The University of Melbourne

Doctoral Thesis

Measurement of the branching fractions of $\bar{B}^0 \to D^{*+}h^-$ decays at the Belle experiment and development of a global particle vertex fitting algorithm for Belle II

Author:
Jo-Frederik Krohn

Supervisor:
Phillip Urquijo

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

in the

School of Physics

May 2, 2021

Declaration of authorship

I, Jo-Frederik Krohn, declare that this thesis titled, "Measurement of the branching fractions of $\bar{B}^0 \to D^{*+}h^-$ decays at the Belle experiment and development of a global particle vertex fitting algorithm for Belle II" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- This work does not contain more than 100000 words.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: May 2, 2021

THE UNIVERSITY OF MELBOURNE

Abstract

Faculty of Science
School of Physics

Doctor of Philosophy

Measurement of the branching fractions of $\bar{B}^0 \to D^{*+}h^-$ decays at the Belle experiment and development of a global particle vertex fitting algorithm for Belle II

by Jo-Frederik KROHN

In this thesis high precision QCD-factorisation tests were performed. The branching ratios $\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-) = (2.67 \pm 0.02 \pm 0.09) \times 10^{-3}$ and $\mathcal{B}(\bar{B}^0 \to D^{*+}K^-) = (2.27 \pm 0.06 \pm 0.08) \times 10^{-4}$ were measured using $(771.58 \pm 10.56) \times 10^6$ B-meson pairs recorded by the Belle experiment. Both values are in tension with the theoretical expectation.

The ratio of the branching ratio is measured in a way that allows for the cancellation of the systematic uncertainties arising from the D^*-meson reconstruction; the value of $\mathcal{R}_{K/\pi} = \mathcal{B}(\bar{B} \to D^{*+}K^-)/\mathcal{B}(\bar{B} \to D^{*+}\pi^-) = (8.41 \pm 0.24 \pm 0.13) \times 10^{-2}$ was found. Both $\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-)$ and $\mathcal{B}(\bar{B}^0 \to D^{*+}K^-)$ have shown deviations from the prediction, this suggests that the estimation of the Feynman diagrams contributing to the predictions may be inaccurate.

The new measured branching ratios were used to perform a high precision QCD factorisation test by measuring ratios with respect to semi-leptonic branching ratios at fixed momentum transfers for different particle species. A deviation for the ratio $\Gamma(\bar{B}^0 \to D^{*+}h^-)/d\Gamma(\bar{B}^0 \to D^{*+}l^-\bar{\nu})/dq^2$ of 16% from theoretical predictions was found, suggesting large non-factorisable contributions and/or new physics contributions.

Furthermore, $SU(3)$-symmetry was tested by measuring ratios for pions and kaons of $a_1^2(K)/a_1^2(\pi) = 1.05 \pm 0.05$ as well as for different particle species. The found value is consistent with unity and therefore no evidence for $SU(3)$-symmetry breaking effects was found in this test to 5% precision. Thus, for $\mathcal{R}_{K/\pi}$ one can rule out $SU(3)$-symmetry breaking effects as an explanation for the deviation.

Finally, a new vertex fitting algorithm and its implementation for the Belle II software framework was reported. It can improve D^*-meson reconstruction, is computationally very efficient and is now the standard vertex fitting tool of the Belle II experiment.

Acknowledgements

I especially want to thank my supervisor A/Prof. Phillip Urquijo. His support and mentorship have made this work possible. He was also the one that inspired me to leave my home country and perform my research overseas. My special thanks goes to Daniel and Marcel, both of whom brilliantly contributed with their collaboration for my work. I also would like to thank all members of the particle physics groups at the University of Melbourne. There are too many amazing people list all of them here. My special thanks I would like to give, in order of proximity to my desk: M. Milesi, F. Scutti, C. MacQueen and G. Caria. You guys made the sometimes dry work at the university very fun and interesting. Finally, I want to thank my family in Germany and N. Lin for their constant support and love. Lastly and very importantly, I want to mention that I was awarded a Melbourne Research Scholarship by the University of Melbourne for which I am extremely grateful.

Contents

Dedicated to Elisabeth Töpper †*29.06.2018*

“ 一顆明珠 - Ikka no myōju - [The whole universe is in its ten directions is] One Bright Pearl.”

Gensa Shibi (China, 835-907) [1]

Chapter 1

Introduction

Particle Physics is a field of science that aims to understand nature and the universe as a whole. This is done following a reductionist philosophy that many thinkers of different cultural backgrounds laid the conceptual ground work. The commonly accepted framework for this description at small scales, the Standard Model (SM), is not able to reduce the universe to a single symmetric object like Gensa Shibi, who laid the ground stones for Japanese Zen-Buddhism, suggested. Instead, it uses a set of building blocks and their interactions based on symmetries. But the goal is still a reduction to more fundamental objects. Its predictions have been very successfully confirmed over the last five decades and only a few fundamental discoveries that can not be explained by the SM, like for example gravitational waves, have been made. However, it has been known to be incomplete from the point on it was conceived, since it does not include gravity. Gravity is an effect that is weaker than the forces of the SM by a factor of $\approx 10^{40}$. This explains why the SM, even though incomplete, is still a very effective theory.

The publicly well known collider experiments of recent years, such ATLAS or CMS at the Large Hadron Collider, search for new physics beyond the SM at higher and higher energy scales via direct production. Less known are precision experiments that search for new physics by measuring known SM effects with a high precision in order to find smaller and rarer deviations that may point to new physics effects. The context of this thesis are experiments of the latter type.

The Belle and Belle II experiments, among other aims, aim to measure CP violation of the $B-\bar{B}$-meson system. The precision of the experiments improves as the size of the accumulated data grows. Consequently, the theory tools and assumptions have to be tested with a high precision, in order to allow improvements to match the experimental side. One such assumption is QCD factorisation, which is used to calculate hadronic contributions to the direct and mixing CP-violation observables $\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}$ and $\mathcal{A}_{\mathrm{CP}}^{\mathrm{mix}}$. In the context of this thesis the branching ratios of the hadronic B-decays $B^0 \to D^{*-}\pi^+$ and $B^0 \to D^{*-}K^+$ will be measured. These allow us to measure key observables to test QCD factorisation as a tool to predict amplitudes in hadronic B-decays with unprecedented precision.

Experimental work in Particle Physics takes place in large international collaborations. The work presented in this thesis builds on top of work that other people have provided, such as the collider team that operated the collider, data acquisition, the sub-detector groups, physics performance groups, and the software team, to name a few. In order to be allowed to analyse data, one has to contribute to the collaboration. The algorithm development during this thesis, the last chapter, can be seen as this contribution, in addition to data acquisition shifts for Belle II. It is also a fundamental tool for physics analysis at Belle II for background suppression and CP-violation measurements.

The time-frame of this thesis falls between the end of the Belle experiment and the start of the Belle II experiment. As a consequence, most of the high profile analyses like the measurement of CP-violation in the B-meson system have already been performed at Belle. As for Belle II, the experiment is still in the detector calibration stage and the data sets are very small. Therefore this thesis includes work done for both experiments. A Belle dataset will be analysed with the main motivation to lay the groundwork for high precision CP-violation measurements in the future, once the full Belle II dataset is available. The interpretation of the results of these analyses are limited by the knowledge of hadronic effects, which are effects of the strong interaction. Many of these effects have been known since the early 1990s, however, the theoretical assumptions made in order to calculate hadronic B-decays, have never been challenged with high precision tests, despite the datasets being available for a long time. Therefore it is important to provide this input before the availability of the Belle II data sets, to give the theory community time to prepare high precision predictions and to maximise the usefulness of both.

This thesis is structured as follows. A general introduction will be given in Chapter 1. Chapter 2 will give an overview of the Standard Model and the theoretical motivations behind the measurements performed in Chapter 5. In Chapter 3 the Belle and Belle II experiments and detector details will be discussed. Furthermore, the analysed dataset will be discussed. Chapter 4 will give an overview over the statistical tools that are used to analyse the data. Chapter 5 will discuss the physics analysis. The branching ratios of $B^0 \to D^{*-}\pi^+$ and $B^0 \to D^{*-}K^+$ will be measured with unprecedented precision and the ratio of the branching ratios will be measured in a direct way that allows for the cancellation of the systematic uncertainties. This has never been done before. Chapter 6 will provide a phenomenological analysis of the results in collaboration with Daniel Ferlewicz, who kindly provided momentum transfer dependent branching ratio results of $B^0 \to D^{*-}l^+\nu$ decays. These allow for use of the branching ratios measured in this thesis to extract physics parameters and to test QCD factorisation. In Chapter 7 an algorithm developed for the Belle II experiment that allows for improvements in the reconstruction precision and software performance will be presented.

This thesis uses the following conventions. Energies and momenta are given in natural units, i.e. the speed of light and the Planck Constant are set to unity, $c = 1, \hbar = 1$. Metric units are obtained by multiplying with the corresponding factors for energies GeV $\to$ GeV/c^2 and for momenta GeV $\to$ GeV/c. Furthermore, implicit charge conjugation is implied, reconstructing π^- always implies reconstructing π^+ and for decays, a statement made for $D^0 \to K^-\pi^+$ implies that $\bar{D}^0 \to K^+\pi^-$ is included. Decays are sometimes abbreviated as $D^0(K^-\pi^+) = D^0 \to K^-\pi^+$ in longer decay chains for viewing convenience.

Chapter 2

Theoretical motivations

In this chapter the fundamental theoretical principles of particle physics that are needed to understand the motivation behind the measurement in this thesis will be explained. The first section will focus on the Standard Model of Particle Physics followed by a discussion of modern QCD calculations, subsequently new physics models will be briefly discussed. The standard model section largely follows the description as in Ref. [2], while the section describing hadronic B-decays is based on Ref. [3].

2.1 The Standard Model of particle physics

The Standard Model of Particle Physics (SM) is a theory that describes nature at high energy densities. It is following a reductionist description of nature using a minimal set of building blocks called particles and their interactions, based on symmetries and gauge invariance. Particles are characterised by their masses and quantum numbers, such as spin and charges. Generally, there are two sets of particles that can be distinguished, particles with spin $s = \pm\frac{1}{2}$, called fermions and particles with integer spin, named bosons. Irreducible, fundamental bosons are mediating forces between fermions. Forces arise from underlying symmetries. The known forces are, the strong force, following a non-abelian $SU(3)$ group with eight gluons as the gauge bosons, the weak force, following an non-abelian $SU(2)_L$ group with the two charged $W^{\pm}$ and neutral Z^0 gauge bosons and the electro-magnetic force, following an abelian $U(1)$ group with the γ as the only gauge boson. The SM therefore is a quantum filed theory with the following gauge group $G_{SM} = SU(3) \times SU(2)_L \times U(1)$. Fermions can be further categorised by whether or not they interact by the strong force. Those that do are called quarks, the others leptons. Measurements of the Z^0 invariant mass have shown that there are three generations of quarks and leptons in a SM with massless neutrinos. Dynamics of the theory are describe via a Lagrange density $\mathcal{L}(\phi)$ that is gauge invariant. Gauge invariance forbids mass terms for the bosons, suggesting that all gauge bosons should be massless. However, measurements of the $SU(2)_L$ gauge bosons have shown that these have masses. Therefore another mechanism, the Higgs mechanism, is needed to explain this. The Higgs mechanism introduces a spontaneous symmetry breaking [4, 5], giving mass to all particles, including fermions. This requires the introduction of a new particle, the Higgs Boson, which was proven to exist by the CMS and ATLAS collaborations in 2012 [6, 7]. This was considered the greatest success of the SM in recent years. A summary of all particles contained in the SM and their quantum numbers can be found in Fig. 2.1. However, the SM does not qualify as a theory for everything as it does not provide an explanation for phenomenons such as Dark Matter or gravity. Furthermore, it assumes neutrino to be massless, which has already been proven false [8].

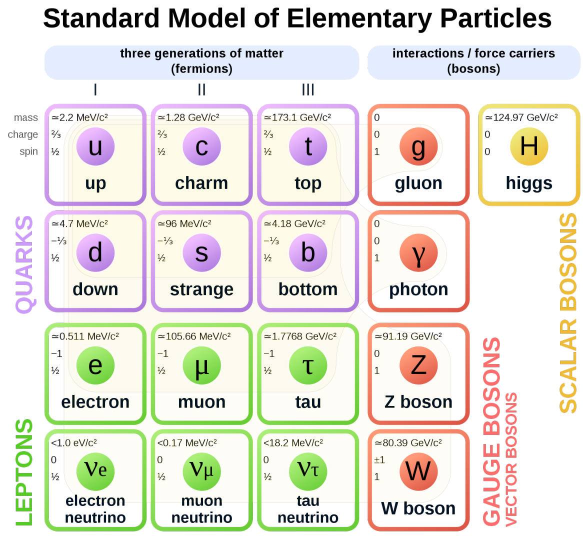


FIGURE 2.1: Overview of the particles contained in the Standard Model and their parameters.

2.1.1 The weak interaction

The strong and electro-magnetic interaction conserve lepton and quark flavour; as a consequence, many particles can only decay via the weak force. An example decay is the muon decay $\mu \to e\bar{\nu}_e\nu_\mu$. The chirality of the weak interaction was experimentally found to be only left handed and therefore the right handed operator $P_R = \frac{1}{2}(1+\gamma^5)$ does not appear in the SM Lagrangian of the weak interaction. Consequently, only left handed neutrinos and right handed anti-neutrinos can participate in the interaction. Furthermore, by decomposing the Hamiltonian into an vector $\bar{\psi}\gamma^\alpha\psi$ and a axial-vector term $\bar{\psi}\gamma^\alpha\gamma^5\psi$, one can show that the axial part is invariant under parity operations, while the vector part it not. This means that when comparing particles and anti-particles, or more rigorously, applying the operators charge and parity, C, and P, the combination, CP, can be violated, an effect known as CP-violation. This will be discussed in more detail later. The global phase of the wave function as given by the operators $\theta = CPT$, where T is the time operator, cannot be violated. Additionally, the weak force allows quark flavours to mix. The mixing strengths are described by a complex unitary matrix called the Cabibbo-Kobayashi-Maskawa (CKM) matrix. This matrix connects mass eigenstates $\{d, s, b\}$ and the respective weak force eigenstates $\{d', s', b'\}$ for down type quarks as

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{2.1}$$

Using unitarity constraints one can derive all parameters from a set of four fundamental parameters. These are three angles in the complex plane and a complex phase. The existence of CP-violation can be directly attributed to the existence of the complex phase. It has been measured successfully by the Belle [9] and BaBar [10] experiments and a Nobel Prize was awarded to Kobayashi and Maskawa in 2008 for its formulation and therefore prediction of the existence of a third generation of quarks.

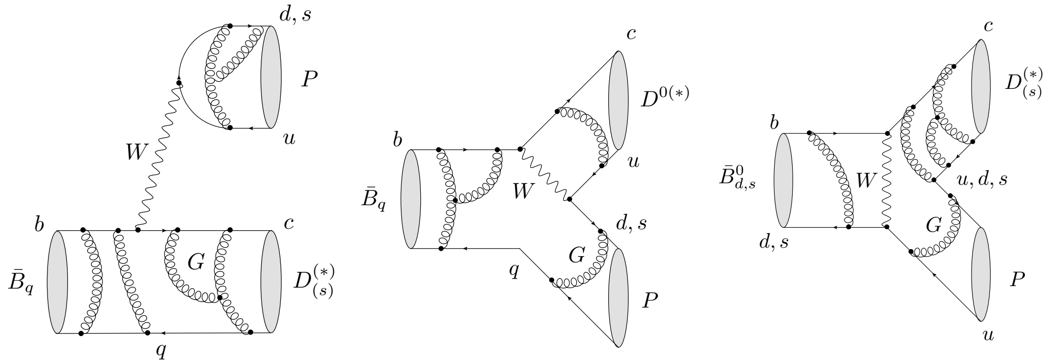


FIGURE 2.2: Hadronic decay topologies of the B-meson into a heavy charm meson and a hadron. The colour allowed tree (T), colour suppressed (C) and the exchange topology (E) are all possible [11].

2.1.2 The strong interaction

Quantum chromodynamics (QCD) is often referred to as the strong interaction, because its effects are much stronger than the electro-magnetic interaction. This allows, for example a proton consisting of three quarks of the same electro-magnetic charge to be stable. The charge of QCD is named colour. The major difference between QCD and all other interactions is that the underlying gauge group is non-abelian, which allows the gauge bosons to self interact. As a consequence the gauge boson themselves have to be colour charged and colour charged particles can only exist in colour neutral objects. This means that single quarks and gluons cannot exist on their own, an effect known as confinement. Colour neutral particles containing two quarks are called mesons, while baryons consist of three quarks, both are categorised as hadrons. Furthermore renormalising the theory and calculating loop correction becomes substantially more difficult. There are two approaches to calculating dynamics, the small mass limit and Heavy Quark Effective Theory (HQET) for heavy quarks. The former is suitable for calculations only involving u, d, s quarks and the latter for c, b, t quarks.

2.2 $\bar{B}_d \to D^{(*)}h$ decays

Hadronic B-meson decays such as $\bar{B}_d \to D^{(*)}h$, where h can be any hadron are interesting for a variety of aspects. Their branching ratios are large, which makes them a good testing ground for theoretical predictions; CP-asymmetries can be measured, their polarisations can be probed for new physics currents and they can be measured in direct searches for new physics.

In this section the phenomenology and motivation for the measurements in this thesis will be discussed. The details of the theoretical calculations of the observables will follow.

2.2.1 Phenomenology

Hadronic decays of the B-meson involving a heavy charm meson, $D^{(*)}_{(s)}$, can occur via three topologies. These are the colour allowed tree (T), colour suppressed (C) or the exchange topology (E), as depicted in Fig. 2.2. The relative amplitudes of these depend on, the quark content of $D^{(*)}_{(s)}$ and h. For $\bar{B}_d \to D^{*+}K^-$ only (T) is allowed, while for $\bar{B}_d \to D^{*+}\pi^-$, both (T) and (E) are possible.

The branching fraction can be written as a function of an amplitude A, a phase-space factor $\Phi^d_{D^{*+}}$ and the lifetime of the B-meson, τ, such that

$$\mathcal{B}(\bar{B}_d \to D^{*+}h^-) = \left|A(\bar{B}_d \to D^{*+}h^-)\right|^2 \Phi^d_{D^{*+}h^-}\tau_{B_d}. \tag{2.2}$$

The amplitude directly relates to the respective diagram

$$A(\bar{B}_d \to D^{*+}K^-) = T^*_K, \tag{2.3}$$

and

$$A(\bar{B}_d \to D^{*+}\pi^-) = T^*_\pi + E^*_\pi, \tag{2.4}$$

where the * denotes the presence of a D^*. Doubly Cabibbo suppressed contributions from $b \to u$ transitions are ignored due to their size of less then 1%. Therefore, by taking the ratio of the branching fractions one can asses the size of any $SU(3)$ symmetry-breaking affects and the size of (E) as

$$\frac{\mathcal{B}(\bar{B}_d \to D^{*+}K^-)}{\mathcal{B}(\bar{B}_d \to D^{*+}\pi^-)} = \left|\frac{T^*_K}{T^*_\pi + E^*_\pi}\right|^2. \tag{2.5}$$

It can be assumed that $T^*_\pi = T^*_K$ if $SU(3)$ symmetry holds. Performing a high precision measurement of this ratio is the first motivation for this thesis and results are presented in Sec. 5.12.

