. 6.9)

$$A_{CP}^{RAW}(t) \approx A_{CP}(t) + A_\mu + A_{prod}.$$

However, a time dependence can be introduced, if there is a dependence of the asymmetry related to the kinematics. Different momenta spectra in different decay-time bins can cause a decay-time dependent asymmetry. To study this effects, the $D^0$ and $\mu$ transverse momentum spectra are plotted in 6 different decay time bins (Fig. 8.3, Fig. 8.4):

$$
\begin{aligned}
bin_0 && &\in [0, 530) \text{ in fs} && (8.1) \\
bin_1 && &\in [530, 832) \text{ in fs} \\
bin_2 && &\in [832, 1132) \text{ in fs} \\
bin_3 && &\in [1132, 1508) \text{ in fs} \\
bin_4 && &\in [1508, 1773) \text{ in fs} \\
bin_5 && &\in [1773, 5000) \text{ in fs}
\end{aligned}
$$

The distributions are rather similar, however slightly lower transverse momenta for the $D^0$ and $\mu$ are observed for higher decay time bins. This could be caused by effect from the trigger, reconstruction or selection. To evaluate the impact of this difference we reweight the events in each decay time bin $i$ to match one reference bin.

$$w_{\text{bin}=i}(p_t, t) = \frac{P_{\text{ref bin}}(p_t)}{P_{\text{bin}=i}(p_t)} \qquad (8.2)$$

These weights are directly calculated from the histograms in Fig. 8.4 and Fig. 8.3. Applying these weights to the data and repeating the fit for $A_\Gamma$ on the reweighted samples results in a shift $\Delta A_\Gamma$:

$$\Delta A_\Gamma = |A_\Gamma^{\text{reweight}} - A_\Gamma^{\text{default}}| \qquad (8.3)$$

These shifts are determined using in turn as reference the distribution of $D^0$ and $\mu$ transverse momentum in each of the 6 decay time bins listed above. Beside the transverse momentum of the $D^0$-meson and the muon, further distributions (e.g. momentum, $\phi$) for reweighting have been studied. However, negligible effects have been observed for them. The largest difference observed among all $D^0$ and $\mu$ transverse momentum reweightings are

$$
\begin{aligned}
D^0 \to K^- K^+ : \ & \Delta A_\Gamma(\text{kinematics}) && \le 0.0174\% && (8.4) \\
D^0 \to \pi^- \pi^+ : \ & \Delta A_\Gamma(\text{kinematics}) && \le 0.0400\% \\
D^0 \to K^- \pi^+ : \ & \Delta A_\Gamma(\text{kinematics}) && \le 0.0098\%
\end{aligned}
$$

All the results of reweighting for the three different decays are listed in Appendix A.4 Tabs. A.1, A.2, A.3 and A.4, A.5, A.6. Since these numbers are statistically dominated, we assign the largest difference from the (higher-statistics) $D^0 \to K^- \pi^+$ mode as the systematic uncertainty. This is justified by the fact that the production asymmetry is basically the same in each mode, and the fact that the detection asymmetry is largest in the $D^0 \to K^- \pi^+$ mode. Further, as will be shown in the Section 8.11 by enlarging the detection asymmetries when using different magnet polarities and also looking at different ranges of B decay times shows consistency between the results.
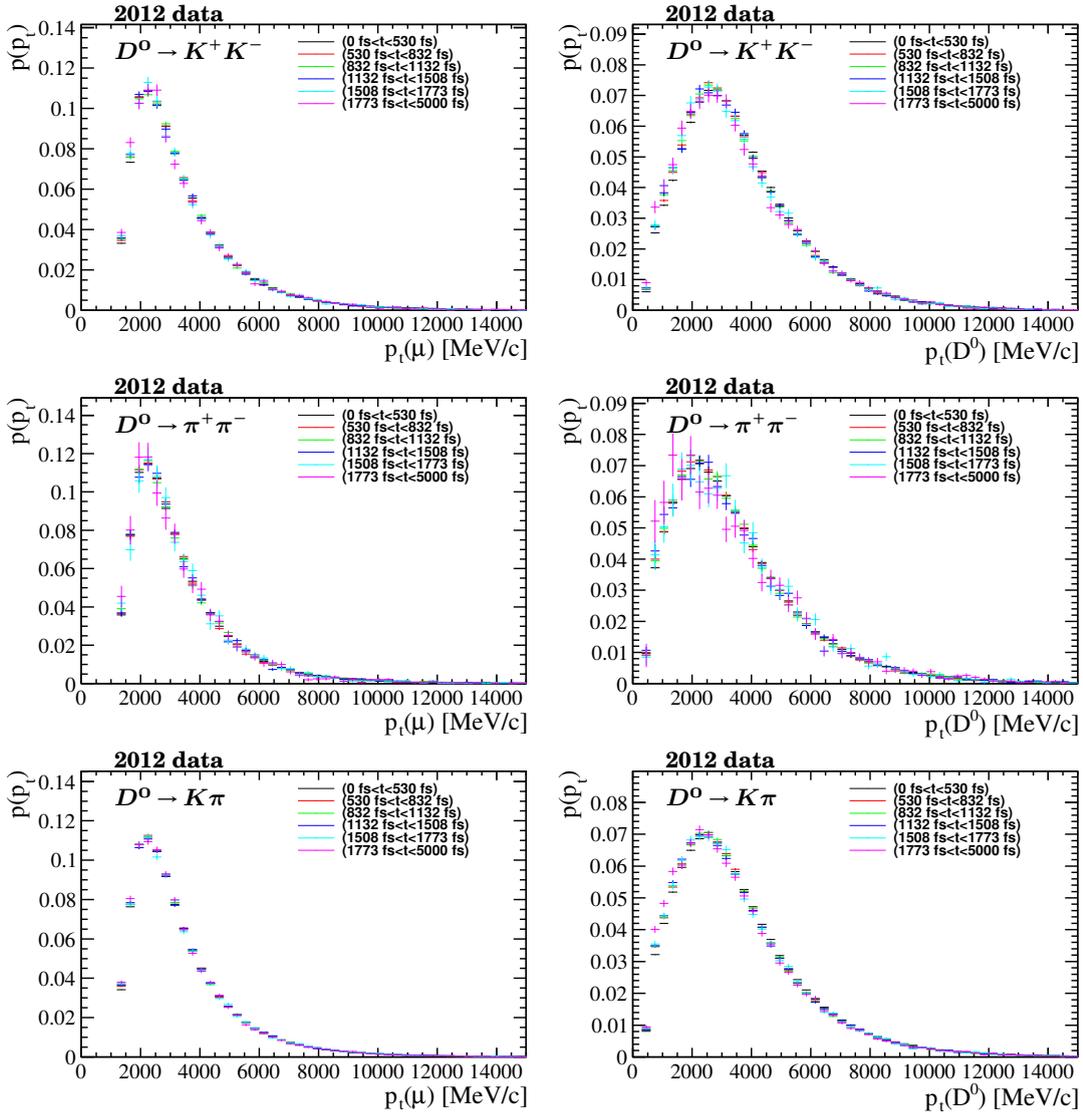
**Figure 8.3:** *Transverse momentum spectra for $D^0$ and $\mu$ in different $D^0$ decay-time bins for 2012 data.*
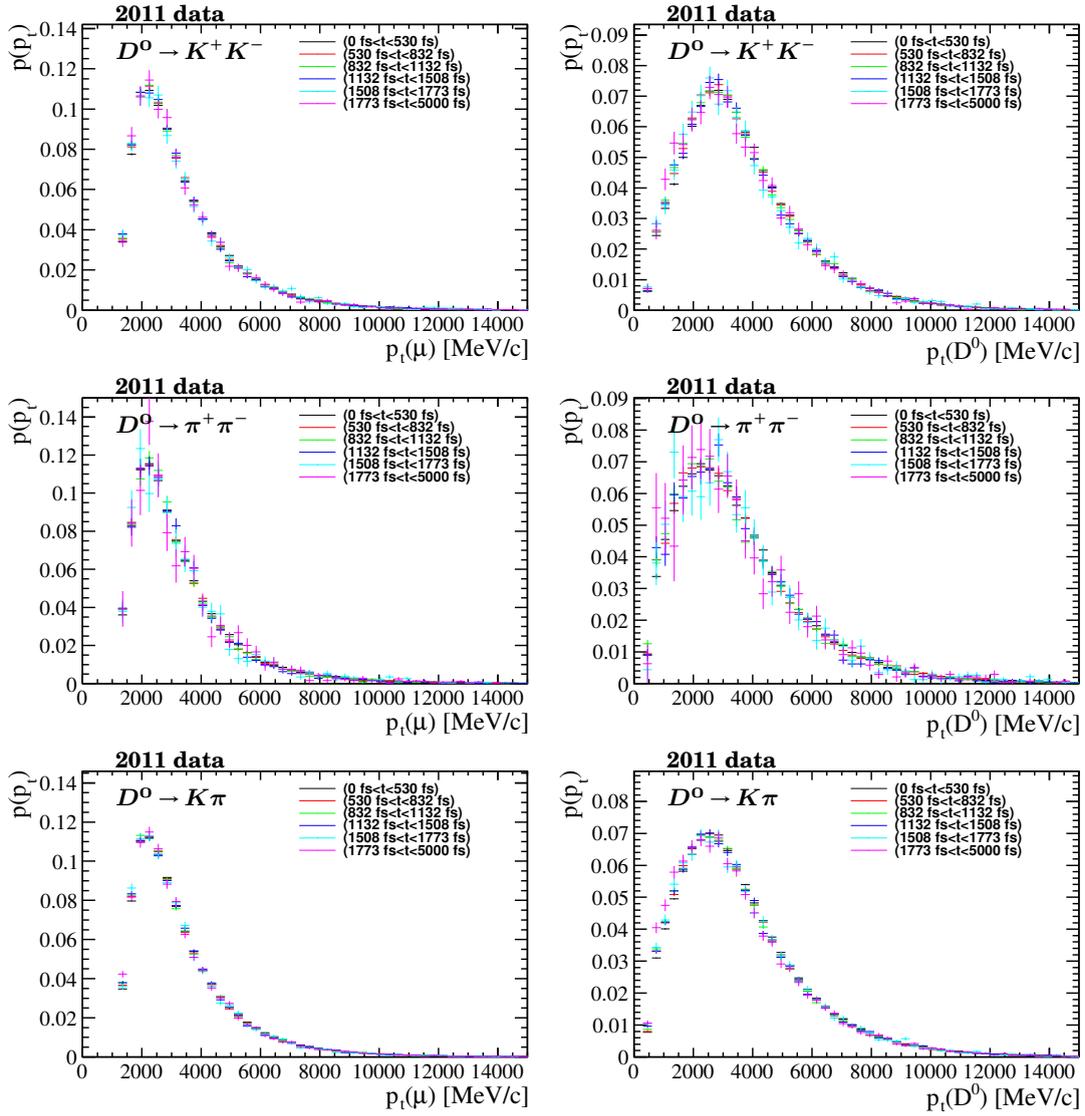
**Figure 8.4:** *Transverse momentum spectra for $D^0$ and $\mu$ in different $D^0$ decay-time bins for 2011 data.*

## 8.3   Decay time resolution bias

The decay time resolution $\sigma_t$ is expected to have a tiny impact on the measurement if it is small enough $\frac{\sigma_t}{\tau(D^0)} << \mathcal{O}(1)$. Although, the resolution is the same for $D^0$ and $\overline{D}^0$ events and it does affect the observed time scale and therefore, the measured $A_\Gamma$ value.

Simulated events with $B^0$ and $B^+$ decays have been studied to determine the parameters of the resolution function. Both the reconstructed $D^0$ and the muon are required to be associated with a true $D^0$ and a true muon. Both particles are however not requested to be the direct decay products of the same $B$ meson. Thus, as well candidates where the $D^0$ comes from one $B$ meson and the muon from the opposite $B$ are considered for the resolution determination. These fake candidates have a wrong flavour tag and are thus as well taken into account in the systematic uncertainties assigned to the mistag rate studied in one of the following sections. Their contribution is in any case found to be small. Additionally, $B \to D^0 \tau X$ decays with subsequent $\tau \to \mu \nu_\tau \overline{\nu_\mu}$ are included. They give the correct tag but the decay time resolution is worse and introduces a bias on the decay time. These type of events are studied separately in Appendix A.13. They have as well only a minor impact on the measured asymmetry.

The decay time resolution model used in here consists of the sum of four Gaussians. The center of the first three is fixed to the same value, while the fourth one is used to describe asymmetric tails, see Appendix B.5. The majority of the events (90%) is described by the first two Gaussians, calculated using their widths and fractions $\sigma_{av.} = \sqrt{\sum_i f_i \sigma_i^2}$, with about $\sigma_{av.} \approx 103\,\text{fs}$ average decay time resolution.

The resolution parameters derived from these Monte Carlo samples are the basis for the subsequent systematic studies. Figs. 8.5 show the obtained parameters for both channels. The average resolution for $D^0 \to K^- K^+$ and $D^0 \to \pi^- \pi^+$ events is very similar. The extracted parameters show that in the $D^0 \to \pi^- \pi^+$ channel the core resolution is smaller but larger tails are included. The extracted resolutions are varied by $\pm 10\%$ to mimic potential data/MC discrepancies.

The impact of the resolution model is studied in different ways. First, the theoretical decay rates are convoluted with the resolution function and the effect on the measured $A_\Gamma$ value is derived mathematically. This method is further confirmed through a toy Monte Carlo studies that are performed, as described in previous Section 8.1 . Finally, also a data-driven approach is used that shows consistent results. These studies are described in the following.

For the extraction of the impact of the resolution on $A_\Gamma$ measurement, the following parameters are used.
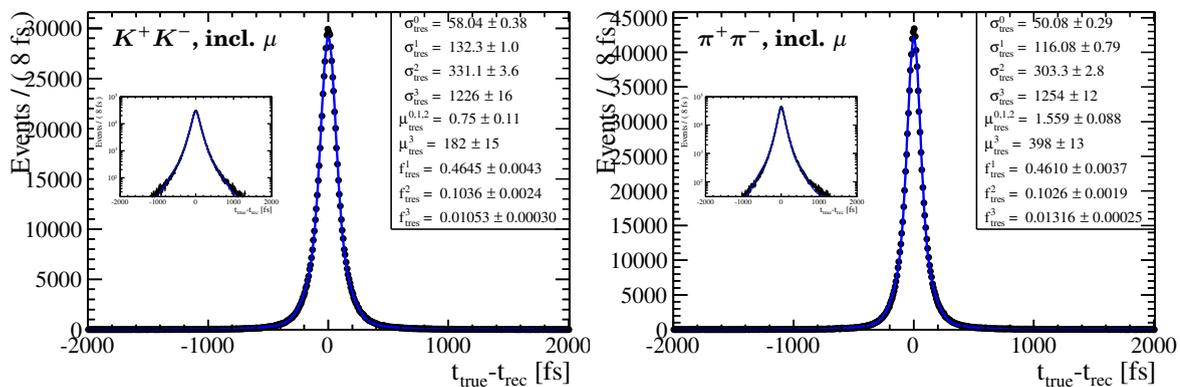


**Figure 8.5:** *Resolution determined on full MC simulation, for $D^0 \to K^- K^+$ decays (left) and $D^0 \to \pi^- \pi^+$ decays (right).*

| $A_\Gamma$ | $A_{CP}^{dir} + A_\mu + A_{prod.}$[1] | $\tau(D^0)$ | $scale_{\sigma_t}$ |
|---|---|---|---|
| $-0.3...0.3\%$ | $-0.05\%$ | $410.1\,\mathrm{fs}$ | $0.9, 1.0, 1.1$ |

In the next step, the theoretical decay rates of $D^0$ and $\overline{D}^0$ are convoluted with the default resolution function. Additionally, the widths of all four Gaussians are scaled by 1.1 and 0.9. Afterwards, the raw asymmetry can be calculated for all three scenarios based on these functions. To mimic additionally the effect from binning in decay time, the theoretical decay rates convoluted with the resolution function are integrated within a decay time bin for the $D^0$ and $\overline{D}^0$ distribution. The asymmetries are thus calculated per bin from these yields. Then, a linear fit is performed to the decay-time dependent asymmetry. This method will in the following be referred to as binned fit.

Figure 8.6 shows these results separately for $D^0 \to K^+K^-$ and $D^0 \to \pi^+\pi^-$ decays with a generated value of $A_\Gamma$ of 0.1%. As both channels have very similar resolution functions, hardly any difference in the size of the bias on $A_\Gamma$ is found.

The core resolution has the main impact on the raw asymmetries at low decay times and the tails of the resolution have the main impact at higher decay times. Due to the underlying exponential distribution, more events are shifted from lower to higher decay times due to resolution effects, as the other way around. The absolute bias

$$\Delta A_\Gamma = A_\Gamma^{estimator} - A_\Gamma^{true} \tag{8.5}$$

depends on the true value of $A_\Gamma$. To calculate this dependence, a binned fit to the asymmetry is performed with different $A_\Gamma$ values as input. The bias is found to be proportional to $A_\Gamma$. The difference in the resolution function for the both channels under investigation turned out to result in a 15% effect on the bias (Fig. 8.7).

Additionally, 1000 toys (with 4M events each) with all parameters as determined from data are generated for different $A_\Gamma$ values. This results can be found in Appendix A.9. Fig. 8.7 the determined values and their uncertainties. The observed biases on $A_\Gamma$ value are consistent and compatible with the calculation of the first method.

---

[1]Thereby the absolute shift of $A_{CP}^{dir} + A_\mu + A_{prod.}$ has negligible impact on the $A_\Gamma$ extraction, since it shifts the asymmetries in all time bins and has no impact on the resolution effect that shifts events between the time bins.
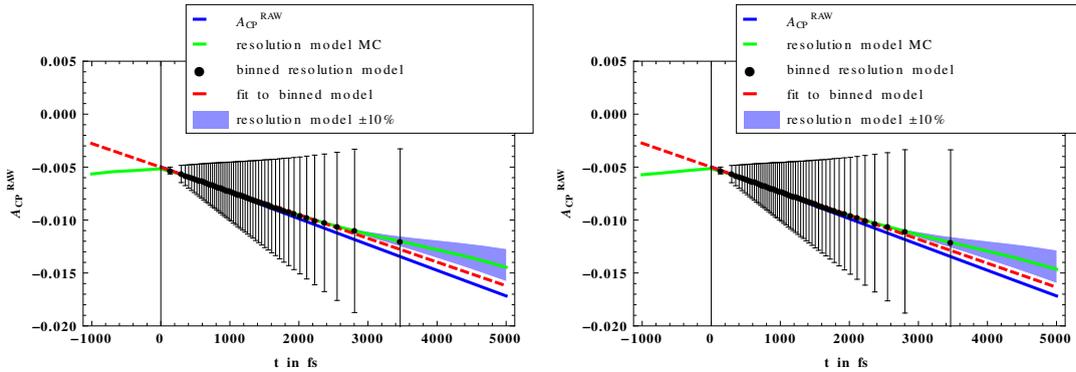


**Figure 8.6:** $A_{CP}^{RAW}$ dependence for different resolution models for $D^0 \to K^-K^+$ (left) and $D^0 \to \pi^-\pi^+$ decays (right). In blue - the initial asymmetry is shown, assuming no decay time resolution effects. In green - the resolution function derived on MC is applied. The blue band denotes the uncertainty changing the resolution by $\pm$ 10%. The black points include additionally the effect of binning in decay time for the default resolution model. In red - the fit to these data points is displayed. A true $A_\Gamma$ value of 0.1% has been chosen for all distribution shown in this plot.
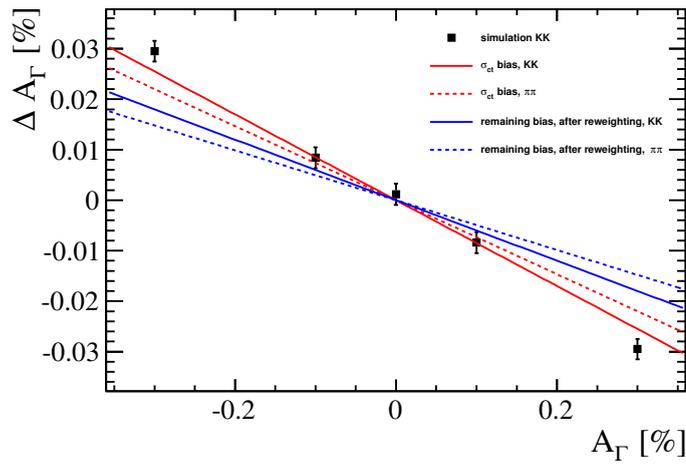
**Figure 8.7:** *The bias on $\Delta A_\Gamma$ due to resolution effects as function of $A_\Gamma$ value. Points are extracted using 1000 toy Monte Carlo runs with 4M events each, using the second method. In red: the bias extracted from a binned fit on the asymmetry function, using the first method (for details see text). In blue: the remaining bias after applying the resolution weights, using the third method. The solid lines show the behavior for $D^0 \to K^- K^+$ and dashed for the $D^0 \to \pi^- \pi^+$ events.*
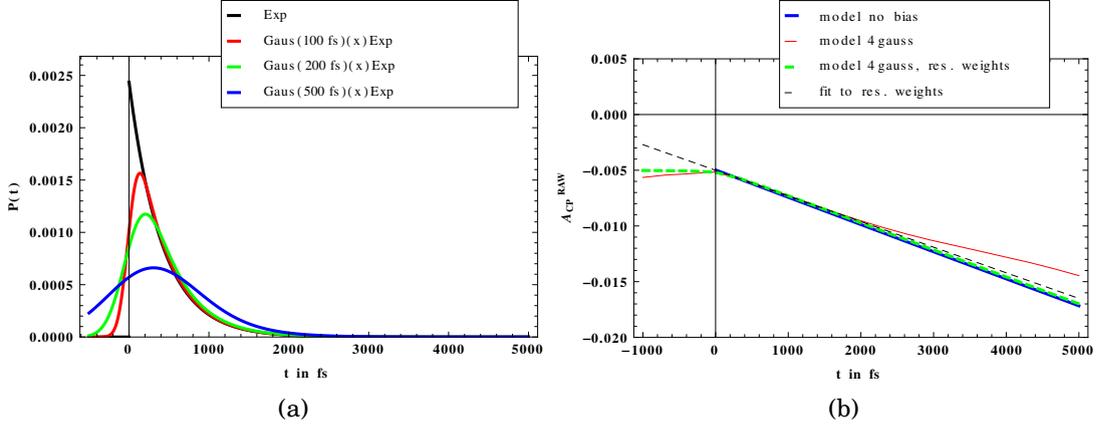
**Figure 8.8:** *(a) Decay time distribution for different resolution: in black exponential, in red $100 fs$ resolu-
tion, in green $200 fs$ resolution and in blue $500 fs$. (b) Effect of resolution in weights applied on
a toy Monte Carlo sample.*

Assuming that the per-event estimate of the uncertainties from the B decay tree fitter are
correct, a data-driven correction method to compensate for part of the resolution effects can be
applied. The $D^0$ decay time PDF-functions are shown for different uncertainties in Fig. 8.8a.

Therefore, depending on the resolution one can give to each event a weight and correct the
distribution to match back to an exponential distribution. This means, events that are potentially
shifted from low decay times to high decay times due to resolution effects should obtain a smaller
weight compared to precisely measured events at high decay times.

$$\text{weight}(t, \sigma_t) = \frac{\exp}{\text{gauss} \otimes \exp} \tag{8.6}$$

$$= \frac{\exp(-\frac{t}{\tau})}{\exp(\frac{\sigma_t^2 - 2t}{2\tau})\text{erfc}(\frac{\frac{\sigma_t^2}{\tau} - t}{\sqrt{2}\sigma_t})/2}$$

$$= \frac{2}{\exp(\frac{\sigma_t^2}{2\tau^2})\text{erfc}(\frac{\frac{\sigma_t^2}{\tau} - t}{\sqrt{2}\sigma_t})}$$

$$\text{with erfc}(x) = 1 - \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \exp(-t^2)dt \tag{8.7}$$

Notice, the cancellation of the time dependence $\frac{t}{\tau}$ that can be different for $D^0$ and $\overline{D}^0$. This
is the reason for the bias at high decay times. The term $\exp(-\frac{t}{\tau})$ is not present anymore in the
weight function. The remaining term contains no $\frac{t}{\tau}$ but a $\frac{\sigma_t}{\tau}$ term. This term is responsible for the
bias on the asymmetry at low decay times. Using true effective lifetimes for $D^0$ and $\overline{D}^0$ separately
to derive the correction, the bias would be completely removed. However, we are blinded to this
value. We use $\tau \approx 410.1 fs$ [54] instead and thus keep a bias. The impact of the weighting using
our resolution model is shown in Fig. 8.8b. Where one can see that the tail at high decay times is
removed but also the remaining resolution effect is shown. For comparison, the bias using these
weights are compared to the bias applying no correction in Fig. 8.7. The following Tables 8.1, 8.2
and 8.3 show the result obtained from resolution reweighting of 2012 and 2011 datasets. In
all three channels the difference between applying the weight and applying no weights is small.
The shift in $A_\Gamma$ in $D^0 \to K^- K^+$ decays is $\Delta A_\Gamma = -0.0094 \pm 0.0210\%$. In the $D^0 \to \pi^- \pi^+$ channel it
is $\Delta A_\Gamma = 0.0386 \pm 0.0627\%$. For the control channel $D^0 \to K^- \pi^+$ the resolution has an effect of

**Table 8.1:** *Impact of resolution on $A_\Gamma$ after applying the resolution weights on 2011 and 2012 data in $D^0 \to K^- K^+$ channel. $\Delta A_\Gamma$ denotes the difference to the default $A_\Gamma$ weighted mean of 2012 and 2011 data without reweighting. The default $A_\Gamma$ is set to 0 since it is blinded.*

|  | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ |
|---|---|---|---|---|
| resolution 2012 | $-0.579 \pm 0.14$ | $-0.029 \pm 0.094$ | $-0.551 \pm 0.118$ | $0.009 \pm 0.080$ |
| resolution 2011 | $-0.482 \pm 0.22$ | $0.0422 \pm 0.151$ | | |
| weighted mean from 2011 and 2012 fits | | | $-0.545 \pm 0.114$ | $0 \pm 0.077$ |
| $\Delta$ with corrected uncertainty $\sigma_\Delta = \sqrt{\sigma_{weights}^2 - \sigma_{mean}^2}$ | | | $-0.006 \pm 0.0304$ | $0.009 \pm 0.021$ |

**Table 8.2:** *Impact of resolution on $A_\Gamma$ after applying the resolution weights on 2011 and 2012 data in $D^0 \to \pi^- \pi^+$ channel. $\Delta A_\Gamma$ denotes the difference to the default $A_\Gamma$ weighted mean of 2012 and 2011 data without reweighting. The default $A_\Gamma$ is set to 0 since it is blinded.*

|  | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ |
|---|---|---|---|---|
| resolution 2012 | $-0.422 \pm 0.281$ | $0.167 \pm 0.187$ | $-0.655 \pm 0.235$ | $0.039 \pm 0.158$ |
| resolution 2011 | $-1.2 \pm 0.431$ | $-0.279 \pm 0.295$ | | |
| weighted mean from 2011 and 2012 fits | | | $-0.766 \pm 0.214$ | $0 \pm 0.145$ |
| $\Delta$ with corrected uncertainty $\sigma_\Delta = \sqrt{\sigma_{weights}^2 - \sigma_{mean}^2}$ | | | $-0.111 \pm 0.097$ | $0.039 \pm 0.063$ |

**Table 8.3:** *Impact of resolution on $A_\Gamma$ after applying the resolution weights on 2011 and 2012 data in $D^0 \to K^- \pi^+$ channel.*

|  | $A_{CP}^{RAW}(0)[\%]$ | $A_\Gamma[\%]$ | $A_{CP}^{RAW}(0)[\%]$ | $A_\Gamma[\%]$ |
|---|---|---|---|---|
| resolution 2012 | $-1.49 \pm 0.0576$ | $0.0381 \pm 0.0382$ | $-1.55 \pm 0.049$ | $0.006 \pm 0.032$ |
| resolution 2011 | $-1.68 \pm 0.0907$ | $-0.0761 \pm 0.0613$ | | |
| weighted mean from 2011 and 2012 fits | | | $-1.51 \pm 0.048$ | $0.00927 \pm 0.0318$ |
| $\Delta$ with corrected uncertainty $\sigma_\Delta = \sqrt{\sigma_{weights}^2 - \sigma_{mean}^2}$ | | | $-0.04 \pm 0.010$ | $0.00308 \pm 0.006$ |

$\Delta A_\Gamma = 0.0031 \pm 0.0060\%$ on $A_\Gamma$. The absolute values determined on data are consistent with the expected impact on $A_\Gamma$ from simulation.

As pointed out earlier, this method only corrects for effects at high decay times (where the contribution from the unknown effective lifetime $\tau$ are negligible). Hence the total bias on $A_\Gamma$ due to resolution effect coming from effects at low decay times is extracted from Fig. 8.7, depending on the true $A_\Gamma$ value. Within the $x$-axis range, the bias in Fig. 8.7 can be compared with the extracted shifts on data by reweighting. They are well in agreement with our expectations within the uncertainties. However, the uncertainties extracted from simulation are more precise.

The final systematic contribution from resolution effect is determined from simulation. Fig. 8.9 shows the extracted linear dependence. Additionally to the previous figure, also the uncertainties
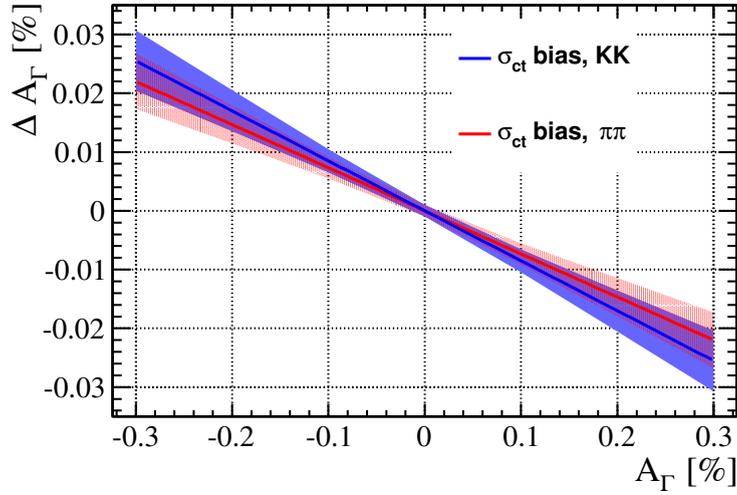


**Figure 8.9:** *The bias on $\Delta A_\Gamma$ due to resolution effects as function of $A_\Gamma$ value. In red the bias extracted from a binned fit on the asymmetry function in the $D^0 \to K^- K^+$ channel, in blue for $D^0 \to \pi^- \pi^+$ channel. The uncertainty bands show the impact by changing the resolution by $\pm 10\%$*

from changing the resolution by $\pm 10\%$ are shown.

The linear bias dependence on the true $A_\Gamma$ value is parametrised in the following way

$$A_\Gamma^{\text{estimator}} = (1+u)A_\Gamma^{\text{true}} \tag{8.8}$$

$$\Delta A_\Gamma = A_\Gamma^{\text{estimator}} - A_\Gamma^{\text{true}} = u A_\Gamma^{\text{true}}. \tag{8.9}$$

This notation is also further used in the next sections. Using this parametrization, one obtains the following final multiplicative systematic parameters for the resolution.