Another category of very interesting tests of hadronic effects are measurements of a_1, a parameter describing the deviation from naive factorisation of the two meson currents, as discussed in Sec. 2.2.6. It is part of the hadronic branching ratios and can be isolated by taking the ratio with respect to semi-leptonic decays. These only have strong interactions between the quarks and gluons within the $B \to D^{(*)}$ current, as shown in Fig. 2.3. By measuring the ratio at the squared momentum transfer of the respective light meson $q^2 = m(K/\pi)$, these hadronic effects cancel, so that

$$R^* = \frac{\Gamma(\bar{B}^0_d \to D^{*+}h^-)}{d\Gamma(\bar{B}^0_d \to D^{*+}l^-\bar{\nu})/dq^2|_{q^2=m_h}} = 6\pi^2\,\tau_B\,|V_{uq}|\,f_h^2\,|a_1(D^*h)|^2\,X_h, \tag{2.6}$$

where $X_h = 1 + \mathcal{O}(m_h^2/m_B^2)$ and f_h is a meson decay constant. Thus, using this ratio, the a_1 contribution can be measured with high precision, because the CKM matrix element V_{uq} and the meson decay constant f_h are known with high accuracy. Performing the measurement of a_1 is the second phenomenological aim of this thesis. The results are discussed in Sec. 6.1.

2.2.2 Hadronic effects in direct CP-violation

The fact that for some decays the invariance of charge, C and parity, P, operators in combination is violated, can be attributed to a complex phase ϕ_i that enters a process via the CKM matrix. It changes sign under CP. The most general way one can write the amplitudes for such a process is using two complex numbers

$$A = |M_1|\,e^{i\theta_1}e^{i\phi_1} + |M_2|\,e^{i\theta_2}e^{i\phi_2} \tag{2.7}$$

and for the respective anti-particle

$$\bar{A} = |M_1|\,e^{i\theta_1}e^{-i\phi_1} + |M_2|\,e^{i\theta_2}e^{-i\phi_2}, \tag{2.8}$$

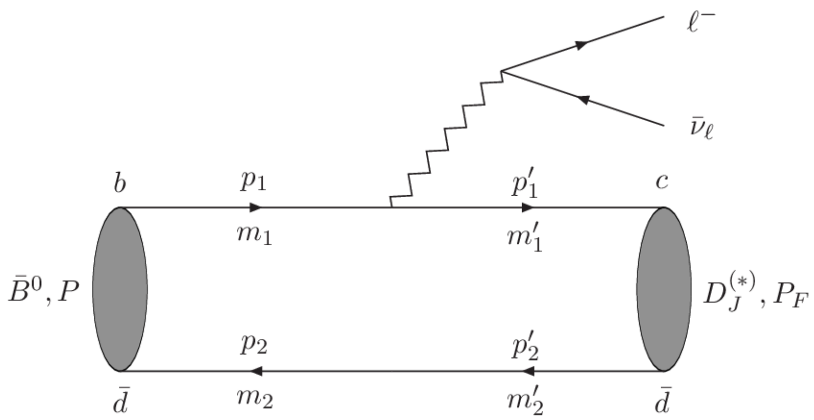


FIGURE 2.3: Semi-leptonic decay $\bar{B}_d^0 \to D^{*+}l^-\bar{\nu}$. This topology only has hadronic interactions taking place within the $B \to D$ current.

where M is the absolute value and the exponential terms are the phase, where θ is an additional strong phase. The observed physical amplitudes may be different for particles and anti-particles, if the phases are such that

$$|A|^2 - |\bar{A}|^2 = -4\,|M_1|\,|M_2|\sin(\theta_1 - \theta_2)\sin(\phi_1 - \phi_2), \tag{2.9}$$

is non-zero. This requires that $CP\ (A) \neq \bar{A}$, if $\phi_1 - \phi_2 \neq 0$ or $\theta_1 - \theta_2 \neq 0$ and thus CP is violated. However, it was found that for most processes $|\theta| < 5 \times 10^{-10}$, which makes the θ contribution negligible [12]. One can match these on the amplitudes as in Eq. 2.3 and Eq. 2.4 with the replacements $T^* \to \lambda\,|T|\,e^{i\gamma}$, where $|T|$ or $|E|$, is the hadronic matrix element, γ is a combination of the phases and λ is the CKM matrix element. This means that a large uncertainty on $|T|$ would always result in a large uncertainty on γ. The Belle II experiment aims for a high precision measurement of the phases. However, on the experimental side the improvements that can be made are limited to improving the decay vertex resolution and reducing the statistical uncertainty by collecting more data. As presented above, an orthogonal strategy to improve the precision is to narrow down the uncertainty on the hadronic matrix elements. Even though $\bar{B}_d^0 \to D^{*+}h^-$ may not be a suitable channel for direct CP-violation measurements, it is very well suited to test the theoretical framework used to calculate the matrix elements of other decays like $B_d^0 \to K_s J/\psi$, one of the golden modes to study CP-violation [13]. Furthermore, the measurement of this particular channel relies on $SU(3)$-symmetry, meaning that the calculation of the branching ratios do not fundamentally change, apart from a CKM matrix element, form factor and decay constant, when one of the quarks is replaced by a quark of a different flavour. In this particular example one uses the non-CP-violating decay $B_d^0 \to \pi^0 J/\psi$ as a normalisation mode. Thus exchanging an s quark with a d quark. However, the $SU(3)$-symmetry assumption has never been challenged with a high precision test.

2.2.3 The weak Effective Hamiltonian

Decays that change quark flavour can only proceed via the weak interaction. A very efficient way of describing these decays is via a generalisation of Fermi-Theory, describing interactions at low energies with an effective field theory. Low energies here refers to scales below M_W. It can be regarded as an expansion of local effective operators,

Q_i, and their Wilson coefficients, C_i, at some scale μ, such that

$$\mathcal{H}_{\text{eff}} = \frac{G_F}{\sqrt{2}} \sum_i C_i(\mu) Q_i(\mu), \tag{2.10}$$

where G_F is the Fermi constant. Amplitudes can be calculated via the scalar product of the inital state, i, and final state, f, as

$$A(i \to f) = \langle f | \mathcal{H}_{eff} | i \rangle = \frac{G_F}{\sqrt{2}} \sum_i C_i(\mu) \langle f | Q_i(\mu) | i \rangle . \tag{2.11}$$

To obtain a scale independent expression, the scale dependence of $C_i(\mu)$ must somehow cancel in $\langle f | Q_i(\mu) | i \rangle$. This is usually done by choosing the scale such that short distance effects are contained in $C_i(\mu)$ and long distance effects are put into $Q_i(\mu)$. As a consequence the challenging part will be the calculation of $\langle f | Q_i(\mu) | i \rangle$. The following will give an overview of how to calculate both to arrive at the amplitudes.

2.2.4 Calculation of the amplitudes

The decays of B-mesons contain at least one heavy quark and are usually viewed as the decay of this heavy quark in a four-point interaction, where the second quark is just a spectator that does not participate. Hence these are called spectator quarks. The interaction is described by an effective Hamiltonian and correction factors to account for the presence of the second quark and higher order interactions. The amplitude for two body decay of the B-meson into two mesons $B \to M_1 M_2$ is given by

$$\mathcal{A}(B \to M_1 M_2) = \frac{G_F}{\sqrt{2}} \sum_i \lambda_i C_i(\mu) \langle M_1 M_2 | \mathcal{O}_i | \bar{B} \rangle (\mu). \tag{2.12}$$

The virtual QCD effects above the scale m_b renormalise the local operators $\mathcal{O}_i$ of the weak Hamiltonian and are well understood [14]. The CKM matrix elements are λ_i and the coefficient function $C_i(\mu)$ accounts for QCD effects for scales above $\mu \sim m_b$. The explicit effective Hamiltonian for $\bar{B}^0_d \to D^{*+} h^-$ can be written as

$$H_{eff} = \frac{G_{\text{F}}}{\sqrt{2}} V^*_{\text{ud}} V_{\text{cb}} (C_0 O_0 + C_8 O_8), \tag{2.13}$$

with the singlet operator

$$O_0 = \bar{c}\gamma^\mu (1 - \gamma_5) b \bar{d} \gamma_\mu (1 - \gamma_5) u, \tag{2.14}$$

and the octet operator

$$O_8 = \bar{c}\gamma^\mu (1 - \gamma_5) T^A b \bar{d} \gamma_\mu (1 - \gamma_5) T_A u. \tag{2.15}$$

2.2.5 Wilson coefficients

The Wilson coefficients, C_i, at next-to-leading order (NLO) are given by

$$\begin{aligned} C_0 &= \frac{N_c+1}{2N_c}C_+ + \frac{N_c-1}{2N_c}C_-, \\ C_8 &= C_+ - C_-, \\ C_\pm(\mu) &= \tilde{C}_\pm(\mu)\left(1+\frac{\alpha_s(\mu)}{4\pi}B_\pm\right), \\ \tilde{C}_\pm(\mu) &= \left[\frac{\alpha_s(M_W)}{\alpha_s(\mu)}\right]^{d_\pm}\left[1+\frac{\alpha_s(M_W)-\alpha_s(\mu)}{4\pi}(B_\pm - J_\pm)\right], \end{aligned} \tag{2.16}$$

with NLO running coupling α_s defined as

$$\alpha_s(\mu) = \frac{4\pi}{\beta_0 \ln(\mu^2/\Lambda^2_{QCD})}\left[1-\frac{\beta_1}{\beta_0^2}\frac{\ln(\mu^2/\Lambda^2_{QCD})}{\ln(\mu^2/\Lambda^2_{QCD})}\right], \tag{2.17}$$

where

$$\beta_0 = \frac{11N_c-2f}{3}, \quad \beta_1 = \frac{34}{3}N_c^2 - \frac{10}{3}N_c f - 2C_F f, \quad C_F = \frac{N_c^2-1}{2N_c}, \tag{2.18}$$

and

$$d_\pm = \begin{cases} \frac{6}{23} \\ -\frac{12}{23} \end{cases} \qquad B_\pm - J_\pm = \begin{cases} \frac{6473}{3174} \\ -\frac{9371}{1587}. \end{cases} \tag{2.19}$$

Here $N_c = 5$ are the number of colours and $f = 5$ the number of flavours. Calculating the matrix elements is more complicated and explained in the next section.

2.2.6 Using QCD factorisation to calculate the matrix elements

QCD factorisation refers to the process of approximating the matrix elements in a way that factorises into components that are calculable. For example for $B \to D\pi$ decays, one factorises the decay current $j_1 \otimes j_2$ into a $B \to D$ and a vacuum current $0 \to \pi$:

$$\langle D\pi|\, j_1 \otimes j_2\, |B\rangle = \langle D|\, j_1\, |B\rangle \langle \pi|\, j_2\, |0\rangle\,. \tag{2.20}$$

More generally, for a heavy meson M_1 and a light meson M_2 at leading order in Λ_{QCD}/m_b the transition matrix elements of the operators O_i can be expressed by a form factor and a convolution integral;

$$\langle M_1 M_2|\, \mathcal{O}_i\, |\bar{B}\rangle = \sum_j F_j^{B\to M_1}(m_2^2)\int_0^1 du\, T^I_{ij}(u)\Phi_{M_2}(u), \tag{2.21}$$

where $F_j^{B\to M_{1,2}}(m^2_{1,2})$ is a form factor describing the $B \to M_{1,2}$ transition, $\Phi_{M_2}(u)$ is the light cone distribution amplitude of the quarks in the light meson and $T^I_{ij}(u)$ is a scattering kernel, described in the normalised phase space u. The meaning of these quantities in the context of Feynman graphs is depicted in Fig. 2.4. This assumption ignores the annihilation diagram, i.e. the one including T_i^{II}. The second diagram contributes at order $(\Lambda_{QCD}/m_b)^2 \sim 0.0025$ and thus is far away from experimental sensitivity.

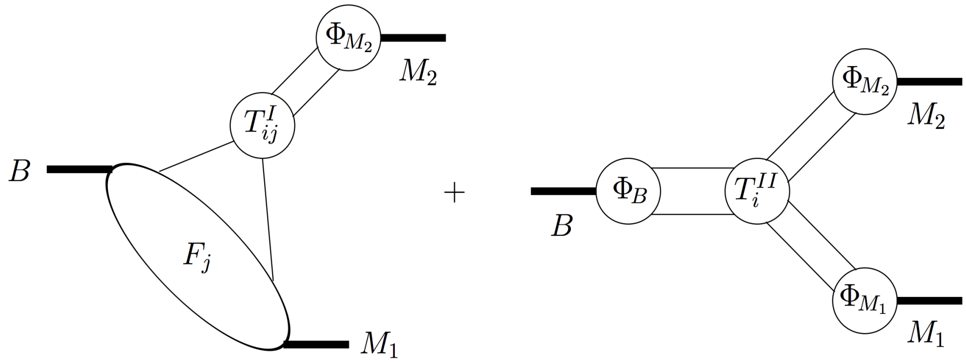


FIGURE 2.4: QCD factorisation terms and their meaning in the context of a Feynman diagram [3]. The first depiction is for type I decays where M_1 is heavy and M_2 is a light meson. The second "annihilation" diagram contributes only at $(\Lambda_{QCD}/m_b)^2$ to these.

In the following the definitions of these functions and their calculation will be discussed using the NLO Wilson coefficients. The descriptions in next-to-next-to-leading order (NNLO) are available and used in Ref. [15], but will not be discussed further due to their complexity.

The form factor can be calculated assuming the dominant contribution is driven by soft processes. One can approximate it by an overlap integral of light-cone wave functions such that

$$F^{B\to D}(0)|_{q^2=0} \sim \int \frac{d\xi d^2k_\perp}{16\pi^3} \Psi_B(\xi, k_\perp)\Psi_D(\xi'(\xi), k_\perp), \tag{2.22}$$

at momentum transfer $q^2 = 0$. Furthermore, one can show that this scales as a constant. For a more rigorous discussion in the context of semi-leptonic decays see Ref. [16, 17, 18]. The scattering kernels $T^I_{0,8}$ for $M_1 = D^*$ depend on the symmetries in $\phi_L(M_2)$ and are given by

$$T_0(u, z) = 1 + \mathcal{O}(\alpha_s^2), \tag{2.23}$$

for the singlet operator, and

$$T_8(u, z) = \frac{\alpha_s}{4\pi}\frac{C_F}{2N_c}\left[-6\ln\frac{\mu^2}{m_b^2} - B + F(u, -z)\right] + \mathcal{O}(\alpha_s^2), \tag{2.24}$$

for the octet. The function B is scheme dependent and either $B = 11$ in the naive dimensional regularisation scheme (NDR) or $B = 9$ in the “t-Hooft-Veltman” scheme. The light cone distribution function $\phi_L(u)$ is symmetric for $h^- = \pi^-$ under $u \leftrightarrow \bar{u}$. In that case the loop function $F(u, z)$ is given by

$$F(u, z) = 3\ln z^2 - 7f(u, z) + f(u, 1/z), \tag{2.25}$$

and

$$f(u, z) = -\frac{u(1-z^2)(3(1-u(1-z^2))+z)}{(1-u(1-z^2))^2}\cdot\ln(u(1-z^2)) - \frac{z}{1-u(1-z^2)}. \tag{2.26}$$

In the case that $\phi_L(u)$ is not symmetric under $u \leftrightarrow \bar{u}$ the expressions read

$$F(u,z) = (3 + \ln\frac{u}{\bar{u}})\ln z^2 - 7f(u,z) + f(\bar{u}, 1/z), \tag{2.27}$$

where

$$\begin{aligned} f(u,z) = &-\frac{u(1-z^2)(3(1-u(1-z^2))+z)}{(1-u(1-z^2))^2} \cdot \ln(u(1-z^2)) - \frac{z}{1-u(1-z^2)} \\ &+ 2\left[\frac{\ln(u(1-z^2))}{1-u(1-z^2)} - \ln(u(1-z^2)) - \mathrm{Li}_2(1-u(1-z^2)) - \{u \leftrightarrow \bar{u}\}\right], \end{aligned} \tag{2.28}$$

and $\mathrm{Li}_2(x) = -\int_0^x dt/t \ln(1-t)$. This is the case for for $h^- = K^-$. For the operators one can write using the factorisation formula

$$\langle D^{*+}L^-|\, O_{0,8}\, |\bar{B}^0{}_d\rangle = \langle D^{*+}|\, \bar{c}\gamma^\mu(1-\gamma^5)b\, |\bar{B}^0{}_d\rangle \cdot if_L q_\mu \int_0^1 du T_{0,8}(u,z)\phi_L(u), \tag{2.29}$$

where f_L is a meson decay constant. It is important to note that at zeroth order only O_0 will contribute, otherwise the light quark pair would not produce a colour singlet at zeroth order of α_s and likewise at first order only O_8 can contribute.

Assuming Eq. 2.29 holds, one can join everything together and factorise corrections into a factor a_1 and define a transition operator for D^* and a light meson L,

$$\mathcal{T}_{B^0_d \to D^{*+}L^-} = -\frac{G_F}{\sqrt{2}} V^*_{uq} V_{cb} a_1(D^*L) Q_A, \tag{2.30}$$

with the operator

$$Q_A = \bar{c}\gamma^\mu\gamma_5 b \otimes \bar{d}\gamma_\mu(1-\gamma_5)u, \tag{2.31}$$

and a_1 defined as

$$\begin{aligned} a_1(D^*L) = &\frac{N_c+1}{2N_c}\tilde{C}_+(\mu) + \frac{N_c-1}{2N_c}\tilde{C}_-(\mu) \\ &+ \frac{\alpha_s}{4\pi}\frac{C_f}{2N_c} C_8(\mu)\left[-6\ln\frac{\mu^2}{m_b^2} + \int_0^1 du\, F(u,-z)\phi_L(u)\right]. \end{aligned} \tag{2.32}$$

Therefore the entire factorisation hypothesis is contained in this observable. It will be measured in this thesis and may give information about the validity of the factorisation approach, see Chapter 6.1.

To calculate the convolution integral in Eq. 2.32, the light cone distribution functions are expanded into Gegenbauer polynomials

$$\phi_L(u) = 6u(1-u)\left[1 + \sum_{n=1}^{\infty} \alpha_n^L(\mu) C_n^{3/2}(2u-1)\right], \tag{2.33}$$

with $C_1^{3/2} = 3x$ and $C_2^{3/2} = \frac{3}{2}(5x^2-1)$ and the series is taken only to second order. The theoretical results for a_1 can be found in Table 2.1.

TABLE 2.1: Theoretical results for a_1. With $\alpha_1^{\pi} = 0$ and $\alpha_1^K < 1$.

Parameter	Value	Source
a_1^{LO}	1.049	[3]
a_1^{NLO}	$1.054^{+0.018}_{-0.017} - (0.015^{+0.013}_{-0.007})i$	[3]
$a_1^{NNLO}(D^{*+}\pi^-)$	$1.071^{+0.013}_{-0.012} - (0.032^{+0.016}_{-0.010})i$	[15]
$a_1^{NNLO}(D^{*+}K^-)$	$1.068^{+0.010}_{-0.012} - (0.034^{+0.017}_{-0.011})i$	[15]

TABLE 2.2: Theoretical results for branching fractions.

Mode	LO	NLO	NNLO	Magnitude	Ref.
$\bar{B}_d^0 \to D^{*+}\pi^-$	3.15	$3.32^{+0.44}_{-0.42}$	$3.93^{+0.43}_{-0.42}$	10^{-3}	[15]
$\bar{B}_d^0 \to D^{*+}K^-$	2.17	$2.50^{+0.39}_{-0.36}$	$2.59^{+0.39}_{-0.37}$	10^{-4}	[15]
$\bar{B}_d^0 \to D^{*+}K^-$	NA	NA	$3.27^{+0.39}_{-0.34}$	10^{-4}	[19]
$\frac{BR(\bar{B}_d^0 \to D^{*+}K^-)}{BR(\bar{B}_d^0 \to D^{*+}\pi^-)}$	0.75	$0.75^{+0.02}_{-0.02}$	$0.75^{+0.02}_{-0.02}$	10^{-1}	[15]

2.2.7 Matrix elements and branching ratios

The matrix elements can be found from the product of the factor a_1, the form factor A_0, earlier called $F^{B\to D}$, and a meson decay constant f_h, such that

$$\mathcal{A}(\bar{B}_d^0 \to D^{*+}\pi^-) = -i\frac{G_F}{\sqrt{2}}V_{ud}^* V_{cb} a_1(D^*\pi) f_\pi A_0(m_\pi^2) 2m_{D^*}\epsilon^* \cdot p, \quad (2.34)$$

and

$$\mathcal{A}(\bar{B}_d^0 \to D^{*+}K^-) = -i\frac{G_F}{\sqrt{2}}V_{us}^* V_{cb} a_1(D^*K) f_K A_0(m_K^2) 2m_{D^*}\epsilon^* \cdot p, \quad (2.35)$$

where ϵ is the polarisation vector of the D^* and p the four-momentum. The branching ratios subsequently are given by

$$\Gamma(\bar{B}_d^0 \to D^{*+}\pi^-) = \frac{G_F^2|\vec{q}|^3}{4\pi}|V_{ud}^* V_{cb}|a_1(D^*\pi) f_\pi^2 A_0(m_\pi^2) \quad (2.36)$$

and

$$\Gamma(\bar{B}_d^0 \to D^{*+}K^-) = \frac{G_F^2|\vec{q}|^3}{4\pi}|V_{us}^* V_{cb}|a_1(D^*K) f_K^2 A_0(m_K^2). \quad (2.37)$$

The numerical predictions are listed in Table 2.2. The 2019 predictions for $\mathcal{B}(\bar{B}_d^0 \to D^{*+}K^-)$ determined in Ref. [19] show a larger deviation from measurements of 3σ from PDG values.