**Table 8.4:** *Extracted multiplicative parameter $u$.*

| Channel | Multiplicative parameter $u$ |
|---|---|
| $D^0 \to K^- K^+$ | $(-0.085 \pm 0.017)$ |
| $D^0 \to \pi^- \pi^+$ | $(-0.073 \pm 0.016)$ |
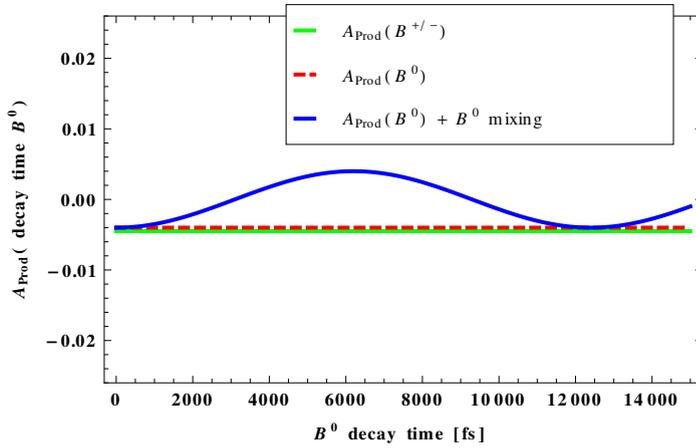
**Figure 8.10:** *$D^0$ production asymmetry varies with $B$ decay time in a toy Monte Carlo.*

# 8.4 Acceptance factorization and contribution from $B$ mixing

The asymmetry in the number of $D^0$ mesons at $D^0$ decay time $t = 0$ is inherited from the $B$ production asymmetry. For $D^0$ mesons from charged $B$ meson decays this asymmetry is expected to be constant with $B$ decay time. Further, for $D^0$ mesons from neutral meson decays it varies in time as shown in Fig. 8.10 due to $B^0$-meson oscillations. Therefore, $B$ decay time acceptance has an impact on the average $D^0$ production asymmetry.

From physics we expect that $B$ decay time and $D^0$ decay time are uncorrelated. But for example due to requirements on the displacement of the $D^0$ daughters from the primary vertex a bias can be introduced if the acceptances do not factorize:

$$acc(t_B, t_D) \neq acc(t_B) \times acc(t_D).$$

To study this effect, $B$ decay time distributions in bins of $D$ decay times are determined in MC (Fig. 8.11). There are small differences observed.

Now these distributions are used as an input to a toy with $B^0$ decays only. It is generated with the following parameters:

| $A_\Gamma$ | $A_{CP}^{dir}$ | $A_{prod.} + A_\mu$ | $\tau_B$ | $\Delta m$ |
|---|---|---|---|---|
| 0.1% | -0.05% | -0.04% | 1523 fs | $0.774/\tau$ |

where $\Delta m$ is the $B^0 - \overline{B^0}$ mixing frequency. Fig. 8.12, shows the impact of the detector acceptance and trigger selection. Mostly, asymmetries are shifted towards 0 as expected due to the oscillating behaviour. After the overall binned $\chi^2$ fit a small bias $\Delta_{factorize} \sim 0.0066\%$ on $A_\Gamma$ is visible. Including the charged $B$ mesons, this effect even reduces.
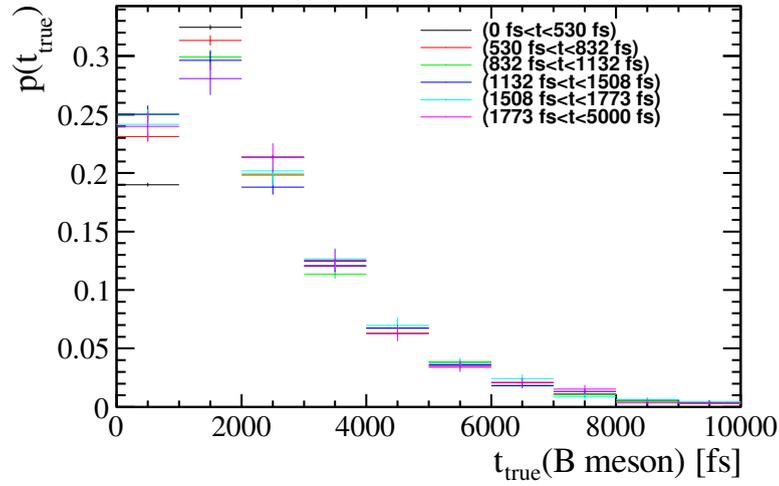
**Figure 8.11:** *Different B decay time distributions calculated for different $D^0$ decay time bins.*
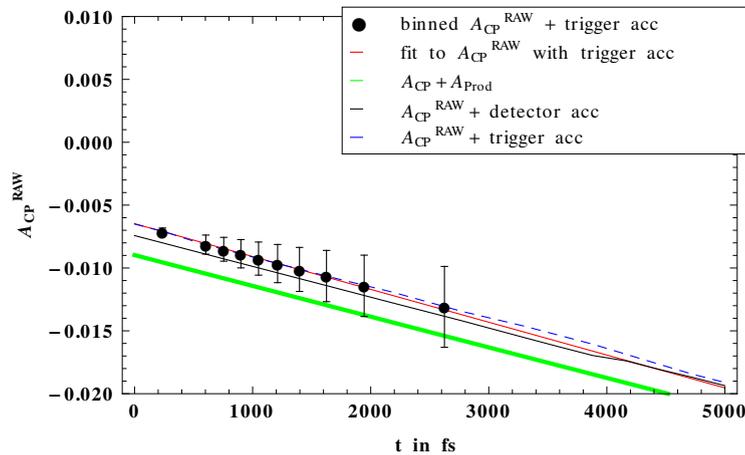


**Figure 8.12:** *The green curve is the theoretical raw asymmetry curve without any acceptance and trigger effects. The black curve takes the acceptances due to the detector geometry into account (The step in this distribution is a result of limited MC statistics for the determination of the acceptance correction.). For the blue curve additionally the trigger acceptance is added. The black point is the same distribution as the blue one however sampled in decay time bins. The red line is a linear fit to the black points. The difference between the input value of $A_\Gamma$ and the slope of the red curve are taken to quantify the systematics related to acceptances.*

## 8.5 Mistag decay time dependence

The raw asymmetry depends also on the mistag probabilities.

$$A_{CP}^{Raw} \approx (1 - 2\bar{\omega})(A_{CP} + A_\mu + A_P) - \Delta\omega$$

Thereby, following definitions have been introduced in section 3.2:

$$\omega^{\bar{D}^0} = P(\bar{D}_{tag}^0 | D^0 \wedge B_{rec}) \text{ and } \omega^{D^0} = P(D_{tag}^0 | \bar{D}^0 \wedge B_{rec})$$

Small time independent mistag probabilities ($\mathcal{O}(1\%)$) have a very tiny effect on the measurement. In this section, time dependence of the mistag probabilities are studied. Both time-dependent factors, the damping factor $\bar{\omega}(t)$ and also an offset mistag difference term $\Delta\omega(t)$, can have an impact on the measurement of $A_\Gamma$.

Therefore, mistag probabilities $\omega^{\bar{D}^0}$ and $\omega^{D^0}$ were determined in the flavour-specific $D^0 \to K^-\pi^+$ channel

$$B \to \bar{D}^0(\to \underbrace{K^+\pi^-}_{\text{confirmation (Den.)}}) \underbrace{\mu^+}_{\text{used as tag (Num.)}} X.$$

From the charge of the $D$ daughters $\bar{D}^0 \to K^+\pi^-$ (or $D^0 \to K^-\pi^+$) one can check the $D^0$ flavour and determine the mistag probabilities. The measured mistag probabilities in this channel are defined therefore as:

$$\omega_{measured}^{\bar{D}^0} = \frac{N(\mu_{tag}^- \wedge (K^+\pi^-))}{N(K^+\pi^-)} \tag{8.10}$$

$$\omega_{measured}^{D^0} = \frac{N(\mu_{tag}^+ \wedge (K^-\pi^+))}{N(K^-\pi^+)}$$

Hereby, one has not only to take the Cabbibo-favoured right-sign (RS: $D^0 \to K^-\pi^+$) but also the wrong-sign component (WS: $\bar{D}^0 \to K^-\pi^+$) into account, coming from doubly Cabbibo-suppressed decays and $D^0$ mixing; for details, see Appendix B.3.

### 8.5.1   Average mistag $\omega(t)$

The result of the average mistag probability extraction is shown in Fig. 8.13. To subtract the WS decays, we use the results from the $D^0$ mixing analysis [77]. The wrong-sign ratio is

$$R(t) = R_D + \sqrt{R_D}\, y' \frac{t}{\tau} + \frac{x'^2 + y'^2}{4} \frac{t}{\tau} \, , \tag{8.11}$$

with $R_D = (3.568 \pm 0.058 \pm 0.033) \times 10^{-3}$, $y' = (4.8 \pm 0.8 \pm 0.5) \times 10^{-3}$, $x'^2 = (5.5 \pm 4.2 \pm 2.6) \times 10^{-5}$. The wrong sign ratio has a relative big effect for small decay times. At larger decay times the mistag probability dominates. At low decay times we have efficient cuts to get rid of mis tagged $D^0$ and at higher lifetimes more mistags, relative to the right signal tags, can be associated. This is related to the fact that most mistag candidates are combinations of a $D^0$ from a $B$ decay which is combined with a $\mu$ from the opposite side $B$ decay. A study described in Appendix B.4 shows that this combination results relatively more often in long than in short (falsely) reconstructed $D^0$ decay times. This can be simply explained by the exponential decay of the signal, whereas the background decreases less steeply as a function of $D^0$ decay time. To model the time dependence of the mistag probability an exponential with an offset, $a * exp(\lambda t) + c$, is used. Additionally, 1-$\sigma$ bands from the fit are calculated and displayed in Fig. 8.13.
The extracted parameters are:

| parameter | value | statistical uncertainty |
|---|---|---|
| a | $7.39 \cdot 10^{-3}$ | $0.54 \cdot 10^{-3}$ |
| $\lambda\,[\mathrm{fs}^{-1}]$ | $7.70 \cdot 10^{-4}$ | $0.29 \cdot 10^{-4}$ |
| c | $-2.04 \cdot 10^{-3}$ | $6.03 \cdot 10^{-3}$ |

To study the impact of $\omega$ time dependence on $A_\Gamma$, we are using the model parameters listed in Tab. 8.5.

From the fit to the binned raw asymmetry the impact on $A_\Gamma$ from the extracted average mistag probability is calculated. The result is shown in Fig. 8.14. The time dependent average mistag probability for the scenario with $A_\Gamma = 0.1\%$ and $D^0 \rightarrow K^- K^+$ channel from Tab. 8.5 is used. Afterwards, a toy simulation and a binned $\chi^2$ fit is performed to extract $A_\Gamma^{\text{estimator}}$.

This can be done scanning through the parameters in Tab. 8.5. Following Fig. 8.15 is obtained, which shows the difference $\Delta A_\Gamma = A_\Gamma^{\text{estimator}} - A_\Gamma^{\text{true}}$ vs. the generated $A_\Gamma^{\text{true}}$ value for different
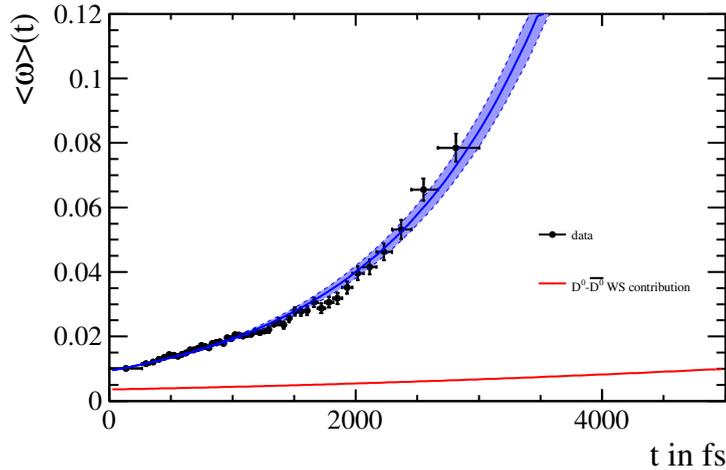


**Figure 8.13:** *Average mistag probability extracted from the $K\pi$ channel. In red the expected contribution from WS decays is shown.*

**Table 8.5:** *Parameters used for different scenarios*

| $A_\Gamma$ | $A_{CP}^{dir} + A_{prod.} + A_\mu$ | $\bar\omega(t)$ |
|---|---|---|
| -0.3%..0.3% | -0.545% and -0.768% ($\pm$ 1-$\sigma$ from $KK$ and $\pi\pi$) | $a * exp(\lambda t) + c$ ($\pm$ 1-$\sigma$ bands) |
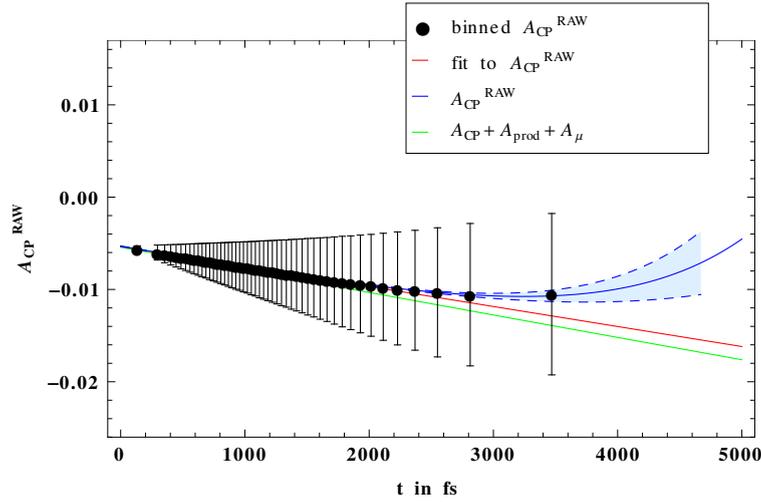


**Figure 8.14:** *Influence of the mistag time dependence on the raw asymmetry and the extracted $A_\Gamma$. In green - the initial asymmetries. In blue - raw asymmetry with average mistag probability from exponential fit, together with the 1$\sigma$ bands from the fit. In red - binned fit to the raw asymmetry.*

channels. The uncertainty bands are obtained from the uncertainty at each point on the average mistag probability parametrization and the raw asymmetry offset extracted from data.

This extraction of the bias on $A_\Gamma$ from the time-dependent average mistag probability can be separated into $A_\Gamma$ independent and a multiplicative part.

$$A_{CP}^{RAW}(t) \approx (1 - 2\bar\omega(t))(A_{CP}(t) + A_\mu + A_P) \qquad (8.12)$$

$$= (A_{CP}^{RAW}(0) - A_\Gamma \frac{t}{\tau})(1 - 2\bar\omega(t))$$

$$= \underbrace{A_{CP}^{RAW}(0)(1 - 2\bar\omega(t))}_{\text{independent.}} - \underbrace{A_\Gamma \frac{t}{\tau}(1 - 2\bar\omega(t))}_{\text{multiplicative}}$$

The independent part depends on the raw asymmetry offset at t=0 (means: on the sum $A_{CP}^{dir} + A_{prod.} + A_\mu$). The multiplicative part depends on the true $A_\Gamma$ value and shows a linear $A_\Gamma$ dependence (see Fig. 8.15). Following parametrisations of this parts can be used.

$$A_\Gamma^{\text{estimator}} = \qquad\qquad v + (1+u)\, A_\Gamma^{\text{true}} \qquad (8.13)$$

$$\Delta A_\Gamma = A_\Gamma^{\text{estimator}} - A_\Gamma^{\text{true}} = \qquad\qquad v + u\, A_\Gamma^{\text{true}}$$

Therefore, the additive offset $v$ and the multiplicative parameter $u$ can be extracted from the fits separately for $KK$ and $\pi\pi$ channels. The result of the fit is listed in the Tab. 8.6. Similarly, for the $D^0 \to K^-\pi^+$ channel (we have here only additive part) one can calculate the additive offset $v_{K\pi} = 0.0166\%$. The impact of the multiplicative part is evaluated and is also listed in the final
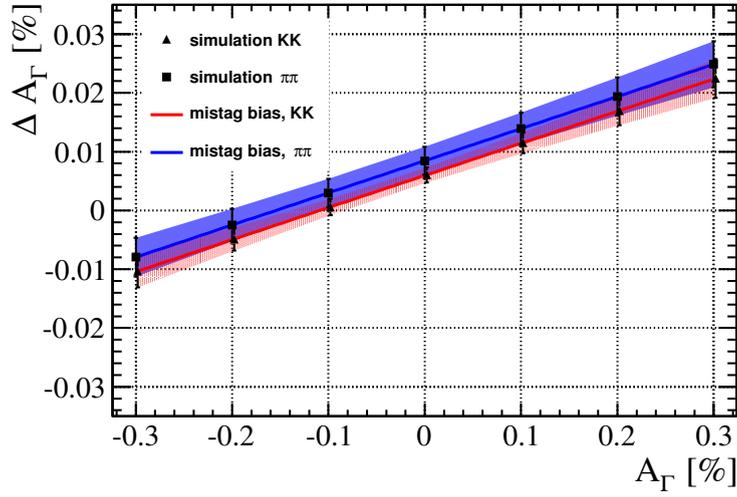
**Figure 8.15:** *The bias on $\Delta A_\Gamma$ due to average mistag as function of $A_\Gamma$ value. In red: for the $D^0 \to K^- K^+$ events. In blue: for the $D^0 \to \pi^- \pi^+$ events. The lines show the linear fit model, which parameters are listed in Tab. 8.6.*

**Table 8.6:** *Extracted additive offset $v$ and multiplicative parameter $u$.*

| Channel | Additive offset $v$ [%] | Multiplicative parameter $u$ |
|---|---|---|
| $D^0 \to \pi^- \pi^+$ | $(8.45 \pm 1.08)10^{-3}$ | $(0.0547 \pm 0.0062)$ |
| $D^0 \to K^- K^+$ | $(6.06 \pm 0.71)10^{-3}$ | $(0.0547 \pm 0.0095)$ |

systematics table. Additionally, the direct fit of equation Ref. 8.12 can be performed on data. This is summarized in Appendix A.10 and gives fully consistent results compared to this simulation toy study.
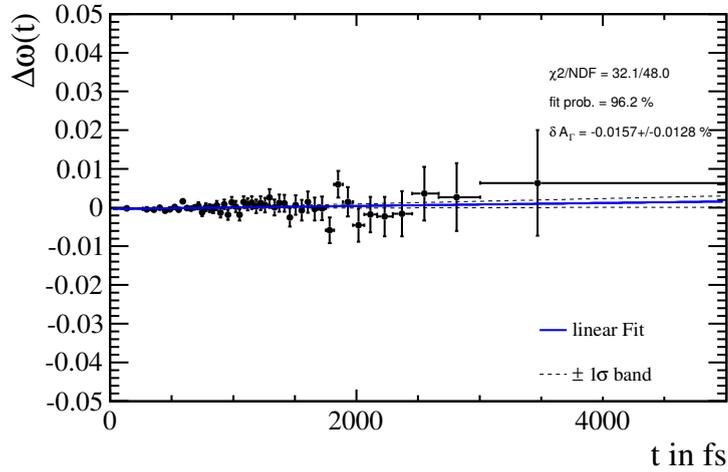
**Figure 8.16:** *Difference in mistag probability time dependence, extracted from the $K\pi$ channel.*

### 8.5.2  Mistag difference $\Delta\omega(t)$

The time dependence of the mistag difference $\Delta\omega(t)$ is determined on data as well. There is no significant CP violation measured and expected in $D^0 \to K^-\pi^+$ channel (Refs. [60, 77]). Thus, one can assume

$$\Delta\omega_{measured}(t) \approx \Delta\omega(t).$$

The result of the binned linear $\chi^2$ fit on $\Delta\omega(t)$ is shown in Fig. 8.16. The extracted value is $\Delta A_\Gamma = 0.0157 \pm 0.0153\%$. It is consistent with having no time dependence. $\Delta\omega(t)$ would directly contribute to the slope $A_\Gamma$ and therefore we use $\Delta A_\Gamma^{\Delta\omega(t)} = 0.0157\%$ as uncertainty for the systematics.

## 8.6  Input lifetime knowledge

In Sec. 6 the procedure of $A_\Gamma$ extraction is explained and Eq. 6.10 is used. Thereby, the fitted slope is $\frac{-A_\Gamma}{\tau_{D^0}}$. We are using $\tau = 410.1 \pm 1.5\,\text{fs}$ from PDG [54] for the final $A_\Gamma$ determination. The uncertainty on the input lifetime value results in the following uncertainty on $A_\Gamma$

$$\Delta A_\Gamma^{\tau\,\text{knowlege}} = slope \cdot \sigma_\tau \tag{8.14}$$

$$= A_\Gamma \cdot \frac{\sigma_\tau}{\tau} \tag{8.15}$$

$$= A_\Gamma \cdot 0.366\%. \tag{8.16}$$

Assuming $A_\Gamma \sim 0.1\%$ we obtain

$$\Delta A_\Gamma^{\tau\,\text{knowlege}} \sim 0.000366\%. \tag{8.17}$$

This systematic uncertainty is tiny compared to our statistical sensitivity.

## 8.7   Uncertainty on the detector length scale

The accuracy of distance and therefore of decay time depends on the knowledge of detector and how well the position of the modules in the VELO is known. The uncertainty on the longitudinal-scale (z-scale) of LHCb can be calculated as described in Ref. [78] and is determined to be

$$\sigma_{z-scale} = 0.022\%. \tag{8.18}$$

This uncertainty scales directly with $A_\Gamma$. Therefore, an uncertainty due to this effect can be obtained in a similar way as explained in Sec.8.6:

$$\Delta_{A_\Gamma}^{z-scale} = A_\Gamma \cdot 0.022\%. \tag{8.19}$$

Assuming $A_\Gamma \sim 0.1\%$ we obtain:

$$\Delta_{A_\Gamma}^{z-scale} \sim 0.000022\%. \tag{8.20}$$

This systematic uncertainty is again tiny compared to our statistical sensitivity.

## 8.8   Model dependence

Several fit models for the signal and background in the $D^0$ invariant mass distribution have been tested. For the signal these are

- default: sum of three Gaussian PDFs,

- sum of two Gaussian PDFs,

- and sum of a Gaussian and a Crystal BallPDFs,

and for the background

- default: Exponential PDF,

- polynomial 1st orderPDFs,

- and polynomial 2nd orderPDFs.

The best description in terms of $\chi^2/ndf$ of the mass fits is our default model (Sec. 6). For systematic studies all combinations of signal and background models are investigated. The results for the different channels are shown in the Figs. 8.17. The detailed Tabs. A.7, A.8 and A.9 can be found in Appendix A.6.1. The differences in $A_\Gamma$ using different models to describe the $D^0$ invariant mass distribution relative to our default fit are tiny. The largest one is taken as a systematic uncertainty:

- $D^0 \to K^- K^+$: $\Delta A_\Gamma$(mass model) $= 0.0111\%$

- $D^0 \to \pi^- \pi^+$: $\Delta A_\Gamma$(mass model) $= 0.0074\%$

- $D^0 \to K^- \pi^+$: $\Delta A_\Gamma$(mass model) $= 0.0100\%$.
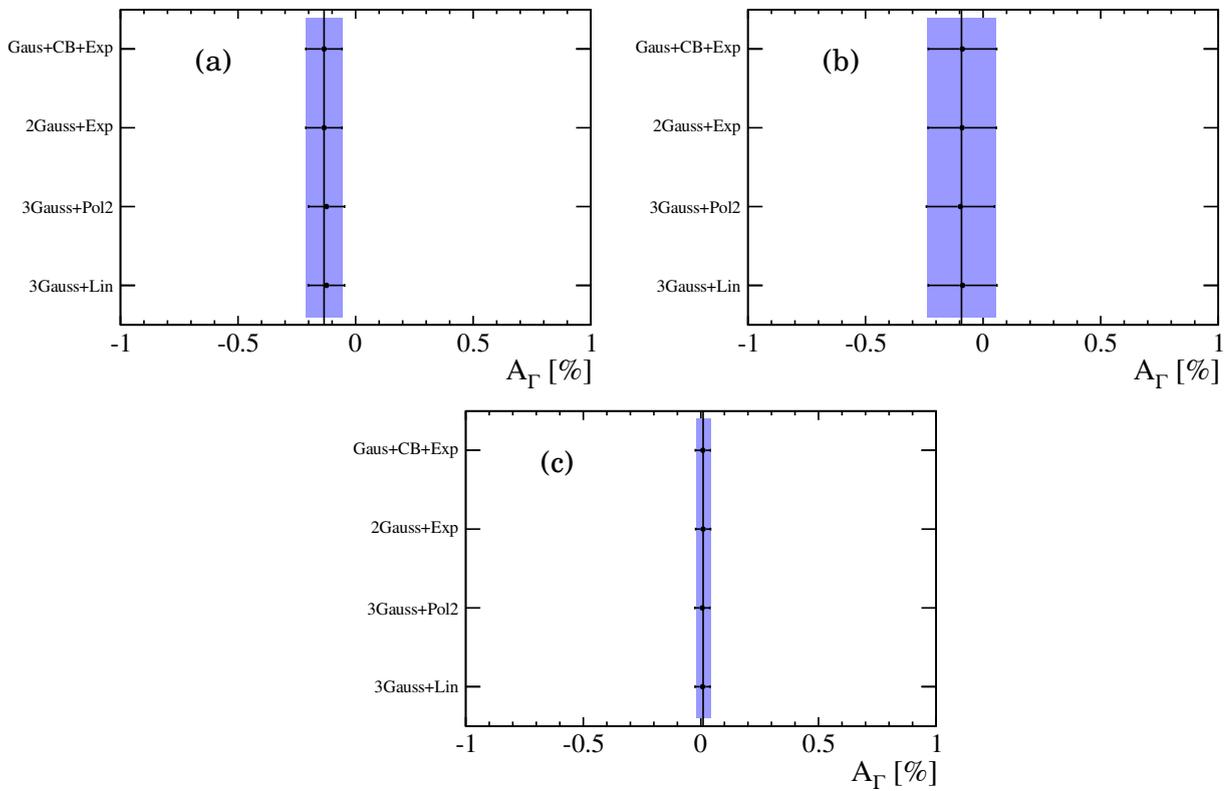
**Figure 8.17:** *Fit results performed on the same data set are shown. Different models for $A_\Gamma$ are used. Shown are (a) $D^0 \to K^- K^+$, (b) $D^0 \to \pi^- \pi^+$ and (c) $D^0 \to K^- \pi^+$ decays. The vertical line and blue band indicate the default $A_\Gamma$ value together with its statistical uncertainty only.*

## 8.9  Selection biases

Any time-dependent asymmetry coming from a momentum-dependent (detection or production) asymmetry in combination with a momentum-dependent efficiency, is already accounted for in the reweighting systematic as explained in Section 8.2. As shown in Chapter 5, Fig. 5.4, the Hlt2 selection has the largest effect on the decay time acceptance. A possible bias on $A_\Gamma$ that could arrise if Hlt2MuTopoNBody lines have different, time-dependent efficiencies when positive or negative muons are involved.

Any possible effect from the Hlt2MuTopoNBody triggers is expected to be small, as the IP requirements are independent of (muon) charge. The trigger lines have been also commonly used and confirmed in asymmetry studies, e.g. in the analysis [71], [72]. Asymmetries in these trigger lines have been studied in Ref. [79]. This method uses the HLT2 topological N body and SingleMuon trigger where no asymmetry from Hlt2 is observed within 0.01%. We propagate this number as a systematic uncertainty.

## 8.10   Binning of decay time

Already in previous, sections different binning schemes were studied and we expect tiny impact on the measurement from the toys. However, this is also studied on data. We observe, very tiny differences applying again the full analysis chain on different binning. The results for default 50 time bins and for alternatively finer (180 bins) or coarser (26 bins) decay time binning are shown in Figs. 8.18 for different channels and summarized in the tables in Appendix A.6.2. No systematic uncertainty due to the choice of the binning is assigned.
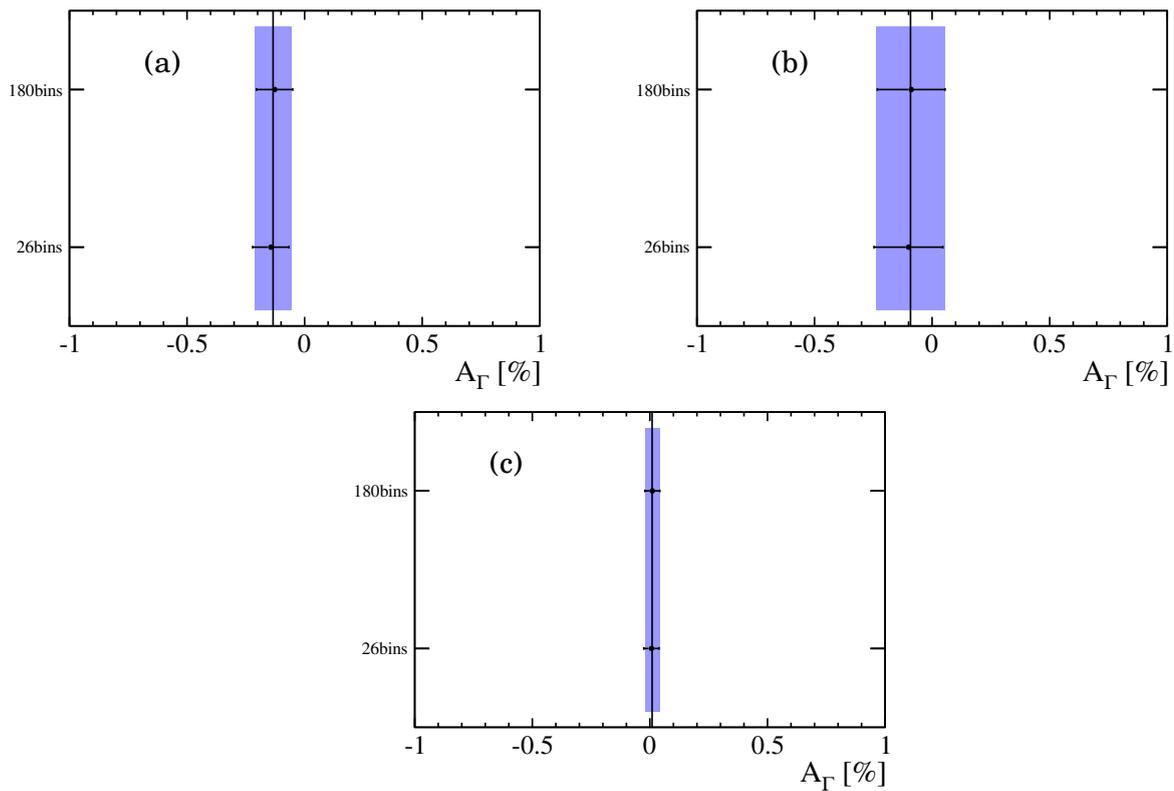


**Figure 8.18:** *The data sets are subdivided in different run periods 2011/2012 and up/down magnet polarities. Measured $A_\Gamma$ values are shown for (a) $D^0 \to K^- K^+$, (b) $D^0 \to \pi^- \pi^+$ and (c) $D^0 \to K^- \pi^+$ decays. The vertical line and blue band indicate the combined $A_\Gamma$ value together with its statistical uncertainty only.*

## 8.11 Additional cross checks and other sources of systematics

This section concentrates on several cross check on data and demonstrates the robustness of the analysis.