2.2.8 New physics models

A deviation of the branching ratio measurements from the above presented values would mostly likely be due to large non-factorisable contributions. However, another possibility could be contributions from new physics. Two possible models are suggested in Ref. [20]. These can enter the predictions via the Wilson coefficients at some scale Λ_{NP}. However, current measurements constraint their sizes significantly [21, 22, 23, 24, 25]. The first plausible model could be a modification of the Lagrangian by flavour changing neutral currents (FCNC) via Yuakawa couplings, Y^u, so that

$$\mathcal{L} = \frac{1}{2\Lambda_{\mathrm{NP}}^2}\left[\bar{Q}_L^i(\delta_{ij}) + a(Y^u Y^{u\dagger})\gamma^\mu Q_L^j\right]^2, \quad (2.38)$$

with $Y^u = V^\dagger \mathrm{diag}(y_u, y_c, y_t)$ and a dimensionless coupling a. The resulting modification to C_0, Eq.2.16, would read

$$C_0^{q,\mathrm{MFV}} = -\frac{1}{\Lambda_{\mathrm{NP}}^2}\frac{\sqrt{2}}{4G_f}\left[1 + a y_t \frac{V_{ts}^*}{V_{cb}}\right]. \tag{2.39}$$

However, their size is constrained to

$$-0.002 \lesssim \frac{C_0^{q,\mathrm{MFV}}(m_b)}{C_0^{q,\mathrm{SM}}(m_b)}. \tag{2.40}$$

Another possibility is an extension of the electroweak gauge group to $SU(2)_1 \times SU(2)_2 \times U(1)_Y$. This would imply the existence of two new massive gauge bosons Z'^0 and $W'^\pm$ and a set of vector-like fermions, V. Their interactions are given by

$$\mathcal{L} = +\frac{g_{ij}}{2} Z'_\mu \bar{d}^i_L \gamma^\mu d^j_L - \frac{(V g V^\dagger)_{ij}}{2} Z'_\mu \bar{u}^i_L \gamma^\mu u^j_L - \frac{(Vg)_{ij}}{\sqrt{2}} W'_\mu \bar{u}^i_L \gamma^\mu d^j_L + h.c.. \tag{2.41}$$

Setting $M_{W'^\pm} = M_{Z'} = M_V$ for simplicity and integrating out both $M_{W'^\pm}$, one obtains the modification of the Wilson coefficient as

$$C_0^{q,W'}(M_V) = \frac{1}{4\sqrt{2} G_F M_V^2} \frac{(Vg)_{23}(Vg)^*_{1q}}{V_{cb} V^*_{uq}}. \tag{2.42}$$

The flavour structure can be chosen to be

$$g_{ij} = \begin{pmatrix} g_{11} & 0 & 0 \\ 0 & g_{22} & g_{23} \\ 0 & g_{23} & g_{33} \end{pmatrix}, \tag{2.43}$$

where in principle another entry, g_{ij}, may be chosen to be zero. The size of the NP contributions of this model depends on this choice, $-0.05 \lesssim C_0^{W'}/C_0^{SM}$ for $g_{23} = 0$, $-0.01 \lesssim C_0^{W'}/C_0^{SM}$ for $g_{33} = 0$ and $-0.1 \lesssim C_0^{W'}/C_0^{SM}$ if all are non-zero. This allows deviations of around 10% from the SM expectations toward smaller values.

Chapter 3

Experimental setup

The data analysed in this thesis are based on e^+e^--collisions provided by the KEKB and SuperKEKB colliders, recorded by the Belle and Belle II detectors, located in Tsukuba, Japan. The Belle experiment utilised the Belle detector and KEKB collider from 1999 to 2010. It was then upgraded to the Belle II experiment and the SuperKEKB collider, which recorded its first collision data in April 2018. This chapter will summarise the experimental setups and briefly discuss the fundamentals of particle collider experiments.

3.1 Experiments with particle colliders

Particle colliders allow us to create and study environments of very high energy density as they were present in the primordial universe. This environment is needed to understand the fundamental interactions and symmetries of the universe. There are two basic principles that are exploited. The first is the widely know Einstein relation between the energy and invariant mass of a particle at rest,

$$E = m \cdot c^2, \tag{3.1}$$

where m is the invariant mass of the particle and c the speed of light. By accelerating particles to high energies and colliding them one can generate new particles based on the available energy. The second principle is the duality, first described by De Broglie, between particles of momentum p and waves of wavelength λ, such that

$$\lambda = \frac{h}{p}, \tag{3.2}$$

where $h = 6.62 \cdot 10^{-34}$ Js is the Planck constant. This relation connects the energy of a particle to a length scale; the larger the energy and hence the momentum of a particle, the smaller the structure it can probe. For example the smallest structure a microscope with a lamp shining through an object can resolve is given by the wavelength of the light. Particle colliders are characterised by their centre of mass energy (CMS), $\sqrt{s}$, found to be

$$\sqrt{s} \approx \sqrt{4E_1E_2}, \tag{3.3}$$

assuming the collisions are head on and large enough energies that particle masses can be neglected. The indices refer to the colliding particles. However, in reality the particle collisions take place at a certain angle and are not between single particles but bunches of particles that have a certain distribution.

The Belle and Belle II experiments operate at a collision energy slightly above the mass of the $\Upsilon(4S)_{b\bar{b}}$-resonance, $m(\Upsilon(4S)_{b\bar{b}}) = 10.579$ GeV [26], to maximise the probability of producing $\Upsilon(4S)_{b\bar{b}}$ particles. The cross section of $e^+e^- \rightarrow \mathit{Hadrons}$ as

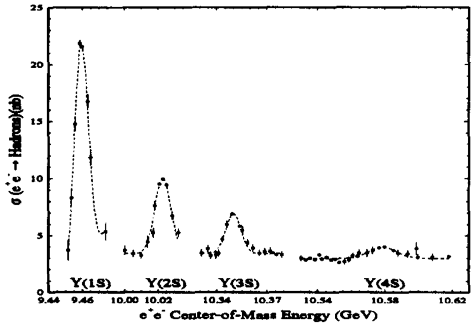


FIGURE 3.1: Cross section of e^+e^- collisions decaying into hadrons as a function of CMS energy [27].

a function of the beam energy is depicted in Fig. 3.1. $\Upsilon(4S)$-resonances decay $> 97\%$ of the time as $\Upsilon(4s) \to \bar{B}B$, allowing us to study the interactions of the heavy b-quark contained in the B-meson. These types of experimental setups are therefore called B-factories.

The amount of data a collision experiment collects over time depends on the instantaneous luminosity of the collider. The instantaneous luminosity, L, is a measure of the number of interactions, R, per unit time and the effective cross section of the process, σ, given by

$$L = \sigma^{-1}\frac{dR}{dt}. \tag{3.4}$$

Particle colliders collide clouds of particles with each other to increase the chance of particles hitting each other; these clouds are called bunches. Ignoring the crossing angle, offsets between the particle beams, potential differences in bunch size and the Hour-Glass effect, as discussed in Ref. [28], one can write the instantaneous luminosity as

$$L = \frac{N_1 N_2 f N_b}{4\pi\sigma_x\sigma_y}, \tag{3.5}$$

where $N_{1,2}$ is the number of particles in the colliding bunches, f is the collision frequency, N_b is the number of bunches in each beam and the bunch sizes are σ_x and σ_y. The integrated luminosity, $\mathcal{L}_{\text{int}}$, measures the luminosity collected in a time span $T_2 - T_1$ as

$$\mathcal{L}_{int} = \int_{T_1}^{T_2} L\, dt. \tag{3.6}$$

3.2 The KEKB and SuperKEKB collider

The KEKB collider was an asymmetric e^+e^--collider. It provided a maximum peak luminosity of $L = 2.11 \cdot 10^{34}$ cm^{-2}s^{-1} and a total integrated luminosity of $\mathcal{L}_{int} = 710\,\text{fb}^{-1}$ at the $\Upsilon(4S)_{b\bar{b}}$-resonance. The colliding beams were configured with energies of 3.5 GeV for the low energy e^+-ring (LER) and 8 GeV for the high energy e^--ring (HER) [29]. Due to the asymmetry of the beam energies, the generated B-mesons were produced with a relativistic boost that that allows for the study of time dependent CP-violation. The design instantaneous luminosity target of SuperKEKB is $8 \cdot 10^{35}$ cm^{-2}s^{-1}, more than 10 times higher than KEKB. And a total integrated luminosity of 50 ab^{-1} will be provided at the $\Upsilon(4S)_{b\bar{b}}$-resonance. The asymmetry of the beam energies has been reduced to 4 GeV in the LER and 7 GeV in the HER. Furthermore, the interaction region was redesigned by introducing a nano-beam

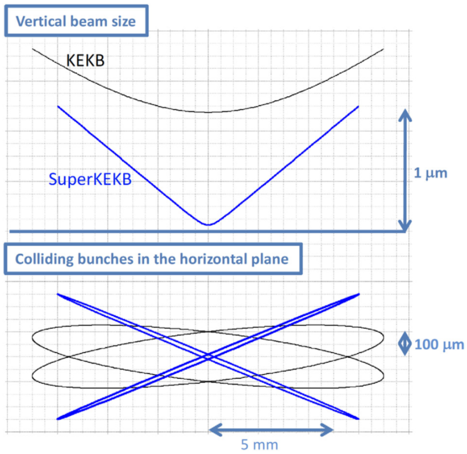


FIGURE 3.2: Schematic depiction of the interaction region comparing KEKB and SuperKEKB.

scheme that decreases the cross-sectional area, $\sigma_{x,y}$ in Eq.3.5, in which the collisions take place, depicted in Fig. 3.2, to increase the instantaneous luminosity. A schematic depiction of the two colliders is displayed in Fig. 3.3.

3.3 The Belle and Belle II detector

The Belle [30] and Belle II [31] detectors are optimised for the detection of particles in an energy range of approximately 20 MeV up to 5 GeV in a hermetic design, surrounding the interaction points of the KEKB and SuperKEKB colliders, respectively. The Belle detector was upgraded to the Belle II detector. The most important differences are the silicon vertex detectors, that were upgraded to a 2-layer Pixel Detector (PXD) and a 4-layer silicon strip detector (SVD), and the replacement of the time of flight detectors (TOF) in favour of a larger tracking volume and Cherenkov Angle based detectors for Kaon-Pion separation (TOP and ARICH). The Central Drift Chamber (CDC) was completely replaced with a new chamber with supercell structure and finer spacing at inner radii. Most other systems like the K_L^0 and Muon detector, the Electromagnetic Calorimeter (ECL) and the Solenoid Magnet only underwent smaller changes. However, to achieve a nano-beam scheme the final focussing magnets, QCS in Fig. 3.4, had to be made larger, decreasing the solid angle coverage of the detectors forward and backward in tracking acceptance. The individual components of the detector will be discussed in some detail in the following, with more detailed descriptions in the technical documentation of the experiments.

3.3.1 The Belle detector

SVD The innermost layer of the detector is the Silicon Vertex Detector which is used for precise tracking of charged particles close to the interaction point. It is designed for the measurement of the decay vertices of short lived particles such as B-mesons. Siliconstrip sensors were used to measure the three dimensional flight paths of charged particles and together with a magnetic field, one can infer their momentum by measuring the curvature of the particle track. Two configurations were used, in the first data taking period between 1999-2004 it was operated with 3 sensor layers and a polar angular coverage between $23° < \theta < 139°$ degrees. This configuration is

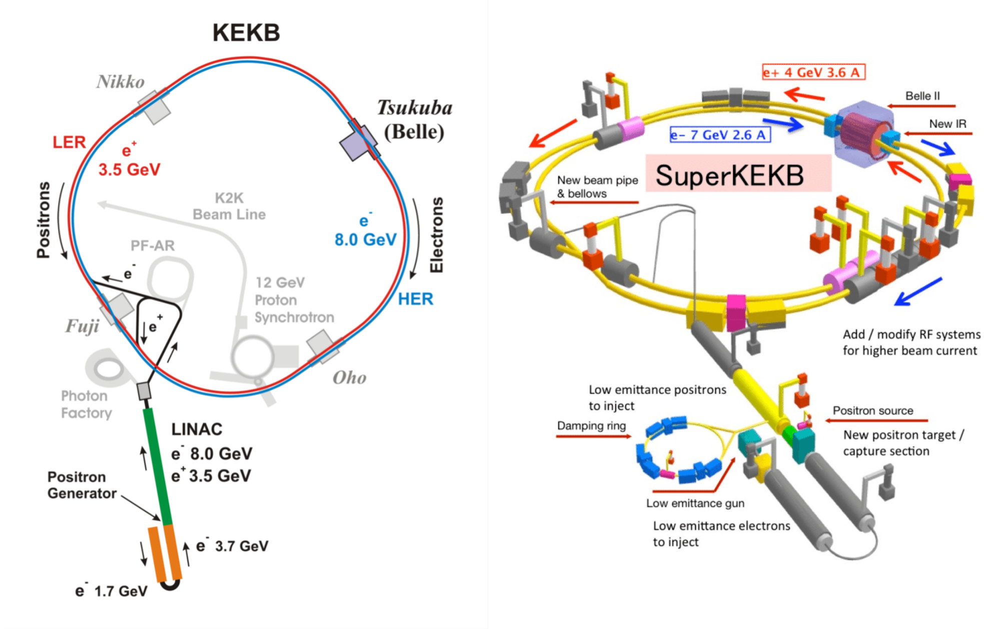


FIGURE 3.3: Schematic depiction of KEKB (left) and SuperKEKB (right). In addition to the upgraded optics and redesigned interaction region, SuperKEKB has been equipped with a damping ring.

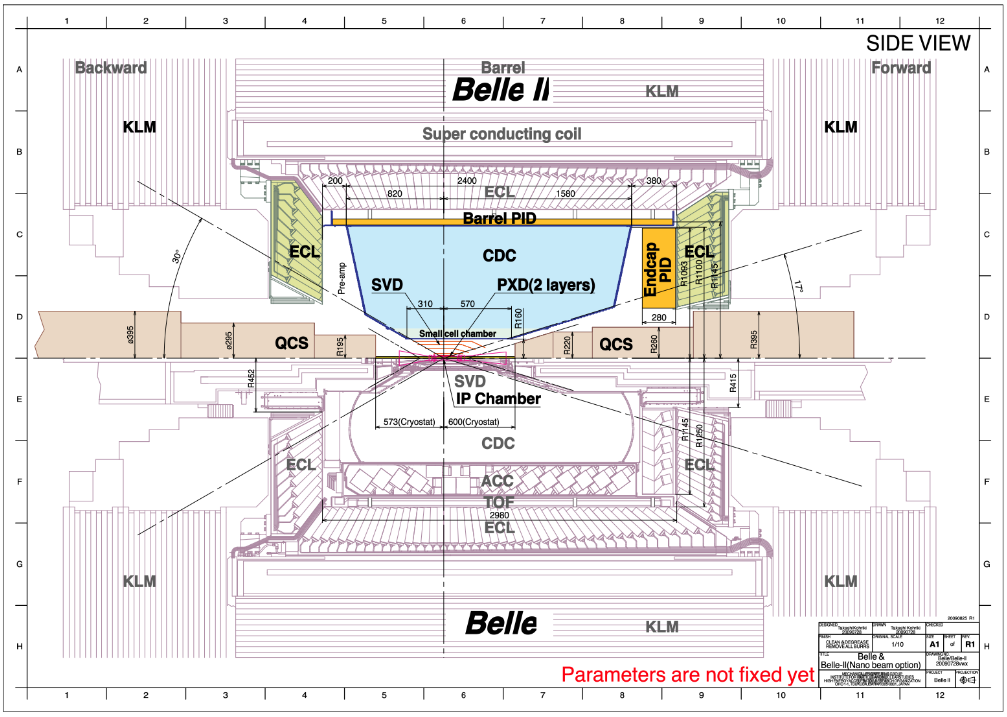


FIGURE 3.4: Depiction of the Belle (below) and Belle II (above) detector.

referred to as SVD1. In 2004 it was upgraded to 4 layers of sensors and with a polar angular coverage of $17° < \theta < 150°$ degrees, referred to as SVD2.

CDC The Central Drift Chamber is the largest tracking detector. It measures the transversal flight path of particles and their momenta. Furthermore, by independently measuring the specific energy loss of a track, dE/dx, it is possible to infer the mass of traversing particles. The CDC consists of 50 cylindrical layers of drift cells organised in 11 super-layers. Each cell consist of a sense wire and field wires. The angular coverage of the CDC is $17° < \theta < 150°$ degrees along the axis of the beam. The detector is filled with a saturated gas and a high voltage is applied between the sense and field wires. As charged particles pass through they ionize the gas along their trajectory. The freed charge will be accelerated by the potential, ionizing more of the gas and thus creating an avalanche of charge that can be measured in each sense wire. By connecting the triggered drift cells, one can infer the particle trajectory, which combined with a magnetic field allows for a calculation of the energy loss, dE/dx, and the particles momentum.

ACC The Aerogel Cherenkov Counter surrounds the CDC in the barrel and in the forward direction. Its purpose is to measure the Cherenkov light emitted by high momentum particles passing through an aerogel radiator. The ACC is a threshold counter, tuned such that only certain particles emit radiation at the typical production energies at KEKB. This allows for the differentiation of pions, kaons and protons better than just purely using the CDC. The ACC consists of 960 cells of $12 \times 12 \times 12$ cm^3 size.

TOF The Time Of Flight (TOF) detector is the next layer following the ACC and surrounds it in the barrel region in a polar angle range of $34° < \theta < 120°$ and the forward end caps in a polar angle of $14° < \theta < 33°$. It consists of 128 plastic scintillator counters read out by photomultipliers with a time resolution of 100 ps. It is used to measure the velocity of low momentum particles by referencing it to the collision time. It is part of the particle identification system together with the dE/dx-measurement and the ACC.

ECL The Electromagnetic Calorimeter is used to measure the energy of charged and neutral particles by measuring the amount of energy deposited via electromagnetic interactions with the detector material. It is designed to fully absorb most of the electrons and photons. The ECL consists of 6634 CsI(Tl) crystals in the barrel and a total of 2112 in the end-caps. Each crystal has the shape of a truncated pyramid aligned towards the interaction point. The average crystal size is 6×6 cm^2 with a length of 30 cm, equivalent to > 16 electromagnetic and 0.8 hadronic radiation lengths.

Magnet A super-conducting Niobium-Titanium-Copper magnet provides an almost homogenous magnetic field of 1.5 T inside of the barrel of the detector, oriented along the beam axis. It is used to deflect the flight paths of charged particles into helices, such that their momentum can be measured. The magnet is actively cooled by liquid helium to a temperature at which the material becomes super-conducting.

KLM The outermost detector is the K_L^0 and muon detector. It is used to measure K_L^0 and muons and its stopping material serves as the flux return yoke of the magnet.

It is a sampling calorimeter interlacing 15 detector layers and 14 layers of iron that serve as stopping material. The detectors are instrumented with Glass Resistive Plate Chambers (RPC) that provide a binary hit or no hit measurement. This allows for a crude estimation of particle energies based on the number of layers penetrated by a particle.

Data acquisition of the Belle detector It is not possible to store the data of all recorded collisions, because the amount data is in the hundreds of TB per day and only few types of events are interesting from a physics point of view. For example, the rate of interesting physics events such as $e^+e^- \to \bar{q}q$ or $e^+e^- \to \bar{\tau}\tau$ is relatively small compared to the number of $e^+e^- \to e^+e^-(\gamma)$ reactions. In order to filter these, a dedicated real time data acquisition system (DAQ) is used. The Belle DAQ used four level of triggers to decide whether to store an event. The first two levels are hardware triggers forming the global decision logic (GDL) of the detector. They source low level detector information and decide within 2.2μs after a collision if it should be stored. For hadronic events there are four triggers, two based on the number of tracks in an event, one based on the total energy deposited in the ECL. The software based level 3 trigger decides based on the impact parameters of the tracks in an event, if these are relevant physics events, selecting only events with most of the tracks originating from the collision point. This helps to remove tracks coming from beam gas interactions or synchrotron radiation. It is done prior to writing the data onto tape, using faster less precise versions of the reconstruction software. Finally, the level four trigger filters the recorded events offline, using the full high precision software reconstruction. The total efficiency for hadronic events is very close to 100%, while much more than half of the background events are removed.

3.3.2 The Belle II detector

PXD and SVD The innermost detector is the Pixel detector. It consists of two layers of pixelated, depleted field effect transistors (PXD) of 50 μm thickness. It is followed by four layers of double-sided silicon strip sensors (SVD). Together they form the vertex detector, allowing for measurements of tracks of charged particles very close to the interaction point, where secondary decay vertices of B- and charm-meson will be located. A schematic depiction is displayed in Fig. 3.5.

CDC The CDC has been redesigned to fit the geometry of Belle II. It consists of 8+1 super layers of 6 (8) wires each. The innermost super layer is more densely packed to increase the resolution close to the SVD. A simulated physics event with multiple particles leaving tracks in the CDC is displayed in Fig. 3.6.

ECL The ECL has not been redesigned in Belle II. However, the readout electronics and software have been modernised to allow for better rejection of beam background and improved particle identification.

TOP Belle II is equipped with a Time Of Propagation quartz Cherenov detector made up of quartz crystal bars of $2750 \times 400 \times 20$ mm^3 dimension. They work by internal light reflection at a mirror at one end of the bar and a position vs. time measurement is performed with a silicon photomultiplier array, as displayed in Fig. 3.7. It is used to measure Cherenkov angles for different particle species, travelling with different velocities and is specifically designed to distinguish kaons from pions. The TOP replaces the Belle ACC and TOF in the barrel.

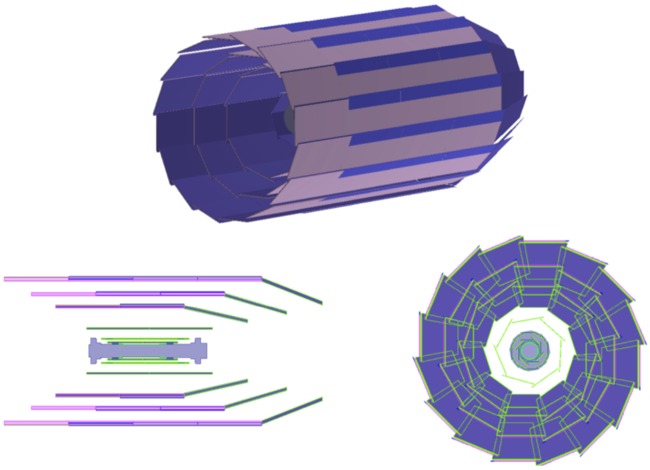


FIGURE 3.5: Depiction of the vertex detector of the Belle II detector. The two innermost layers are the PXD while the surrounding layers define the SVD.

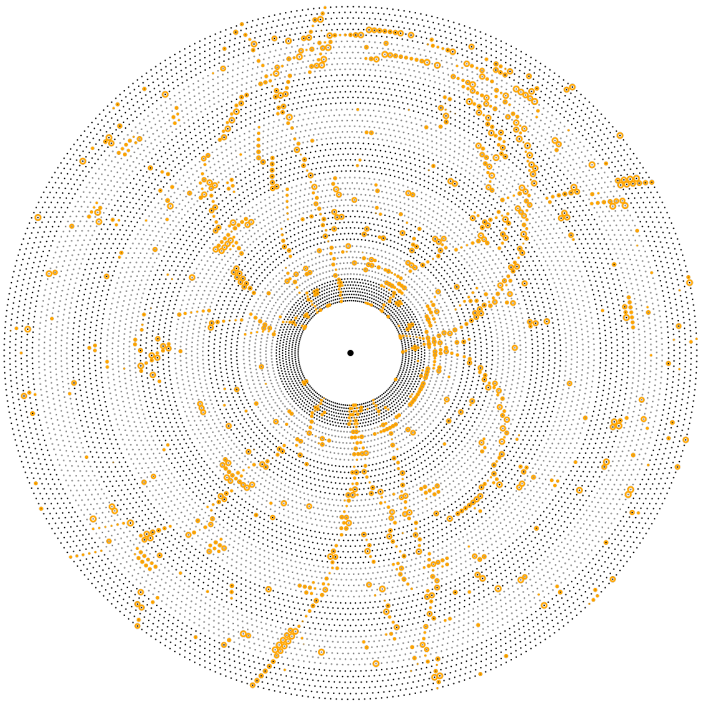


FIGURE 3.6: Schematic view of a simulated collision event, overlaid with background noise, in the CDC. The $X-Y$ plane is shown, along the detector axis of symmetry. Each dot corresponds to a simulated hit from either background or physics. The tracking algorithms will combine the individual hits to to form particle track objects.

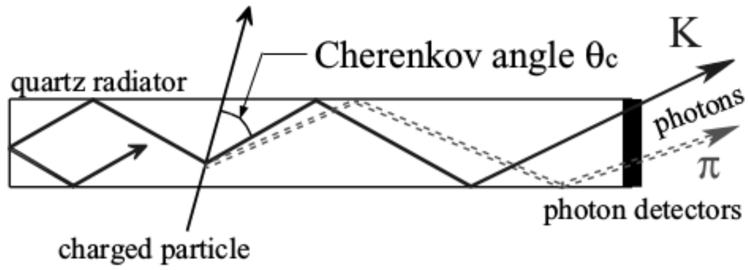


FIGURE 3.7: A particle interaction with the quartz of the TOP detector. Depending on the particle species the Cherenkov angles and path lengths in the detectors will be different.

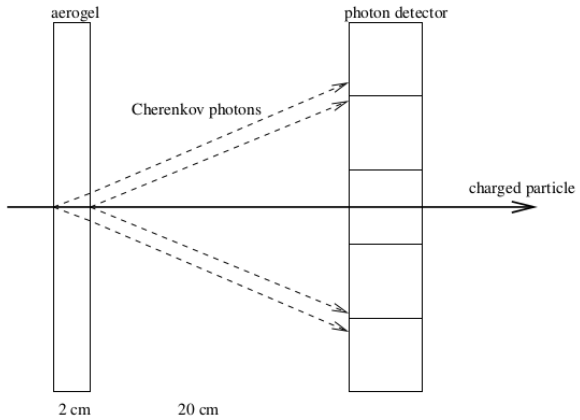


FIGURE 3.8: Schematic depiction of the ARICH detector. A charged particle will radiate Cherenkov light in the aerogel. The interference between radiation caused by different materials can be analysed to tag the responsible particle species.

ARICH In the forward direction a proximity focusing ring image Aerogel Cherenkov detector replaces the Belle ACC. It consists of a non-homogenous radiator of 20 mm width that causes charged particles to radiate Cherenkov photons. The interference rings of these are projected on a screen and the radius is used to determine the particle type. A schematic depiction can be found in Fig. 3.8. Its main purpose is to help separate kaons and pions.

Solenoid magnet Belle II uses the same magnet and solenoid as Belle with a magnetic field strength of 1.2 T.

KLM The KLM of Belle II in the barrel is the same design as used for Belle, with the exception that the innermost RPC layer was replaced with plastic scintillators as these are less sensitive to neutrons from beam background. The end-cap RPCs have been entirely replaced with plastic scintillators for the same reason.

3.4 Data samples

The data samples used for this thesis originate from the Belle and Belle II experiment. For algorithm development Monte Carlo simulated data (MC) was generated using the Belle II software framework (BASF2), simulating specific particle decays reconstructed with the Belle II detector. The physics analysis uses run dependent MC samples, mimicking the different experimental conditions during the lifetime of the Belle experiment, and physics data recorded during the runtime of the Belle experiment. The MC samples were provided by the Belle collaboration, generated with the Belle software framework (BASF). It was converted and analysed in BASF2 using the B2Bii [32] conversion tool.

3.4.1 Simulated data

Both software frameworks are used to simulate physics events in the environment of the respective detector. This allows for study and optimisation of an analysis for any known physics process and furthermore study the reconstruction efficiency. The simulation is performed over several steps. First, the particle four momenta are simulated, calculating phase space and angular distributions according to their quantum numbers. This is done using the EVTGEN [33] package. The next step is to simulate hadronisation into bound quark states using the PYTHIA [34] package where decays are not explicitly defined. Simulation of final state radiation is done using PHOTOS [35]. After that the interaction with the detectors is simulated using GEANT3 [36] (Belle) and GEANT4 [37] (Belle II).

3.4.2 Data taken by the Belle experiment

In eleven years the Belle experiment collected a total integrated luminosity of 1040 fb^{-1} at energies corresponding to the masses of the $\Upsilon(1S)$ through $\Upsilon(5S)$ resonances as well as about 60 MeV below the $\Upsilon(4S)$, as listed in Tab. 3.1 and depicted in Fig. 3.9. For more details see Ref. [38]. However, this thesis will only use the data taken at the

TABLE 3.1: Data collected by the Belle experiment between 1999 and 2010.

Type	$\mathcal{L}_{\text{integrated}}$ (fb^{-1})	Number of particles
$\Upsilon(1S)$	5.7	$102 \cdot 10^6$
$\Upsilon(2S)$	24.9	$158 \cdot 10^6$
$\Upsilon(3S)$	2.9	$11 \cdot 10^6$
$\Upsilon(4S)$ SVD1	140	$152 \cdot 10^6$ ($\bar{B}B$)
$\Upsilon(4S)$ SVD1 (off resonance)	15.6	N/A
$\Upsilon(4S)$ SVD2	571	$620 \cdot 10^6$ ($\bar{B}B$)
$\Upsilon(4S)$ SVD2 (off resonance)	73.8	N/A
$\Upsilon(5S)$	121	$7.1 \cdot 10^6$ ($\bar{B}_s B_s$)

$\Upsilon(4S)$ resonance.

3.5 Measurement models

The process from raw measurements to objects with physical parameters, such as momentum and energy are described in this section. Final states that are electromagnetically charged are treated differently to those that are not.

3.5.1 Charged final states

Charged particles are identified by the tracking detectors. Their momenta are measured based on their behaviour in magnetic fields and their energy is calculated from

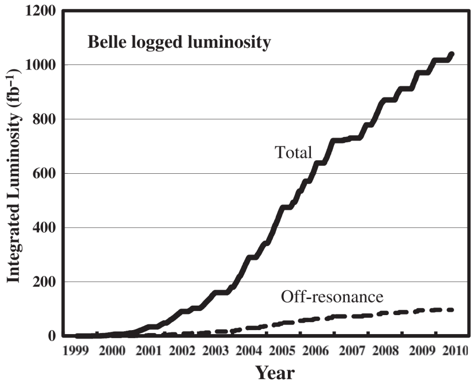


FIGURE 3.9: Integrated luminosity recorded by the Belle experiment.

their momenta based on a mass hypothesis, equivalent to a particle hypothesis. The trajectories of charged particles moving in a homogenous magnetic field can be described by helices. The helix only describes trajectories; to obtain a point on the helix an additional parameter, χ, describing the arclength is needed. A dedicated track finding and fitting algorithm combines the hits in the CDC and vertex detectors, measures the helix and extracts the five helix parameters and the arclength projected on the x-y plane, l. The measurement model becomes $m_{\text{track}} = \{d_0, \phi_0, \omega, z_0, \tan\lambda, l\}$. The exact definition on how to transform from Cartesian space differs slightly between Belle and Belle II. For simplicity, only the Belle II parametrisation will be discussed in detail here. See Fig. 3.10 for a depiction of the geometric meaning of the helix parameters and Tab. 3.2 for a description and some helpful relations. All parameters are always relative to the perigee point, the point of closest approach (POCA) to the coordinate systems origin. The coordinate systems origin is arbitrary, however changing it requires the helix to be transported [39]. The transformation from helix to Cartesian space point is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\phi_0 & -\sin\phi_0 \\ \sin\phi_0 & \cos\phi_0 \end{pmatrix} \begin{pmatrix} -\sin\chi/\omega \\ -(1-\cos\chi)/\omega - d_0 \end{pmatrix}, \tag{3.7}$$

and

$$z = l \cdot \tan\lambda - z_0, \tag{3.8}$$

where $\chi = -l \cdot \omega$. The arc length l is a priori unknown and the coordinate system is chosen such that $l = 0$. The momentum components are calculated by

$$\begin{pmatrix} p_x \\ p_y \\ p_z \\ E \end{pmatrix} = \begin{pmatrix} p_t \cos\phi_0 \\ p_t \sin\phi_0 \\ p_t \tan\lambda \\ \sqrt{m^2 - p^2} \end{pmatrix}, \tag{3.9}$$

with transverse momentum p_t. The transformation from Cartesian to helix parameters is given by

$$\begin{pmatrix} d_0 \\ \phi_0 \\ \omega \\ z_0 \\ \tan\lambda \end{pmatrix} = \begin{pmatrix} A(1+U)^{-1} \\ \operatorname{atan2}(p_y, p_x) - \operatorname{atan2}(\omega \cdot \Delta_{\parallel}, 1 + \omega \cdot \Delta_{\perp}) \\ a \cdot q/p_t \\ z + l \cdot \tan\lambda \\ p_z/p_t \end{pmatrix}, \tag{3.10}$$

TABLE 3.2: Definitions of the Belle II helix parametrisation and their dependencies.

d_0	The distance of closest approach to the z axis (POCA) signed with the z component of the angular momentum w.r.t. to the origin.
ϕ_0	Azimuthal angle of the momentum at the POCA.
ω	Scaled inverse of the track momentum.
z_0	Distance in the z-coordinate to the pivotal point. The pivotal point is the perigee that is the POCA.
$\tan\lambda$	The angle of the momentum at the POCA with respect to the $x-y$ plane.
l	Arclength projected onto the $x-y$ plane.
$\phi = \mathrm{atan2}(p_y, p_x)$	Angle of the $x-y$ plane to the helix.
$\Delta_\parallel = -x\cdot\cos\phi - y\cdot\sin\phi$ $\Delta_\perp = -y\cdot\cos\phi + x\cdot\sin\phi$ $A = 2\cdot\Delta_\perp + \omega\cdot(\Delta_\perp^2 + \Delta_\parallel^2)$	Quantities used to transport the coordinate system, such that $\{x, y, z\}$ points to the perigee.
$U = \sqrt{1+\omega\cdot A}$	
λ	Angle between the $x-y$ plane and the helix.
$l = \mathrm{atan2}(\omega\cdot\Delta_\parallel, 1+\omega\cdot\Delta_\perp)$	Arc length from the perigee to a point.
$p_x,\ p_y,\ p_z$	Momenta along the $x,\ y,\ z$ axes.
q	Charge of the particle.
B_z	Magnetic field strength in z-direction.
$a = B_z/c$	Magnetic field strength in the z direction divided by the speed of light.
$p_t = \lvert a/\omega\rvert$	Transverse momentum.
$p_t = \sqrt{p_x^2 + p_y^2}$	
$p_{t0} = \sqrt{p_{x0}^2 + p_{y0}^2}$	Pseudo transverse momentum.
$p_{x0} = p_x - a\cdot q\cdot y$	Pseudo momentum along the x direction.
$p_{y0} = p_y + a\cdot q\cdot x$	Pseudo momentum along the y direction.
$r^2 = x^2 + y^2$	Radius squared.
$\beta = 1 + \frac{p_{t0}}{p_t}$	Quantity used for reading convenience.

where the transformations of the parameters can be found in Tab. 3.2.

Muons need special treatment. They are minimum ionising particles, therefore can traverse long paths through the detector, which makes it easy to separate them from other particles at high momenta. To do this, tracks are extrapolated into the ECL and KLM and a muon likelihood is assigned based on how well their trajectories align with a KLM cluster and the amount of energy deposited in the ECL. To achieve better separation between kaons and pions an additional likelihood is calculated based on the ARICH and TOP, and ACC and TOF measurements, in Belle and Belle II respectively.

3.5.2 Neutral final states

Electromagnetically interacting neutral final states, i.e. γ, are identified using the ECL. A clustering algorithm finds energy depositions in the ECL and groups hit crystals to form clusters, as depicted in Fig. 3.11. The cluster consists of a fixed size of 25 (Belle) and 21 (Belle II) crystals around the averaged crystal position weighted by the energy depositions. This allows for the calculation of variables like the energy

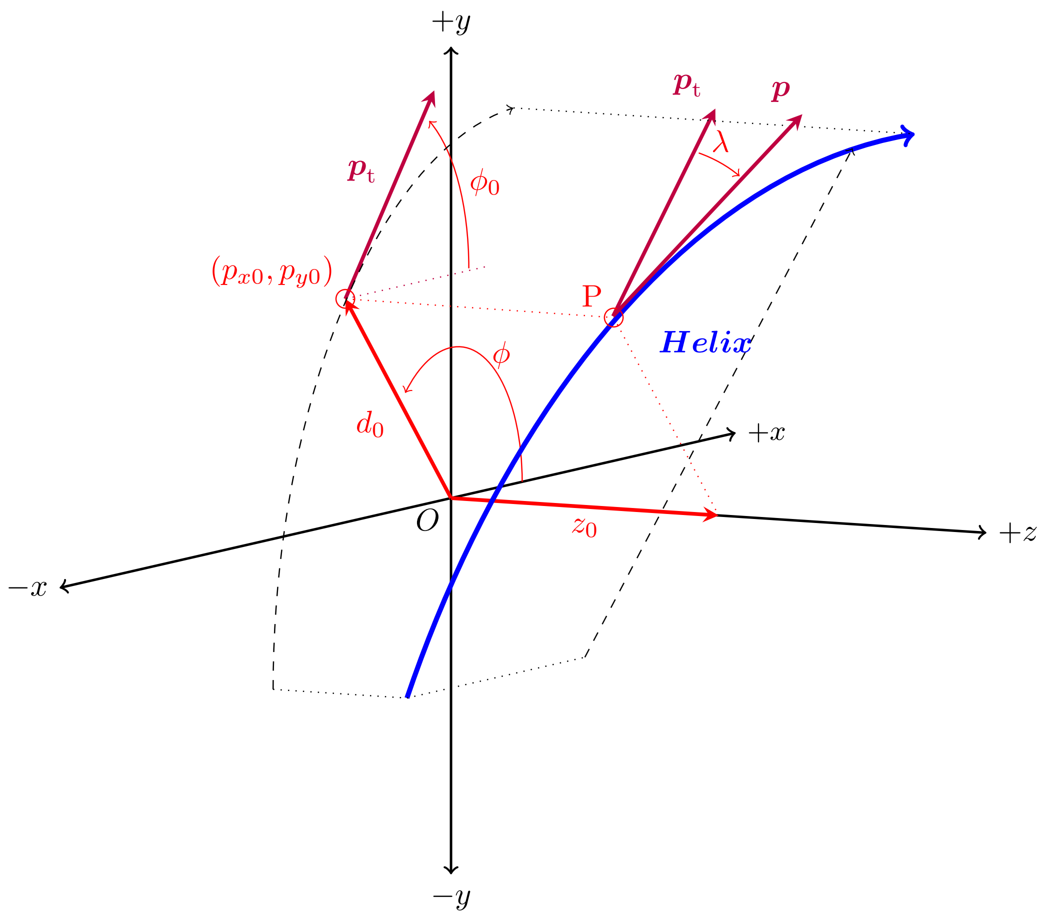


FIGURE 3.10: Parameters of the track helix, adapted from Ref. [40].