The dataset is split into different subsamples to test if there is some dependence on $A_\Gamma$. One should keep in mind that for statistical reasons the probability to find in this case at least one 2-$\sigma$ or large deviation, dividing dataset into 2011 and 2012 data and these into 6 additional samples, is bigger than $> 1 - 0.954^{6 \times 2} \approx 43\%$ and the look-elsewhere effect is not negligible if one starts to split the data set in too many different subsamples. The following quantities are used to split up the sample:

- 2011, 2012 data

- Magnet polarities

- Number of primary vertices

- $B$ decay time

In all three channels the extracted differences in $A_\Gamma$ are compatible within $2\sigma$. The results are summarized in the following tables and figures. [2]

---

[2]$\Delta A_\Gamma$ denotes thereby the difference between this value and the weighted mean of the full data sample in each channel

### 8.11.1 Dependence of $A_\Gamma$ on magnet polarity
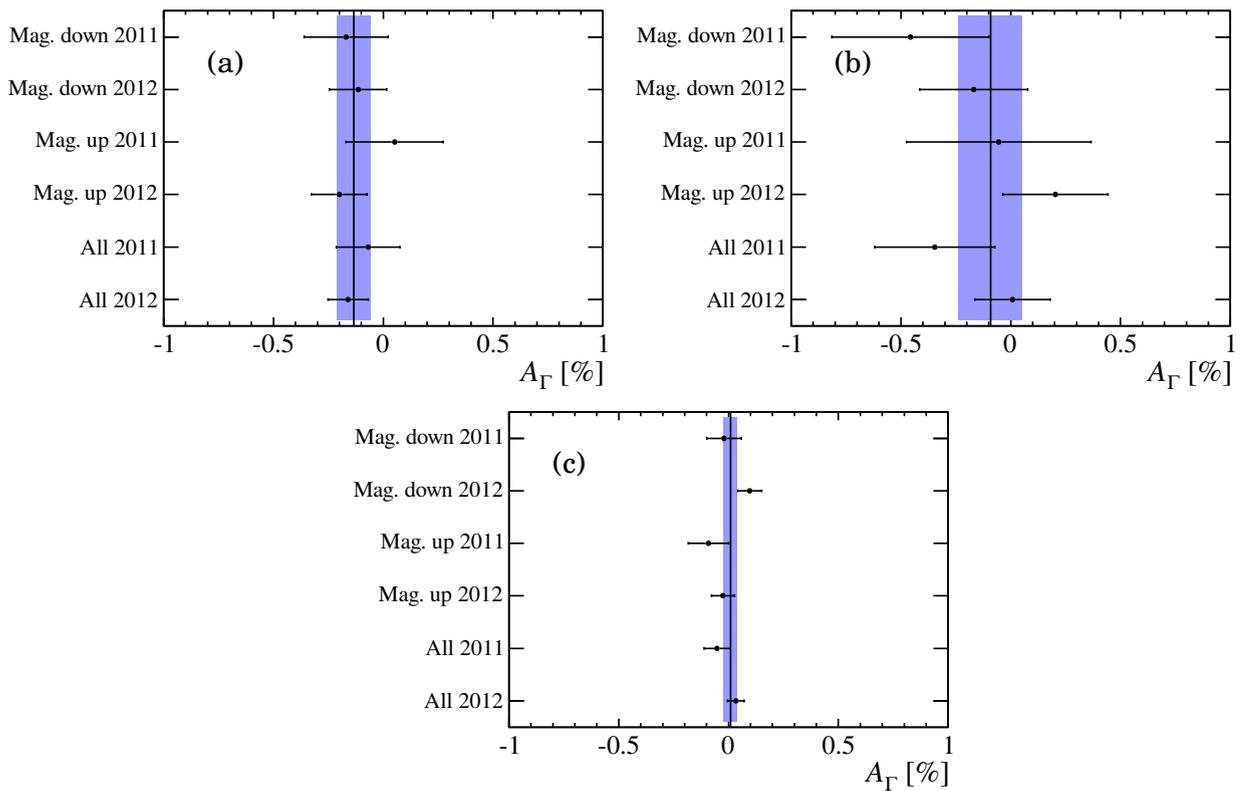


**Figure 8.19:** *The data sets are subdivided in different run periods 2011/2012 and up/down magnet polarities. Measured $A_\Gamma$ values are shown for (a) $D^0 \to K^- K^+$, (b) $D^0 \to \pi^- \pi^+$ and (c) $D^0 \to K^- \pi^+$ decays. The vertical line and blue band indicate the combined $A_\Gamma$ value together with its statistical uncertainty only.*

### 8.11.2 Dependence of $A_\Gamma$ on number of reconstructed primary vertices
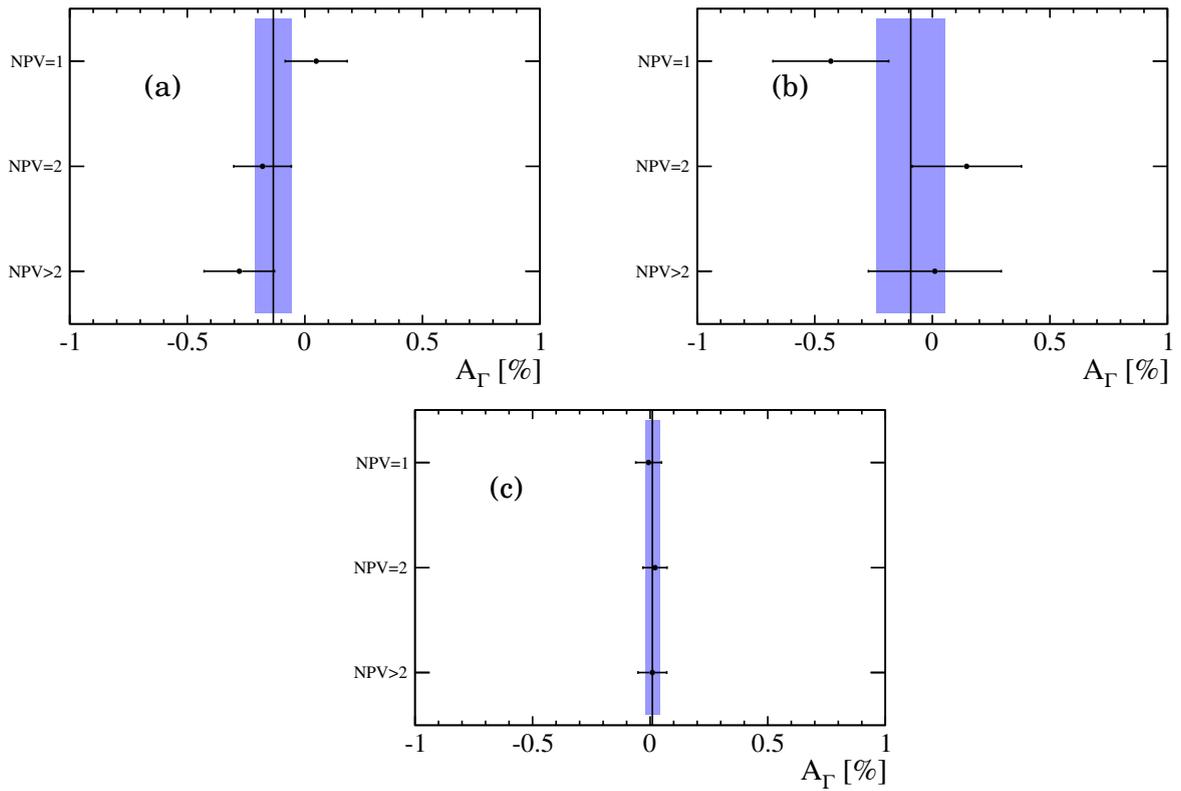


**Figure 8.20:** *Fit results split in data sets of different number of reconstructed primary vertices. Measured $A_\Gamma$ values are shown for (a) $D^0 \to K^- K^+$, (b) $D^0 \to \pi^- \pi^+$ and (c) $D^0 \to K^- \pi^+$ decays. The vertical line and blue band indicate the combined $A_\Gamma$ value together with its statistical uncertainty only.*

### 8.11.3  *B* decay time

The $B$ decay time distribution extracted from the data is shown in Fig 8.21. The sample was split into two parts with about similar statistical significance:

- low $B$ decay time: $t(B) \leq 2000\,\mathrm{fs}$

- high $B$ decay time: $t(B) > 2000\,\mathrm{fs}$.
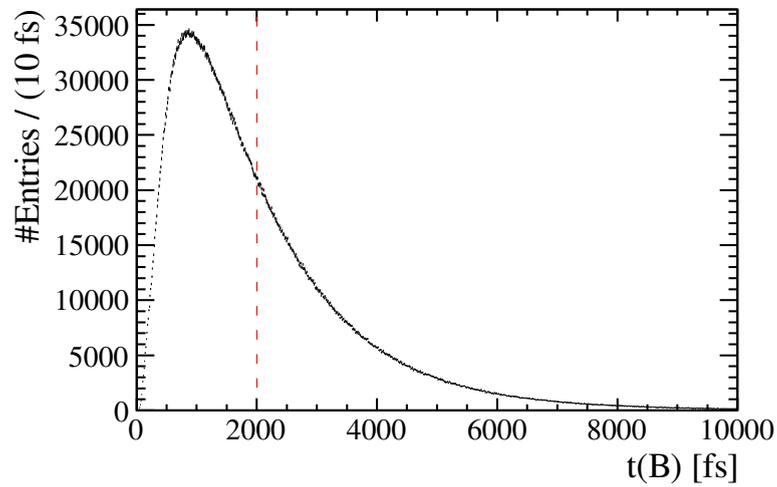
**Figure 8.21:** *B decay time distribution, extracted from data in the $D^0 \to K^- \pi^+$ channel. The dashed red line visualizes the used separation between the low and high decay time regions.*
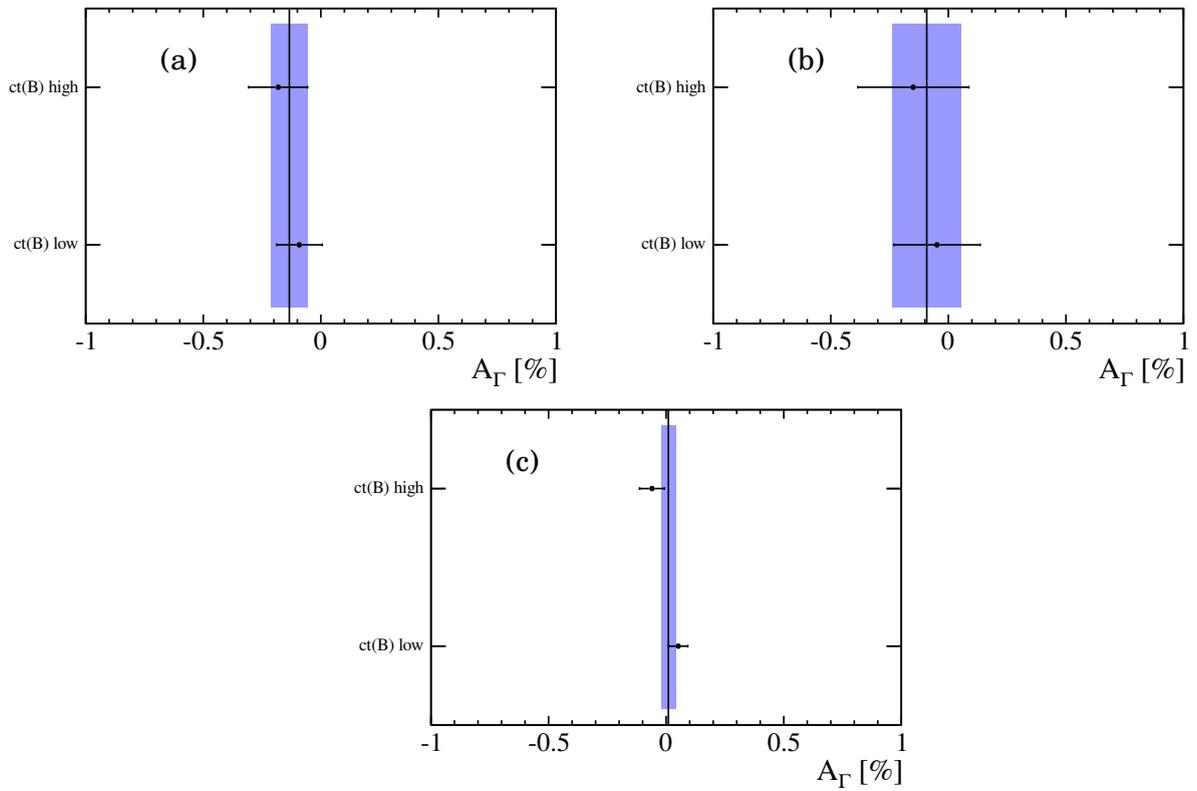
**Figure 8.22:** *Fit results for different B decay time regions. Measured $A_\Gamma$ values are shown for (a) $D^0 \to K^- K^+$, (b) $D^0 \to \pi^- \pi^+$ and (c) $D^0 \to K^- \pi^+$ decays. The vertical line and blue band indicate the combined $A_\Gamma$ value together with its statistical uncertainty only.*

## 8.12   Uncertainty scale factor

As shown in the previous sections, some systematic sources scale with the value of $A_\Gamma$. This section discusses how to deal with such a multiplicative scale factor. Therefore, construction of the Neyman belt is described in the following, more details can be found in Ref. [54].

For an unknown true parameter $x_t$, there are some outcomes of the experiments that we call $x_m$. This parameter represents the measured value. For a probability density function of both variables $\mathscr{PDF}(x_m, x_t)dx_m$, one can now define the following belt

$$P(x_m; x_{t\,\text{low}} < x_t < x_{t\,\text{up}}) = 1 - \alpha = P(x_{m\,\text{low}} < x_m < x_{m\,\text{up}}; x_t) = \int_{x_{m\,\text{low}}}^{x_{m\,\text{up}}} \mathscr{PDF}(x_m, x_t)dx_m. \quad (8.21)$$

Thereby, $(1-\alpha)$ sets the confidence level (CL). [3] It is typically chosen to be 68% for $1\sigma$ intervals for a Gaussian distribution. The Neyman belt provides the confidence intervals for the true $x_t$ observable.

As described in Sections 8.3 8.5, the uncertainty in this analysis has a component that scales linearly with the known true value $x_t$, e.g. see Eq. 8.8, 8.13. The uncertainty can split in an additive ($\sigma_{\text{add}}$) and a multiplicative part ($x_t \sigma_{\text{mult}}$) that scales linearly with $x_t$. The total uncertainty on $x_m$ can be calculated using Gaussian uncertainty propagation, $\sigma_{\text{tot}}^2 = \sigma_{\text{add}}^2 + x_t^2 \sigma_{\text{mult}}^2$. Using this approach, the confidence belt can be directly constructed. The 68% CL upper and lower limits are written in terms of the additive and multiplicative uncertainties, $\sigma_{\text{add}}$ and $\sigma_{\text{mult}}$, as

$$x_{m\,\text{low}} = x_t - \sqrt{\sigma_{\text{add}}^2 + x_t^2 \sigma_{\text{mult}}^2}$$
$$x_{m\,\text{upp}} = x_t + \sqrt{\sigma_{\text{add}}^2 + x_t^2 \sigma_{\text{mult}}^2}, \quad (8.22)$$

This Eq. 8.22 can now be inverted to get an analytical function. These limits are shown in Fig. 8.23 for different $x_t$ values. [4] The confidence interval can be now obtained from this Neyman belt for different observed value of $x_m$.

In order to present a single systematic uncertainty, the effect of the scale uncertainty is evaluated with this Neyman construction. For the measured value $x_m$, $A_\Gamma$ and its statistical uncertainty are used in the evaluation. They are determined in Chapter 7. [5] Additionally, the additive and multiplicative systematic uncertainties that are studied in this chapter are summarized in the next section in Tab. 8.7.

Evaluating the Neyman construction gives a slightly asymmetric systematic uncertainty, which is $^{+0.026}_{-0.034}\%$ for the $D^0 \to K^- K^+$ decay channel and $^{+0.025}_{-0.033}\%$ for the $D^0 \to \pi^- \pi^+$ decay channel. Except for the contribution from the mass fit model, all contributions to the systematic uncertainty are fully correlated, resulting in an overall correlation coefficient of 89% between the systematic uncertainties of $A_\Gamma(K^- K^+)$ and $A_\Gamma(\pi^- \pi^+)$.

---

[3]An additional criteria, e.g. central interval coverage, with exclusion level of $\alpha/2$ at both uncertainties sides, is required, as explained in Ref. [54].

[4]Technically, to make this plot, we have numerically inverted equation 8.22.

[5]Notice that also statistical uncertainty needs to be treated as additive, but can be later subtracted to obtain only the systematic part.
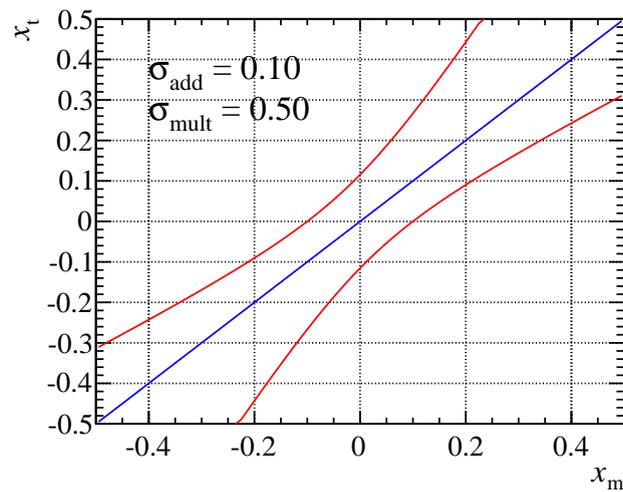
**Figure 8.23:** *Example of a Neyman confidence belt for arbitrary values of the additive and multiplicative uncertainties. For an observed value of $x_t$, the confidence interval can be read off from this plot.*

## 8.13   Summary of the systematics

The validated systematics are summarized in Tab. 8.7. The main systematic effects were directly validated with data and also cross checked with simulation. Several cross checks have demonstrated the robustness of this analysis. Compared to the statistical uncertainties the systematical are factor 2-3 smaller and do not dominate the total uncertainty of this measurement.

**Table 8.7:** *Summary of the systematic contributions.*

| Systematic source | $D^0 \to K^- K^+$ | | $D^0 \to \pi^- \pi^+$ | |
|---|---|---|---|---|
| | $\sigma_{\text{add}}$ | $\sigma_{\text{mult}}$ | $\sigma_{\text{add}}$ | $\sigma_{\text{mult}}$ |
| Fitter estimator (Sect. 8.1) | 0.0020% | | 0.0020% | |
| Muon and prod. asymmetries (Sect. 8.2) | 0.0098% | | 0.0098% | |
| $D^0$ decay time resolution (Sect. 8.3) | | 0.085 | | 0.073 |
| $B$-mixing (Sect. 8.4) | 0.0066% | | 0.0066% | |
| $\omega$ time dependence (Sect. 8.5) | 0.0061% | 0.0547 | 0.0084% | 0.0547 |
| $\Delta\omega$ time dependence (Sect. 8.5.2) | 0.0157% | | 0.0157% | |
| $D^0$ lifetime knowledge (Sect. 8.6) | | 0.00366 | | 0.00366 |
| Detector length scale (Sect. 8.7) | | 0.00022 | | 0.00022 |
| Mass model (Sect. 8.8) | 0.0111% | | 0.0074% | |
| Selection bias (Sect. 8.9) | 0.0100% | | 0.0100% | |
| **Quadratic sum** | 0.0255% | 0.101 | 0.0248% | 0.091 |
| **Total syst. uncertainty** (Sect. 8.12) | $^{+0.026}_{-0.034}\%$ | | $^{+0.025}_{-0.033}\%$ | |
| Statistical uncertainty (Table 7.1) | 0.077% | | 0.145% | |

This section presents combined results determined from 2011 and 2012 data sets. They correspond to an integrated luminosity of $3.0\ \mathrm{fb}^{-1}$. The indirect CP asymmetries are measured in the two main channels $D^0 \to K^- K^+$, $D^0 \to \pi^- \pi^+$ and determined to be

$$A_\Gamma(K^- K^+) = (-0.134 \pm 0.077(stat.)^{+0.026}_{-0.034}(syst.))\% \, ,$$
$$A_\Gamma(\pi^- \pi^+) = (-0.092 \pm 0.145(stat.)^{+0.025}_{-0.033}(syst.))\%.$$

The difference between the two channels is

$$\Delta A_\Gamma = A_\Gamma(K^- K^+) - A_\Gamma(\pi^- \pi^+) = (-0.042 \pm 0.164(stat.) \pm 0.013(syst.))\% \, ,$$

and is compatible with no difference in indirect CP violation and therefore with the universality of the measured indirect CP violation. The correlation of the systematic uncertainties is considered in the uncertainty calculation.

Additionally, a high statistic channel $D^0 \to K^- \pi^+$ is analyzed, where no indirect asymmetry is expected. The indirect CP asymmetry parameter in this channel $A_\Gamma(K^- \pi^+) = (0.009 \pm 0.032(stat.))$ is compatible with no indirect CP violation.

The combined measured asymmetries in bins of decay time are shown in Fig. 9.2 and show the result of the time-dependent fit. The fits have good $\chi^2$-probabilities of 54.3% $(D^0 \to K^- K^+)$, 30.8% $(D^0 \to \pi^- \pi^+)$ and 14.5% $(D^0 \to K^- \pi^+)$, also different shape tests of the pulls confirm the agreement with the fit model.

The result of both channels can be combined, assuming the universality of the indirect *CP* violation in $D^0$ decays , Ref. [59]. Both measurements for $D^0 \to K^- K^+$ and $D^0 \to \pi^- \pi^+$ decays are statistically uncorrelated. Accounting for the 89% correlation of the systematic uncertainties the weighted average becomes

$$\text{This analysis}\ : A_\Gamma = (-0.125 \pm 0.073)\% \, .$$

An overview of the current measurements of $A_\Gamma$ is shown in Fig. 9.1. It shows that this muon-tagged analysis has one of the smallest uncertainties. Previously, LHCb published an independent pion-tagged LHCb measurement using the dataset of $1.0\,\mathrm{fb}^{-1}$, see Ref. [22]. However, the pion-tagged results in which the $D^0$ are produced promptly in the proton-proton collisions are
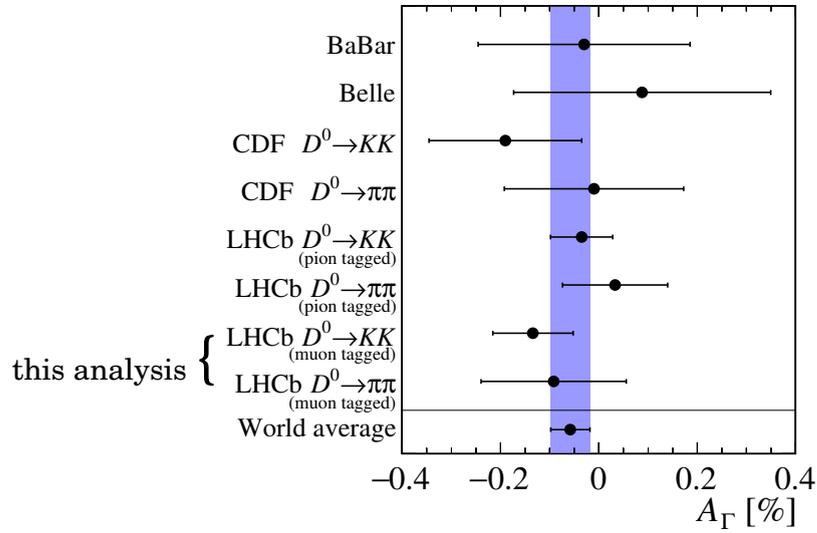
**Figure 9.1:** *Overview of $A_\Gamma$ measurements in $D^0 \to K^- K^+$ and $D^0 \to \pi^- \pi^+$ decays. The results are shown from BaBar Ref. [62], Belle Ref. [61], CDF Ref. [80], LHCb (pion-tagged) Ref. [22] and this thesis (LHCb, muon-tagged). The average of all $A_\Gamma$ measurements, combining the results for $D^0 \to K^- K^+$ and $D^0 \to \pi^- \pi^+$ decays, is also indicated.*

completely independent. The measurement presented in this thesis has a similar sensitivity and uncorrelated systematic uncertainties. Both measurements together dominate the current knowledge of $A_\Gamma$. Combining LHCb results together with the $A_\Gamma$ measurements from BaBar [62], Belle [61] and CDF [80], the world average becomes

$$\text{World average}: A_\Gamma = (-0.058 \pm 0.040)\% \ .$$
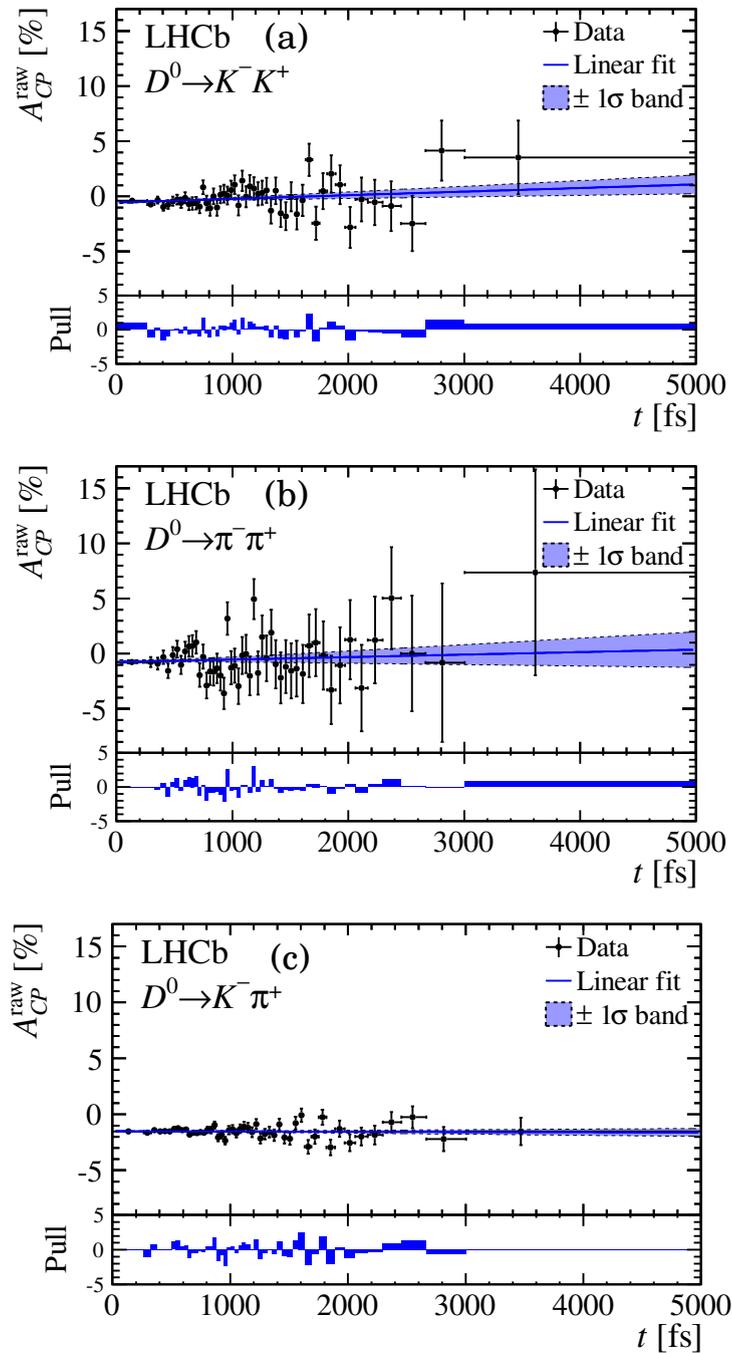
**Figure 9.2:** *Raw CP asymmetry as function of $D^0$ decay time for (a) $D^0 \to K^- K^+$, (b) $D^0 \to \pi^- \pi^+$ and (c) $D^0 \to K^- \pi^+$ candidates . The result of the $\chi^2$ fit is shown as a (blue) solid line with the $\pm 1$ standard-deviation ($\sigma$) band indicated by the dashed lines. Underneath each plot the pull defined as $\frac{data-fit}{\sigma_{data}}$ in each time bin is shown.*

# 10
# Conclusion and perspective

In conclusion, this thesis presents a robust method for the extraction of the time-dependent asymmetries in $D^0 \to K^- K^+$ and $D^0 \to \pi^- \pi^+$ decays. The indirect $CP$ violation parameter $A_\Gamma$, defined as the difference in effective lifetimes between $D^0$ and $\bar{D}^0$ meson, is extracted. The significant detail of this analysis is that it takes advantage of $D^0$ mesons originating from B-meson decays where the muon is exploited as a tag to determine the charm-meson flavour.

The data used in this thesis is recorded with the LHCb experiment in pp collisions and corresponds to an integrated luminosity of $3.0\,\mathrm{fb}^{-1}$. The decay topology of the boosted B mesons and the muon in the final state are essential ingredients for an efficient trigger and offline selection at LHCb. The selection of the analysis is optimized by maximizing the significance and covering sufficiently the low and high $D^0$ decay time regions. The introduced raw asymmetries of the final states are very carefully and extensively studied and the interesting underlying time dependence of the asymmetry of the charm mesons is extracted. The systematic uncertainties have been validated and several checks on data-taking conditions, detector configuration and extraction method have been performed. The final results are determined to be

$$A_\Gamma(K^- K^+) = (-0.134 \pm 0.077(stat.)^{+0.026}_{-0.034}(syst.))\% \,,$$

$$A_\Gamma(\pi^- \pi^+) = (-0.092 \pm 0.145(stat.)^{+0.025}_{-0.033}(syst.))\%,$$

where the first uncertainty is statistical and the second systematic. They are consistent with the hypothesis of no indirect $CP$ violation. This measurement contributes significantly to the world average. The results of this thesis have been submitted for publication to the Journal of High Energy Physics (JHEP), Ref. [1].

In the future, with a rising center-of-mass energy at the hadron colliders, the b production cross section will increase more sharply than that of the charm production cross section. Thus this analysis will gain further significance compared to the study of promptly produced $D^0$ mesons. With a higher precision, one can further nail down the size of indirect CP violation, which will be an important search for the future. With more statistics, reducing the systematic uncertainties will become more important. Therefore, an additional $D^0$ flavour tag from the hadron in the $\bar{B}^0 \to D^{*+}(\to D^0 \pi^+) \mu^- \nu$ decay can be used to reduce the systematic uncertainty and further increase the sensitivity.

LHCb now provides, the world's most precise measurement of indirect $CP$ violation in this channels. This measurement can be used together with the other measured observables to obtain the effective parametrization in the charm sector. Additionally, these results can be used for the validation of theoretical methods and have an impact on the understanding of the perturbative and non-perturbative physics and the New Physics contribution in this sector.

# A
# Appendix

## A.1 Blinding

The analysis is done blinded, see Ref. [76] for general blinding technique. The blinding used in this analysis is described in the following. The important quantities for this analysis are the slopes of the raw asymmetries as function of decay time in the decay channels $D^0 \to K^- K^+$ and $D^0 \to \pi^- \pi^+$. The following function is fitted to the raw asymmetries

$$A_{CP}^{RAW}(t) = A_{CP}^{RAW}(0) - A_{\Gamma}^{initial} \cdot t/\tau. \tag{A.1}$$

$A_{\Gamma}^{initial}$ is the extracted value of the fit and is hidden to the author, while the offset $A_{CP}^{RAW}(0)$ is not blinded. Afterwards the data points in each bin [1] are corrected with this function

$$
\begin{aligned}
A_{CP}^{RAW,blind}(t_i) &= A_{CP}^{RAW}(t_i) + A_{\Gamma}^{initial} \cdot t_i/\tau \\
&= A_{CP}^{RAW}(0).
\end{aligned}
$$

This means that the data points are "blind" for the slope but not for the offset. Finally, the blinded value of $A_{\Gamma}$ (forced here to be $= 0$) is extracted with a second $\chi^2$ fit to the $A_{CP}^{RAW,blind}(t_i)$ data points to determine the sensitivity of each channel. With this blinding strategy one can as well judge the quality of the fit. Thereby, $A_{\Gamma}$ is different for both channels and therefore, one is also blinded to the difference in $A_{\Gamma}$ between the two channels. After the full validation of the blinded analysis the unblinding was performed. In the following, already the unblinded results are presented.

---

[1] The binning scheme is introduced in appendix B.12.

## A.2 Reweighting technique

For every single event a weight

$$W_i(a, b, c, ...) \tag{A.2}$$

can be used. This is utilised for different purposes, e.g. to subtract background or to reweight variables $a, b, c$ and to match the kinematic distribution. This means that in first order the expected yield would be $\sum w_i$.

However, the statistical power by applying the reweighting decreases even if the sum of weight would increase. Therefore, the uncertainty on the extracted parameter(in our case the slope in decay time) would be slightly underestimated. All the weights need a global correction factor for a correct error estimate.

$$W^{norm} = \frac{\sum W_i}{\sum W_i^2} \tag{A.3}$$

The correct weight has to be corrected with the effective weight factor:

$$W_i^{corr} = W^{norm} \cdot W_i \tag{A.4}$$

This leads to the following the corrected effective yield ( $N^{corr}$) result:

$$N^{corr} = \sum W_i^{corr} = W^{norm} \sum W_i \tag{A.5}$$

The weights $W_i$ and the correction factor $W^{norm}$ have to be calculated and applied on the datasets to obtain the correct estimate of a quantity, in our case this quantity is $A_\Gamma$.

### A.2.1 Weights for exponential function calculation

To show the weighting factor on a concrete example we define first the Likelihood that needs to be maximized. The same method can be calculated for different distribution. However here we want to show that this works first of all for an exponential distribution. In case of using weights every term gets a weight and is defined then as:

$$-\ln L(t_i) = -\sum w_i \ln(f(t_i|\tau)) \tag{A.6}$$

$$= -\sum w_i \ln(\frac{1}{\tau} \exp^{-\frac{t}{\tau}})$$

$$= \sum w_i \ln(\tau) + \sum \frac{w_i t_i}{\tau}$$

Searching for the maximum one finds $\tau$ to be:

$$\frac{\partial \ln L(t_i)}{\partial \tau}|_{<\tau>} = 0 \tag{A.7}$$

$$<\tau> = \frac{\sum w_i t_i}{\sum w_i}. \tag{A.8}$$

Now we want to obtain the variance $V[\tau]$. In the case without the weights this would be simply

$$V[\tau] = \int_0^\infty \frac{1}{\tau} \exp^{-\frac{t}{\tau}} (t - \tau)^2 dt \tag{A.9}$$

$$= \tau^2. \tag{A.10}$$

Without weights this result one gets also through second derivative near the maximum:

$$V_L[<\tau>] = \frac{1}{\frac{\partial^2 \ln L(t_i)}{\partial \tau^2}}|_{<\tau>}. \tag{A.11}$$

However this is not fully correct since the weights can have an arbitrary scaling factor in the likelihood leading to the same minimum but a different Variance. The correct variance one can calculate using the assumption that the $t_i$ values are uncorrelated one can use the central limit theorem applied on A.8 without weights and also obtains A.9. With weights this gives us a slightly different result.

$$V_{CLT}[<\tau>] = \sum (\frac{\partial \tau}{t_i})^2 \cdot V[t_i] \tag{A.12}$$

$$= \sum_i (\frac{w_i}{\sum_j w_j})^2 \cdot V[t_i] \tag{A.13}$$

$$\approx \tau^2 \cdot \sum_i (\frac{w_i}{\sum_j w_j})^2 \tag{A.14}$$

On the other side, evaluating the variance from the likelihood one obtains:

$$V_L[\tau] = \frac{1}{\frac{\partial^2 \ln L(t_i)}{\partial \tau^2}}|_{<\tau>} \tag{A.15}$$

$$= \frac{\tau^2}{\sum_i w_i}. \tag{A.16}$$

Both equations above are not equal because the variance estimator from the likelihood does not take weights properly into account and needs a correction factor. A correction factor $W^{norm}$ is needed, that can be applied on every weight $w_j$ and transform in this way $V_L[\tau]$ to $V_{CLT}[<\tau>]$. This factor one obtains with the following calculation

$$W^{norm} = \frac{V_L[<\tau>]}{V_{CLT}[<\tau>]} \tag{A.17}$$

$$= \frac{\frac{1}{\sum_j w_j}}{\sum_i (\frac{w_i}{\sum_j w_j})^2} \tag{A.18}$$

$$= \frac{\sum_j w_j}{\sum_i w_i^2}. \tag{A.19}$$

This result can be extended using other distribution.

## A.2.2 Test with general weights applied on data

Also to show the robustness and power of the reweigthing technique, an arbitrary fake $A_\Gamma$ with respect to the blinded value can be introduced through weights. Following decay time and muon charge dependent per event weights can be used to achieve this

$$\tau = 410.1 fs \tag{A.20}$$

$$\tau_\pm^{eff} = \frac{\tau}{(1 \pm A_\Gamma)} \tag{A.21}$$

$$W_i(t, \pm) = \frac{\frac{1}{\tau_\pm^{eff}} \exp(-\frac{t}{\tau_\pm^{eff}})}{\frac{1}{\tau} \exp(-\frac{t}{\tau})}. \tag{A.22}$$

Beside the consistency between 2011 and 2012 datasets in all the 3 channels, this test shows in Fig. A.1 that the fitter can extract the correct $A_\Gamma = 0.3\%$ values introduced through weights and a possible bias can be studied and properly extracted with this method.
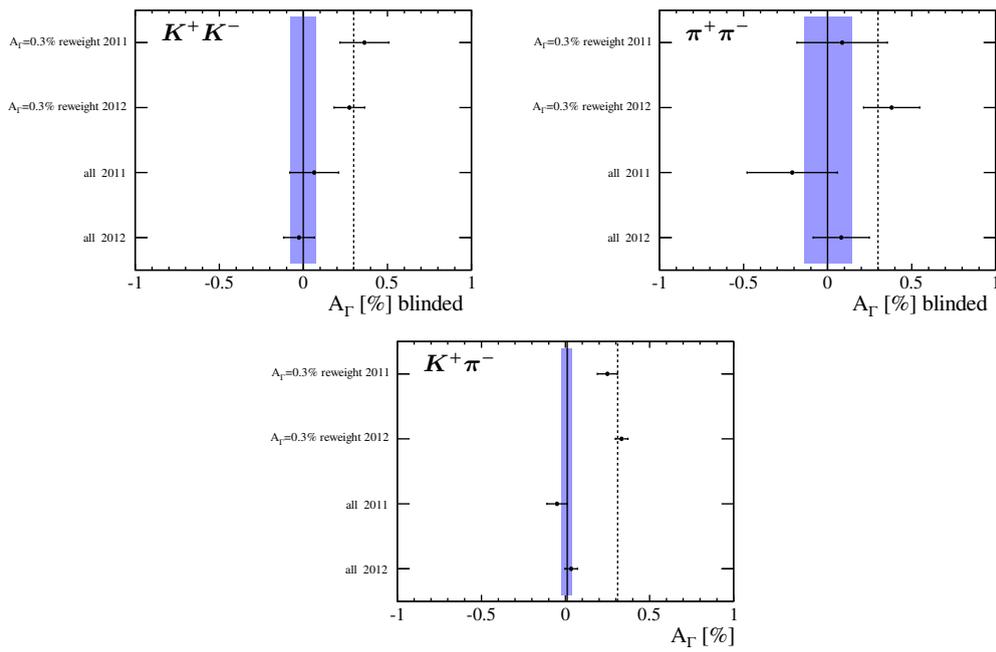
**Figure A.1:** *Data introduced $A_\Gamma = 0.3\%$ value through reweighting and the blinded $A_\Gamma = 0$ is shown. The line shows the combined mean $A_\Gamma$ value and the blue rectangle its $1\sigma$ confidence region. Left: for the $D^0 \to K^- K^+$ (blinded) channel, right: for the $\pi\pi$(blinded) channel, bottom center: for the $K\pi$(unblinded) channel.*
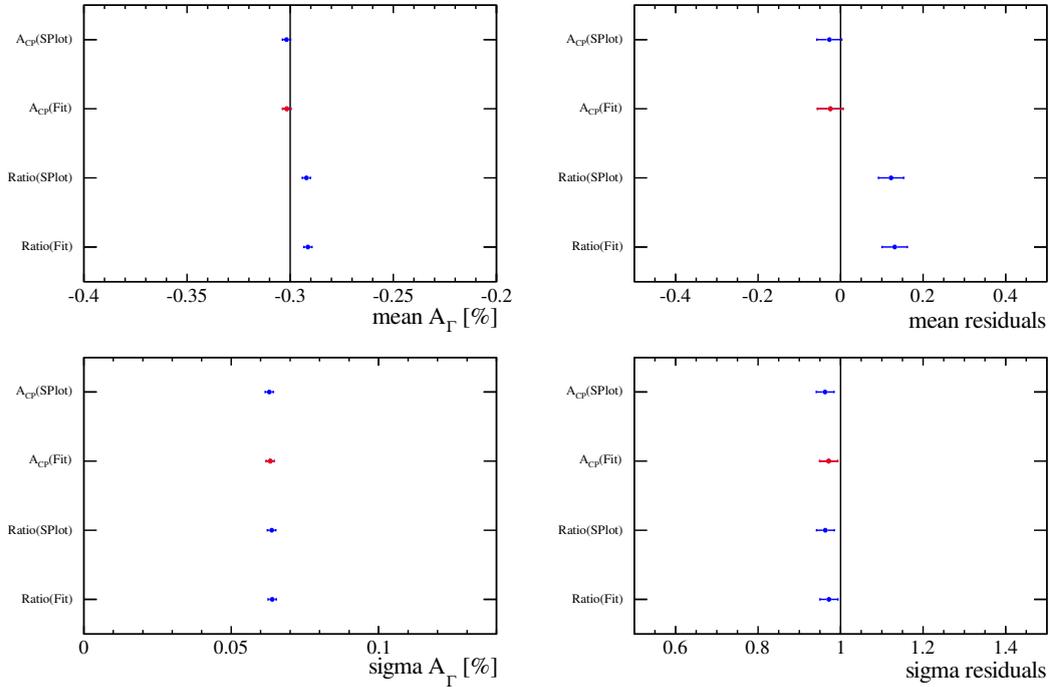
**Figure A.2:** *Comparison of different methods using 50 bins. The default fit method is highlighted in red. The vertical lines represent the input value used in the generation. Note that the results of the different methods are correlated as they use the same toy data sets as input.*

## A.3   Alternative $A_\Gamma$ extraction methods

Figure A.2 compares the result of our default method to several other methods. All methods have very similar precision. Also their residuals have the correct width. The ratio method has a small bias.

The same bias on $A_\Gamma$ is also seen in signal-only toy MC. The reason is that the ratio distribution gets very asymmetric with lower statistics, e.g. it can not become negative and has a tail to higher values. If the statistics varies in different time bins this can lead therefore to a bias on $A_\Gamma$. Choosing more bins (180 instead of 50) makes this effect larger which is shown in Fig. A.3. The bias in the ratio method is smaller than the statistical uncertainty on the extracted value. However, it becomes larger reducing the yields per bin, as statistical fluctuation become more pronounced.

## A.4   Detection and production asymmetry tables
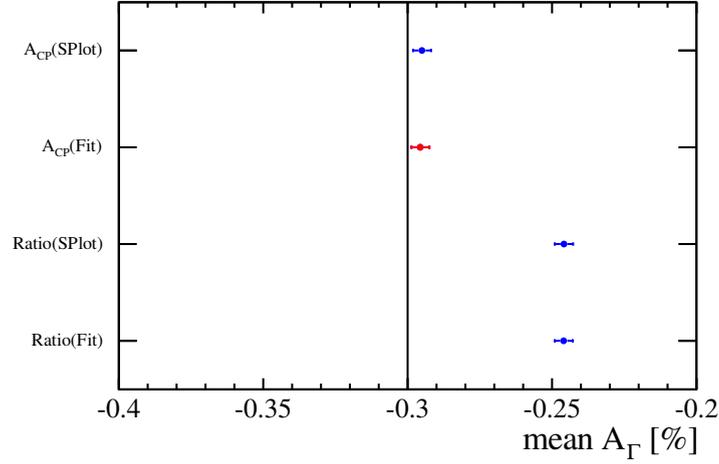
**Figure A.3:** *Comparison of different methods using 180 bins. The default fit method is highlighted in red. The vertical lines represent the input value used in the generation. Note that the results of the different methods are correlated as they use the same toy data sets as input.*

**Table A.1:** $\Delta A_{CP}^{RAW}$ and $\Delta A_\Gamma$ shifts for different time bins in $D^0 \to K^- K^+$ channel, extracted by reweighting using the transverse momentum of the muon.

| | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ |
|---|---|---|
| reweight MuPt bin1 | $-0.539 \pm 0.114$ | $0.006 \pm 0.077$ |
| reweight MuPt bin2 | $-0.539 \pm 0.114$ | $0.006 \pm 0.077$ |
| reweight MuPt bin3 | $-0.538 \pm 0.114$ | $0.006 \pm 0.077$ |
| reweight MuPt bin4 | $-0.539 \pm 0.114$ | $0.006 \pm 0.077$ |
| reweight MuPt bin5 | $-0.542 \pm 0.114$ | $0.002 \pm 0.077$ |
| reweight MuPt bin6 | $-0.543 \pm 0.115$ | $0.003 \pm 0.077$ |
| weighted mean from 2011 and 2012 fits | $-0.545 \pm 0.114$ | $0 \pm 0.077$ |
| max $\Delta A_\Gamma$ with $\sigma_\Delta = \sqrt{\sigma_{weights}^2 - \sigma_{mean}^2}$ | | $0.006 \pm 0.002$ |

**Table A.2:** $\Delta A_{CP}^{RAW}$ and $\Delta A_\Gamma$ shifts for different time bins in $D^0 \to \pi^- \pi^+$ channel, extracted by reweighting using the transverse momentum of the muon.

| | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ |
|---|---|---|
| reweight MuPt bin1 | $-0.737 \pm 0.214$ | $0.025 \pm 0.146$ |
| reweight MuPt bin2 | $-0.743 \pm 0.214$ | $0.024 \pm 0.146$ |
| reweight MuPt bin3 | $-0.75 \pm 0.214$ | $0.012 \pm 0.146$ |
| reweight MuPt bin4 | $-0.728 \pm 0.214$ | $0.033 \pm 0.146$ |
| reweight MuPt bin5 | $-0.787 \pm 0.218$ | $0.040 \pm 0.151$ |
| reweight MuPt bin6 | $-0.715 \pm 0.217$ | $0.040 \pm 0.149$ |
| weighted mean from 2011 and 2012 fits | $-0.763 \pm 0.214$ | $0 \pm 0.145$ |
| max $\Delta A_\Gamma$ with $\sigma_\Delta = \sqrt{\sigma_{weights}^2 - \sigma_{mean}^2}$ | | $0.040 \pm 0.042$ |

**Table A.3:** $\Delta A_{CP}^{RAW}$ and $\Delta A_\Gamma$ shifts for different time bins in $D^0 \to K^-\pi^+$ channel, extracted by reweighting using the transverse momentum of the muon.

|  | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ |
|---|---|---|
| reweight MuPt bin1 | $-1.5 \pm 0.0474$ | $0.0064 \pm 0.0318$ |
| reweight MuPt bin2 | $-1.5 \pm 0.0477$ | $0.0051 \pm 0.0321$ |
| reweight MuPt bin3 | $-1.5 \pm 0.0476$ | $0.0056 \pm 0.0321$ |
| reweight MuPt bin4 | $-1.5 \pm 0.0477$ | $0.0034 \pm 0.0321$ |
| reweight MuPt bin5 | $-1.5 \pm 0.0475$ | $0.0058 \pm 0.0318$ |
| reweight MuPt bin6 | $-1.5 \pm 0.0475$ | $0.0040 \pm 0.0318$ |
| weighted mean from 2011 and 2012 fits | $-1.51 \pm 0.0475$ | $0 \pm 0.0318$ |
| max $\Delta A_\Gamma$ with $\sigma_\Delta = \sqrt{\sigma_{weights}^2 - \sigma_{mean}^2}$ | | $0.0064 \pm 0.001$ |

**Table A.4:** $\Delta A_{CP}^{RAW}$ and $\Delta A_\Gamma$ shifts for different time bins in $D^0 \to K^-K^+$ channel, extracted by reweighting using the transverse momentum of the $D^0$ meson.

|  | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ |
|---|---|---|
| reweight D0Pt bin1 | $-0.547 \pm 0.114$ | $0.005 \pm 0.077$ |
| reweight D0Pt bin2 | $-0.588 \pm 0.121$ | $0.017 \pm 0.079$ |
| reweight D0Pt bin3 | $-0.589 \pm 0.121$ | $0.017 \pm 0.079$ |
| reweight D0Pt bin4 | $-0.584 \pm 0.121$ | $0.014 \pm 0.079$ |
| reweight D0Pt bin5 | $-0.586 \pm 0.121$ | $0.012 \pm 0.079$ |
| reweight D0Pt bin6 | $-0.585 \pm 0.121$ | $0.008 \pm 0.079$ |
| weighted mean from 2011 and 2012 fits | $-0.545 \pm 0.114$ | $0 \pm 0.077$ |
| max $\Delta A_\Gamma$ with $\sigma_\Delta = \sqrt{\sigma_{weights}^2 - \sigma_{mean}^2}$ | | $0.0174 \pm 0.0176$ |

**Table A.5:** $\delta A_{CP}^{RAW}$ and $\delta A_\Gamma$ shifts for different time bins in $D^0 \to \pi^-\pi^+$ channel, extracted by reweighting using the transverse momentum of the $D^0$ meson.

|  | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ |
|---|---|---|
| reweight D0Pt bin1 | $-0.775 \pm 0.214$ | $0.008 \pm 0.146$ |
| reweight D0Pt bin2 | $-0.778 \pm 0.213$ | $0.010 \pm 0.145$ |
| reweight D0Pt bin3 | $-0.78 \pm 0.213$ | $0.015 \pm 0.145$ |
| reweight D0Pt bin4 | $-0.75 \pm 0.214$ | $0.006 \pm 0.146$ |
| reweight D0Pt bin5 | $-0.724 \pm 0.216$ | $0.012 \pm 0.147$ |
| reweight D0Pt bin6 | $-0.746 \pm 0.215$ | $0.012 \pm 0.147$ |
| weighted mean from 2011 and 2012 fits | $-0.763 \pm 0.214$ | $0 \pm 0.145$ |
| max $\Delta A_\Gamma$ with $\sigma_\Delta = \sqrt{\sigma_{weights}^2 - \sigma_{mean}^2}$ | | $0.015 \pm 0.015$ |

**Table A.6:** $\Delta A_{CP}^{RAW}$ and $\Delta A_\Gamma$ shifts for different time bins in $D^0 \to K^- \pi^+$ channel, extracted by reweighting using the transverse momentum of the $D^0$ meson.

|  | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ |
|---|---|---|
| reweight D0Pt bin1 | $-1.5 \pm 0.0477$ | $0.0098 \pm 0.0322$ |
| reweight D0Pt bin2 | $-1.49 \pm 0.0475$ | $0.0098 \pm 0.0318$ |
| reweight D0Pt bin3 | $-1.5 \pm 0.0475$ | $0.0089 \pm 0.0318$ |
| reweight D0Pt bin4 | $-1.5 \pm 0.0475$ | $0.0072 \pm 0.0318$ |
| reweight D0Pt bin5 | $-1.5 \pm 0.0475$ | $0.0088 \pm 0.0318$ |
| reweight D0Pt bin6 | $-1.5 \pm 0.0475$ | $0.0079 \pm 0.0318$ |
| weighted mean from 2011 and 2012 fits | $-1.51 \pm 0.0475$ | $0 \pm 0.0318$ |
| max $\Delta A_\Gamma$ with $\sigma_\Delta = \sqrt{\sigma_{weights}^2 - \sigma_{mean}^2}$ | | $0.0098 \pm 0.0051$ |

## A.5 $D^0/\overline{D}^0$ mass fits parameter evolution

As presented in Chapter 7, the decay-time evolution of several fit parameters is presented here. Following $D^0$-and $\overline{D}^0$-mass parameters, which are the Gaussian means $\mu_{1,2}$, the average mass resolution

$$< \sigma_{M_{D^0}} >= \sqrt{c_1 \sigma_1^2 + c_2 \sigma_2^2 + c_3 \sigma_3^2},$$

the average of the combinatorial background parameter for $D^0$ and $\overline{D}^0$ decays $\lambda = \frac{\alpha^+ + \alpha^-}{2}$, the signal and background yields as well as their ratio $N_{sig}/N_{bkg}$, are visualizes in Figs. A.4-A.9. Further, the bin-to-bin fluctuation is rather small, which confirms a rather stable fit model.
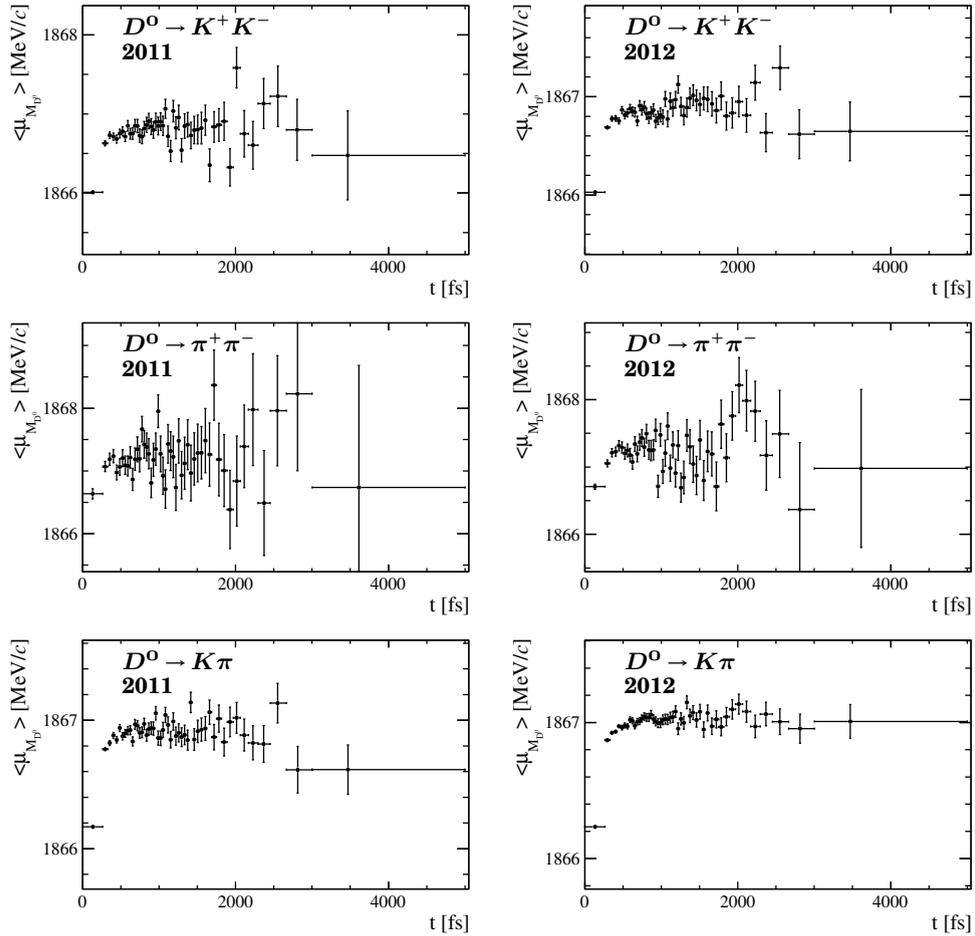