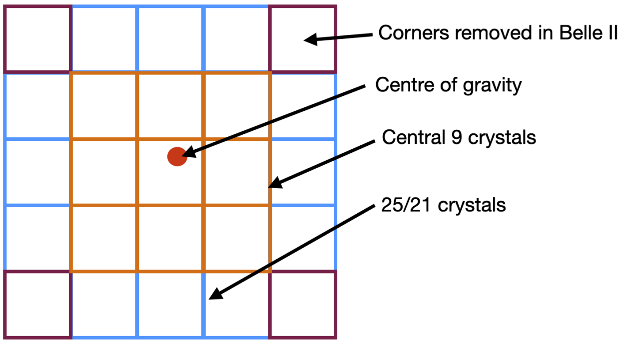


FIGURE 3.11: Schematic view of a cluster in the ECL.

deposition contained in the central crystal divided by the central nine crystals and the central nine crystals divided by the energy deposition in the total 25 (21) crystals. This information is used to distinguish between hadronic and electromagnetic showers. Further information about the underlying interaction can be obtained from the time evolution of light collection in the clusters, measured only in Belle II. Other hadronic neutral particles like neutrons or K_L^0 are measured using the KLM. However, the granularity of the KLM is large and usually only a single cell is hit by a neutral particle. The measurement model for neutral particles is the four-vector pointing from the origin to the centre of gravity of the cluster, $m = \{p_x, p_y, p_z, E\}$. More details can be found in Ref. [13].

Chapter 4

Statistical modelling methods

Experiments in particle physics never have the exact same outcome. This is due to the statistical nature of quantum mechanics and noise from various sources when performing an experiment. In order to interpret the outcome of an experiment and infer some quantities such as parameters of the SM, one has to perform a large number of experiments and use statistical modelling methods. In the context of this thesis these experiments are e^+e^--collision events recorded by Belle and Belle II.

Physics parameters are often measured in counting experiments, such as i.e. counting the number, $N(\bar{B}^0 \to D^{*+}h^-)$, of B-mesons decaying as $\bar{B}^0 \to D^{*+}h^-$ defines the branching ratio $\mathcal{B}(\bar{B}^0 \to D^{*+}h^-) = N(\bar{B}^0 \to D^{*+}h^-)/N(B^0)$. However, this requires isolating signal from background with a certain degree of belief. This is done using probability density function fitted to distributions to obtain parameters such was events numbers. In this chapter the basic principles of probability, probability density functions and parameter estimations will be discussed, following Ref. [41].

4.1 Probability

Probability is the chance that an event takes place, for example an experiment having a certain outcome. The probability, $\mathcal{P}$, of an event taking place is denoted as $\mathcal{P}(E)$. It has to satisfy a set of axioms, that is, it can take place, it is a real number and that an event takes place

$$\mathcal{P}(E) \in \mathbb{R}, \mathcal{P}(E) \geq 0,\ \forall E \in \Omega \tag{4.1}$$

and

$$\mathcal{P}(\Omega) = 1, \tag{4.2}$$

where Ω are all possible events. Furthermore, the joint probability of two independent events $\mathcal{A}$ and $\mathcal{B}$ taking place is given by the sum of their probabilities

$$\mathcal{P}(\mathcal{A} \cup \mathcal{B}) = \mathcal{P}(\mathcal{A}) + \mathcal{P}(\mathcal{B}). \tag{4.3}$$

The underlying probabilities $\mathcal{P}(E)$ can be defined using symmetries and constraints of the model. For k realisations of an event, E, with n possible outcomes one can define

$$\mathcal{P}(E) = \frac{k}{n}. \tag{4.4}$$

So that for example in a fair dice roll the probability of rolling a number less than three results in $\mathcal{P}(E) = 1/3$, with $k = 2$ and $n = 6$. If these are unknown one can measure them and infer the probability from the frequency of occurrence, such that

$$\mathcal{P}(E) = \lim_{n\to\infty} \frac{k}{n}. \tag{4.5}$$

This procedure is known as the Frequentist approach, with the obvious drawback that performing an infinite number of measurements is impossible and perhaps, the biggest drawback is that it does not allow to predict events unless all underlying symmetries are known. This can be overcome using conditional probabilities and Bayes Theorem

$$\mathcal{P}(\mathcal{A}|\mathcal{B}) = \mathcal{P}(\mathcal{B}|\mathcal{A})\frac{\mathcal{P}(\mathcal{A})}{\mathcal{P}(\mathcal{B})}, \tag{4.6}$$

where $\mathcal{P}(\mathcal{A}|\mathcal{B})$ describes the probability of $\mathcal{A}$ occurring under the condition that $\mathcal{B}$ has occurred. This allows for extrapolation from measurable quantities to unmeasurable quantities in the Frequentist sense. It is for example not possible to infer a certain signal yield from a physics data set as the absolute truth is unknown and what one sees in the data could just be background. But if one obtains efficiencies and misidentification rates from simulated data where truth is known, one can use these a a prior to interpret a physics data sample and get an estimator of the true number of signal candidates.

4.1.1 Probability density functions

The distributions of random variables, x, are called probability density functions (PDF). These random variables may be discrete or continuous. In the following it is assumed that the variables are continuous; if they are discrete, one has to replace the integrals with sums, all relations remain true. Let x be a random variable and $f(x)$ its PDF, we define it to be positive definite and normalised, so that

$$f(x) \geq 0 \,\text{and} \int_{-\infty}^{\infty} f(x)\,dx = 1. \tag{4.7}$$

The probability to find x in an interval $a < x < b$ can then be found by integration

$$\mathcal{P}(x|a,b) = \int_a^b f(x)\,dx. \tag{4.8}$$

The first central moment of a PDF corresponds to the expectation value, μ,

$$\mu = \langle x \rangle = E[x] = \int_{-\infty}^{\infty} x \cdot f(x)\,dx. \tag{4.9}$$

The second moment corresponds to the variance σ

$$var(x) = \sigma^2 = E[(x-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)\,dx. \tag{4.10}$$

The same conditions apply for n dimensional PDFs $f(x_1, x_2, ..., x_n)$ where

$$\mu_i = E[x_i] = \int_{-\infty}^{\infty} x_i \cdot f(x_1, x_2, ..., x_n)\,dx^n, \tag{4.11}$$

and

$$\sigma_i^2 = E[(x_i - \mu_i)^2] = \int_{-\infty}^{\infty} (x_i - \mu_i)^2 \cdot f(x_1, x_2, ..., x_n)\,dx^n. \tag{4.12}$$

Note that now correlation terms of the form

$$\sigma_{x_i x_j} = E[(x_i - \mu_i)(x_j - \mu_j)] = \int_{-\infty}^{\infty} (x_i - \mu_i)(x_j - \mu_j) \cdot f(x_1, x_2, ..., x_n)\,dx^n. \tag{4.13}$$

arise. These are referred to as covariance terms and form a matrix V, setting $V_{ij} = \sigma_{x_i x_j}$, see Sec.4.1.4. The definitions of the PDFs used in this thesis can be found in Appendix C.

4.1.2 Cumulative distribution

The cumulative distribution (CDF) $F_x(y)$ for a random variable x evaluated at a point y describes the probability that x will take the a value less than or equal to y. It is defined via the integral of its PDF, f_x, such that

$$F_x(y) = \int_{-\infty}^{y} f_x(t)dt. \tag{4.14}$$

It is monotonically rising and hence always has an inverse F^{-1}.

4.1.3 Quantiles

A quantile α is the value of a random variable x_α such that $P(x \leq x_\alpha) = \alpha$. It can be calculated from the inverse CDF as

$$F^{-1}(\alpha) = x_\alpha. \tag{4.15}$$

The special value $F^{-1}(0.5)$ is called the median of the distribution. For a Gaussian PDF the following approximations can be made $F^{-1}(0.16) \approx \mu - \sigma$, $F^{-1}(0.84) \approx \mu + \sigma$ and $|F^{-1}(0.84) - F^{-1}(0.16)| \approx 2\sigma$. These quantities can be useful when comparing non-Gaussian distributions, such as the sum of multiple Gaussian distributions with different widths and shared mean to a single Gaussian distribution.

4.1.4 Covariance and correlation

Covariance is a linear measure for relatedness of the variance of two random variables x_1, x_2. Covariance is defined as

$$cov(x_1, x_2) = E[(x_1 - E[x_1])(x_2 - E[x_2])] = E[x_1 \cdot x_2] - E[x_1]E[x_2], \in [0, \infty]. \tag{4.16}$$

A more convenient measure is the normalised (Pearson) correlation coefficients

$$corr(x_1, x_2) = \frac{cov(x_1, x_2)}{\sqrt{var(x_1)var(x_2)}}, \in [-1, 1], \tag{4.17}$$

which has an additional upper bound. If $\boldsymbol{x}$ is a d-dimensional random vector, one can define a symmetric, positive definite matrix of covariances between the parameters as

$$cov(\boldsymbol{x}) = V(\boldsymbol{x}) = E[(\boldsymbol{x} - E[\boldsymbol{x}])(\boldsymbol{x} - E[\boldsymbol{x}])^T]. \tag{4.18}$$

4.2 Parameter estimation

A common problem is to estimate parameters of a model, such as the true mass of a particle, given a distribution of measurement results that are assumed to follow a PDF. The estimator $\hat{a}$ for the value of a parameter, a_0, serves as a substitute to connect the real world and the model. An estimator has to fulfil the following axioms to qualify as a parameter estimator;

- Consistency: $\lim_{N\to\infty} \hat{a} = a_0$, where N is the number of measurements.

- Unbiasedness: $E[\hat{a}] = a_0$.
- Efficiency: minimum variance.
- Robustness: Robust against wrong or incorrect data.

4.2.1 Maximum likelihood estimator

A common estimator for a parameter, a, measured in a set of N independent observations $X = x_1, x_2, ..., x_n$ is the value of a for which the likelihood defined as the product of the individual PDFs,

$$L(\boldsymbol{x}|a) = \prod_{i=1}^{N} f(x_i, a), \tag{4.19}$$

has its maximum. Instead of maximising $L(\boldsymbol{x}|a)$ it is easier to equivalently minimise $g(\boldsymbol{x}|a) = \ln L(\boldsymbol{x}|a)$. That means we seek a value of $\hat{a}$ for which

$$\frac{\partial}{\partial a} \sum_{i=1}^{N} \ln f(x_i, a) = 0. \tag{4.20}$$

This equation is called a likelihood equation. It can be shown that an estimator built from this equation fulfils the necessary axioms. This method is used in this thesis to determine the most probable values of the normalisations of model PDFs for data samples.

4.2.2 Chi squared estimator

In the special case of Gaussian PDFs we set $a = \mu$ and define the χ^2 function

$$\chi^2 = \sum_{i-1}^{N} \frac{(x_i - \mu)^2}{\sigma_i}. \tag{4.21}$$

After dropping constants originating from the PDFs, the likelihood equation becomes

$$\frac{\partial}{\partial a} \chi^2 = 0. \tag{4.22}$$

This function will serve as the starting point to derive a Kalman Filter in section 7.2.

Chapter 5

Measurement of $\bar{B}^0 \to D^{*+}h^-$

Hadronic B-decays such as $\bar{B}^0 \to D^{*+}h^-$, where $h = \pi$ or $h = K$, are interesting for a variety of reasons. They have large branching fractions and therefore large data samples containing them are available. The Belle experiment is designed for clean reconstruction of high purity samples, due to the low track multiplicity in the events and no overlap between neighbouring events, a challenge at the LHC. This allows for high precision tests of the SM, particularly effects of the strong force. In this chapter the branching ratios of $\bar{B}^0 \to D^{*+}\pi^-$ and $\bar{B}^0 \to D^{*+}K^-$ will be measured with unprecendented precision. The decay $\bar{B}^0 \to D^{*+}\pi^-$ involves the CKM matrix element V_{ud}, while $\bar{B}^0 \to D^{*+}K^-$ involves V_{us}. Since $V_{ud} \approx 5V_{us}$ the decay $\bar{B}^0 \to D^{*+}K^-$ is referred to as Cabibbo suppressed and much lower values for the branching ratio are expected. The measurement strategy is to use only one D^{*+}-decay channel, $D^{*+} \to D^0\pi$, and two D^0-decay channels, the clean but smaller branching ratio $D^0 \to K^-\pi^+$ channel and the less clean but larger branching ratio $D^0 \to K^-2\pi^+\pi^-$ channel. The alternative channel $D^0 \to K^-\pi^+\pi^0$ channel was not considered due to large systematic uncertainties associated with the reconstruction of π^0. The measurement is optimised on MC and then performed on data with the same selection criteria. Differences between MC and data will be accounted for with corrections to the reconstruction efficiencies. These corrections are partially provided by the experiment performance groups. However, these do not exist for the low momentum phase-space of kaons and pions. Therefore, for pions these were measured here using a $K_S \to \pi^+\pi^-$-sample and for kaons the affected candidates were removed from the analysis.

5.1 Previous measurements

In this section we will give an overview of the most recent measurements and their uncertainties. There are many measurements of the branching ratio of the abundant Cabbibo favoured decay mode $\bar{B}^0 \to D^{*+}\pi^-$, we list the most recent ones in Table 5.1. There are many more older measurements by different experiments such as CLEO [42, 43, 44], CLEO-II [45], OPAL [46] and ARGUS [47, 48, 49]. Due to their large statistical uncertainties we do not discuss them here. The measurements of the Cabibbo suppressed decay mode $\bar{B}^0 \to D^{*+}K^-$ can be found in Table 5.2. Note that some of the the referenced papers do not contain the actual numbers of the branching ratios, these were taken from the PDG [50], where these papers are cited. The published measurements of the ratio $\mathcal{R}_{K/\pi} = \mathcal{B}(\bar{B}^0 \to D^{*+}K^-)/\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-)$ are tabulated in Table 5.3. It has to be mentioned that apart from Belle, the ratio is measured only in indirect ways by taking similar ratios of different channels such as $\mathcal{B}(\bar{B}^0 \to D^*K\pi\pi)/\mathcal{B}(\bar{B}^0 \to D^*\pi\pi\pi)$ in the LHCb case.

Table 5.1: Recent measurements of $\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-)$. The first cited value corresponds to the measured central value, the second to the statistical uncertainty and the third to the systematic uncertainty.

$\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-)$	Experiment	Year	Data sample
$(2.99 \pm 0.23 \pm 0.24) \times 10^{-3}$	BaBar [51]	2006	231×10^6 $B\bar{B}$-pairs
$(2.79 \pm 0.08 \pm 0.17) \times 10^{-3}$	BaBar [52]	2007	65×10^6 $B\bar{B}$-pairs
$(2.81 \pm 0.11 \pm 0.21) \times 10^{-3}$	CLEO-II [53]	1998	3.1 fb^{-1}

Table 5.2: Recent measurements of $\mathcal{B}(\bar{B}^0 \to D^{*+}K^-)$. The first cited value corresponds to the measured central value, the second to the statistical uncertainty and the third to the systematic uncertainty. Note that these numbers are taken from PDG and are not directly listed in the attributed papers.

$\mathcal{B}(\bar{B}^0 \to D^{*+}K^-)$	Experiment	Year	Data sample
$(2.13 \pm 0.12 \pm 0.10) \times 10^{-4}$	BaBar [54]	2006	226×10^6 $B\bar{B}$-pairs
$(2.00 \pm 0.40 \pm 0.10) \times 10^{-4}$	Belle [55]	2001	10.4 fb^{-1}

Table 5.3: Recent measurements of $\mathcal{R}_{K/\pi}$. The first cited value corresponds to the measured central value, the second to the statistical uncertainty and the third to the systematic uncertainty.

$\mathcal{R}_{K/\pi}$	Experiment	Year	Data sample
$(7.76 \pm 0.34 \pm 0.29) \times 10^{-2}$	BaBar [54]	2005	226×10^6 $B\bar{B}$-pairs
$(7.40 \pm 1.50 \pm 0.60) \times 10^{-2}$	Belle [55]	2001	10.4 fb^{-1}
$(7.76 \pm 0.34 \pm 0.26) \times 10^{-2}$	LHCb [56]	2013	1 fb^{-1}

5.2 Datasets

The measurements in this analysis are optimised using simulated Monte Carlo (MC) data. The MC data set corresponds to five times the full Belle data set. The data can be divided into five sub sets with the size of the Belle data set. In the following, one stream will mean a dataset containing both SVD1 and SVD2 configuration, meaning for example stream 1 and stream 10. The dataset for both decay modes is reconstructed only once with the pion mass assigned to the hadron. The set is then split into $\bar{B}^0 \to D^{*+}\pi^-$, pion enriched, and $\bar{B}^0 \to D^{*+}K^-$, kaon enriched bin, using the likelihood ratio $\mathcal{L}_{K/\pi}$. All other preselection criteria will be the same used for both datasets. For the reconstruction of the D^0-mesons, the channels $D^0 \to K^-\pi^+$ and $D^0 \to K^-2\pi^+\pi^-$ were chosen. D^{*+}-mesons are reconstructed from $D^{*+} \to D^0\pi^+$ with minimal requirements on the pion, which in the following will be referred to as "slow-π", due to its typical low momentum.

5.2.1 Categories

The MC data are categorised by their different physics process components. These are "uds" for decays into continuum events, $e^+e^- \to q\bar{q}$, containing u-, d- or s-quark pairs and "charm" for continuum events containing $c\bar{c}$-quark pairs. Background from $e^+e^- \to \Upsilon(4S) \to B\bar{B}$ processes are referred to by the decay process of the B-meson. Due to the finite efficiency of the K/π-selection there are $\bar{B}^0 \to D^{*+}\pi^-$ backgrounds in $\bar{B}^0 \to D^{*+}K^-$ and similarly $\bar{B}^0 \to D^{*+}K^-$ in $\bar{B}^0 \to D^{*+}\pi^-$, they will be treated as background. Another very common type of background is the semi-leptonic decay $\bar{B}^0 \to D^{*+}l^-\bar{\nu}$, these have large branching fractions and the lepton is sometimes mistakenly reconstructed as a hadron. Furthermore, there are $\bar{B}^0 \to D^{*+}\rho^-$ decays with $\rho^- \to \pi^0 h^-$, where the π^0 is not reconstructed and background from $\bar{B}^+ \to D^0\pi^-$, where an additional pion is reconstructed to incorrectly build a $D^{*+} \to D^0\pi^-$ decay. Background from various B-decays that do not fit in the above mentioned categories are as follows: where the D^0 is wrongly reconstructed are referred to as "fake D^0", where a correctly reconstructed D^0 is combined with an incorrect pion are named "comb. π_s", for combinatorial slow-π; and where a correctly reconstructed D^0 originating from a true D^{*0} was falsely combined with a pion to form a D^{*+} are called "$D^{*0} \to D^0 X$". All those that do not fit any of these categories are named "charged" for $\Upsilon(4S) \to B^- B^+$ processes and "mixed" for $\Upsilon(4S) \to B^0\bar{B}^0$ processes.

5.2.2 Simulated decay channels

The MC is made of most known decay channels of most known B-decays. The branching ratios are taken from the PDG and decays are simulated using these values as central values smeared out by their widths, which are also taken from PDG. However, many branching fractions have not been measured with a high precision and some are only known to exist and their branching ratio values have a large uncertainty. For $D^0 \to K^-2\pi^+\pi^-$ in particular, there are multiple ways the particles in the reconstructed chain can decay that result in the same final state. A list of the channels simulated and reconstructed in this analysis and their respective branching fractions (as simulated) can be found in Table 5.4. The values are taken from the inputs of the last major simulation campaign undertaken by the Belle experiment.