**Figure A.4:** *Time dependence of fit parameters (explanation see text) for 2011 (left) and 2012 (right) data.*
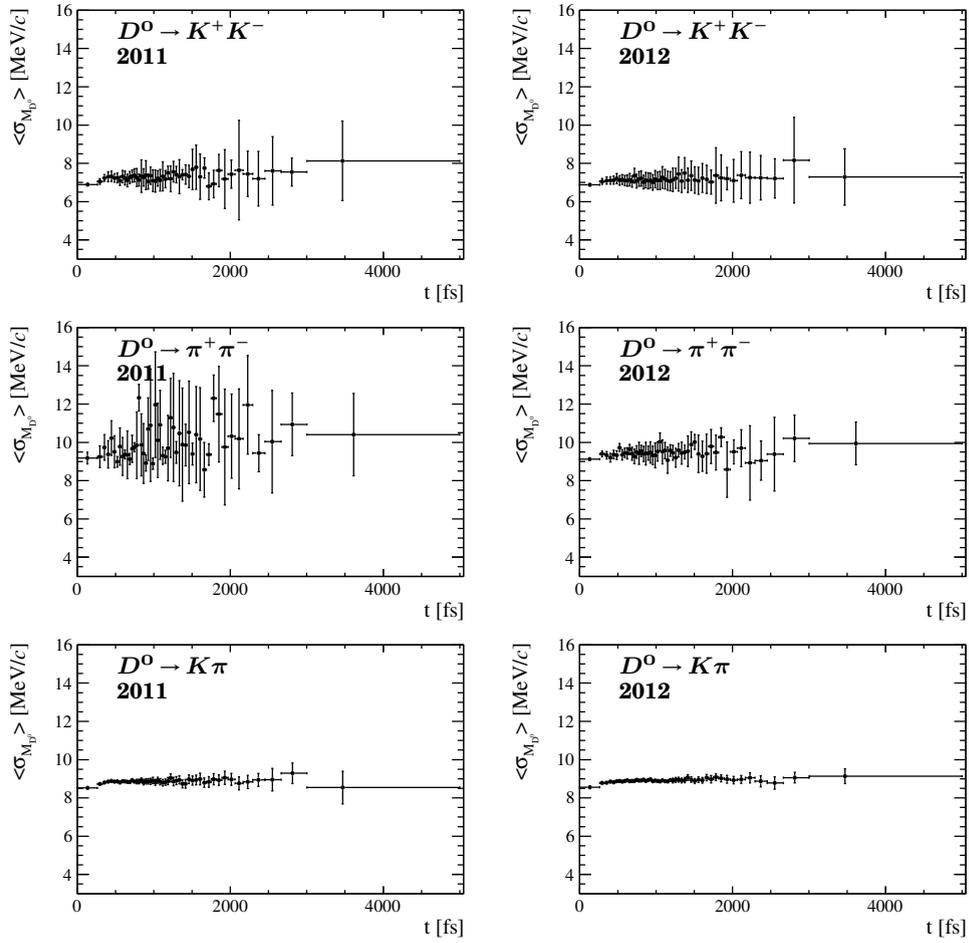
**Figure A.5:** *Time dependence of fit parameters (explanation see text) for 2011 (left) and 2012 (right) data.*
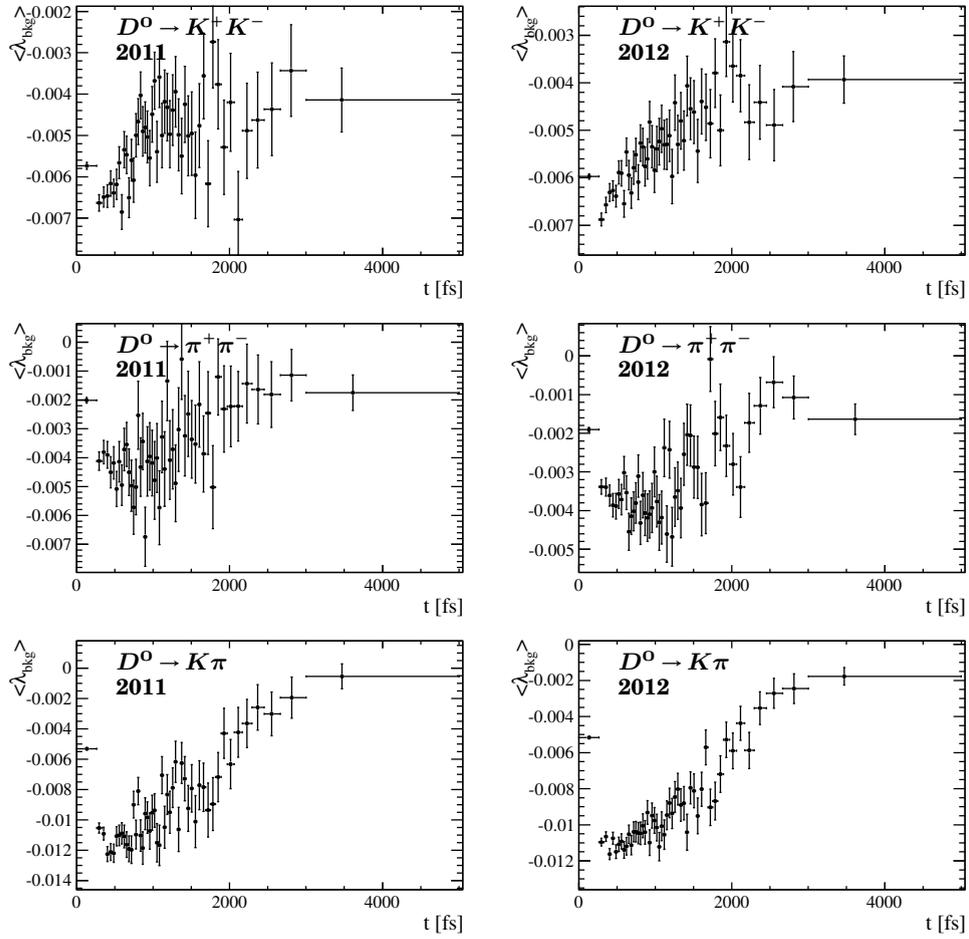
**Figure A.6:** *Time dependence of fit parameters (explanation see text) for 2011 (left) and 2012 (right) data.*
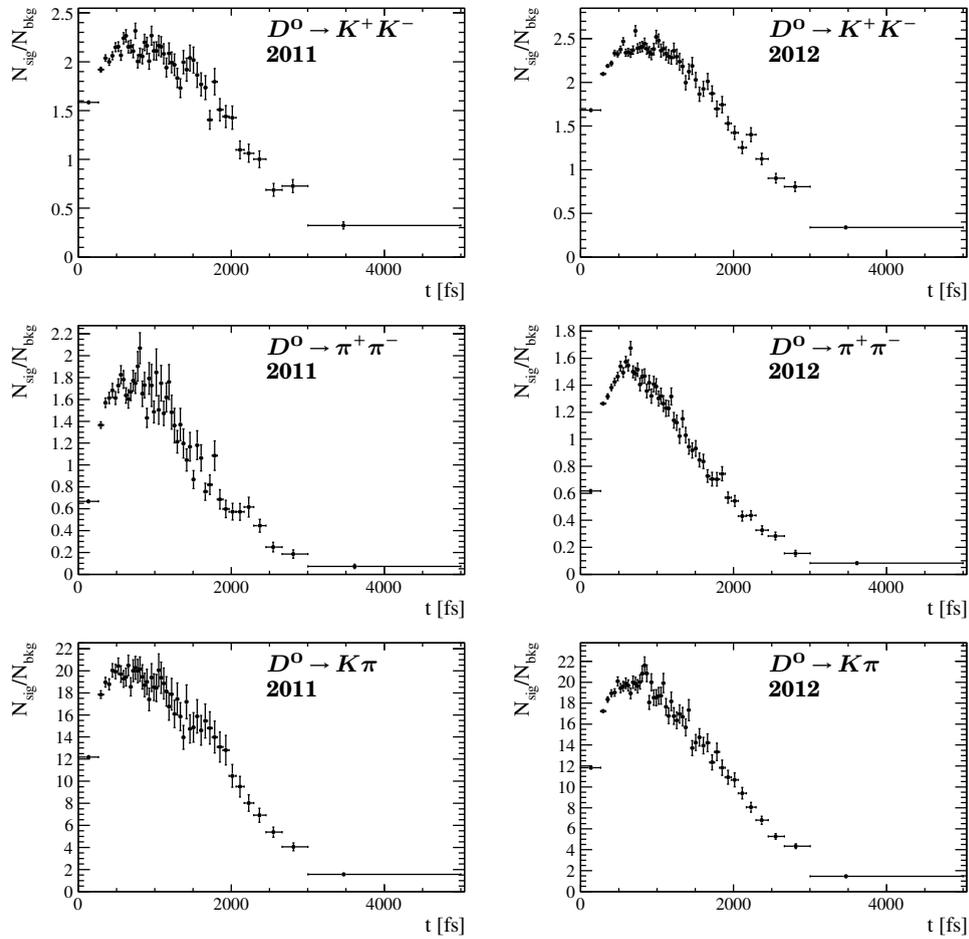
**Figure A.7:** *Time dependence of fit parameters (explanation see text) for 2011 (left) and 2012 (right) data.*
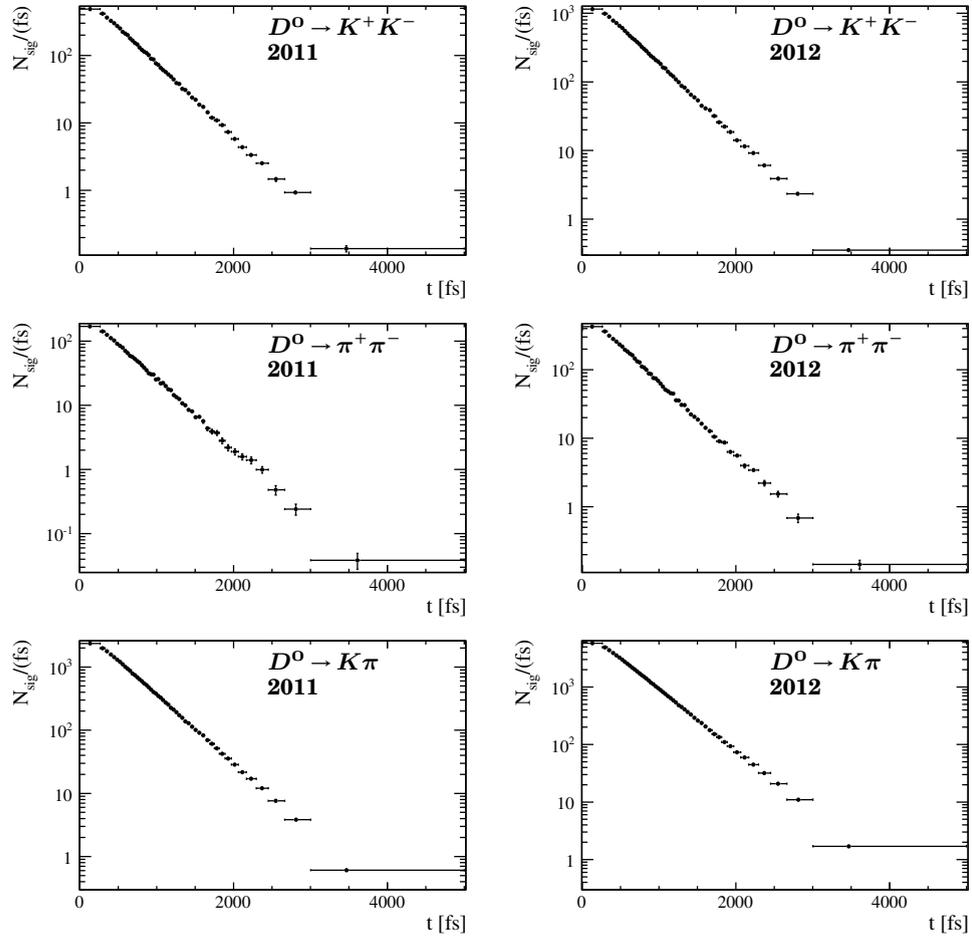
**Figure A.8:** *Time dependence of fit parameters (explanation see text) for 2011 (left) and 2012 (right) data.*
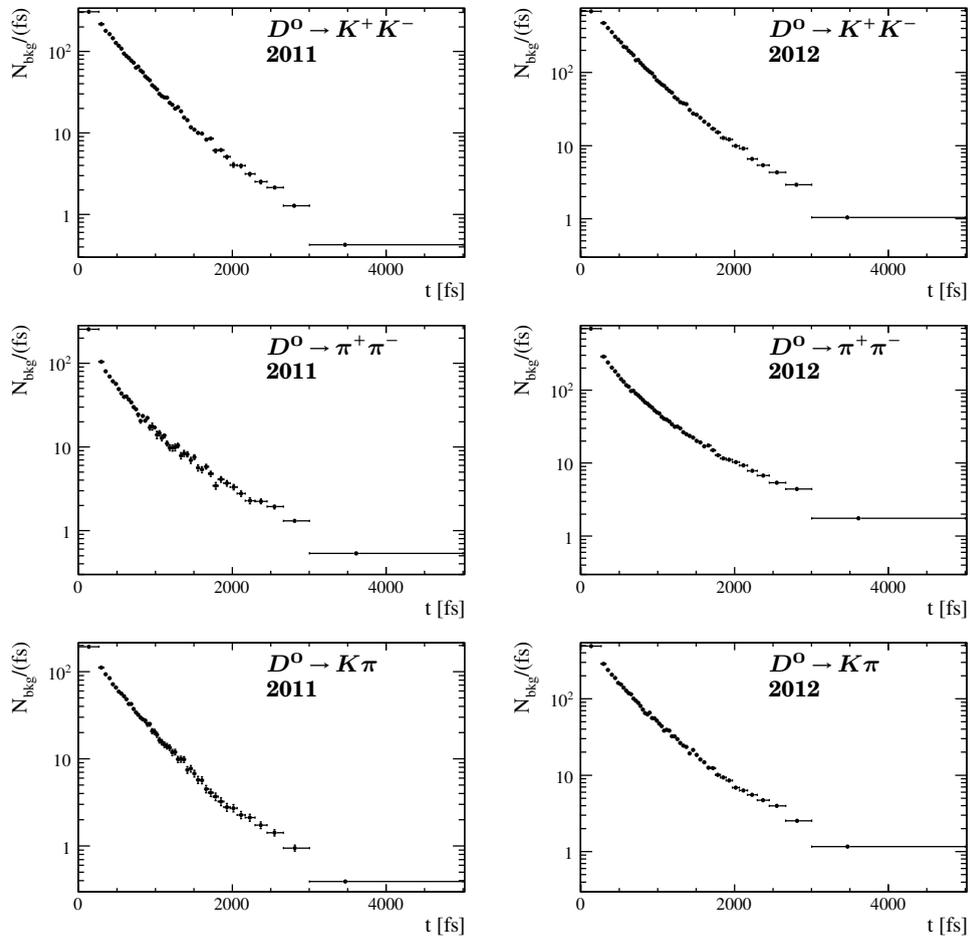
**Figure A.9:** *Time dependence of fit parameters (explanation see text) for 2011 (left) and 2012 (right) data.*

## A.6 Tables for cross check

### A.6.1 Model dependence

**Table A.7:** *Model dependence in $D^0 \to K^- K^+$ channel.*

| | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ |
|---|---|---|---|---|
| 3Gauss+Lin 2012 | $-0.577 \pm 0.136$ | $-0.0146 \pm 0.0911$ | $-0.539 \pm 0.115$ | $0.010 \pm 0.077$ |
| 3Gauss+Lin 2011 | $-0.446 \pm 0.213$ | $0.0738 \pm 0.145$ | | |
| 3Gauss+Pol2 2012 | $-0.576 \pm 0.135$ | $-0.0141 \pm 0.0902$ | $-0.537 \pm 0.114$ | $0.011 \pm 0.076$ |
| 3Gauss+Pol2 2011 | $-0.442 \pm 0.212$ | $0.0752 \pm 0.144$ | | |
| 2Gauss+Exp 2012 | $-0.589 \pm 0.136$ | $-0.0253 \pm 0.0909$ | $-0.547 \pm 0.114$ | $0 \pm 0.077$ |
| 2Gauss+Exp 2011 | $-0.444 \pm 0.213$ | $0.0634 \pm 0.145$ | | |
| Gaus+CB+Exp 2012 | $-0.587 \pm 0.136$ | $-0.0253 \pm 0.0909$ | $-0.546 \pm 0.114$ | $0 \pm 0.077$ |
| Gaus+CB+Exp 2011 | $-0.444 \pm 0.213$ | $0.0638 \pm 0.145$ | | |
| weighted mean from 2011 and 2012 fits | | | $-0.545 \pm 0.114$ | $0 \pm 0.077$ |

**Table A.8:** *Model dependence in $D^0 \to \pi^- \pi^+$ channel.*

| | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ |
|---|---|---|---|---|
| 3Gauss+Lin 2012 | $-0.617 \pm 0.254$ | $0.102 \pm 0.172$ | $-0.768 \pm 0.214$ | $0.007 \pm 0.146$ |
| 3Gauss+Lin 2011 | $-1.14 \pm 0.4$ | $-0.233 \pm 0.274$ | | |
| 3Gauss+Pol2 2012 | $-0.584 \pm 0.252$ | $0.105 \pm 0.17$ | $-0.799 \pm 0.213$ | $-0.002 \pm 0.144$ |
| 3Gauss+Pol2 2011 | $-1.33 \pm 0.397$ | $-0.274 \pm 0.272$ | | |
| 2Gauss+Exp 2012 | $-0.591 \pm 0.254$ | $0.107 \pm 0.172$ | $-0.749 \pm 0.214$ | $0.005 \pm 0.145$ |
| 2Gauss+Exp 2011 | $-1.14 \pm 0.399$ | $-0.254 \pm 0.274$ | | |
| Gaus+CB+Exp 2012 | $-0.598 \pm 0.254$ | $0.108 \pm 0.172$ | $-0.755 \pm 0.214$ | $0.006 \pm 0.145$ |
| Gaus+CB+Exp 2011 | $-1.14 \pm 0.399$ | $-0.253 \pm 0.273$ | | |
| weighted mean from 2011 and 2012 fits | | | $-0.763 \pm 0.214$ | $0 \pm 0.145$ |

**Table A.9:** *Model dependence in $D^0 \rightarrow K^- \pi^+$ channel.*

| | $A_{CP}^{RAW}(0)[\%]$ | $A_\Gamma[\%]$ | $A_{CP}^{RAW}(0)[\%]$ | $A_\Gamma[\%]$ |
|---|---|---|---|---|
| 3Gauss+Lin 2012 | $-1.45 \pm 0.0563$ | $0.0305 \pm 0.0375$ | $-1.51 \pm 0.0476$ | $0.0071 \pm 0.0318$ |
| 3Gauss+Lin 2011 | $-1.66 \pm 0.0887$ | $-0.0528 \pm 0.06$ | | |
| 3Gauss+Pol2 2012 | $-1.48 \pm 0.0562$ | $0.0297 \pm 0.0374$ | $-1.55 \pm 0.0474$ | $0.0061 \pm 0.0317$ |
| 3Gauss+Pol2 2011 | $-1.73 \pm 0.0886$ | $-0.054 \pm 0.0599$ | | |
| 2Gauss+Exp 2012 | $-1.45 \pm 0.0563$ | $0.0343 \pm 0.0375$ | $-1.5 \pm 0.0475$ | $0.0100 \pm 0.0318$ |
| 2Gauss+Exp 2011 | $-1.65 \pm 0.0886$ | $-0.0521 \pm 0.06$ | | |
| Gaus+CB+Exp 2012 | $-1.45 \pm 0.0563$ | $0.0319 \pm 0.0375$ | $-1.51 \pm 0.0475$ | $0.0084 \pm 0.0318$ |
| Gaus+CB+Exp 2011 | $-1.65 \pm 0.0886$ | $-0.0513 \pm 0.0599$ | | |
| weighted mean from 2011 and 2012 fits | | | $-1.51 \pm 0.0475$ | $0.00926 \pm 0.0318$ |

## A.6.2 Binning of decay time

**Table A.10:** *Binning dependence in $D^0 \to K^- K^+$ channel.*

| | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ |
|---|---|---|---|---|
| 26bins 2012 | $-0.596 \pm 0.136$ | $-0.0356 \pm 0.0914$ | $-0.554 \pm 0.115$ | $-0.0098 \pm 0.0773$ |
| 26bins 2011 | $-0.452 \pm 0.213$ | $0.0555 \pm 0.145$ | | |
| 180bins 2012 | $-0.589 \pm 0.135$ | $-0.0286 \pm 0.0905$ | $-0.536 \pm 0.114$ | $0.0066 \pm 0.0766$ |
| 180bins 2011 | $-0.408 \pm 0.212$ | $0.0957 \pm 0.144$ | | |
| weighted mean from 2011 and 2012 fits | | | $-0.545 \pm 0.114$ | $0 \pm 0.077$ |

**Table A.11:** *Binning dependence in $D^0 \to \pi^- \pi^+$ channel.*

| | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ |
|---|---|---|---|---|
| 26bins 2012 | $-0.617 \pm 0.255$ | $0.0974 \pm 0.173$ | $-0.77 \pm 0.215$ | $-0.006 \pm 0.146$ |
| 26bins 2011 | $-1.15 \pm 0.401$ | $-0.27 \pm 0.276$ | | |
| 180bins 2012 | $-0.606 \pm 0.253$ | $0.0895 \pm 0.17$ | $-0.757 \pm 0.213$ | $0.005 \pm 0.144$ |
| 180bins 2011 | $-1.13 \pm 0.397$ | $-0.207 \pm 0.271$ | | |
| weighted mean from 2011 and 2012 fits | | | $-0.763 \pm 0.214$ | $0 \pm 0.145$ |

**Table A.12:** *Binning dependence in $D^0 \to K^- \pi^+$ channel.*

| | $A_{CP}^{RAW}(0)[\%]$ | $A_\Gamma[\%]$ | $A_{CP}^{RAW}(0)[\%]$ | $A_\Gamma[\%]$ |
|---|---|---|---|---|
| 26bins 2012 | $-1.45 \pm 0.0564$ | $0.0323 \pm 0.0377$ | $-1.51 \pm 0.0476$ | $0.0066 \pm 0.0320$ |
| 26bins 2011 | $-1.66 \pm 0.0889$ | $-0.0591 \pm 0.0603$ | | |
| 180bins 2012 | $-1.45 \pm 0.0562$ | $0.0353 \pm 0.0374$ | $-1.5 \pm 0.0474$ | $0.0109 \pm 0.0317$ |
| 180bins 2011 | $-1.65 \pm 0.0883$ | $-0.0512 \pm 0.0596$ | | |
| weighted mean from 2011 and 2012 fits | | | $-1.51 \pm 0.0475$ | $0.0093 \pm 0.0318$ |

### A.6.3 Dependence of $A_\Gamma$ on magnet polarity

**Table A.13:** *$A_\Gamma$ fit result for different magnet polarities and years for $D^0 \to K^- K^+$ decays.*

|  | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ |
|---|---|---|---|---|
| Mag. up 2012 | $-0.688 \pm 0.189$ | $-0.0663 \pm 0.127$ | $-0.582 \pm 0.163$ | $-0.004 \pm 0.110$ |
| Mag. up 2011 | $-0.267 \pm 0.326$ | $0.186 \pm 0.221$ | | |
| Mag. down 2012 | $-0.477 \pm 0.195$ | $0.0194 \pm 0.131$ | $-0.511 \pm 0.16$ | $0.002 \pm 0.108$ |
| Mag. down 2011 | $-0.581 \pm 0.28$ | $-0.0349 \pm 0.191$ | | |
| All 2012 | $-0.586 \pm 0.136$ | $-0.0261 \pm 0.0909$ | $-0.545 \pm 0.114$ | $0 \pm 0.077$ |
| All 2011 | $-0.442 \pm 0.213$ | $0.066 \pm 0.145$ | | |

**Table A.14:** *$A_\Gamma$ fit result for different magnet polarities and years for $D^0 \to \pi^- \pi^+$ decays.*

|  | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ |
|---|---|---|---|---|
| Mag. up 2012 | $-0.195 \pm 0.353$ | $0.297 \pm 0.239$ | $-0.499 \pm 0.306$ | $0.233 \pm 0.208$ |
| Mag. up 2011 | $-1.42 \pm 0.616$ | $0.031 \pm 0.421$ | | |
| Mag. down 2012 | $-1.03 \pm 0.364$ | $-0.0746 \pm 0.246$ | $-0.977 \pm 0.299$ | $-0.167 \pm 0.203$ |
| Mag. down 2011 | $-0.865 \pm 0.523$ | $-0.363 \pm 0.358$ | | |
| All 2012 | $-0.603 \pm 0.254$ | $0.103 \pm 0.172$ | $-0.763 \pm 0.214$ | $0 \pm 0.145$ |
| All 2011 | $-1.16 \pm 0.399$ | $-0.263 \pm 0.273$ | | |

**Table A.15:** *$A_\Gamma$ fit result for different magnet polarities and years for $D^0 \to K^- \pi^+$ decays.*

|  | $A_{CP}^{RAW}(0)[\%]$ | $A_\Gamma[\%]$ | $A_{CP}^{RAW}(0)[\%]$ | $A_\Gamma[\%]$ |
|---|---|---|---|---|
| Mag. up 2012 | $-1.37 \pm 0.0782$ | $-0.0256 \pm 0.0522$ | $-1.49 \pm 0.0678$ | $-0.0417 \pm 0.0454$ |
| Mag. up 2011 | $-1.87 \pm 0.136$ | $-0.0914 \pm 0.0919$ | | |
| Mag. down 2012 | $-1.53 \pm 0.0811$ | $0.0963 \pm 0.0541$ | $-1.52 \pm 0.0666$ | $0.0591 \pm 0.0447$ |
| Mag. down 2011 | $-1.48 \pm 0.117$ | $-0.0204 \pm 0.0791$ | | |
| All 2012 | $-1.45 \pm 0.0562$ | $0.0332 \pm 0.0375$ | $-1.51 \pm 0.0475$ | $0.0094 \pm 0.0318$ |
| All 2011 | $-1.65 \pm 0.0886$ | $-0.0519 \pm 0.06$ | | |

### A.6.4 *B* decay time

**Table A.16:** *Fit results for $A_\Gamma$ for different B decay time regions in the $D^0 \to K^- K^+$ channel.*

|  | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ | $A_{CP}^{RAW}(0)[\%]$ | $<\Delta A_\Gamma>[\%]$ |
|---|---|---|---|---|
| low *B* decay time 2012 | $-0.428 \pm 0.175$ | $0.0253 \pm 0.115$ | $-0.463 \pm 0.148$ | $0.043 \pm 0.097$ |
| low *B* decay time 2011 | $-0.549 \pm 0.274$ | $0.0881 \pm 0.183$ | | |
| high *B* decay time 2012 | $-0.8 \pm 0.215$ | $-0.0823 \pm 0.148$ | $-0.647 \pm 0.181$ | $-0.047 \pm 0.126$ |
| high *B* decay time 2011 | $-0.273 \pm 0.336$ | $0.0402 \pm 0.235$ | | |
| weighted mean from 2011 and 2012 fits | | | $-0.545 \pm 0.114$ | $0 \pm 0.077$ |

**Table A.17:** *Fit results for $A_\Gamma$ for different B decay time regions in the $D^0 \to \pi^- \pi^+$ channel.*

|  | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ |
|---|---|---|---|---|
| low *B* decay time 2012 | $-0.629 \pm 0.326$ | $0.116 \pm 0.217$ | $-0.682 \pm 0.275$ | $0.046 \pm 0.184$ |
| low *B* decay time 2011 | $-0.813 \pm 0.512$ | $-0.133 \pm 0.346$ | | |
| high *B* decay time 2012 | $-0.576 \pm 0.402$ | $0.0853 \pm 0.279$ | $-0.919 \pm 0.339$ | $-0.055 \pm 0.236$ |
| high *B* decay time 2011 | $-1.77 \pm 0.633$ | $-0.406 \pm 0.441$ | | |
| weighted mean from 2011 and 2012 fits | | | $-0.763 \pm 0.214$ | $0 \pm 0.145$ |

**Table A.18:** *Fit results for $A_\Gamma$ for different B decay time regions in the $D^0 \to K^- \pi^+$ channel.*

|  | $A_{CP}^{RAW}(0)[\%]$ | $A_\Gamma[\%]$ | $A_{CP}^{RAW}(0)[\%]$ | $A_\Gamma[\%]$ |
|---|---|---|---|---|
| low *B* decay time 2012 | $-1.37 \pm 0.0725$ | $0.0899 \pm 0.0475$ | $-1.43 \pm 0.0612$ | $0.0515 \pm 0.0403$ |
| low *B* decay time 2011 | $-1.57 \pm 0.114$ | $-0.0467 \pm 0.076$ | | |
| high *B* decay time 2012 | $-1.57 \pm 0.0893$ | $-0.0644 \pm 0.0612$ | $-1.63 \pm 0.0754$ | $-0.0604 \pm 0.0519$ |
| high *B* decay time 2011 | $-1.76 \pm 0.141$ | $-0.0505 \pm 0.0975$ | | |
| weighted mean from 2011 and 2012 fits | | | $-1.51 \pm 0.0475$ | $0.00926 \pm 0.0318$ |

### A.6.5 Dependence of $A_\Gamma$ on number of reconstructed primary vertices

**Table A.19:** $A_\Gamma$ *fit result for different number of reconstructed primary vertices in* $D^0 \to \pi^-\pi^+$ *decays.*

|  | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ |
|---|---|---|---|---|
| $NPV > 2$ 2012 | $-0.0761 \pm 0.47$ | $0.172 \pm 0.321$ | $-0.139 \pm 0.412$ | $0.105 \pm 0.282$ |
| $NPV > 2$ 2011 | $-0.349 \pm 0.858$ | $-0.122 \pm 0.591$ | | |
| $NPV = 2$ 2012 | $-0.473 \pm 0.406$ | $0.232 \pm 0.274$ | $-0.408 \pm 0.343$ | $0.24 \pm 0.232$ |
| $NPV = 2$ 2011 | $-0.246 \pm 0.641$ | $0.262 \pm 0.44$ | | |
| $NPV = 1$ 2012 | $-1.33 \pm 0.449$ | $-0.134 \pm 0.301$ | $-1.7 \pm 0.366$ | $-0.337 \pm 0.246$ |
| $NPV = 1$ 2011 | $-2.42 \pm 0.632$ | $-0.751 \pm 0.428$ | | |
| weighted mean from 2011 and 2012 fits | | | $-0.763 \pm 0.214$ | $0 \pm 0.145$ |

**Table A.20:** $A_\Gamma$ *fit result for different number of reconstructed primary vertices in* $D^0 \to K^-K^+$ *decays.*

|  | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ |
|---|---|---|---|---|
| $NPV > 2$ 2012 | $-0.67 \pm 0.251$ | $-0.176 \pm 0.169$ | $-0.596 \pm 0.22$ | $-0.144 \pm 0.149$ |
| $NPV > 2$ 2011 | $-0.349 \pm 0.458$ | $-0.031 \pm 0.317$ | | |
| $NPV = 2$ 2012 | $-0.463 \pm 0.216$ | $0.00215 \pm 0.145$ | $-0.541 \pm 0.182$ | $-0.046 \pm 0.123$ |
| $NPV = 2$ 2011 | $-0.732 \pm 0.34$ | $-0.167 \pm 0.231$ | | |
| $NPV = 1$ 2012 | $-0.645 \pm 0.242$ | $0.098 \pm 0.161$ | $-0.494 \pm 0.197$ | $0.182 \pm 0.132$ |
| $NPV = 1$ 2011 | $-0.201 \pm 0.339$ | $0.352 \pm 0.229$ | | |
| weighted mean from 2011 and 2012 fits | | | $-0.545 \pm 0.114$ | $0 \pm 0.077$ |

**Table A.21:** $A_\Gamma$ *fit result for different number of reconstructed primary vertices in $D^0 \to K^- \pi^+$ decays.*

| | $A_{CP}^{RAW}(0)[\%]$ | $A_\Gamma[\%]$ | $A_{CP}^{RAW}(0)[\%]$ | $A_\Gamma[\%]$ |
|---|---|---|---|---|
| $NPV > 2$ 2012 | $-1.52 \pm 0.103$ | $-0.00876 \pm 0.0692$ | $-1.51 \pm 0.0908$ | $0.0091 \pm 0.0611$ |
| $NPV > 2$ 2011 | $-1.44 \pm 0.19$ | $0.0726 \pm 0.13$ | | |
| $NPV = 2$ 2012 | $-1.36 \pm 0.0897$ | $0.0648 \pm 0.0598$ | $-1.5 \pm 0.0758$ | $0.0207 \pm 0.0508$ |
| $NPV = 2$ 2011 | $-1.87 \pm 0.142$ | $-0.0936 \pm 0.0963$ | | |
| $NPV = 1$ 2012 | $-1.5 \pm 0.101$ | $0.029 \pm 0.0672$ | $-1.51 \pm 0.0822$ | $-0.0067 \pm 0.0548$ |
| $NPV = 1$ 2011 | $-1.55 \pm 0.141$ | $-0.0776 \pm 0.0946$ | | |
| weighted mean from 2011 and 2012 fits | | | $-1.51 \pm 0.0475$ | $0.0093 \pm 0.0318$ |

## A.7 MC samples

To study efficiencies in several selection steps and also their dependencies versus true decay time a Monte Carlo sample is generated. In the MC12 production the events are generated with `Pythia8(Sim08a)` and go through the full detector simulation. Thereby, the events are reconstructed with Reco14a and the trigger with the TCK `Trig0x409f0045` is used. [2] This sample is also used for the determination of the resolution for the $D^0$ decay times and are summarized in the following Tab. A.22. The sample accepts signal $B^+$ or $B^0$ (and charge-conjugation ) decays

**Table A.22:** *Generated and fully simulated MC12 samples*

| Signal | Event Type | Decay descriptor Name | Event Number |
|---|---|---|---|
| $B^0 \to (D^0 \to \pi^- \pi^+)\mu + X$ | 11874414 | `Bd_Dstmunu,pipi` | 10M |
| $B^0 \to (D^0 \to K^- K^+)\mu + X$ | 11874424 | `Bd_Dstmunu,KK` | 10M |
| $B^0 \to (D^0 \to K^- \pi^+)\mu + X$ | 11874004 | `Bd_Dstmunu,Kpi` | 20M |
| $B^+ \to (D^0 \to \pi^- \pi^+)\mu + X$ | 12873412 | `Bu_D0munu,pipi` | 10M |
| $B^+ \to (D^0 \to K^- K^+)\mu + X$ | 12873422 | `Bu_D0munu,KK` | 10M |
| $B^+ \to (D^0 \to K^- \pi^+)\mu + X$ | 12873002 | `Bu_D0munu,Kpi` | 10M |

with the muon in the LHCb acceptance. Thereby, the $D^0$ mesons and muons can originate from different decay modes. In the decay descriptor [3] the $B^+$ mesons are forced to decay in $D^0$ directly or over other possible charm resonance where the branching fractions are taken from PDG [54]. Also the prompt charm decays are generated in the same way. For the $B^0$ the transition to $D^0$ is only possible over higher charm resonances. Not only B decays to charm and muon decays are contained in the sample, but also tau decays (with a fraction of about 5% compared to muons). To enlarge the taus sample they are forced to decay in muons and can also be studied using this simulation.

---

[2] Also MC11 samples have been studied and give compatible results.