5.3 Selection criteria

In order reduce the dataset to a manageable size and suppress the background to a controllable level, a set of pre selection criteria are applied. This will improve the statistical uncertainty of the final measurement. The pre-selection criteria are applied in a top down approach for viewing convenience. However, for the signal reconstruction the order of the selection does not matter, as long as the per-event candidate ranking and selection is done as the last step. A summary of the pre-selection can be found in Table 5.5, the details and variable definitions are discussed throughout this section.

5.3.1 Kaon and pion selection

To separate kaons from pions and vice versa, a likelihood ratio requirement defined as

$$\mathcal{L}_{K/\pi} = \frac{\mathcal{L}_K}{\mathcal{L}_\pi + \mathcal{L}_K}, \tag{5.1}$$

is set to $\mathcal{L}_{K/\pi} > 0.6$ for kaon, and $\mathcal{L}_{K/\pi} < 0.6$ for pion track candidates, excluding low momentum (slow) pions from D^{*+}-meson decays. The individual likelihoods are determined by combining measurements of sub-detectors hit by the particle. No lepton identification requirements are set, because the expected systematic uncertainty on these is expected to be large. As a result there will be a small contamination from semi-leptonic decays in the pion enriched sample. However, these are fully suppressed in the signal region, see Sec. 5.4.

TABLE 5.4: Simulated branching ratios for the signal processes. The first fraction refers to the specified decay, the second fraction is the fraction into the reconstructed final state, taking intermediate resonances into account.

D^0-meson decay	Fraction	Fraction into the final state
$D^0 \to K^-\pi^+$	3.82%	3.82%
$D^0 \to K^-2\pi^+\pi^-$ (non-resonant)	1.17%	1.17%
$D^0 \to K_S^0 K^-\pi^+$	3.2%	3.2%
$D^0 \to K^-a_1^+$	7.5%	2.39%
$D^0 \to K^{*0}\pi^-\pi^+$	1.06%	0.7%
$D^0 \to \bar{K}^{*0}\omega^0$	1.1%	0.01%
$D^0 \to \bar{K}^{*0}\rho^0$	1.1%	0.72%
$D^0 \to K_1^-\pi^+$	1.1%	0.01%
$D^0 \to K^-\pi^+\omega^0$	0.4%	0.006%
$D^0 \to K^-\pi^+\rho^0$	0.46%	0.45%
$D^0 \to K^-2\pi^+\pi^-$ (resonant)		6.74%
$D^0 \to K^-2\pi^+\pi^-$ (total)		7.91%

Other branching ratios	
$a_1^+ \to \rho^0\pi^+$	35%
$a_1^+ \to f^0(600)\pi^+$	20%
$\rho^0 \to \pi^+\pi^-$	98.9%
$\rho^0 \to \pi^+\pi^-\gamma$	0.99%
$\eta' \to \rho^0\gamma$	3%
$\omega^0 \to \pi^+\pi^-$	1.53%
$f^0(600) \to \pi^+\pi^-$	66%
$\bar{K}_1^- \to K^-\pi^+\pi^-$	1.4%
$\bar{K}^{*0} \to K^-\pi^+$	66.51%
$\bar{D}^{*+} \to D^0\pi^+$	67.70%
$\bar{B}^0 \to D^{*+}\pi^-$	0.276%
$\bar{B}^0 \to D^{*+}K^-$	0.021%

TABLE 5.5: Pre selection cuts.

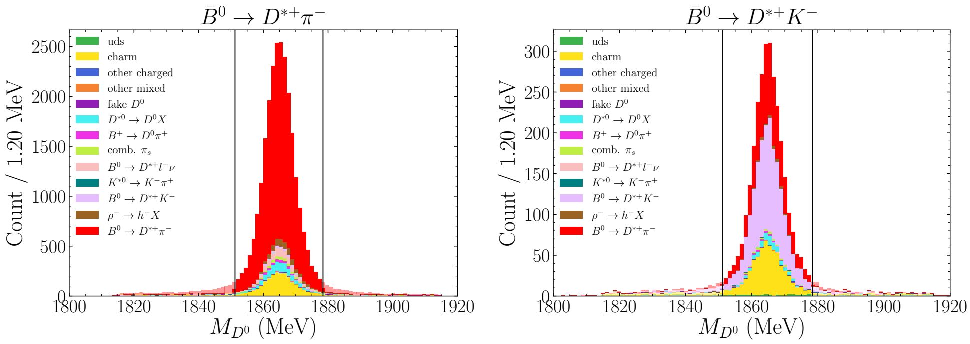
Particle candidate	Requirement
all tracks	$\lvert dz\rvert < 2$ cm, $\lvert dr\rvert < 4$ cm
fast π^-	$\mathcal{L}_{K/\pi} < 0.6$
slow π^-	no requirements
K^-	$\mathcal{L}_{K/\pi} > 0.6$
D^0	$M - 3\sigma_M < M < M + 3\sigma_M$
D^{*+}	$\Delta M - 3\sigma_{\Delta M} < \Delta M < \Delta M + 3\sigma_{\Delta M}$
B^0	$M_{\mathrm{bc}} > 5.27$ GeV
B^0	-150 MeV $< \Delta E <$ 125 MeV

5.3.2 Selection of D^0 candidates

Clean D^0 candidates with a high signal to background ratio, are selected in a window of 3σ around the fitted central value of the D^0-meson mass. This is done because the track momentum resolution is expected to be different in data and MC. By selecting the window as a function of the width it can be assured that the efficiency does not change. For MC the window is obtained by fitting the MC distribution with a double Gaussian function for the signal and a Chebishev Polynomial of rank one for the background. The width is defined as the weighted average of the individual Gaussian widths weighted by the fraction of their normalisation in the total signal, such that

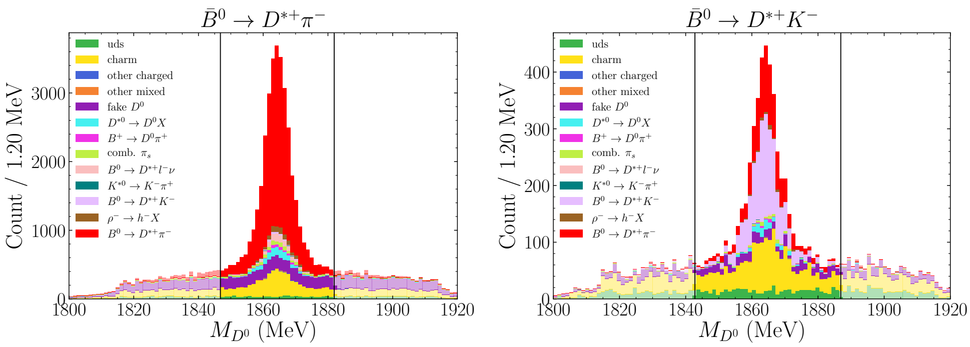
$$\sigma_{\text{combined}} = \frac{1}{N_1 + N_2}(N_1\sigma_1 + N_2\sigma_2), \tag{5.2}$$

where N_i are the individual yields and σ_i the widths of the respective Gaussian PDFs. The evaluation was done using the pion sample because of the larger dataset size. The resulting distributions are depicted in Fig. 5.1 for the $D^0 \to K^-\pi^+$ mode and Fig. 5.2 for the $D^0 \to K^-2\pi^+\pi^-$ mode. The selection is applied after selecting kaons and pions. The results of the fits are displayed in Fig. 5.3. For data the same fit fucntion is used. The fit results are depicted in Fig. 5.4.



(A) (B)

FIGURE 5.1: Selected $D^0 \to K^-\pi^+$ invariant mass window for a) $\bar{B} \to D^{*+}\pi^-$ and b) $\bar{B} \to D^{*+}K^-$.



(A) (B)

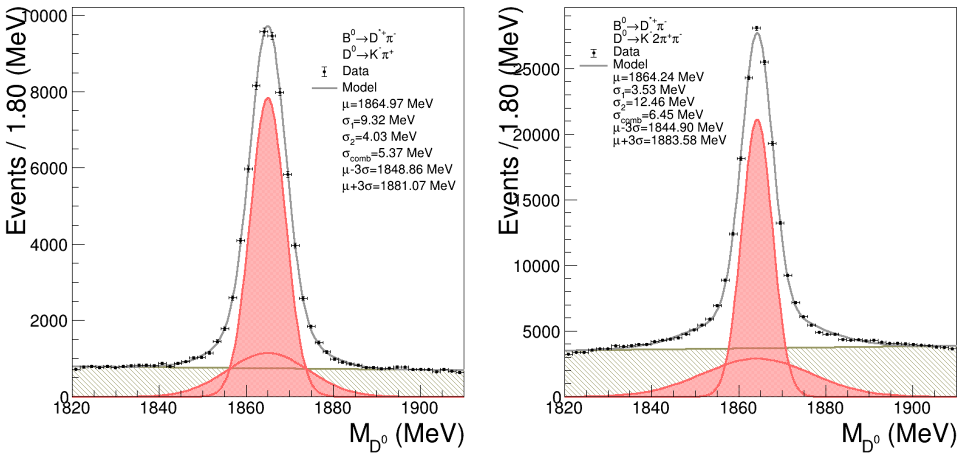
FIGURE 5.2: Selected $D^0 \to K^-2\pi^+\pi^-$ mass window for a) $\bar{B} \to D^{*+}\pi^-$ and b) $\bar{B} \to D^{*+}K^-$.

5.3.3 Selection of D^{*+} candidates

The candidate selection for D^{*+} candidates is done similarly as for D^0 candidates. The separation variable is ΔM, defined as

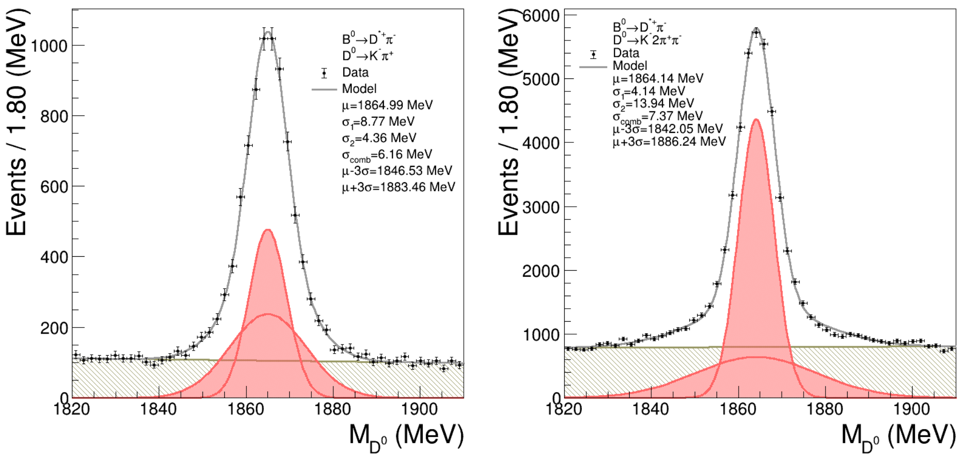
$$\Delta M = M_{D^{*+}} - M_{D^0}. \tag{5.3}$$

This variable is chosen over the invariant mass of the D^{*+}, $M_{D^{*+}}$, as it yields a narrower peak and the background can be constraint better close to the production threshold at the charged pion mass. The fit PDF is a weighted average of a multi



(A) (B)

FIGURE 5.3: Invariant mass window fits in MC for $D^0 \to K^-\pi^+$ a) and $D^0 \to K^-2\pi^+\pi^-$ b). Both fits were done on the $\bar{B} \to D^{*+}\pi^-$ bin due to larger dataset size.



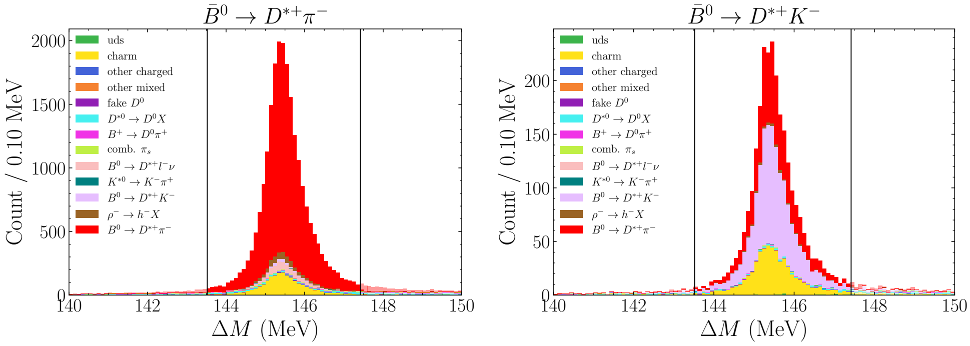
(A) (B)

FIGURE 5.4: Invariant mass window fits in data for $D^0 \to K^-\pi^+$ a) and $D^0 \to K^-2\pi^+\pi^-$ b). Both fits were done in the $\bar{B} \to D^{*+}\pi^-$ channel in the signal region.

Gaussian width for MC and a Gaussian with an added bifurcated Gaussian for real data. So that for MC the fit function reads

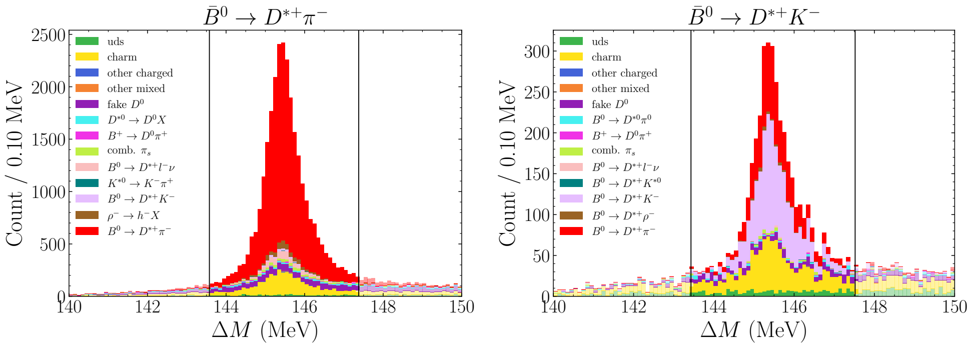
$$F = N_{\text{sig.}}\left[r_1 \mathrm{G}_1(\mu, \sigma_1) + r_2 \mathrm{G}_{\text{bifur.}}(\mu, \sigma_l, \sigma_r)\right] + \mathrm{D}(a, b, c), \tag{5.4}$$

where N are the respective signal and background yields, r_i the ratios of the PDFs $\mathrm{G}(\mu, \sigma)$ are Gaussian PDF, $\mathrm{G}_{\text{bifur.}}$ is a bifurcated Gaussian PDF and $\mathrm{D}(a, b, c)$ is an empirical background function tailored for ΔM, with the threshold fixed to the π^+ invariant mass, $\Delta M_0 = 139.57$ MeV, see Appendix C for the definition. The resulting distributions for MC are depicted in Fig. 5.5 for the $D^0 \to K^-\pi^+$ mode and Fig. 5.6 for the $D^0 \to K^-2\pi^+\pi^-$ mode. The selection is applied after selection of D^0-mesons. The results of the fits for MC are displayed in Fig. 5.7. For data these can be found in Fig. 5.8.



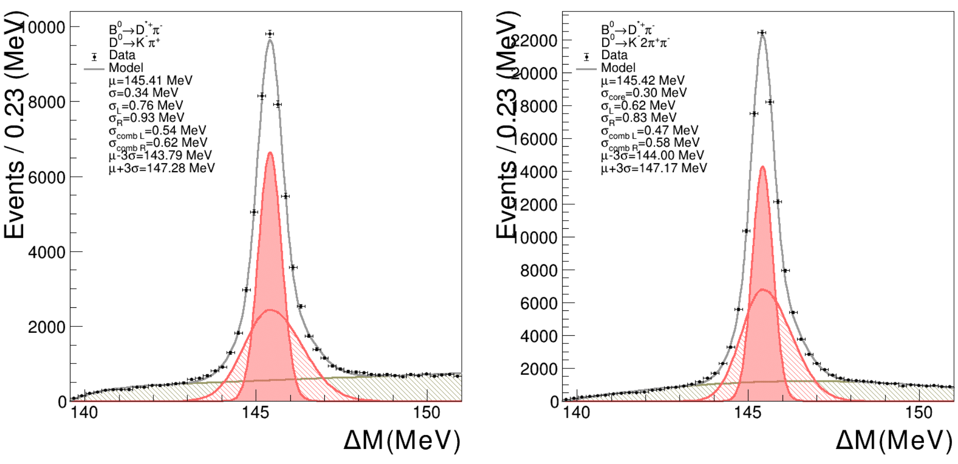
(A) (B)

FIGURE 5.5: Selected ΔM window for the $D^0 \to K^-\pi^+$ mode for a) $\bar{B}^0 \to D^{*+}\pi^-$and b) $\bar{B}^0 \to D^{*+}K^-$.



(A) (B)

FIGURE 5.6: Selected ΔM window for the $D^0 \to K^-2\pi^+\pi^-$ mode for a) $\bar{B}^0 \to D^{*+}\pi^-$and b) $\bar{B}^0 \to D^{*+}K^-$.



(A) (B)

FIGURE 5.7: Fits to ΔM in MC for $D^0 \to K^-\pi^+$ a) and $D^0 \to K^-2\pi^+\pi^-$ b). Both fits were done in $\bar{B}^0 \to D^{*+}\pi^-$due to larger available signal yields.

5.3.4 Charm meson selection summary

The resulting selections for charm mesons are listed in Table 5.6. The MC criteria will be used to optimise the analysis. The window obtained from data will be used

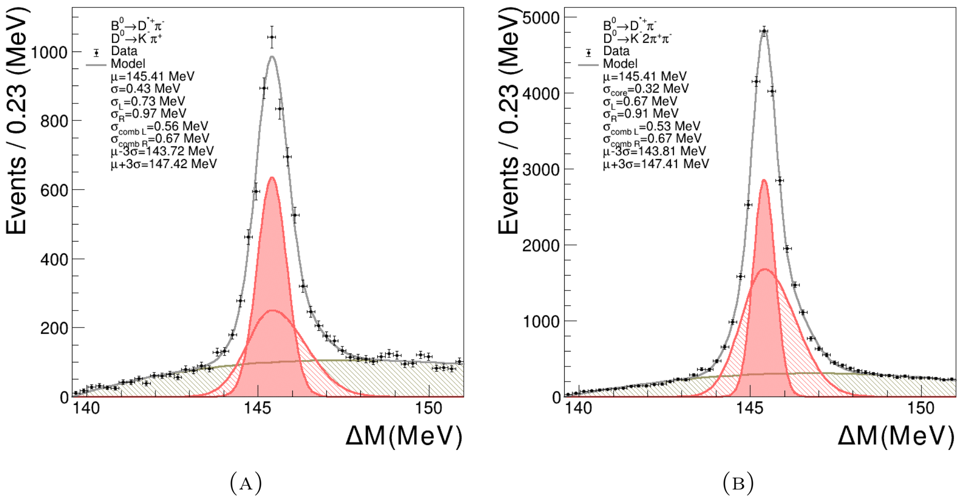


FIGURE 5.8: Fits to ΔM in data for $D^0 \to K^-\pi^+$ a) and $D^0 \to K^-2\pi^+\pi^-$ b). Both fits were done on data in $\bar{B}^0 \to D^{*+}\pi^-$in the signal box.

on data, assuming the same efficiency as for MC. Overall quiet similar results are obtained in MC and data; for the $D^0 \to K^-\pi^+$ channel it was found that the D^0 mass is a bit more narrow, which can be explained by a mis-measurement of the third Gaussian component of the PDF.

TABLE 5.6: Summary of the invariant mass selection criteria and their resolutions, defined by the combined width of the peaks.