[3] see e.g. `$DECFILESROOT/dkfiles/Bu_D0munu,KK=cocktail,D0muInAcc,BRcorr1.dec`

# A.8    $A_{CP}^{indir}$ calculation

The calculation of $A_{CP}^{indir}$ described in Sec. 3 can be followed here step by step.

$$A_{CP}^{indir} = -2|R_m|^2 \frac{\left[(1+|\lambda_f|^2)\,x\,\Im(\lambda_f)+(1-|\lambda_f|^2)\,y\,\Re(\lambda_f)\right]}{\left(|R_m|^2+|\lambda_f|^2\right)^2} \tag{A.23}$$

$$A_{CP}^{indir} = -2(1+A_m)\frac{\left[(1+(1+A_m)(1+A_d))\,x\,\Im(\lambda_f)+(1-(1+A_m)(1+A_d))\,y\,\Re(\lambda_f)\right]}{((1+A_m)+(1+A_d)(1+A_m))^2}.$$

Using

$$|\lambda_f|^2 = |R_m|^2|R_f|^2 = (1+A_m)(1+A_d)$$

one gets

$$A_{CP}^{indir} = -2(1+A_m)\frac{\left[(1+(1+A_m)(1+A_d))\,x\,\Im(\lambda_f)+(1-(1+A_m)(1+A_d))\,y\,\Re(\lambda_f)\right]}{((1+A_m)+(1+A_d)(1+A_m))^2} \tag{A.24}$$

$$= -2\frac{\left[(1+(1+A_m)(1+A_d))\,x\,\Im(\lambda_f)+(1-(1+A_m)(1+A_d))\,y\,\Re(\lambda_f)\right]}{(1+(1+A_d))^2(A_m+1)}$$

$$= 2\sqrt{\frac{A_d+1}{A_m+1}}\frac{\left[(1+(1+A_m)(1+A_d))\,x\,(\sin\phi)+(1-(1+A_m)(1+A_d))\,y\,(\cos\phi)\right]}{(1+(1+A_d))^2}.$$

Now we can expand this equation till the second order in $A_m$ and $A_d$, and for small values one can obtain:

$$A_{CP}^{indir} \approx x\,\sin(\phi)-\frac{1}{2}(A_m+A_d)\,y\,cos(\phi). \tag{A.25}$$
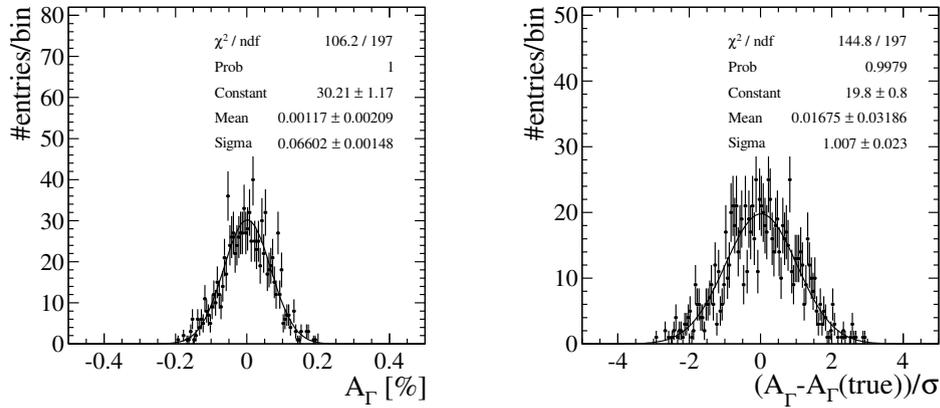
**Figure A.10:** *Toy study for a fit with 50 bins. A generated input value of $A_\Gamma = 0\%$ is used. Measured $A_\Gamma$ distribution (left) and residuals (right).*

## A.9 Toy fit results

This section shows the toy distributions used in Sec. 8.3 for different true $A_\Gamma$ values. The decay time for $D^0$ and $\overline{D}^0$ is generated according to a given value of $A_\Gamma$. Then we smear the generated decay time with the resolution function of Fig. 8.5 and apply an acceptance function as observed in Sec. 5.3. For every $A_\Gamma$ value 1000 toys are generated. Finally, we split the samples using the default binning with 50 bins and fit extracted distributions with a Gaussian to extract the bias.
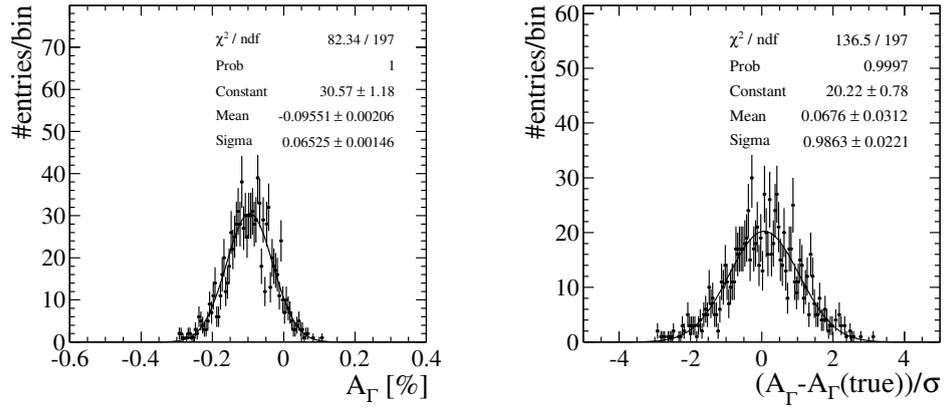
**Figure A.11:** *Toy study for a fit with 50 bins. A generated input value of $A_\Gamma = -0.1\%$ is used. Measured $A_\Gamma$ distribution (left) and residuals (right). A small shift used in the systematic study is visible.*
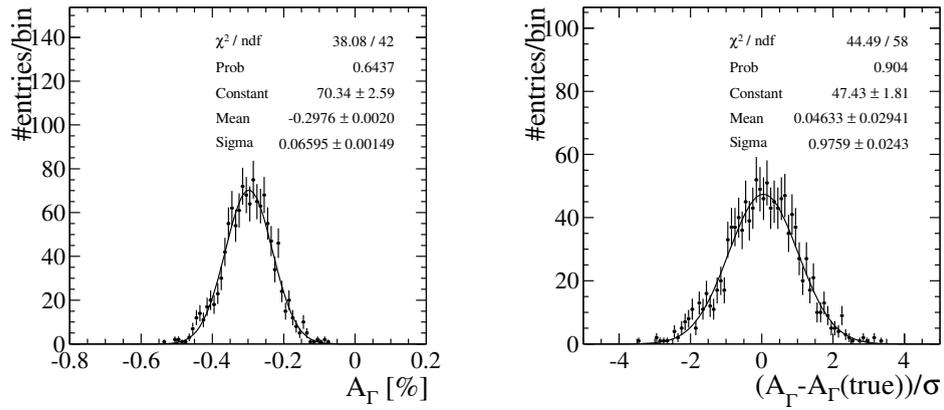


**Figure A.12:** *Toy study for a fit with 50 bins. A generated $A_\Gamma = -0.3\%$ is used. Measured $A_\Gamma$ distribution (left) and residuals (right). A small shift used in the systematic study is visible.*

## A.10   Average mistag $\omega(t)$ used on data

For data the $\bar{\omega}(t)$ term is used directly in the fit. The fit results determined in the fit include the $(1-2\bar{\omega}(t))$ correction term. The fits describe the data well in all the three channels and do not deviate significantly from the fit without $\bar{\omega}(t)$ correction. The results are summarized in Tab. A.23. Therefore one can estimate systematic uncertainty for all the channel separately:

- $D^0 \to K^- K^+$: $|\Delta A_\Gamma(\text{mistag})| = 0.0065\%$

- $D^0 \to \pi^- \pi^+$: $|\Delta A_\Gamma(\text{mistag})| = 0.0144\%$

- $D^0 \to K^- \pi^+$: $|\Delta A_\Gamma(\text{mistag})| = 0.0166\%$

**Table A.23:** *Mistag dependence in all three channels.*

| $D^0 \to \pi^- \pi^+$ | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ |
|---|---|---|---|---|
| omega corr 2012 | $-0.611 \pm 0.26$ | $0.121 \pm 0.179$ | $-0.775 \pm 0.219$ | $0.0144 \pm 0.152$ |
| omega corr 2011 | $-1.18 \pm 0.408$ | $-0.256 \pm 0.285$ | | |
| weighted mean from 2011 and 2012 fits | | | $-0.763 \pm 0.214$ | $0 \pm 0.145$ |
| $D^0 \to K^- K^+$ | | | | |
| omega corr 2012 | $-0.602 \pm 0.139$ | $-0.0245 \pm 0.0954$ | $-0.555 \pm 0.117$ | $0.00646 \pm 0.0807$ |
| omega corr 2011 | $-0.441 \pm 0.218$ | $0.0846 \pm 0.151$ | | |
| weighted mean from 2011 and 2012 fits | | | $-0.545 \pm 0.114$ | $0 \pm 0.077$ |
| $D^0 \to K^- \pi^+$ | $A_{CP}^{RAW}(0)[\%]$ | $\Delta A_\Gamma[\%]$ | $< A_{CP}^{RAW}(0) > [\%]$ | $\Delta A_\Gamma[\%]$ |
| omega corr 2012 | $-1.46 \pm 0.0577$ | $0.0516 \pm 0.0395$ | $-1.52 \pm 0.0487$ | $0.0166 \pm 0.0334$ |
| omega corr 2011 | $-1.67 \pm 0.0909$ | $-0.0395 \pm 0.063$ | | |
| weighted mean from 2011 and 2012 fits | | | $-1.51 \pm 0.0475$ | $0 \pm 0.0318$ |

## A.11 Time integrated fit results

Here the results of the global fit using a simfit are shown in numbers, see Tab. A.24

**Table A.24:** *Parameters determined from simultaneous unbinned likelihood fit in $D^0 \to \pi^- \pi^+$ channel for 2011 dataset*

| Parameter | Value | Uncertainty |
|---|---|---|
| $A^{CP}_{\pi\pi}$ | -0.00750 | 0.00266 |
| $\sigma_1$ | 4.294 | 0.612 |
| $f^1_{sig}$ | 0.0484 | 0.0322 |
| $f^2_{sig}$ | 0.7829 | 0.0877 |
| $s_{\sigma_2}$ | 1.898 | 0.186 |
| $s_{\sigma_3}$ | 3.180 | 0.282 |
| $\mu_1$ | 1866.99 | 0.056 |
| $\delta_\mu$ | -1.649 | 0.899 |
| $\lambda^-_{bkg} \times 10^5$ | -270.5 | 20.3 |
| $\lambda^+_{bkg} \times 10^5$ | -261.1 | 21.7 |
| $\mu^-_{refl.bkg}$ | 1792.73 | 2.31 |
| $\mu^+_{refl.bkg}$ | 1790.38 | 3.26 |
| $\sigma^-_{refl.bkg}$ | 10.97 | 1.43 |
| $\sigma^+_{refl.bkg}$ | 11.78 | 1.78 |
| $f^-_{refl.bkg}$ | 0.06137 | 0.00544 |
| $f^+_{refl.bkg}$ | 0.06500 | 0.00590 |
| $nBkg^-$ | 111653 | 828 |
| $nBkg^+$ | 111990 | 850 |
| $nSig^-$ | 108747 | 826 |

## A.12 Fully truth matched $D^0$ decay-time resolution

Fig. B.6 shows the resolution for decays where B is fully truth matched to the reconstructed B. The results are very similar like explained in Sec. 8.3. The taus have a broader and also asymmetric resolution shifting the reconstructed decay times towards the lower values, as explained in the next appendix section A.13.

**Table A.25:** *Parameters determined from simultaneous unbinned likelihood fit in $D^0 \to \pi^- \pi^+$ channel for 2012 dataset*

| Parameter | Value | Uncertainty |
|---|---|---|
| $A_{\pi\pi}^{CP}$ | -0.00639 | 0.00169 |
| $\sigma_1$ | 6.789 | 0.282 |
| $f_{sig}^1$ | 0.421 | 0.119 |
| $f_{sig}^2$ | 0.614 | 0.173 |
| $s_{\sigma_2}$ | 1.412 | 0.136 |
| $s_{\sigma_3}$ | 1.8103 | 0.0929 |
| $\mu_1$ | 1867.12 | 0.051 |
| $\delta_\mu$ | -1.75 | 1.01 |
| $\lambda_{bkg}^- \times 10^5$ | -258.86 | 9.02 |
| $\lambda_{bkg}^+ \times 10^5$ | -244.46 | 9.52 |
| $\mu_{refl.bkg}^-$ | 1794.22 | 1.04 |
| $\mu_{refl.bkg}^+$ | 1789.79 | 1.91 |
| $\sigma_{refl.bkg}^-$ | 9.977 | 0.675 |
| $\sigma_{refl.bkg}^+$ | 12.197 | 0.947 |
| $f_{refl.bkg}^-$ | 0.05994 | 0.00223 |
| $f_{refl.bkg}^+$ | 0.06231 | 0.00244 |
| $nBkg^-$ | 314596 | 927 |
| $nBkg^+$ | 317090 | 950 |
| $nSig^-$ | 275541 | 906 |

**Table A.26:** *Parameters determined from simultaneous unbinned likelihood fit in $D^0 \to K^- K^+$ channel for 2011 dataset*

| Parameter | Value | Uncertainty |
|---|---|---|
| $A_{KK}^{CP}$ | -0.00362 | 0.00141 |
| $\sigma_1$ | 4.881 | 0.229 |
| $f_{sig}^1$ | 0.3344 | 0.0980 |
| $f_{sig}^2$ | 0.8765 | 0.0270 |
| $s_{\sigma_2}$ | 1.4415 | 0.0281 |
| $s_{\sigma_3}$ | 2.635 | 0.193 |
| $\mu_1$ | 1866.44 | 0.014 |
| $\delta_\mu$ | -1.010 | 0.415 |
| $\lambda_{bkg}^- \times 10^5$ | -584.75 | 8.29 |
| $\lambda_{bkg}^+ \times 10^5$ | -565.62 | 8.24 |
| $nBkg^-$ | 174950 | 681 |
| $nBkg^+$ | 176758 | 685 |
| $nSig^-$ | 315726 | 778 |

**Table A.27:** *Parameters determined from simultaneous unbinned likelihood fit in $D^0 \to K^- K^+$ channel for 2012 dataset*

| Parameter | Value | Uncertainty |
|---|---|---|
| $A_{KK}^{CP}$ | -0.004081 | 0.000900 |
| $\sigma_1$ | 4.551 | 0.277 |
| $f_{sig}^1$ | 0.2013 | 0.0830 |
| $f_{sig}^2$ | 0.8392 | 0.0252 |
| $s_{\sigma_2}$ | 1.4397 | 0.0346 |
| $s_{\sigma_3}$ | 2.5038 | 0.0978 |
| $\mu_1$ | 1866.47 | 0.009 |
| $\delta_\mu$ | -0.287 | 0.126 |
| $\lambda_{bkg}^- \times 10^5$ | -601.68 | 5.50 |
| $\lambda_{bkg}^+ \times 10^5$ | -594.21 | 5.47 |
| $nBkg^-$ | 387248 | 977 |
| $nBkg^+$ | 391627 | 984 |
| $nSig^-$ | 761173 | 1153 |
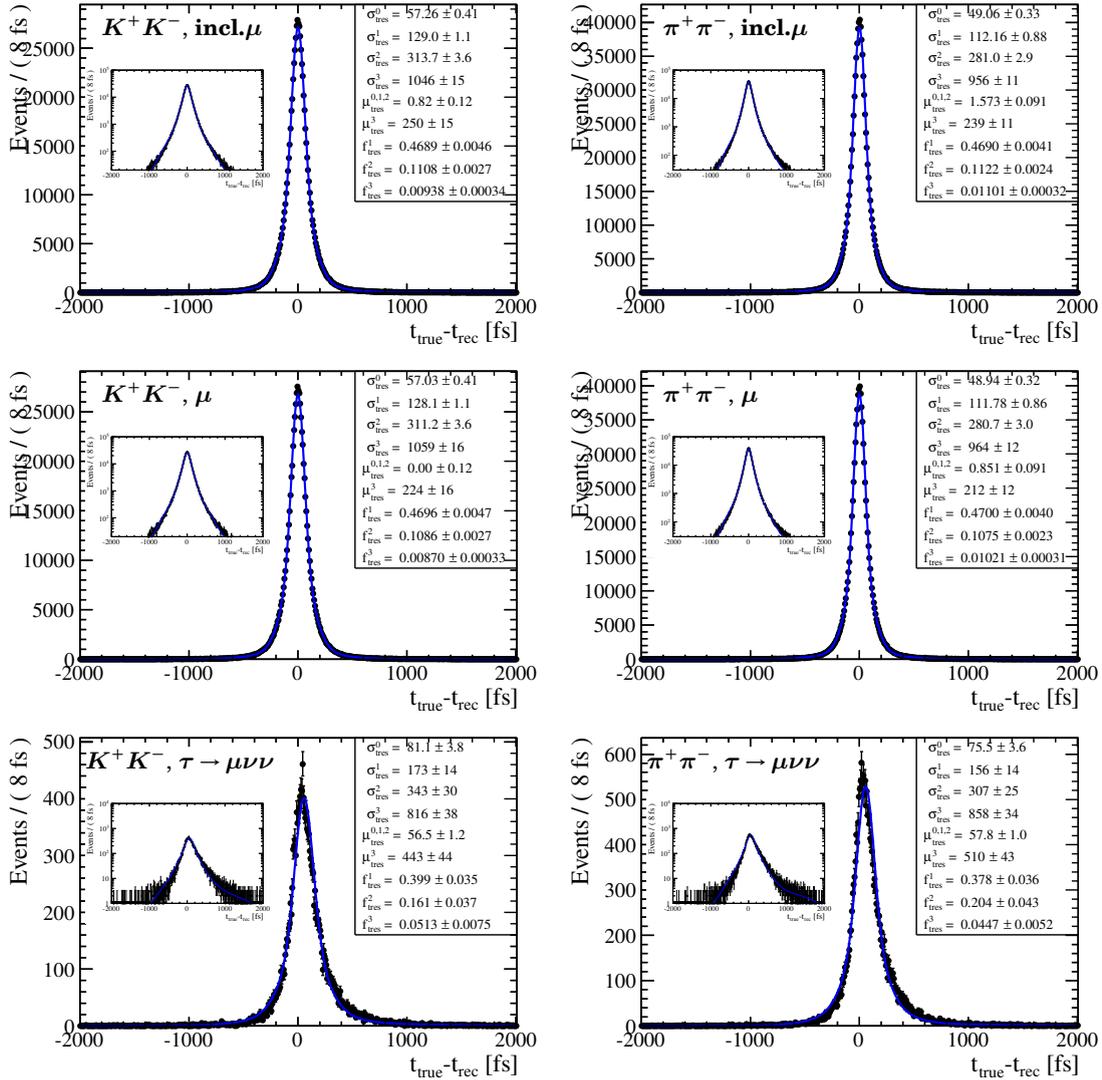
**Figure A.13:** $D^0$ decay-time resolution with core resolution for different subsamples for the two main channels(left:$D^0 \rightarrow K^- K^+$, right:$D^0 \rightarrow \pi^- \pi^+$), inclusive muons on top, only muons in the middle and taus at the bottom.
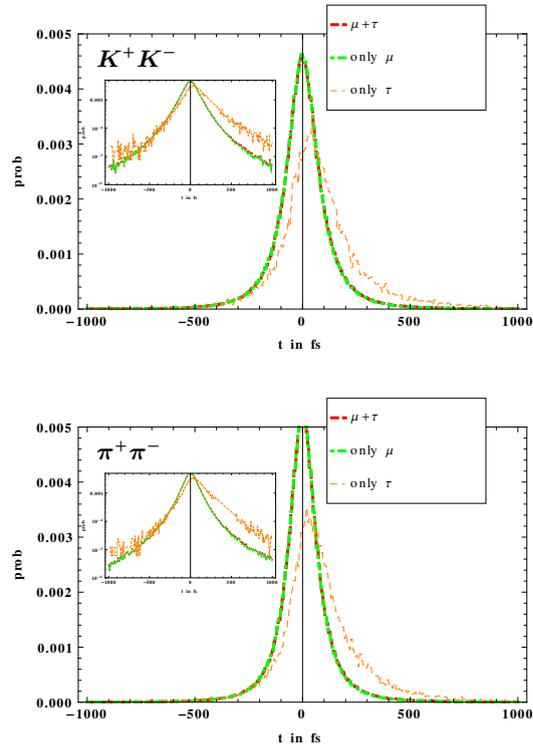
**Figure A.14:** *Resolution determined on MC12 simulation sample, for $D^0 \to K^- K^+$ on top and $D^0 \to \pi^- \pi^+$ on bottom. Inclusive $\mu$ sample (including $\tau$ decays) is used.*

## A.13   Tau decays, lifetime bias

The tau decays can also be used for the tag of the produced $D^0$. Like explained in section 8.3 they have a small impact on $A_\Gamma$. Since not all of them can be reconstructed and also the is an additional shift to lower decay times as explained in Fig.A.16
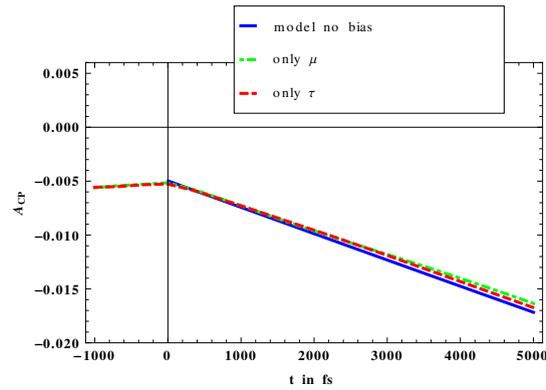
**Figure A.15:** *The tau dacays show no significant bias to the raw Asymmetry.*
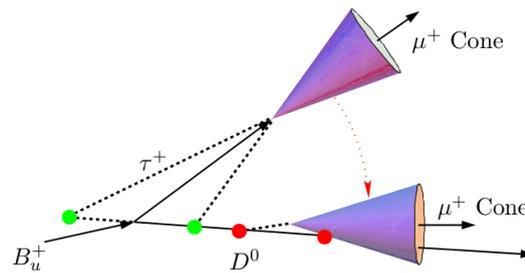


**Figure A.16:** *The $B^+ \to \tau^+(\to \nu_\tau \bar{\nu}_\mu \mu^+) \, \bar{\nu}_\tau \, D^0$ decay. The $\mu$ cone shows the possible distributions of the muon momentum. In case the cone is very closed to the $D^0$ flight direction $\mu$ reconstructed vertex(here shown in red) always underestimates the $D^0$ production Vertex. The measured lifetime is biased to lower values.*

# B Appendix

## B.1  Modified Kolmogorov-Smirnov probability calculation

Compared to the $\chi^2/ndf$-test Kolmogorov-Smirnov test(K-S) test can tests better the agreement between a sample and an expected shape of distributions and if there are some tendencies. It can test if a quantity follows an empirical distribution function, with $f(t)$ as a $\mathscr{PDF}$:

$$F(t) = \int_0^t f(t')dt' \tag{B.1}$$

This can be rewritten in a binned form:

$$F_n(t) = \sum_{t_{bin}<t}^{n} N_{t_{bin}}/n \tag{B.2}$$

where n as a number of bins.

The function that we fit to the asymmetry is not a $\mathscr{PDF}$. However, this procedure can be generalized to the pulls distribution, since the pull distribution satisfies the Glivenko-Cantelli theorem:

$$d_n = \sup_t |\sum_{t_i<t} pull_i/n| \tag{B.3}$$

$$P(\lim_{n\to\infty} d_n = 0) = 1.$$

Therefore, similar to Kolmogorov-Smirnov test, a modified K-S probability can be calculated from the pull distribution [1]. The tested quantity is then

$$x = \sqrt{n} \max_t |F_n(t) - F(t)| \tag{B.4}$$

In our case the hypothesis would be that the pulls are flat distributed. This would be in the case of the pulls

$$F_n(t) - F(t) = \sum_{t_i<t} pull_i/n \tag{B.5}$$

---

[1] Remember that also the binning is chosen in a way that every bin has the same significance on $A_\Gamma$ measurement
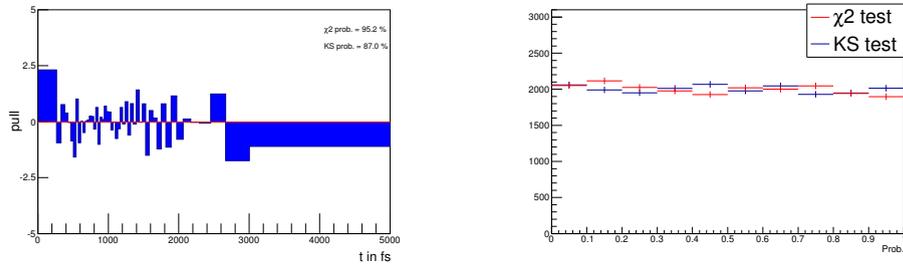
**Figure B.1:** *Left: One toy with generated pulls with center 0. Right: Flat probability distributions for KS and $\chi^2$ test from 20 ktoys.*
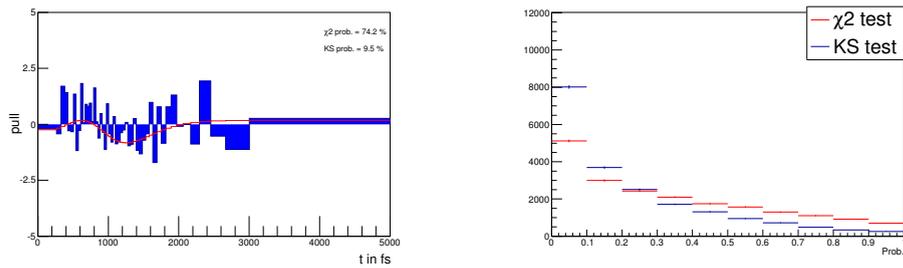


**Figure B.2:** *Left: One toy with generated pulls with center changing with sinus shape. Right: Probability distributions for KS and $\chi^2$ tests from 20 ktoys.*

and therefore the tested quantity is

$$x = \max_t |\sum_{t_i < t} pull_i|/\sqrt{n}. \tag{B.6}$$

This lead as described by Kolmogorov distribution to a probability of obtaining a value $x > x_0$

$$P(x > x_0) = 2 \sum_{j=1}^{\infty} (-1)^{j-1} \cdot exp(-2(jx_0)^2). \tag{B.7}$$

Using this, a probability for the hypothesis, that the pulls are flat distributed in time, can be now calculated directly from the pull distribution.

The free parameters of the fit have an impact on degrees of freedom and we do not expect that the probability distribution is totally flat. However, one can compare the values between different fits with each other. The fully correct approach requires the calibration of probabilities making them flat, using $P'(p) = \int_{Prob=0}^{p} P(Prob)dProb$. The example below shows how this works. The pulls were generated with a Gaussian with width one.

After calibrating, one can obtain the flat distributions. Fig. B.1 shows the flat distributed probabilities. After generation of the pulls with a sinus shape function, KS-test probability distribution prefers more the 0 values. Fig. B.2 shows that KS-test more often rejects the flat hypothesis compared to the $\chi^2$. This example illustrates that the modified KS test works well in discovering some shapes that deviate from the flat distribution.

## B.2   Correction factor for $A_\Gamma$ from mistag and asymmetries

As we saw in previous sections , the raw asymmetry can be written as

$$A_{RAW}^{CP}(t) = (1 - 2\bar{\omega})(A_{CP}^{dir} - A_\Gamma \frac{t}{\tau} + A_{mu} + A_{prod}) - \Delta\omega. \qquad (B.8)$$

If we now consider the asymmetries $A_\Sigma = A_{CP}^{dir} + A_{mu} + A_{prod}$ ,the mistag probabilities $\bar{\omega}$ and mistag differences $\Delta\omega$ to have some time dependence. One can always choose a time window small enough starting at a time $t_0$ that one can parametrize them with Taylor expansion as a linear function:

In case the quantities are time independent we get then directly the slope

$$\frac{dA_{RAW}^{CP}(t)}{dt} = -\frac{A_\Gamma}{\tau}$$

and therefore $A_\Gamma$.

However, in case this quantities are time dependent. Using this one can insert it into Eq.

$$
\begin{aligned}
A_{RAW}^{CP}(t) = &\ A_\Sigma(t_0) - \Delta\omega(t_0) - 2A_\Sigma(t_0)\bar{\omega}(t_0) \\
&+ (-\frac{A_\Gamma}{\tau} + \frac{dA_\Sigma}{dt} - \frac{\Delta\omega}{dt} - 2A_\Sigma(t_0)\frac{d\bar{\omega}}{dt} + 2\frac{A_\Gamma}{\tau}\bar{\omega}(t_0) - 2\frac{dA_\Sigma}{dt}\bar{\omega}(t_0))t \\
&+ (2\frac{A_\Gamma}{\tau}\frac{d\bar{\omega}}{dt} - 2\frac{dA_\Sigma}{dt}\frac{d\bar{\omega}}{dt})t^2
\end{aligned}
$$

and therefore:

$$\frac{dA_{RAW}^{CP}(t)}{dt} = -\frac{A_\Gamma - \Delta_{correction}(t)}{\tau}$$

with

$$
\begin{aligned}
\Delta_{correction}(t) = &-\frac{dA_\Sigma}{dt} - \frac{\Delta\omega}{dt} - 2A_\Sigma(t_0)\frac{d\bar{\omega}}{dt} + 2\frac{A_\Gamma}{\tau}\bar{\omega}(t_0) - 2\frac{dA_\Sigma}{dt}\bar{\omega}(t_0) \\
&+ 4\frac{A_\Gamma}{\tau}\frac{d\bar{\omega}}{dt}t - 4\frac{dA_\Sigma}{dt}\frac{d\bar{\omega}}{dt}
\end{aligned}
$$

and

$$\Delta_{correction}(t) = -\frac{dA_\Sigma}{dt} - \frac{\Delta\omega}{dt} - 2A_\Sigma(t_0)\frac{d\bar{\omega}}{dt} + 2\frac{A_\Gamma}{\tau}\bar{\omega}(t_0) - 2\frac{dA_\Sigma}{dt}\bar{\omega}(t_0)$$

$$\Delta_{correction}(t) = (-\underbrace{\frac{dA_\Sigma}{dt}}_{<0.0005} - \underbrace{\frac{\Delta\omega}{dt}}_{<0.015\%} - \underbrace{2A_\Sigma(t_0)\frac{d\bar{\omega}}{dt}}_{<0.001\%} + \underbrace{2\frac{A_\Gamma}{\tau}\bar{\omega}(t_0)}_{<0.015\%} - \underbrace{2\frac{dA_\Sigma}{dt}\bar{\omega}(t_0)}_{<0.01\cdot0.2\%})$$

## B.3   $\Delta\omega$ and $\omega$ time dependence

$$P(\overline{D}^0_{tag}|D^0 \wedge B_{rec}) = \frac{N(\mu^-_{tag} \wedge D^0 \to K^+\pi^-)}{N(D^0 \to K^+\pi^-)} \tag{B.9}$$

$$= \frac{N(\mu^-_{tag} \wedge (WS + RS)) - N(\mu^-_{tag} \wedge WS)}{N(RS)} \tag{B.10}$$

$$\approx \frac{N(\mu^-_{tag} \wedge (WS + RS))}{N(RS + WS)} - \frac{N(WS)}{N(RS)} \tag{B.11}$$
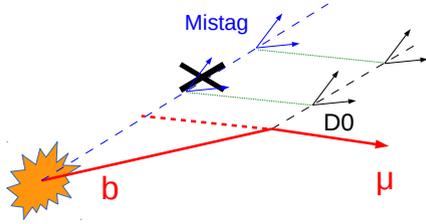
**Figure B.3:** *Relative to the right tagged events, the fraction of mistagged (background) events rises going to larger decay times.*
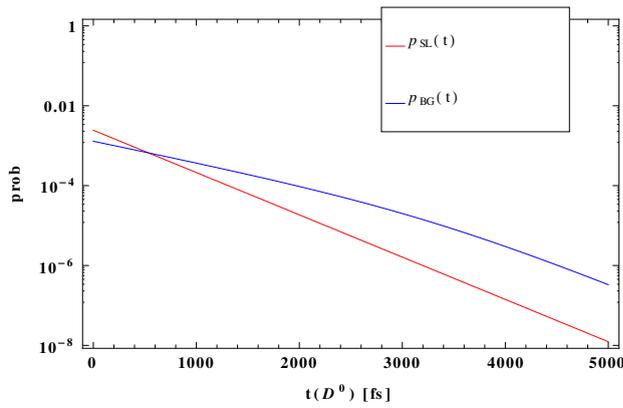


**Figure B.4:** *Due to the fact that the mistagged $D^0$ (background here in blue) is closer to the PV, it has a larger acceptance correction compared to the right tagged (shown in red)*

## B.4  $\omega$ **time dependence toy, explanation**

A possible explanation for the rise of the mistag probabilities with decay time is given in Fig. B.3. It shows that at larger lifetimes more mistags can be associated. To model this, following distribution models (Fig. B.4) for signal (red) and background (blue) were used. Background events have larger acceptances and time shift, e.g. due to mistag association and selection.