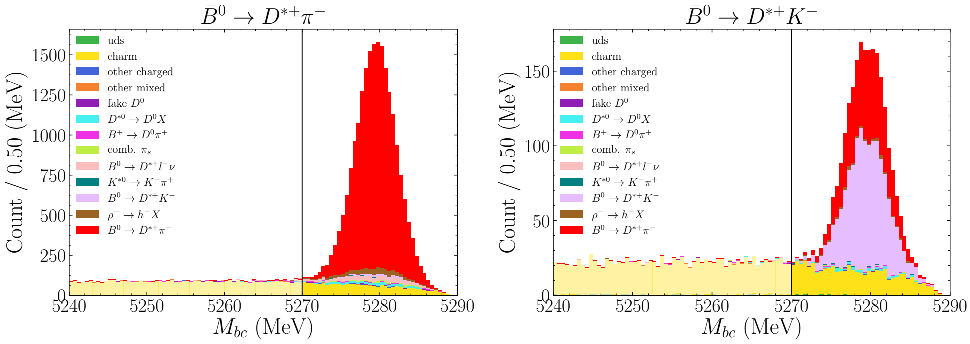
$D^0 \to K^-\pi^+$	Selection	Resolution
D^0 (MC)	$1848.46 < M < 1881.07$ MeV	5.37 MeV
D^0 (data)	$1846.53 < M < 1883.46$ MeV	6.16 MeV
ΔM (MC)	$143.79 < \Delta M < 147.28$ MeV	$\sigma_l = 0.54, \sigma_r = 0.62$ MeV
ΔM (data)	$143.72 < \Delta M < 147.42$ MeV	$\sigma_l = 0.56, \sigma_r = 0.67$ MeV
$D^0 \to K^-2\pi^+\pi^-$		
D^0 (MC)	$1844.90 < M < 1883.58$ MeV	6.45 MeV
D^0 (data)	$1842.05 < M < 1886.24$ MeV	7.37 MeV
ΔM (MC)	$144.00 < \Delta M < 147.17$ MeV	$\sigma_l = 0.47, \sigma_r = 0.58$ MeV
ΔM (data)	$143.81 < \Delta M < 147.41$ MeV	$\sigma_l = 0.53, \sigma_r = 0.67$ MeV

5.3.5 Selection of B^0 candidates

To select good B^0 candidates, we require a threshold on the beam-energy constrained reconstructed mass of a B-meson candidate, defined as

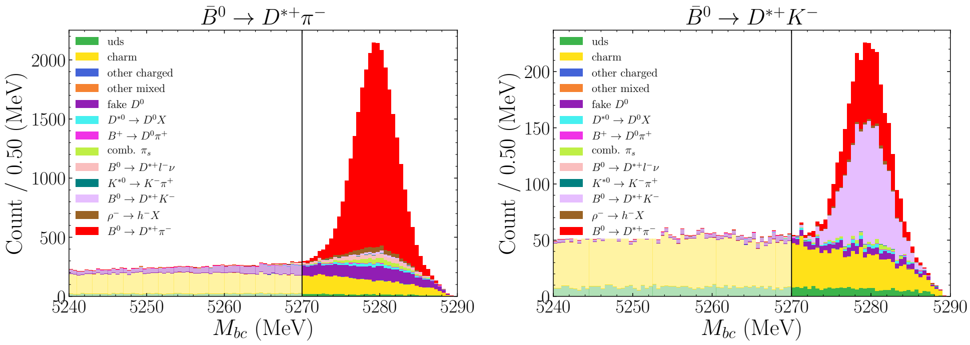
$$M_{\text{bc}} = \sqrt{E^2_{\text{beam}} - \vec{p}^2_{\text{CMS}}}. \tag{5.5}$$

This variable peaks at the B-meson mass and is sensitive to the number of tracks in the event. The backgrounds can be classified into two categories, continuum events and $B\bar{B}$-events which peak at the same position as the signal, see Fig. 5.9 for the $D^0 \to K^-\pi^+$ channel and Fig. 5.10 for the $D^0 \to K^-2\pi^+\pi^-$ channel. In the kaon channel the dominant background process is due to pions that are reconstructed as the bachelor kaon and continuum events. For the pion channel the main backgrounds are from fake D^*-mesons and continuum.



(A) (B)

FIGURE 5.9: Selection of B^0 candidates in a) $\bar{B}^0 \to D^{*+}\pi^-$and b) $\bar{B}^0 \to D^{*+}K^-$for $D^0 \to K^-\pi^+$.



(A) (B)

FIGURE 5.10: Selection of B^0 candidates in a) $\bar{B}^0 \to D^{*+}\pi^-$and b) $\bar{B}^0 \to D^{*+}K^-$for $D^0 \to K^-2\pi^+\pi^-$.

5.3.6 Candidate selection

The reconstructed events contain on average 12 charged tracks. This yields a large number of possible ways to combine them to form B-meson candidates. Even with the full set of selection criteria applied around 2% of the events have more than a single candidate. To select the best candidate candidates are ranked according to a χ^2 defined as

$$\chi^2 = \frac{(\Delta M^{\text{meas}} - \Delta M_{\text{PDG}})^2}{\sigma^2_{\Delta M}} + \frac{(m^{\text{meas}}_{D^0} - m^{\text{PDG}}_{D^0})^2}{\sigma^2_{m_{D^0}}}. \tag{5.6}$$

The candidate with the smallest rank is chosen after all other selections are applied. Alternative methods are to use the p-value of a vertex fit applied to the D^0-meson or to choose a random candidate. A comparison of the different methods was performed using $\bar{B}^0 \to D^{*+}\pi^-$ decays in the $D^0 \to K^-2\pi^+\pi^-$ channel, see Table 5.7 and Table 5.8 for the $D^0 \to K^-\pi^+$ channel. In both cases the method employing the χ^2 as defined in Eq. 5.6 is superior. If two B^0 candidates arise a random one is chosen, however this did not happen in a single case. Since the selection is independent of the mass assigned to the hadron it was not repeated for $\bar{B}^0 \to D^{*+}K^-$.

TABLE 5.7: Comparison of efficiency and background rejection of different methods to rank and select $\bar{B}^0 \to D^{*+}\pi^-$-meson candidates for the $D^0 \to K^-\pi^+$ channel. The background rejection is defined as $1 - \epsilon_{\text{bkg}}$, where ϵ_{bkg} is the background efficiency.

Method	Efficiency	Background rejection
χ^2	0.978	0.054
vertex fit p-value	0.973	0.053
random	0.967	0.050

TABLE 5.8: Comparison of efficiency and background rejection of different methods to rank and select $\bar{B}^0 \to D^{*+}\pi^-$-meson candidates for the $D^0 \to K^-2\pi^+\pi^-$ channel. The background rejection is defined as $1 - \epsilon_{\text{bkg}}$, where ϵ_{bkg} is the background efficiency.

Method	Efficiency	Background rejection
χ^2	0.967	0.103
vertex fit p-value	0.959	0.095
random	0.950	0.087

5.3.7 Cut flow

In order to understand the efficiencies of the pre selection criteria, the yields after a applying the selection were studied. The relative change of the signal and background yields for both D^0-meson channels after consecutively applying the selection criteria is listed in Table 5.9 and Table 5.10. The largest efficiency loss is due to the particle-ID selection and will be corrected by an efficiency correction. This criterion is absolutely necessary as otherwise the $\bar{B}^0 \to D^{*+}\pi^-$ contribution in the $\bar{B}^0 \to D^{*+}K^-$ channel would increase the statistical uncertainty on the extracted signal by a significant amount. All other selections have a reasonable signal rejection rate and will be kept as well.

TABLE 5.9: Relative change to previous selection after consecutively applying the preselection and event criteria, evaluated using MC equivalent of five times the size of the Belle dataset.

	$D^0 \to K^-\pi^+$			
	$\bar{B}^0 \to D^{*+}\pi^-$		$\bar{B}^0 \to D^{*+}K^-$	
Selection	Signal	Background	Signal	Background
ΔE	-	-	-	-
M_{bc}	0.2%	59.8%	0.1%	60.3%
Particle-ID	16.5%	48.3%	16.4%	54.1%
M_{D^0}	3.5%	12.2%	3.6%	11.5%
ΔM	4.5%	22.3%	3.3%	13.5%
Event selection	1.7%	3.1%	2.2%	8.0%

5.4 Signal yield fit variable

The measurement is performed using the energy conservation quantity of the event, ΔE, defined as

$$\Delta E = E_{cms} - E_{beam}/2. \tag{5.7}$$

For the $\bar{B}^0 \to D^{*+}K^-$ mode, see Fig. 5.12, there are essentially four components separable in this variable for the $D^0 \to K^-\pi^+$ channel. The $\bar{B}^0 \to D^{*+}K^-$ signal

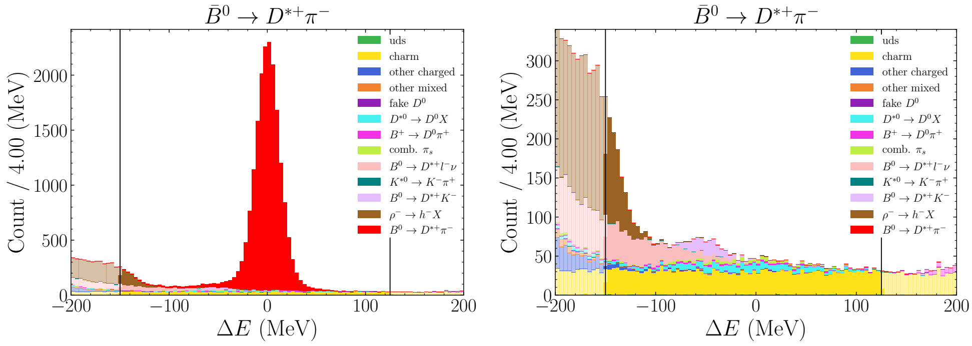
Table 5.10: Relative change to previous selection after consecutively applying the preselection and event criteria, evaluated using MC equivalent of five times the size of the Belle dataset.

	$D^0 \to K^-2\pi^+\pi^-$			
	$\bar{B}^0 \to D^{*+}\pi^-$		$\bar{B}^0 \to D^{*+}K^-$	
Selection	Signal	Background	Signal	Background
ΔE	-	-	-	-
M_{bc}	0.5%	64.1%	0.4%	67.6%
Particle-ID	15.0%	82.0%	15.0%	86.6%
M_{D^0}	6.8%	49.1%	5.8%	45.6%
ΔM	4.7%	39.6%	3.9%	32.8%
Event selection	12.8%	44.8%	12.2%	43.2%

peaks at $\Delta E \approx -48$ MeV, because the kaon was reconstructed with a pion hypothesis. Decays where a π^- is reconstructed instead of a K^- peak at $\Delta E = 0$ MeV. Background from fake D^{*+}, where a true D^0 is combined with a random particle that is not a pion and fake slow π where a true D^0 is combined with a random pion that fakes the slow pion, also peak in the signal region. For the $D^0 \to K^-2\pi^+\pi^-$ channel there is an additional background from fake D^0 that can be divided further into a peaking component, coming from real D^{*+} or D^{*0} and a single final state particle that was misidentified, and a widely distributed component where 4 tracks randomly pass the invariant mass requirement. The $D^0 \to K^-2\pi^+\pi^-$ channel is depicted in Fig. 5.14. The background in the $\bar{B}^0 \to D^{*+}\pi^-$ mode, depicted in Fig. 5.11 for the $D^0 \to K^-\pi^+$ channel and in Fig. 5.13 for the $D^0 \to K^-2\pi^+\pi^-$ channel, are different. There is a component peaking at $\Delta E \approx -50$ MeV where instead of a pion a kaon was reconstructed, a linearly decreasing background from continuum events, as well as wider distributed background from fake D^* and slow pions, as compared to $\bar{B}^0 \to D^{*+}K^-$, peaking at about $\Delta E = 0$ MeV. Contribution from in-flight decaying pions from the decay $D^{*+} \to D^0\pi^+$, where $\pi^+ \to \mu^+\bar{\nu}$, are considered as signal, because these decays typically take place far outside of the central tracking volume. The resolution for data, as in the final fit result and MC is listed in Table 5.11

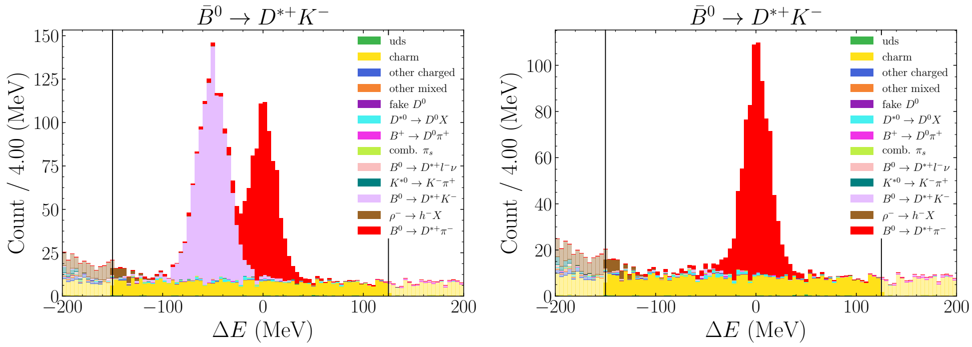
Table 5.11: Summary of the resolution of the signal peaks in ΔE, defined by the combined widths of the peaks.

$D^0 \to K^-\pi^+$	$\bar{B}^0 \to D^{*+}\pi^-$	$\bar{B}^0 \to D^{*+}K^-$
ΔE (MC)	14.48 MeV	17.25 MeV
ΔE (data)	17.07 MeV	19.1 MeV
$D^0 \to K^-2\pi^+\pi^-$		
ΔE (MC)	14.54 MeV	16.62 MeV
ΔE (data)	17.18 MeV	19.20 MeV



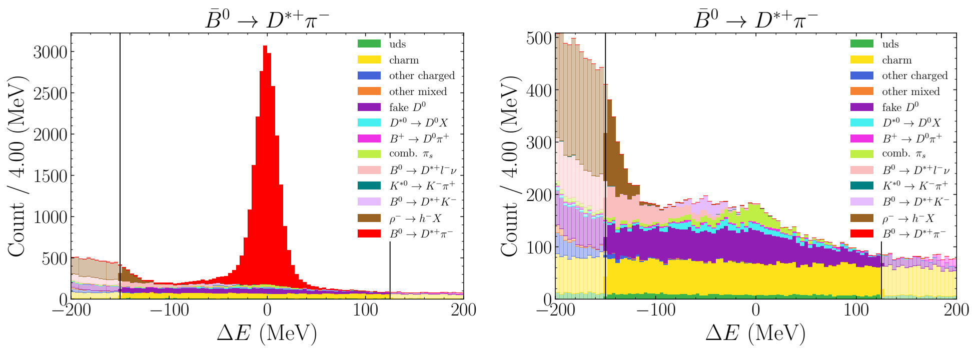
(A) (B)

FIGURE 5.11: Distribution of ΔE with (a) and without (b) signal to improve the visibility of the background components.



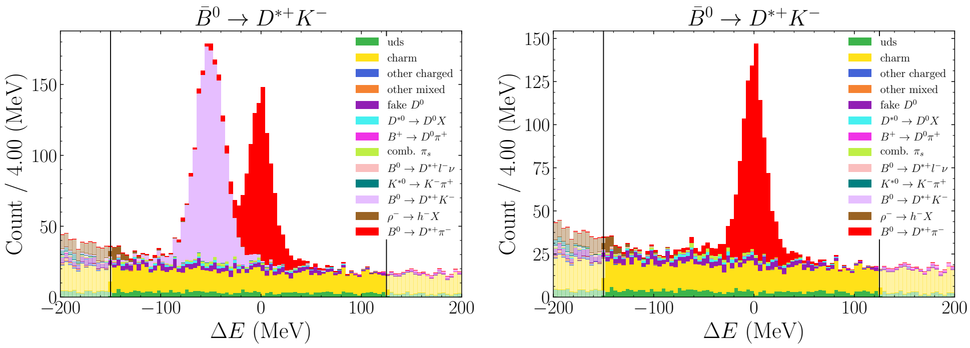
(A) (B)

FIGURE 5.12: Distribution of ΔE in the $D^0 \to K^-\pi^+$ channel with (a) and without (b) signal to improves the visibility of the background components.



(A) (B)

FIGURE 5.13: Distribution of ΔE in the $D^0 \to K^-2\pi^+\pi^-$ channel with (a) and without (b) signal to improve the visibility of the background components.



(A) (B)

FIGURE 5.14: Distribution of ΔE in the $D^0 \to K^-\pi^+$ channel with (a) and without (b) signal to improve the visibility of the background components.

5.5 Signal extraction and PDF models

The measurement strategy is to measure the signal yields in the variable ΔE, as discussed in Sec. 5.4. For the pion and kaon sample it is extracted with a simultaneous fit to the distributions in both samples. This allows the well defined signal peaks to constrain background shapes in the other respective decay mode. The signal shapes will be extracted from MC and will be fixed. A calibration factor will be used to allow only the widths of the distributions to float simultaneously. The same calibration factor will be used for signal in the pion component as for the pion background in the kaon sample and vice versa. The means are free. This section will discuss the PDFs used to parametrise the individual components in each of these variables. The shapes are extracted from MC using a data sample equivalent to fives times the Belle data set, each corresponding to the full Belle dataset.

5.5.1 Signal models

The signal is modelled as oe or two Gaussian peaks with a common mean, and a Crystal Ball shape, see Appendix. C for the PDF definitions. The full signal PDF for $\bar{B}^0 \to D^{*+}\pi^-$ reads

$$F_{\bar{B}^0 \to D^{*+}\pi^-} = N\left[r_1 \mathrm{G}_1(\mu_1, \sigma_1) + r_2 \mathrm{G}_2(\mu_1, \sigma_2) + (1 - r_1 - r_2)\mathrm{Cb}_\pi(\mu_1, \sigma_3, \alpha_\pi, n_\pi)\right], \tag{5.8}$$

and for $\bar{B}^0 \to D^{*+}K^-$

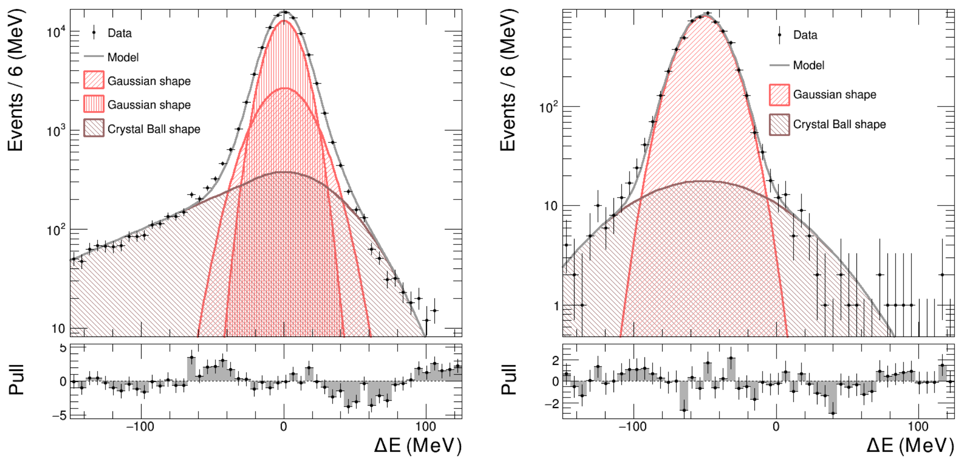
$$F_{\bar{B}^0 \to D^{*+}K^-} = N\left[r_3 \mathrm{G}_3(\mu_2, \sigma_3) + (1 - r_3)\mathrm{Cb}_K(\mu_2, \sigma_4, \alpha_K, n_K)\right], \tag{5.9}$$

where G are Gaussian PDFs, Cb is a Crystal Ball PDF, N is the total yield and r_i are the relative fractions. The floating parameters will be the yield and the mean. The Crystal Ball shape has to be introduced to model parts of the signal distribution that undergo radiative losses and hence have a tail towards lower values of the measurement variable, this effect is stronger in $\bar{B}^0 \to D^{*+}\pi^-$due to the smaller mass of the pion. Furthermore, it was observed that the steep slopes of the Gaussian do not fit the data very well. This is partially caused by either very forward or backward tracks that are not measured very well. Restricting the analysis to tracks that hit the calorimeter in the barrel region reduces this effect. However, this is not done in this analysis to keep the efficiency as high as possible. Introducing an asymmetric Gaussian shape does not improve the PDF in replicating. The $D^0 \to K^-2\pi^+\pi^-$ channel can decay directly into the final state or via various resonances, listed in Table 5.4.The resulting distributions are depicted in Fig. 5.15 for the $D^0 \to K^-\pi^+$ channel and Fig. 5.16 for the $D^0 \to K^-2\pi^+\pi^-$ channel. For the final simultaneous fit the individual widths are replaced $\sigma \to \sigma'$ to introduce a calibration factor, β_h, that is shared and floated for all widths of the signal PDFs simultaneously, such that

$$\sigma_i' = \beta_h \cdot \sigma_i. \tag{5.10}$$

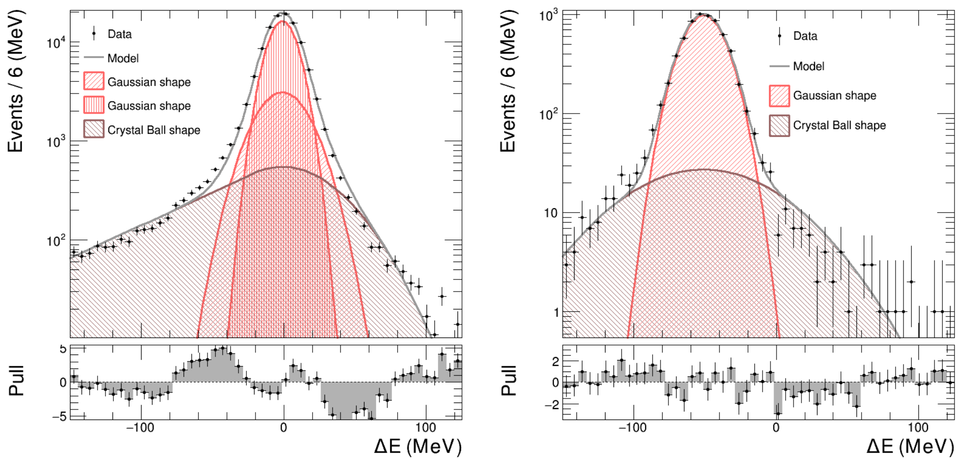
5.5.2 Background models in the $D^0 \to K^-\pi^+$ channel