Using the model above (Fig. B.4) we assigned an efficiency (here 50%) of finding the right tag. In case the right tag is not found we associate it with 1% probability with a wrong one, this results in a time dependent mistag probability (see Fig. B.5).

$$\frac{0.5\ (1-0.5)\ 0.01\ p_{BG}\ (t)}{0.5\ p_{SL}\ (t)\ +\ 0.01\ (1-0.5)\ p_{BG}\ (t)}$$
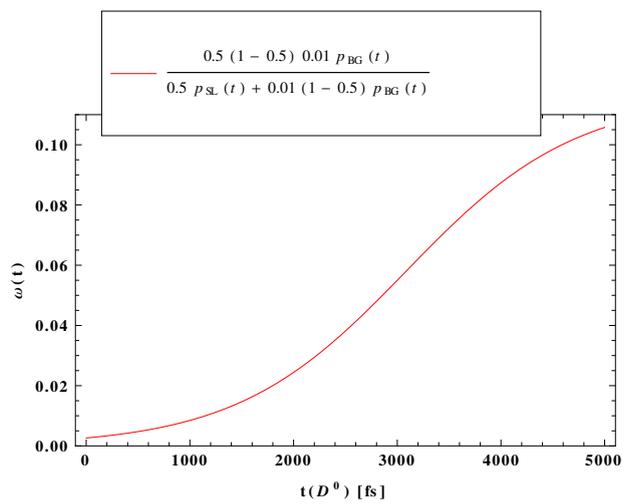
**Figure B.5:** *Resulting mistag probability time dependence looks similar to what we observe, see Sec. 8.5*

# B.5   $D^0$ **decay-time resolution**

To study the relative fractions of the decay-time resolution contribution we are using a sum of 4 Gaussian as a model.

$$\mathscr{PDF}(\Delta t | \vec{\lambda_{sig}}) \equiv \sum_{i=0}^{3} \frac{f_i}{\sqrt{2\pi}\sigma_i} \exp\left\{ -\frac{1}{2}\left( \frac{\Delta t - \mu_i}{\sigma_i} \right)^2 \right\} \tag{B.12}$$

Fig.    B.6 shows the result of the fits, the two main Gaussians have resolutions of $58fs(fraction = 43\%)$ and $132fs(fraction = 46\%)$ . We used the default `LifeTimeFitter` that has no mass constraints and also `DecayTreeFitter`[2] to reconstruct the $D^0$ decay time $t_{rec}$. Both have similar resolutions[3].

---

[2]with $D^0$ mass constraint

[3]Difference between `LifeTimeFitter` and `DecayTreeFitter` is well described by a Gaussian with width=40fs. Therefore using one or another fitter has no impact(below $\mathscr{O}(10^{-6})$) on the result
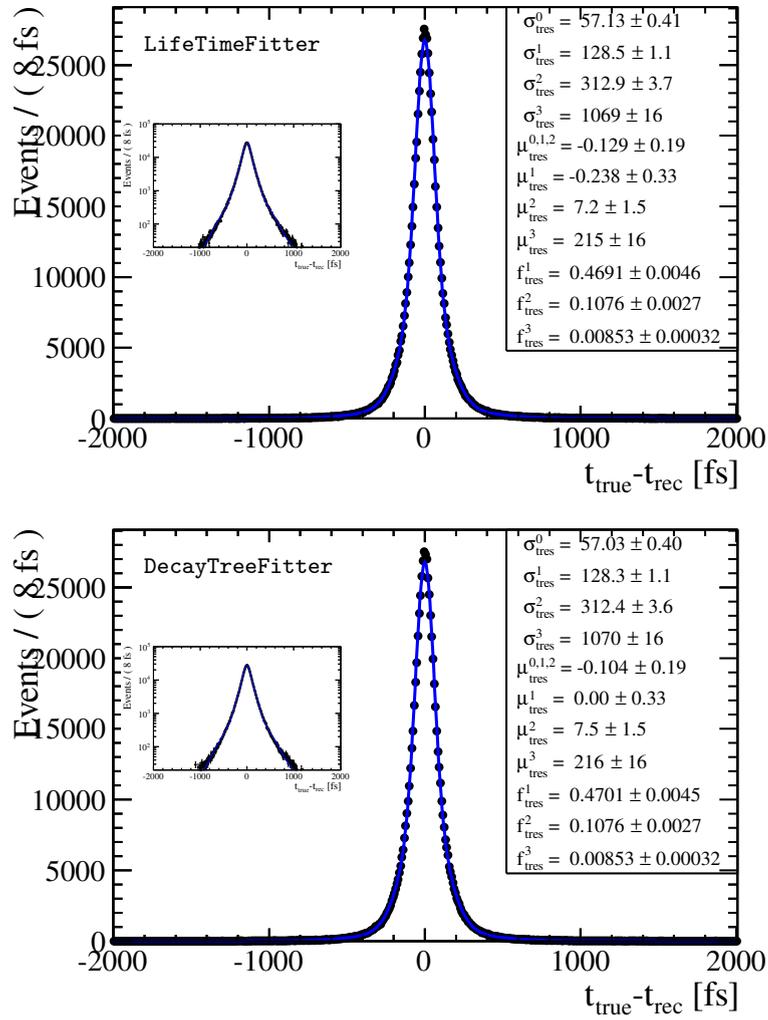
**Figure B.6:** $D^0$ *decay-time resolution with core resolution of* 57$fs$(`LifeTimeFitter` *on top and* `DecayTreeFitter` *bottom*).

## B.6 Fitted function to binned asymmetry

Asymmetry $A^{CP}(t)$ is calculated separately for every bin. This is done calculating the mean $<t>$ using both parts $D^0$ and $\bar{D}^0$. Here we calculate the effect on the slope of the fitted function and check that the conditions for this approximation are satisfied and are negligible for our current purpose.

The ratio in every bin can be calculated with the following equation:

$$r(t|const_r, \mathscr{A}_\Gamma) = const_r(1 + 2a_{CP}^{dir}) - 2\mathscr{A}_\Gamma \frac{t}{\tau} \tag{B.13}$$

$$\mathscr{P}_{D^0}(t) = \frac{1}{\tau_{D^0}} e^{-t/\tau_{D^0}} \tag{B.14}$$

For the yield in a time bin $t_1 < t < t_2$ one can determine

$$n_{D^0}(t_1, t_2) = \int_{t_1}^{t_2} N_{D^0} \mathscr{P}_{D^0}(t') \, dt' = \frac{N_{D^0}}{\tau_{D^0}} \int_t^{t_2} e^{-t_1'/\tau_{D^0}} \, dt'$$

$$= \frac{N_{D^0}}{\tau_{D^0}} \cdot \tau_{D^0} \cdot \left\{ e^{-t_1/\tau_{D^0}} - e^{-t_2/\tau_{D^0}} \right\}$$

$$= N_{D^0} e^{-t_1/\tau_{D^0}} \left\{ 1 - e^{-(t_2-t_1)/\tau_{D^0}} \right\} \tag{B.15}$$

Introducing $\tau = \frac{\tau_{D^0} + \tau_{\bar{D}^0}}{2}$ and $A_\Gamma = \frac{\tau_{D^0} - \tau_{\bar{D}^0}}{\tau_{D^0} + \tau_{\bar{D}^0}}$ For the ratio one obtains:

$$R(t_1, t_2) = \frac{n_{\bar{D}^0}(t_1, t_2)}{n_{D^0}(t_1, t_2)}$$

$$= \frac{N_{\bar{D}^0}}{N_{D^0}} \frac{e^{-t_1/\tau_{\bar{D}^0}} \left\{ 1 - e^{-(t_2-t_1)/\tau_{\bar{D}^0}} \right\}}{e^{-t_1/\tau_{D^0}} \left\{ 1 - e^{-(t_2-t_1)/\tau_{D^0}} \right\}}$$

$$= \frac{N_{\bar{D}^0}}{N_{D^0}} (1 - 2 \cdot \frac{e^{-t_2/\tau}(t_2 + \tau) - e^{-t_1/\tau}(t_1 + \tau)}{e^{-t_2/\tau} - e^{-t_1/\tau}} \frac{A_\Gamma}{\tau} + O([\frac{A_\Gamma <t>}{\tau}]^2))$$

$$= \frac{N_{\bar{D}^0}}{N_{D^0}} (1 - <t>_{t_1}^{t_2} 2 \cdot \frac{A_\Gamma}{\tau}) \tag{B.16}$$

Where $<t>_{t_1}^{t_2} = \frac{\int_{t_1}^{t_2} t e^{-t/\tau} dt}{\int_{t_1}^{t_2} e^{-t/\tau} dt}$ is the avarage decay time of the bin. For the time evolution of the binned Asymmetry $A^{CP}(t)$ one can obtain in a similar way:

$$A^{CP}(t_1, t_2) = \frac{n_{\bar{D}^0}(t_1, t_2) - n_{D^0}(t_1, t_2)}{n_{\bar{D}^0}(t_1, t_2) + n_{D^0}(t_1, t_2)}$$

$$= \frac{(1+A_\Gamma)(e^{-\frac{t_1}{\tau(1+A_\Gamma)}} - e^{-\frac{t_2}{\tau(1+A_\Gamma)}})n_{D^0} - (1-A_\Gamma)(e^{-\frac{t_1}{\tau(1-A_\Gamma)}} - e^{-\frac{t_2}{\tau(1-A_\Gamma)}})n_{\bar{D}^0}}{(1+A_\Gamma)(e^{-\frac{t_1}{\tau(1+A_\Gamma)}} - e^{-\frac{t_2}{\tau(1+A_\Gamma)}})n_{D^0} + (1-A_\Gamma)(e^{-\frac{t_1}{\tau(1-A_\Gamma)}} - e^{-\frac{t_2}{\tau(1-A_\Gamma)}})n_{\bar{D}^0}}$$

$$= (1 - \frac{2N_{\bar{D}^0}}{N_{D^0} + N_{\bar{D}^0}}) + \frac{4N_{\bar{D}^0}N_{D^0}}{(N_{D^0} + N_{\bar{D}^0})^2} \cdot \frac{e^{-t_2/\tau}(t_2 + \tau) - e^{-t_1/\tau}(t_1 + \tau)}{e^{-t_2/\tau} - e^{-t_1/\tau}} \frac{A_\Gamma}{\tau} + O([\frac{A_\Gamma <t>}{\tau}]^2)$$

$$= a_{CP}^{dir} - (1 - a_{CP}^{dir\,2}) \cdot \frac{e^{-t_2/\tau}(t_2 + \tau) - e^{-t_1/\tau}(t_1 + \tau)}{e^{-t_2/\tau} - e^{-t_1/\tau}} + O([\frac{A_\Gamma <t>}{\tau}]^2)$$

$$\approx a_{CP}^{dir} - \frac{e^{-t_2/\tau}(t_2 + \tau) - e^{-t_1/\tau}(t_1 + \tau)}{e^{-t_2/\tau} - e^{-t_1/\tau}} \frac{A_\Gamma}{\tau}$$

$$\approx a_{CP}^{dir} - <t>_{t_1}^{t_2} \cdot \frac{A_\Gamma}{\tau} \tag{B.17}$$

Therefore fitting a linear function to the data using the bin centers is a correct approach. A difference can only be seen for the sensitivity of $A_\Gamma \in O(10^{-6})$ and is negligible for our current sensitivity of $A_\Gamma \in O(10^{-4})$.

Calculating the bias for all the bins after the linear fit of $A_\Gamma$.
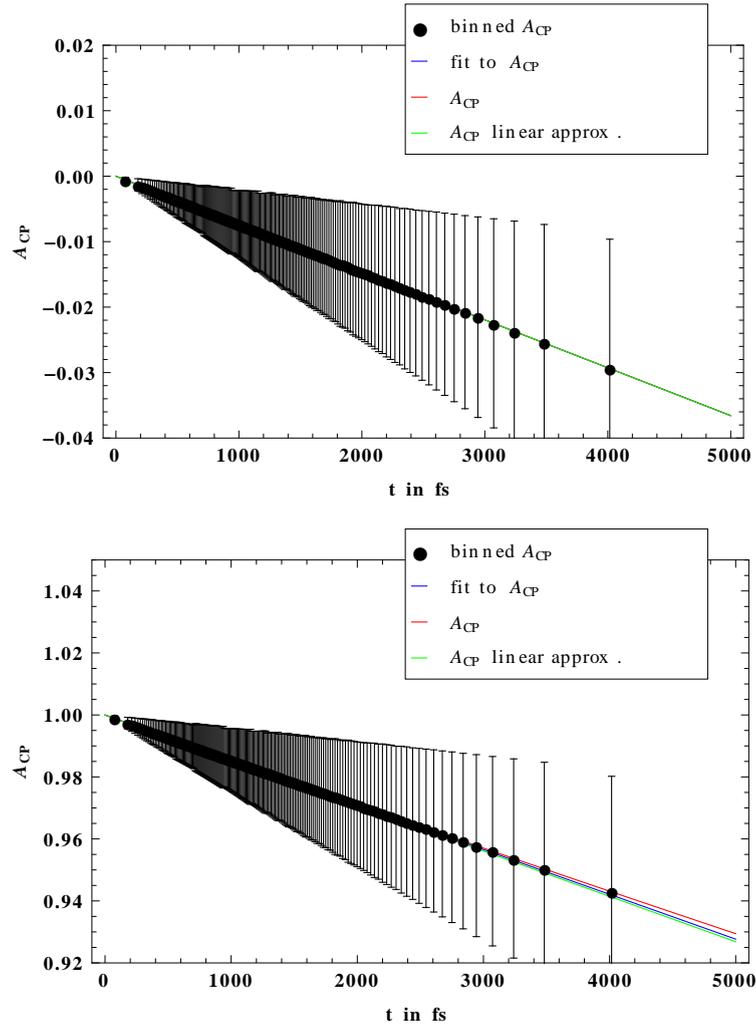
**Figure B.7:** *Asymmetry time dependence for generated $A_\Gamma = 0.3\%$ value is shown. $A^{CP}$ (t) on top and ratio(t) on bottom. Data points do not variate within statistics(only show for error bars visualization) the full model behavior is shown in red, the linear fit in blue and the linear model approximation in red. The difference on $A_\Gamma$ extracted from the $A^{CP}$ (t) linear fit is 0.000016% on $A_\Gamma$. From the ratio(t) linear fit it is 0.00357% bias on $A_\Gamma$ which is far below $O(10^{-4})$ and can be neglected for our current sensitivity.*

## B.7 Mass model definitions

The Crystal Ball (CB) function describes the asymmetric behavior from e.g. photon radiation, was used as a cross check in the systematic studies. Here we give the definition of the CB:

$$\mathscr{PDF}_{CB}(m|\vec{\lambda}) \equiv N \cdot \begin{cases} \exp(-\frac{(x-\mu)^2}{2\sigma^2}), & \text{for } \frac{x-\mu}{\sigma} > -\alpha \\ A \cdot (B - \frac{x-\mu x}{\sigma})^{-n}, & \text{for } \frac{x-\mu}{\sigma} \leqslant -\alpha \end{cases} \tag{B.18}$$

with normalization N, where $\sigma$ is the width of the distribution, $\mu$ the most probable value, $\alpha$ the tail, and $n$ the power of the exponent. So, for full signal fit following $\mathscr{PDF}$ has been used:

$$\mathscr{PDF}_{Signal}(m|\vec{\lambda}) \equiv f \cdot \mathscr{PDF}_{CB}(m|\vec{\lambda}) + (1-f) \cdot \mathscr{PDG}_{CB}(m|\vec{\lambda}) \tag{B.19}$$

## B.8   Variables: signal and background
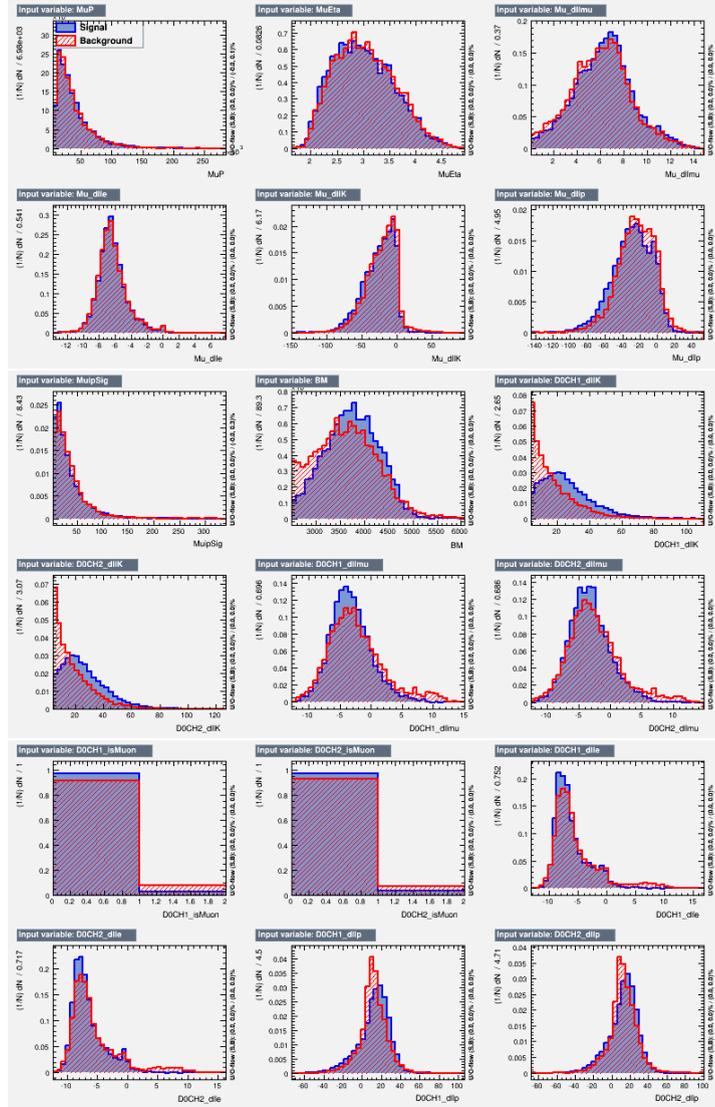


**Figure B.8:** *Distribution for signal and background for different variables in KK channel.*
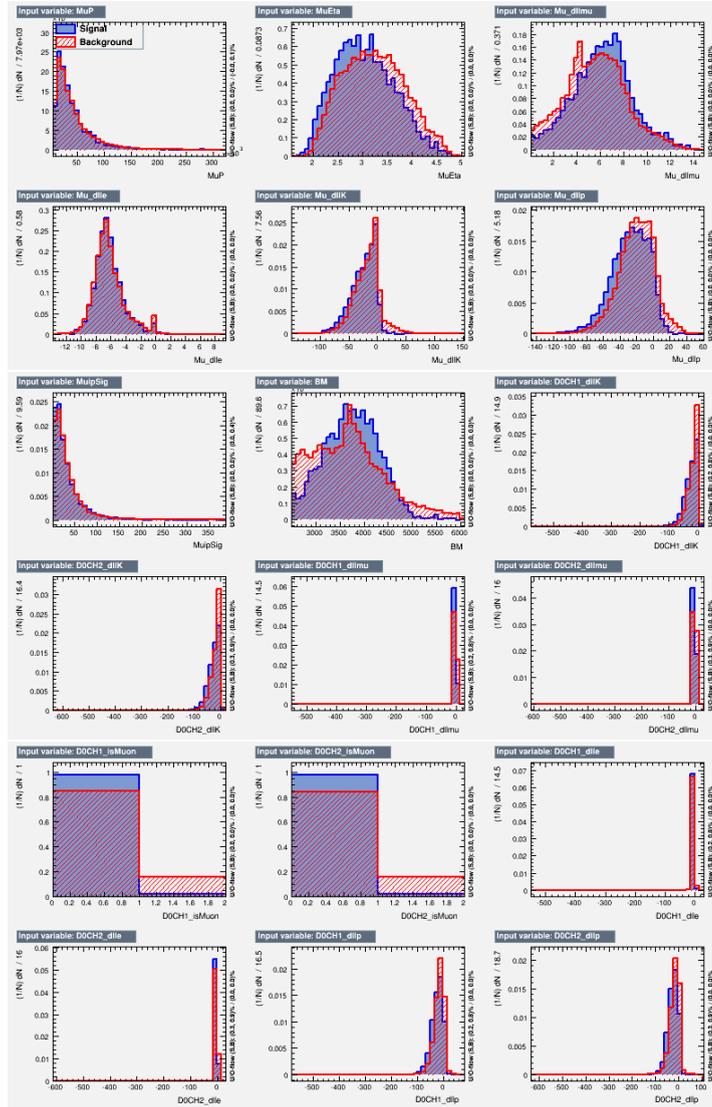
**Figure B.9:** *Distribution for signal and background for different variables in ππ channel*

# B.9  $A_\Gamma$ significance acceptance and range dependence

To study how the time window and acceptance change the significance on $A_\Gamma$ following study has been done. The assumptions used there are the following:

- The Uncertainty of the $A_{CP}(t)$ for uncorrelated quantities in every bin scales with $\sim \frac{1}{\sqrt{N_i}}$, where $N_i$ is the number of entries in bin i.

- a simple exponential decay (with $\tau = 410 fs$) was used

- The offset asymmetry ($A_{CP}$) is assumed do be fixed.

- For qualitatively analyze acceptance impact, acceptance function $acc(t, \beta) = (1 - \beta \frac{t}{\tau})$ with a slope $\beta$ was considered.

- Uncertainty from every bin on $A_\Gamma$ is $\sigma_i = \frac{1}{\sqrt{N_i}} \frac{\tau}{<t>_i}$ , resulting in a total uncertainty combining them via

$$\sigma_{A_\Gamma} = \frac{1}{\sqrt{\Sigma_{i=1}^n 1/\sigma_i}} \tag{B.20}$$

With this Assumptions $\sigma_{A_\Gamma}$ dependence on the $\beta$-factor of the acceptance, the time-window and for Number of equidistant bins is calculated. Fig. B.11 shows this behaviour. Keeping the total number of events constant ($N = const.$), one can see that having a negative $\beta$-factor (means rising acceptance) increases the sensitivity. That is clear if one considers Eq. B.20, where events with higher decay time would have a higher impact on $A_\Gamma$. The time window and bin size, as can be seen from Fig. B.12 should be also sufficiently large ($t_max > 2500 fs$, $n > 20$) not to loose sensitivity. One should consider that this model does not consider all the effects (e.g. is for signal only and does not consider the resolution effects that do not change the behaviour but can change the offsets and absolute values). Therefore, it can be used only for estimating the effect of the binning, time range and acceptance on $A_\Gamma$ significance.
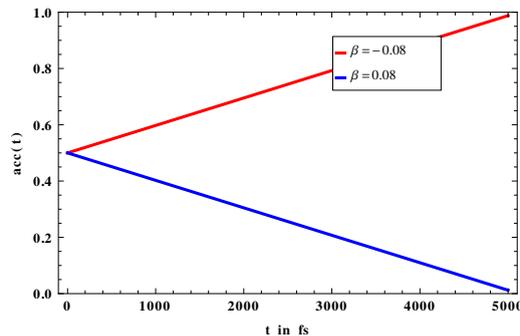


**Figure B.10:** *example of 2 different acceptances used for the study.*
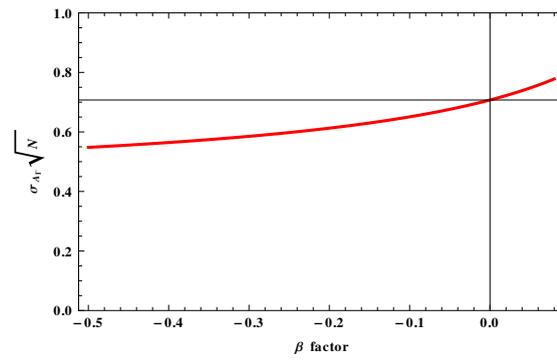
**Figure B.11:** *β-factor(of acceptance) dependence vs. significance*
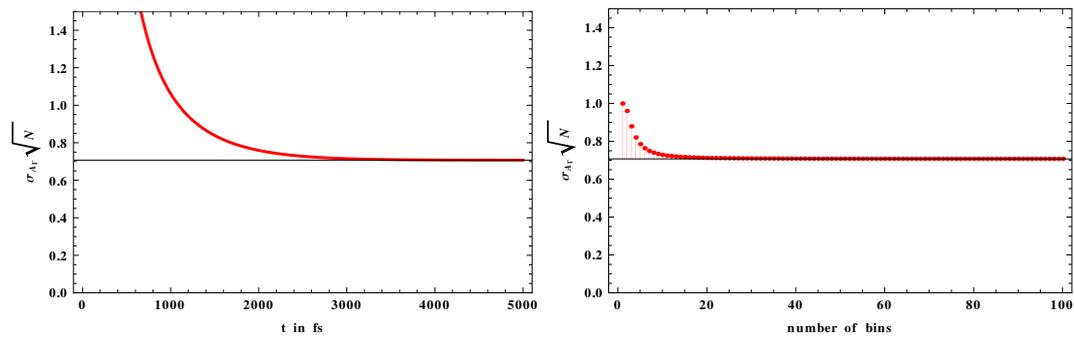


**Figure B.12:** *Significance on $A_\Gamma$ vs. decay-time range and binning dependence.*

## B.10   Mistag definition and raw asymmetry

The time dependent asymmetry $A_{RAW}^{CP}(t)$ can be expressed in the following way:

$$
\begin{aligned}
A_{RAW}^{CP}(t) &= \frac{N(D_{tag}^0, t) - N(\overline{D}_{tag}^0, t)}{N(D_{tag}^0, t) + N(\overline{D}_{tag}^0, t)} \\
&= \frac{P(D_{tag}^0, t) - P(\overline{D}_{tag}^0, t)}{P(D_{tag}^0, t) + P(\overline{D}_{tag}^0, t)}
\end{aligned}
$$

(B.21)

As a simplification we do write explicit the time dependence for every term. And using Bayes theorem introducing conditional probabilities one obtains:

$$
\begin{aligned}
A_{RAW}^{CP}(t) &= \frac{P(D_{tag}^0, t) - P(\overline{D}_{tag}^0, t)}{P(D_{tag}^0, t) + P(\overline{D}_{tag}^0, t)} \\
&= \frac{P(D_{tag}^0|D^0)P(D^0) + P(D_{tag}^0|\overline{D}^0)P(\overline{D}^0) - P(\overline{D}_{tag}^0|\overline{D}^0)P(\overline{D}^0) - P(\overline{D}_{tag}^0|D^0)P(D^0)}{P(D_{tag}^0) + P(\overline{D}_{tag}^0)} \\
&= \frac{(1 - P(\overline{D}_{tag}^0|D^0))P(D^0) + P(D_{tag}^0|\overline{D}^0)P(\overline{D}^0) - (1 - P(D_{tag}^0|\overline{D}^0))P(\overline{D}^0) - P(\overline{D}_{tag}^0|D^0)P(D^0)}{P(D^0) + P(\overline{D}^0)}
\end{aligned}
$$

$$
\begin{aligned}
P(D_{tag}^0, t) &= P(D_{tag}^0 \wedge D^0) + P(D_{tag}^0 \wedge \overline{D}^0) \\
&= P(D_{tag}^0 \wedge D^0 \wedge B_{rec}) + P(D_{tag}^0 \wedge \overline{D}^0 \wedge B_{rec}) \\
&= P(D_{tag}^0|D^0 \wedge B_{rec})P(D^0 \wedge B_{rec}) + P(D_{tag}^0|\overline{D}^0 \wedge B_{rec})P(\overline{D}^0 \wedge B_{rec}) \\
&= (1 - P(\overline{D}_{tag}^0|D^0 \wedge B_{rec}))P(D^0 \wedge B_{rec}) + P(D_{tag}^0|\overline{D}^0 \wedge B_{rec})P(\overline{D}^0 \wedge B_{rec}) \\
&= P(D_{tag}^0|D^0 \wedge B_{rec})(P(D^0 \wedge B_{rec}^-) + P(D^0 \wedge B_{rec}^+)) \\
&\quad + P(D_{tag}^0|\overline{D}^0 \wedge B_{rec})(P(\overline{D}^0 \wedge B_{rec}^-) + P(\overline{D}^0 \wedge B_{rec}^+))
\end{aligned}
$$

$$
\begin{aligned}
P(D_{tag}^0, t) - P(\overline{D}_{tag}^0, t) &= P(D_{tag}^0|D^0 \wedge B_{rec})(P(D^0 \wedge B_{rec}^-) + P(D^0 \wedge B_{rec}^+)) \\
&\quad + P(D_{tag}^0|\overline{D}^0 \wedge B_{rec})(P(\overline{D}^0 \wedge B_{rec}^-) + P(\overline{D}^0 \wedge B_{rec}^+)) \\
&\quad - P(\overline{D}_{tag}^0|\overline{D}^0 \wedge B_{rec})(P(\overline{D}^0 \wedge B_{rec}^-) + P(\overline{D}^0 \wedge B_{rec}^+)) \\
&\quad - P(\overline{D}_{tag}^0|D^0 \wedge B_{rec})(P(D^0 \wedge B_{rec}^-) + P(D^0 \wedge B_{rec}^+))
\end{aligned}
$$

$$
\begin{aligned}
P(D^0_{tag},t) - P(\overline{D}^0_{tag},t) =\ & P(D^0_{tag}|D^0 \wedge B_{rec})P(D^0 \wedge B^-_{rec}) - P(\overline{D}^0_{tag}|D^0 \wedge B_{rec})P(D^0 \wedge B^-_{rec}) \\
& + P(D^0_{tag}|D^0 \wedge B_{rec})(P(D^0 \wedge B^+_{rec})) \\
& + P(D^0_{tag}|\overline{D}^0 \wedge B_{rec})(P(\overline{D}^0 \wedge B^-_{rec}) + P(\overline{D}^0 \wedge B^+_{rec})) \\
& - P(\overline{D}^0_{tag}|\overline{D}^0 \wedge B_{rec})(P(\overline{D}^0 \wedge B^-_{rec}) + P(\overline{D}^0 \wedge B^+_{rec})) \\
& - P(\overline{D}^0_{tag}|D^0 \wedge B_{rec})(P(D^0 \wedge B^+_{rec}))
\end{aligned}
$$

$$
\begin{aligned}
P(D^0_{tag},t) - P(\overline{D}^0_{tag},t) =\ & P(D^0_{tag}|D^0 \wedge B_{rec})P(D^0 \wedge B^-_{rec}) - P(\overline{D}^0_{tag}|D^0 \wedge B_{rec})P(D^0 \wedge B^-_{rec}) \\
& + P(D^0_{tag}|D^0 \wedge B_{rec})(P(D^0 \wedge B^+_{rec})) \\
& - (1 - P(D^0_{tag}|D^0 \wedge B_{rec}))(P(D^0 \wedge B^+_{rec})) \\
& + P(D^0_{tag}|\overline{D}^0 \wedge B_{rec})(P(\overline{D}^0 \wedge B^-_{rec}) + P(\overline{D}^0 \wedge B^+_{rec})) \\
& - P(\overline{D}^0_{tag}|\overline{D}^0 \wedge B_{rec})(P(\overline{D}^0 \wedge B^-_{rec}) + P(\overline{D}^0 \wedge B^+_{rec}))
\end{aligned}
$$

$$
\begin{aligned}
P(D^0_{tag},t) - P(\overline{D}^0_{tag},t) =\ & P(D^0_{tag}|D^0 \wedge B_{rec})P(D^0 \wedge B^-_{rec}) - P(\overline{D}^0_{tag}|D^0 \wedge B_{rec})P(D^0 \wedge B^-_{rec}) \\
& + 2P(D^0_{tag}|D^0 \wedge B_{rec})(P(D^0 \wedge B^+_{rec})) \\
& - (P(D^0 \wedge B^+_{rec})) \\
& + 2 * P(D^0_{tag}|\overline{D}^0 \wedge B_{rec})(P(\overline{D}^0 \wedge B^-_{rec}) + P(\overline{D}^0 \wedge B^+_{rec})) \\
& - (P(\overline{D}^0 \wedge B^-_{rec}) + P(\overline{D}^0 \wedge B^+_{rec}))
\end{aligned}
$$

Now, we can identify in this equation the mistag probability to tag a real $D^0(\overline{D}^0)$ with a wrong tag $\overline{D}^0_{tag}(D^0_{tag})$ as a conditional probability:

$$
\begin{aligned}
\omega^{\overline{D}^0} &= P(\overline{D}^0_{tag}|D^0) \\
\omega^{D^0} &= P(D^0_{tag}|\overline{D}^0)
\end{aligned}
$$

(B.22)

$A^{CP}_{RAW}$ can obtain in the following way:

$$
\begin{aligned}
A^{CP}_{RAW}(t) &= \frac{(1 - \omega^{\overline{D}^0})P(D^0) + \omega^{D^0}P(\overline{D}^0) - (1 - \omega^{D^0})P(\overline{D}^0) - \omega^{\overline{D}^0}P(D^0)}{P(D^0) + P(\overline{D}^0)} \\
&= \frac{(1 - 2\omega^{\overline{D}^0})P(D^0) - (1 - 2\omega^{D^0})P(\overline{D}^0)}{P(D^0) + P(\overline{D}^0)}
\end{aligned}
$$

(B.23)

Defining average mistag probability and mistag probability difference:

$$
\begin{aligned}
\bar{\omega} &= \frac{\omega^{\overline{D}^0} + \omega^{D^0}}{2} \\
\Delta\omega &= \omega^{\overline{D}^0} - \omega^{D^0}
\end{aligned}
$$

(B.24)

The equation above can be parametrized in the following way:

$$
\begin{aligned}
A_{RAW}^{CP}(t) &= \frac{(1-(\omega^{\overline{D}^0}+\omega^{D^0}))P(D^0)-(1-(\omega^{D^0}+\omega^{\overline{D}^0}))P(\overline{D}^0)-(\omega^{\overline{D}^0}-\omega^{D^0})(P(D^0)+P(\overline{D}^0))}{P(D^0)+P(\overline{D}^0)} \\
&= \frac{(1-2\bar{\omega})P(D^0)-(1-2\bar{\omega})P(\overline{D}^0)}{P(D^0)+P(\overline{D}^0)}-\Delta\omega \\
&= (1-2\bar{\omega})\frac{P(D^0)-P(\overline{D}^0)}{P(D^0)+P(\overline{D}^0)}-\Delta\omega
\end{aligned}
$$

$$(\text{B.25})$$

However, in this definition we assumed that all the $D^0$ s coming from $B$ mesons will be also reconstructed. otherwise $P(\overline{D}_{tag}^0|D^0)+P(D_{tag}^0|D^0)\neq 1$. So to make this hold one can do the same calculation as above replacing $D^0 \rightarrow (D^0$ and $B$ was reconstructed).

This brings us to the following equation:

$$
A_{RAW}^{CP}(t)=(1-2\bar{\omega})\frac{P(D^0\wedge B_{rec})-P(\overline{D}^0\wedge B_{rec})}{P(D^0\wedge B_{rec})+P(\overline{D}^0\wedge B_{rec})}-\Delta\omega
$$

$$(\text{B.26})$$

Notice also the change in the $\omega$ definitions:

$$
\begin{aligned}
\omega^{\overline{D}^0} &= P(\overline{D}_{tag}^0|D^0\wedge B_{rec}) \\
\omega^{D^0} &= P(D_{tag}^0|\overline{D}^0\wedge B_{rec})
\end{aligned}
$$

$$(\text{B.27})$$

$P(D^0\wedge B_{rec})$ can be also written in the following way:

$$
P(D^0\wedge B_{rec})=P(D^0\wedge(D^0\text{from}B)\wedge\mu_{rec}\wedge D_{rec}^0)+P(D^0\wedge(D^0\text{notfrom}B)\wedge\mu_{rec}\wedge D_{rec}^0)
$$

$$(\text{B.28})$$

Assuming that we reconstruct $D^0$ only coming from $B$ s one can neglect the second term. Additionally, one can assume that the $\mu$ and $D^0$ reconstruction efficiencies are independent.

$$
\begin{aligned}
P(D^0\wedge B_{rec}) &= P(D^0\wedge(D^0\text{from}B)\wedge\mu_{rec}\wedge D_{rec}^0) \\
&= P(D^0|(D^0\text{from}B))P(B)P(\mu_{rec}|D^0\text{from}B)P(D_{rec}^0|D^0\text{from}B)
\end{aligned}
$$

$$(\text{B.29})$$

So, introducing the following Asymmetries, starting with the muon detection asymmetry one obtains:

$$
A_{mu}=\frac{P(\mu_{rec}^-|D^0\text{from}B)-P(\mu_{rec}^+|\overline{D}^0\text{from}B)}{P(\mu_{rec}^-|D^0\text{from}B)+P(\mu_{rec}^+|\overline{D}^0\text{from}B)}
$$

$$(\text{B.30})$$

and for the Production Asymmetry:

$$
A_{prod}=\frac{P(B)-P(\bar{B})}{P(B)+P(\bar{B})}
$$

$$(\text{B.31})$$

The detection Asymmetry is zero for the CP final states:

$$A_{Ddet} = \frac{P(D^0_{rec}|D^0\mathrm{from}B) - P(\overline{D}^0_{rec}|\overline{D}^0\mathrm{from}B)}{P(D^0_{rec}|D^0\mathrm{from}B) + P(\overline{D}^0_{rec}|\overline{D}^0\mathrm{from}B)}$$
$$= 0 \qquad\qquad\qquad\qquad\qquad\qquad\text{(B.32)}$$

using this one obtains:

$$A^{CP}_{RAW}(t) = (1-2\bar{\omega})\frac{(A^{CP}+1)(A_{mu}+1)(A_{prod}+1) - (1-A^{CP})(1-A_{mu})(1-A_{prod})}{(A^{CP}+1)(A_{mu}+1)(A_{prod}+1) + (1-A^{CP})(1-A_{mu})(1-A_{prod})} - \Delta\omega$$

$$\text{(B.33)}$$

and removing the higher order terms since Asymmetries are in O(1%) range, one obtains:

$$A^{CP}_{RAW}(t) = (1-2\bar{\omega})(A^{CP}+A_{mu}+A_{prod}) - \Delta\omega$$

$$\text{(B.34)}$$

## B.11   Selection optimization

Following variables were considered for optimization of the selection and as input for the training (see Tab. B.1).

| Variable | Separation(KK) | Separation($\pi\pi$) |
|---|---|---|
| reco B mass | 3.3% | 5.7% |
| $\log \frac{\mathscr{L}(\mu)}{\mathscr{L}(\pi)}$ of $D^0$ daughters | 7.1% | 6.9% |
| $\log \frac{\mathscr{L}(K)}{\mathscr{L}(\pi)}$ | 2.6% | 6.7% |
| $\log \frac{\mathscr{L}(p)}{\mathscr{L}(\pi)}$ | 7.6% | 8.2% |
| $\log \frac{\mathscr{L}(e)}{\mathscr{L}(\pi)}$ | 8.7% | 1.9% |
| isMuon flag | 2.7% | 1.08% |
| $\log \frac{\mathscr{L}(p)}{\mathscr{L}(\pi)}$ of muon | 2.0% | 5.0% |
| $\log \frac{\mathscr{L}(K)}{\mathscr{L}(\pi)}$ | 1.8% | 5.9% |
| $\log \frac{\mathscr{L}(e)}{\mathscr{L}(\pi)}$ | $< 1\%$ | 3.5% |
| $\log \frac{\mathscr{L}(\mu)}{\mathscr{L}(\pi)}$ | $< 1\%$ | 2.0% |
| $\chi^2$(IP) | $< 1\%$ | 1.2% |
| $\eta(\mu)$ | $< 1\%$ | 2.7% |
| $p_\text{T}(\mu)$ | $< 1\%$ | $< 1\%$ |
| $p_\text{T}(D^0)$ | $< 1\%$ | $< 1\%$ |

**Table B.1:** *table of used variables including their separation* $< S^2 >= \frac{1}{2} \int \frac{pdf_{sig}(x) - pdf_{bkg}(x)}{pdf_{sig}(x) + pdf_{bkg}(x)} dx$

The training was done with the sWeights determined from data and validated on an independent data sample. All the variables from Tab. B.1 were used for this. The working points(ROC curve) for the background rejection and signal efficiency are shown in Fig. B.13.

The best performance is shown by using Neuronal Network MLP or Boosted Decision Trees (BDTs). Also the cuts introduced so far show a good performance in both channels. For KK an improvement in some $\mathscr{O}(1\%)$ for background rejection using a multivariate is visible.
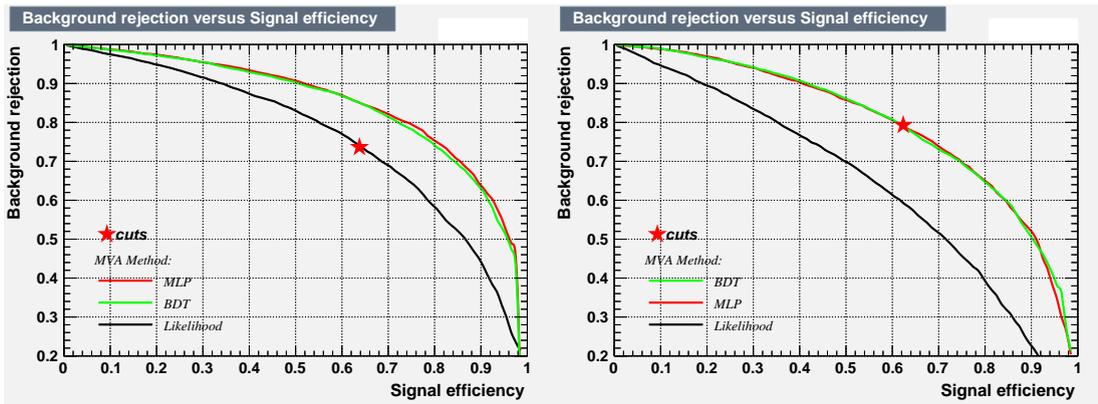
**Figure B.13:** *Significance for different cut applied on the neuronal network signal classification output is shown. It is normalized to the default cuts 5.7 significance value. $D^0 \rightarrow K^- K^+$ (left) and $D^0 \rightarrow \pi^- \pi^+$ (right)*
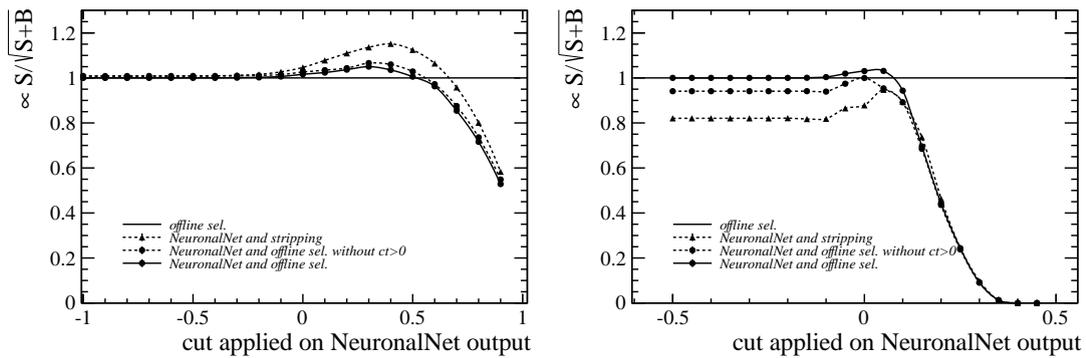


**Figure B.14:** *Significance for different cuts applied on the neuronal network signal classification output is shown. It is normalized to the default cuts 5.7 significance value.*

## B.12   The choice of the binning

A ratio or asymmetry of two exponential distributions is a commonly used function, as generated and shown in Sec. 6. Since the difference in lifetime between this exponential is small compared to the lifetime a linear function can be used to achieve this.

$$f(t) = a + b \cdot t \tag{B.35}$$

In first order, it can be assumed that the offset is already determined by the large statistics in the low lifetime region. This means that a is precisely known offset a can be assumed as fixed and this would result in the following uncertainty of the linear function.

$$\sigma_{f(t)} \approx |\frac{\partial f(t)}{\partial b} \cdot \sqrt{V_{bb}}| \tag{B.36}$$

$$= |t| \cdot \sigma_b \tag{B.37}$$

This means that the uncertainty is growing linearly in in decay time t. One can also do the calculation and consider also the full covariance matrix .

$$\sigma_{f(t)} = \sqrt{(\frac{\partial f(t)}{\partial a})^2 \cdot V_{aa} + 2 \cdot (\frac{\partial f(t)}{\partial b} \frac{\partial f(t)}{\partial a}) \cdot V_{ab} + (\frac{\partial f(t)}{\partial b})^2 \cdot V_{bb}} \tag{B.38}$$

$$= \sqrt{V_{aa} + 2 \cdot V_{ab} \cdot t + V_{bb} \cdot t^2} \tag{B.39}$$

Since we have here a ratio of exponentials distribution, $V_{bb}$ dominates increasing the lifetime and one would obtains an asymptotic convergence towards Eq. B.37. Fig. B.16 shows the $ratio = 1$, together with the slope 0. Fitting a linear function(independent of the binning) on a flat ratio distribution would give you the $1\sigma$ bands B.16 using the covariance matrix from the Fit from Eq. B.39.

Like explained above, fixing the offset would slightly shrink the region covered by $1\sigma$ bands to a region covered by two straight lines. However, one data point can then be directly connected to the slope of the $1\sigma$ band by the error of that bin. Following assumptions are used to simplify this approach:

- Distributions are exponential.

- Offset can be assumed as fixed, which means $1\sigma$ band of the fitted function can be assumed as linear which is true for $\tau << t_{max}$.
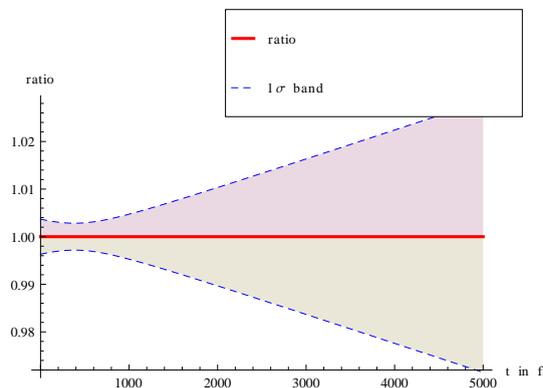


**Figure B.15:** *Ratio(t) with the 1σ bands, determined from a fit of a ratio of two exponential distributions*

- Error of the $ratio_i$ for uncorrelated quantities in every bin scales with $\sim \frac{1}{\sqrt{N_i}}$, where $N_i$ is the number of entries in bin i.

Under these conditions one can require that the error($\sigma_i$) grows linearly with time $<t>_i$, in the same way like $1\sigma$ band of the fit.

$$c = \frac{\sigma_i}{<t>_i} \tag{B.40}$$

Which effectively means that every bin has the same contribution to the slope of a ratio. This leads to the following equation:

$$c = \frac{1/\sqrt{\int\limits_{a_i}^{a_{i+1}} 1/\tau e^{-t/\tau} dt}}{\int\limits_{a_i}^{a_{i+1}} 1/\tau t e^{-t/\tau} dt / \int\limits_{a_i}^{a_{i+1}} 1/\tau e^{-t/\tau} dt}$$

$$c^2 = \frac{\int\limits_{a_i}^{a_{i+1}} 1/\tau e^{-t/\tau} dt}{(\int\limits_{a_i}^{a_{i+1}} 1/\tau t e^{-t/\tau} dt)^2} \tag{B.41}$$

and finally:

$$(e^{-a_i/\tau}(a_i + \tau) - e^{-a_{i+1}/\tau}(a_{i+1} + \tau))^2 c^2 = e^{-a_i/\tau} - e^{-a_{i+1}/\tau} \tag{B.42}$$

Where $[a_i, a_{i+1}]$ is the bin range of bin i.

To solve this one can use either Taylor expansion(depending on your maximal bin size) or solve this directly with numerical methods. For the full solution one has to continue this search recursively: $a_{i+1} = f(a_i)$, starting with $a_0 = 0$. E.g. for 10 bins in range = {0,5175} one gets the following binning: $a_i \in \{0, 530, 687, 832, 978, 1132, 1304, 1508, 1773, 2186, 5175\}$. As a result, the error time evolution is linear, like expected. This procedure can be applied for all kinds of binning and bin ranges.
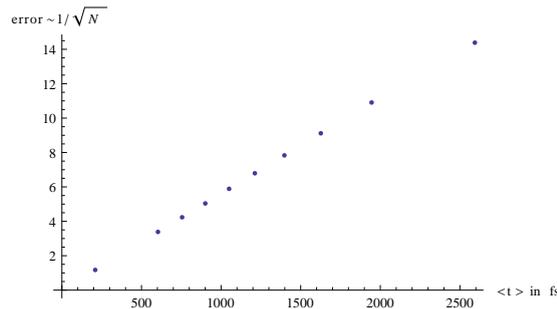


**Figure B.16:** *Time evolution of the error for 10 bins in range= {0,5175}*

# Bibliography

[1] LHCb collaboration, R. Aaij *et al.*, *Measurement of indirect CP asymmetries in $D^0 \to K^- K^+$ and $D^0 \to \pi^- \pi^+$ decays*, arXiv:1501.06777, submitted to JHEP.

[2] K. Olive and P. D. Group, *Review of particle physics*, Chinese Physics C **38** (2014), no. 9 090001.

[3] LHCb collaboration, R. Aaij *et al.*, *Measurement of $D^0$–$\bar{D}^0$ mixing parameters and search for CP violation using $D^0 \to K^+ \pi^-$ decays*, Phys. Rev. Lett. **111** (2013) 251801, arXiv:1309.6534.

[4] A. Jaeger and J. van Tilburg, *Measurement of indirect CP asymmetries in $D^0 \to K^- K^+$ and $D^0 \to \pi^- \pi^+$ decays*, , LHCb-ANA-2014-019, CERN-LHCb-ANA-2014-019.

[5] F. Halzen and A. D. Martin, *QUARKS AND LEPTONS: AN INTRODUCTORY COURSE IN MODERN PARTICLE PHYSICS*, Wiley, New York, USA, 1984.

[6] D. Griffiths and W. InterScience, *Introduction to elementary particles*, Harper & Row New York, 1987.

[7] Y. Ahn, H.-Y. Cheng, and S. Oh, *Wolfenstein Parametrization at Higher Order: Seeming Discrepancies and Their Resolution*, Phys. Lett. **B703** (2011) 571, arXiv:1106.0935.

[8] L. Wolfenstein, *Parametrization of the kobayashi-maskawa matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[9] ZEUS Collaboration, H1 Collaboration, A. Cooper-Sarkar, *PDF Fits at HERA*, PoS **EPS-HEP2011** (2011) 320, arXiv:1112.2107.

[10] E. Norrbin and T. Sjostrand, *Production and hadronization of heavy quarks*, Eur. Phys. J. **C17** (2000) 137, arXiv:hep-ph/0005110.

[11] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.

[12] M. Bahr *et al.*, *Herwig++ Physics and Manual*, Eur. Phys. J. **C58** (2008) 639, arXiv:0803.0883.

[13] M. L. Mangano, P. Nason, and G. Ridolfi, *Heavy-quark correlations in hadron collisions at next-to-leading order*, Nuclear Physics B **373** (1992), no. 2 295 .

[14] private communication, W.J. Stirling, private communication, *PARTON LUMINOSITY AND CROSS SECTION PLOTS*, http://mstwpdf.hepforge.org/plots/plots.html, 2014. Online; accessed 8-September-2014.

[15] ALICE Collaboration, B. Abelev *et al.*, *Measurement of charm production at central rapidity in proton-proton collisions at $\sqrt{s} = 2.76$ TeV*, JHEP **1207** (2012) 191, arXiv:1205.4007.

[16] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, *Parton Fragmentation and String Dynamics*, Phys. Rept. **97** (1983) 31.

[17] E. R. Cazaroto, V. P. Goncalves, F. S. Navarra, and M. Nielsen, *Charm production asymmetry at the lhc*, Journal of Physics: Conference Series **458** (2013), no. 1 012014.

[18] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $\bar{B}^0 - B^0$ and $\bar{B}^0_s - B^0_s$ production asymmetries in $pp$ collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. **B739** (2014) 218, arXiv:1408.0275.

[19] F. Carvalho, F. Durães, F. Navarra, and M. Nielsen, *Does the d - / d + production asymmetry decrease at large xf*, Phys. Rev. Lett. **86** (2001) .

[20] R. Aaij *et al.*, *Measurement of $J\psi$ production in $p - p$ collisions at $\sqrt{s} = 7$ TeV*, Eur. Phys. J. **C71** (2011) 1645, arXiv:1103.0423.

[21] R. Aaij *et al.*, *Prompt charm production in $pp$ collisions at $\sqrt{s} = 7$ TeV*, Nucl. Phys. **B871** (2013) 1, arXiv:1302.2864.

[22] LHCb collaboration, R. Aaij *et al.*, *Measurements of indirect CP asymmetries in $D^0 \to K^- K^+$ and $D^0 \to \pi^- \pi^+$ decays*, Phys. Rev. Lett. **112** (2014) 041801, arXiv:1310.7201.

[23] CERN , CERN document server, *CERN document server*, http://cds.cern.ch, 2014. Online; accessed 6-August-2014.

[24] M. Lamont, *Status of the LHC*, J. Phys. Conf. Ser. **455** (2013) 012001.

[25] M. Lamont, *Status of the lhc*, http://stacks.iop.org/1742-6596/455/i=1/a=012001, 2013.

[26] LHCb collaboration, LHCb speakers bureau, *LHCb material for presentations*, http://lhcb.web.cern.ch/lhcb/speakersbureau/html/Material_for_Presentations.html, 2014. Online; accessed 18-August-2014.

[27] LHCb collaboration public, http://lhcb-public.web.cern.ch/, *LHCb public web page*, http://lhcb-public.web.cern.ch/lhcb-public/en/lhcb-outreach/multimedia/, 2014. Online; accessed 18-August-2014.

[28] LHCb, A. Alves *et al.*, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.

[29] LHCb Collaboration, *LHCb VELO TDR: Vertex locator. Technical design report*, .

[30] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) 09007, arXiv:1405.7808.

[31] M. De Cian *et al.*, *Measurement of the track finding efficiency*, Tech. Rep. LHCb-PUB-2011-025. CERN-LHCb-PUB-2011-025, CERN, Geneva, Apr, 2012.

[32] L. Andricek, G. Lutz, R. H. Richter, and H. Moser, *Radiation-hard strip detectors on oxygenated silicon*, IEEE Trans. Nucl. Sci. **49** (2002) 1117.

[33] N. Van Bakel *et al.*, *Performance of the Beetle readout chip for LHCb*, .

[34] LHCb VELO, E. Jans, *Operational aspects of the VELO cooling system of LHCb*, PoS **Vertex2013** (2013) 038.

[35] LHCb Outer Tracker group, R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) 01002, arXiv:1311.3893.

[36] LHCb, Silicon Tracker UZH Group, *Material for Publications*, http://lhcb.physik.uzh.ch/ST/public/material/index.php, 2014. Online; accessed 11-Oct-2014.

[37] H. Spieler, *Semiconductor detector systems*, Semiconductor Science and Technology, Oxford Univ. Press, Oxford, 2005.

[38] LHCb Silicon Tracker group, J. Luisier, *Performance of LHCb Silicon Tracker Detector in the LHC*, Phys. Procedia **37** (2012) 851.

[39] Collaboration for the LHCb, P. Perret, *First Years of Running for the LHCb Calorimeter System*, arXiv:1407.4289.

[40] LHCb Collaboration, E. P. Olloqui, *LHCb preshower(PS) and scintillating pad detector (SPD): Commissioning, calibration, and monitoring*, J. Phys. Conf. Ser. **160** (2009) 012046.

[41] A. A. Alves, Jr. *et al.*, *Performance of the LHCb muon system*, Journal of Instrumentation **8** (2013) 2022P, arXiv:1211.1346.

[42] LHCb HLT project, J. Albrecht, V. Gligorov, G. Raven, and S. Tolk, *Performance of the LHCb High Level Trigger in 2012*, J. Phys. Conf. Ser. **513** (2014) 012001, arXiv:1310.8544.

[43] E. Aslanides *et al.*, *The level-0 muon trigger for the {LHCb} experiment*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **579** (2007), no. 3 989 .

[44] O. Callot and S. Hansmann-Menzemer, *The Forward Tracking: Algorithm and Performance Studies*, Tech. Rep. LHCb-2007-015. CERN-LHCb-2007-015, CERN, Geneva, May, 2007.

[45] J. P. Baud *et al.*, *The lhcb data management system*, Journal of Physics: Conference Series **396** (2012), no. 3 032023.

[46] LHCb collaboration, *Brunel framework*, http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/, 2014.

[47] LHCb collaboration, *DaVinci framework*, http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/, 2014.

[48] LHCb collaboration, *Gauss framework*, http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/gauss/, 2014.

[49] LHCb collaboration, *Boole framework*, http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole/, 2014.

[50] H.-Y. Cheng and C.-W. Chiang, *Direct CP violation in two-body hadronic charmed meson decays*, Phys. Rev. **D85** (2012) 034036, arXiv:1201.0785.

[51] A. Lenz and T. Rauh, *D-meson lifetimes within the heavy quark expansion*, Phys. Rev. **D88** (2013) 034004, arXiv:1305.3588.

[52] A. Lenz, *What have we learnt in theory from the delta acp saga*, The 6th International Workshop on Charm Physics.

[53] U. Nierste, *Three Lectures on Meson Mixing and CKM phenomenology*, arXiv:0904.1869.

[54] Particle Data Group, J. Beringer *et al.*, *Review of Particle Physics (RPP)*, Phys. Rev. **D86** (2012) 010001.

[55] M. Gersabeck *et al.*, *On the interplay of direct and indirect CP violation in the charm sector*, J. Phys. **G39** (2012) 045005, arXiv:1111.6515.

[56] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and tau-lepton properties as of early 2012*, arXiv:1207.1158.

[57] A. Dighe, D. Ghosh, and B. P. Kodrani, *Non-universality of indirect CP asymmetries in* $D \to \pi\pi, KK$ *decays*, arXiv:1306.3861.

[58] M. Sozzi, *Discrete symmetries and CP violation*, Oxford University Press, 1st edition (2008).

[59] Y. Grossman, A. L. Kagan, and Y. Nir, *New physics and CP violation in singly Cabibbo suppressed D decays*, Phys. Rev. **D75** (2007) 036008, arXiv:hep-ph/0609178.

[60] LHCb collaboration, R. Aaij *et al.*, *Measurements of indirect CP asymmetries in* $D^0 \to K^- K^+$ *and* $D^0 \to \pi^- \pi^+$ *decays*, Phys. Rev. Lett. **112** (2014) 041801, arXiv:1310.7201.

[61] Belle Collaboration, M. Staric, *New Belle results on D0 - D0bar mixing*, arXiv:1212.3478.

[62] BaBar Collaboration, J. Lees *et al.*, *Measurement of* $D^0 - \bar{D}^0$ *Mixing and CP Violation in Two-Body* $D^0$ *Decays*, Phys. Rev. **D87** (2013) 012004, arXiv:1209.3896.

[63] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nuclear Instruments and Methods in Physics Research A **552** (2005) 566, arXiv:physics/0503191.

[64] LHCb collaboration, LHCb, *LHCb event displays*, https://lhcb-reconstruction.web. cern.ch/lhcb-reconstruction/2012/deltaACP, 2014. Online; accessed 7-Nov-2014.

[65] LHCb collaboration, *Moore framework*, http://lhcb-release-area.web.cern.ch/ LHCb-release-area/DOC/moore/, 2014.

[66] M. Needham, *Clone Track Identification using the Kullback-Liebler Distance*, Tech. Rep. LHCb-2008-002. CERN-LHCb-2008-002. LPHE-2008-002, CERN, Geneva, Jan, 2008.

[67] A. Jaeger *et al.*, *Measurement of the track finding efficiency*, Tech. Rep. LHCb-PUB-2011-025. CERN-LHCb-PUB-2011-025, CERN, Geneva, Apr, 2012.

[68] F. Archilli *et al.*, *Performance of the Muon Identification at LHCb*, Journal of Instrumentation **8** (2013) 20P, arXiv:1306.0249.

[69] V. V. Gligorov, C. Thomas, and M. Williams, *The hlt inclusive b triggers*, Tech. Rep. LHCb-PUB-2011-016. CERN-LHCb-PUB-2011-016. LHCb-INT-2011-030, CERN, Geneva, Sep, 2011. LHCb-INT-2011-030.

[70] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, Journal of Instrumentation **8** (2013) 2013P, arXiv:1210.6861.

[71] LHCb collaboration, R. Aaij *et al.*, *Search for direct CP violation in* $D^0 \to h^- h^+$ *modes using semileptonic B decays*, Phys. Lett. **B723** (2013) 33, arXiv:1303.2614.

[72] LHCb collaboration, R. Aaij *et al.*, *Measurement of the semileptonic CP asymmetry in* $B^0 - \overline{B}^0$ *mixing*, Phys. Rev. Lett. **114** (2015) 041601, arXiv:1409.8586.

[73] F. James, *MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1*, .

[74] Wolfram Mathematica 9.0, *Mathematica 9.0*, http://reference.wolfram.com/, 2014. .

[75] M. Pivk and F. R. Le Diberder, *sPlot: a statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.

[76] A. Roodman, *Blind Analysis in Particle Physics*, in *Statistical Problems in Particle Physics, Astrophysics, and Cosmology* (L. Lyons, R. Mount, and R. Reitmeyer, eds.), p. 166, 2003. arXiv:physics/0312102.

[77] LHCb collaboration, 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*, Phys. Rev. Lett. **111** (2013) 251801, arXiv:1309.6534.

[78] LHCb Collaboration, R. Aaij *et al.*, *Precision measurement of the $B_s^0 - \bar{B}_s^0$ oscillation frequency with the decay $B_s^0 \to D_s^-\pi^+$*, New J. Phys. **15** (2013) 053021. 18 p.

[79] M. Vesterinen, *A study of Hlt2 asymmetries using semileptonic B decays and the topological lines*, LHCb-INT-2014-024.

[80] CDF collaboration, T. A. Aaltonen *et al.*, *Measurement of indirect CP-violating asymmetries in $D^0 \to K^+K^-$ and $D^0 \to \pi^+\pi^-$ decays at CDF*, arXiv:1410.5435, submitted to Phys. Rev. Lett.