The background will be modelled with individual PDFs for each physics process, see Sec. 5.4 for distributions of the individual components. They will be grouped by whether or not they peak in ΔM and the groups will be fitted in isolation to extract the shapes of the PDFs using a data sample equivalent of fives times the Belle data set. Note that for the pion and kaon mode the groups can be slightly different as the



(A) (B)

FIGURE 5.15: Signal models for $\bar{B}^0 \to D^{*+}\pi^-$(a) and $\bar{B}^0 \to D^{*+}K^-$(b) in the $D^0 \to K^-\pi^+$ channel. Two (a) and a single (b) Gaussian shapes are used to describe the main signal peak and a Crystal Ball shape describes the radiative tail to negative values of the measurement variable. This tail is more pronounced in $\bar{B}^0 \to D^{*+}\pi^-$due to the larger signal yield.



(A) (B)

FIGURE 5.16: Signal models for $\bar{B}^0 \to D^{*+}\pi^-$(a) and $\bar{B}^0 \to D^{*+}K^-$(b) in the $D^0 \to K^-2\pi^+\pi^-$ channel. Two (a) and a single (b) Gaussian shapes are used to describe the main signal peak and a Crystal Ball shape describes the radiative tail to negative values of the measurement variable. This tail is more pronounced in $\bar{B}^0 \to D^{*+}\pi^-$due to the larger signal yield.

relative abundance of these background contributions can be different and in some cases it is more convenient to join them with other components of similar shape. The results of the fits are listed in Table 5.13 for the kaon component and Table 5.12 for the pion component.

Background from continuum processes such as charm and uds are linear throughout the ΔE spectrum. They are parametrised with a Chebyshev Polynomial of rank one, and the fit result is depicted in Fig. 5.17.

For background originating from decays of $\bar{B}^0 \to D^{*+}\rho^-$ and processes where the

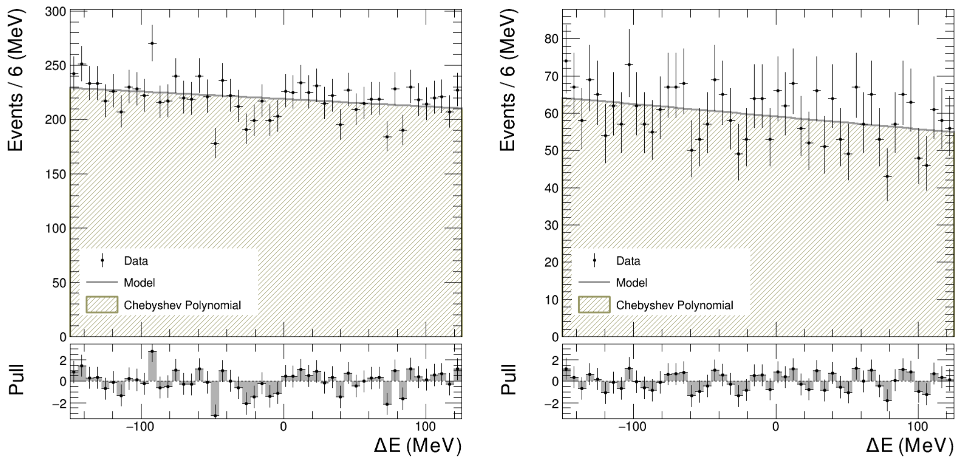
TABLE 5.12: Parameters of all component PDFs, Eq. 5.11, used for $\bar{\mathrm{B}}^0 \to \mathrm{D}^{*+}\pi^-$ in the $D^0 \to K^-\pi^+$ channel.

$\mathbf{\bar{B}^0 \to D^{*+}\pi^-}$	PDF: $F_{\bar{B}^0 \to D^{*+}\pi^-}$
$\sigma_1 = \beta_\pi \cdot 17.92 \pm 0.81$ MeV	$\sigma_2 = \beta_\pi \cdot 10.97 \pm 0.17$ MeV
$\sigma_3 = \beta_\pi \cdot 35.71 \pm 0.71$ MeV	$\mu_{\mathrm{Cb}_\pi} = \mu_{\mathrm{G}_1} = \mu_{\mathrm{G}_2}$
$\alpha_\pi = 0.52 \pm 0.02$	$n_\pi = 99.99 \pm 92.42$
$r_1 = 0.233 \pm 0.035$	$r_2 = 0.083 \pm 0.004$
yield free	
$\mathbf{q\bar{q}}$	PDF: $\mathrm{Ch}_{\mathrm{bkg1}}$
$a_0 = -0.042 \pm 0.016$	yield free
$\boldsymbol{\rho^-} \mathbf{\to h^-}$	PDF: $\mathrm{G}_{\mathrm{bkg2}}$
$\sigma = 15.79 \pm 1.72$ MeV	$\mu = -157.62 \pm 4.12$ MeV
$r_{\mathrm{bkg2}} = 0.310 \pm 0.011$	
$\mathbf{\bar{B}^0 \to D^{*+}l^-}\boldsymbol{\bar{\nu}}$ **, fake** $\boldsymbol{\pi_s}$**,**	PDF: $\mathrm{G}_{\mathrm{bkg3}}$
$\sigma = 131.22 \pm 10.64$ MeV	$\mu = -379.65 \pm 51.17$ MeV
$r_{\mathrm{bkg3}} = 0.482 \pm 0.021$	
$\mathbf{D^{*0}}$**,** $\mathbf{B^+}$**and fake** $\mathbf{D^0}$	PDF: $\mathrm{Cb}_{\mathrm{bkg4}}$
$\sigma = 73.54 \pm 2.37$ MeV	$\mu = -18.72 \pm 2.26$ MeV
$\alpha = 1.12 \pm 0.16$	$n = 0.23 \pm 0.35$
$1 - r_{\mathrm{bkg2}} - r_{\mathrm{bkg3}} = 0.208 \pm 0.036$	
$\mathbf{\bar{B}^0 \to D^{*+}K^-}$	PDF: $F_{\bar{B}^0 \to D^{*+}K^-}$
same PDF as in $\bar{B}^0 \to D^{*+}K^-$ in Table 5.13	yield free

TABLE 5.13: Parameters of all component PDFs, Eq. 5.12, of $\bar{\mathrm{B}}^0 \to \mathrm{D}^{*+}\mathrm{K}^-$ used in the $D^0 \to K^-\pi^+$ channel.

$\mathbf{\bar{B}^0 \to D^{*+}K^-}$	PDF: $F_{\bar{B}^0 \to D^{*+}K^-}$
$\sigma_3 = \beta_K \cdot 15.19 \pm 0.16$ MeV	$\mu \in [-51, -48]$ MeV
$\sigma_4 = \beta_K \cdot 49.71 \pm 2.38$ MeV	$\alpha_K = 3.97 \pm 3.25$
$n_K = 67.45 \pm 57.42$	$r_3 = 0.93 \pm 0.02$
yield free	
$\mathbf{q\bar{q}}$	PDF: $\mathrm{Ch}_{\mathrm{bkg6}}$
$a_0 = -0.077 \pm 0.031$	yield free
Other B, $\boldsymbol{\rho^-} \mathbf{\to h^-}$	PDF: $\mathrm{G}_{\mathrm{bkg6}}$
$\mu_{\mathrm{G}} = -264.33 \pm 110.93$ MeV	$\sigma_{\mathrm{G}} = 40.48 \pm 28.11$ MeV
$r_{\mathrm{bkg6}} = 0.325 \pm 0.048$	
$\mathbf{D^{*0}}$**,** $\mathbf{B^+}$**and fake** $\boldsymbol{\pi_s}$	PDF: $\mathrm{G}_{\mathrm{bkg7}}$ and $\mathrm{G}_{\mathrm{bkg8}}$
$\mu_1 = -109.21 \pm 80.87$ MeV	$\sigma_1 = 144.09 \pm 64.67$ MeV
$\mu_2 = -28.06 \pm 5.52$ MeV	$\sigma_2 = 23.54 \pm 6.85$ MeV
$f_1 = 0.774 \pm 0.111$	$1 - r_6 = 0.674 \pm 0.296$
$\mathbf{\bar{B}^0 \to D^{*+}\pi^-}$	PDF: $F_{\bar{B}^0 \to D^{*+}\pi^-}$
same PDF as in $\bar{B}^0 \to D^{*+}\pi^-$ in Table 5.12	yield free

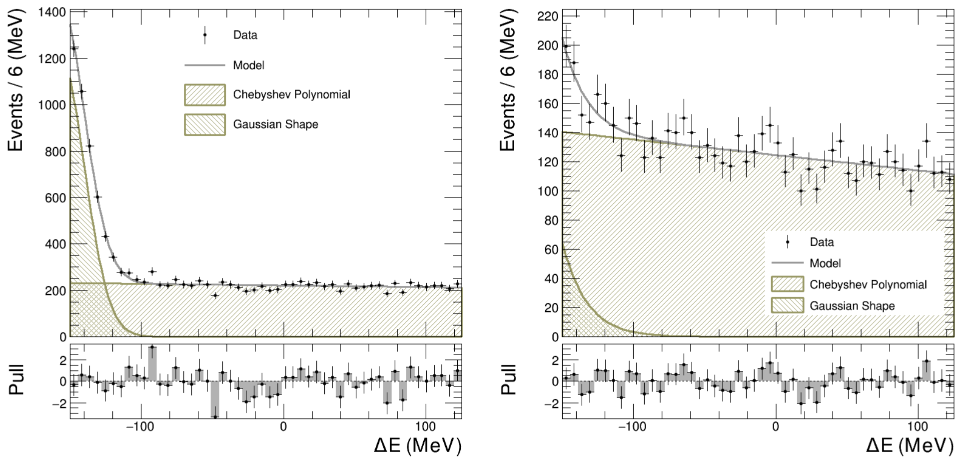
hadron is faked by a random pion that peak towards negative values of the signal variable, a Gaussian shape was chosen. In $\bar{B}^0 \to D^{*+}K^-$background from $\bar{B}^0 \to D^{*+}l^-\nu$ are also added to this PDF, as they have a similar structure and are not abundant enough to justify parametrising them with their own PDF, as is done in $\bar{B}^0 \to D^{*+}\pi^-$. The extracted PDF is shown in Fig. 5.18. The shape was extracted in together with the continuum background using the PDF for these with their parameters fixed to the values found. This was necessary to fit the components without having to restrict the



(A) (B)

FIGURE 5.17: Background model for the continuum background for $\bar{B}^0 \to D^{*+}\pi^-$(a) and $\bar{B}^0 \to D^{*+}K^-$(b) in the $D^0 \to K^-\pi^+$ channel, evaluated using a data sample equivalent of fives times the Belle data set.

fit range, as single outliers are spread out across the measurement variable.

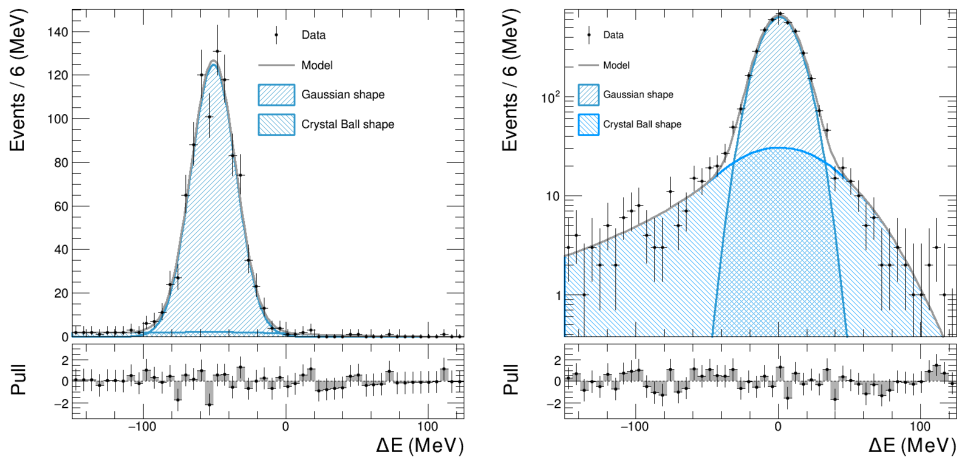


(A) (B)

FIGURE 5.18: Background model for $\bar{B}^0 \to D^{*+}\rho^-$ and $\bar{B}^0 \to D^{*+}l^-\nu$ background for B-meson events in $\bar{B}^0 \to D^{*+}\pi^-$(a), and $\bar{B}^0 \to D^{*+}K^-$(b) in the $D^0 \to K^-\pi^+$ channel.

An important peaking background arises when misidentifying the $\bar{B}^0 \to D^{*+}\pi^-$as a $\bar{B}^0 \to D^{*+}K^-$and vice versa, see Fig. 5.19. They are fitted and shown here to allow fitting all backgrounds in a non-simultaneous fit in order to verify the joined background PDFs. In the simultaneous fit, these will be replaced with the signal PDF of the respective other mode. A calibration factor will be added to the width to adjust for potential data and MC differences.

Background from $B^+ \to D^0h^+$, D^0 from D^{*0}-mesons and wrongly reconstructed D^0 are parametrised with a Crystal Ball shape in $\bar{B}^0 \to D^{*+}\pi^-$, shown in Fig. 5.20a. For $\bar{B}^0 \to D^{*+}K^-$they are all combined and fit with a double Gaussian with the exception of the leptonic background, as depicted in Fig. 5.20b. The central peak is

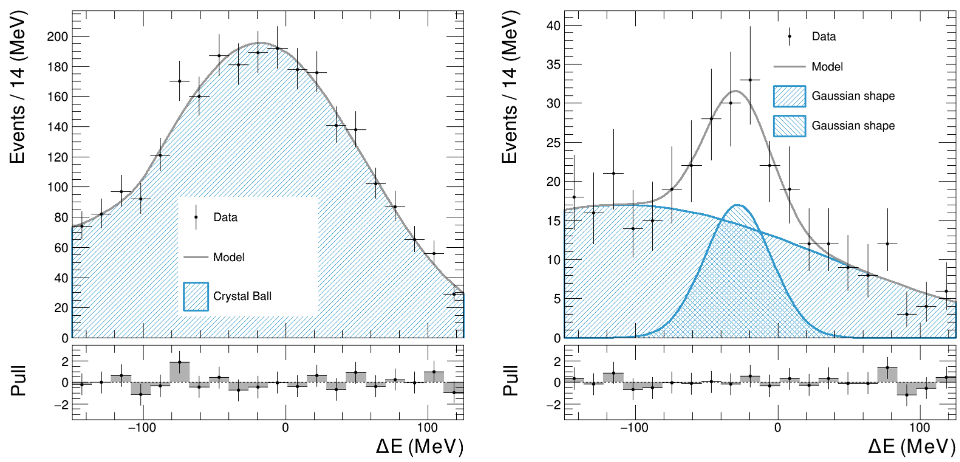


(A) (B)

FIGURE 5.19: Model for the $\bar{B} \to D^{*+}K^-$ contribution in $\bar{B}^0 \to D^{*+}\pi^-$(a), and the $\bar{B} \to D^{*+}\pi^-$ contribution in $\bar{B}^0 \to D^{*+}K^-$(b). Both of these will be replaced by the respective signal model of the other mode in the simultaneous fit.

mainly caused by D^{*0}-background, the other background sources are not abundant enough to extract meaningful shapes.

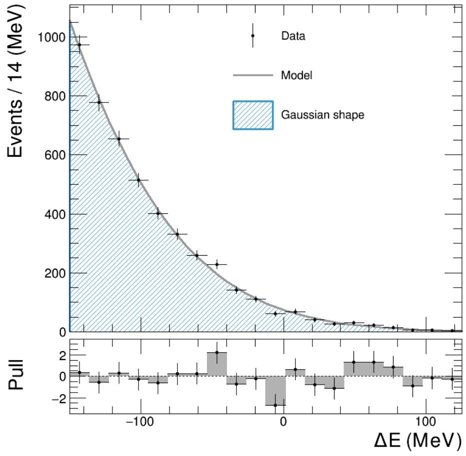
Background from semi-leptonic decays such as $\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu$, true D^0-mesons with a wrongly assigned slow pion and all uncategorised background is combined in $\bar{B}^0 \to D^{*+}\pi^-$and parametrised with a single Gaussian that has its central value far out towards negative values in ΔE, depicted in Fig. 5.21.



(A) (B)

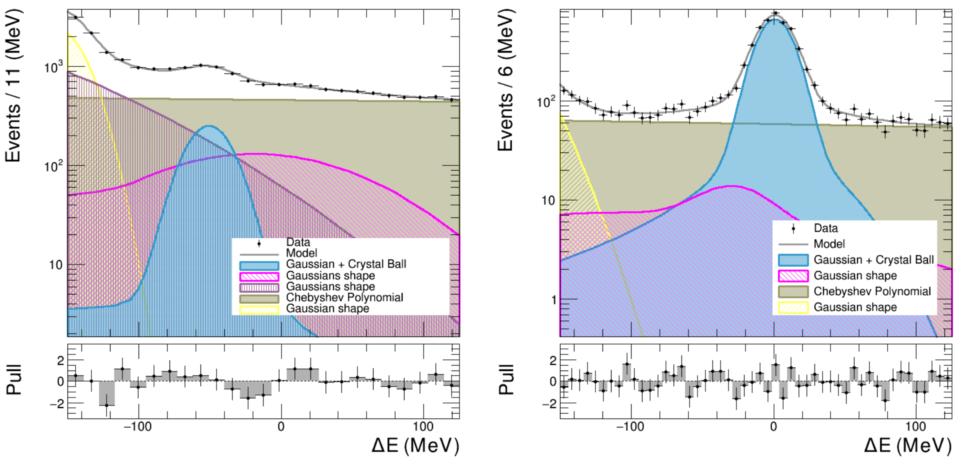
FIGURE 5.20: a) Model for background from D^{*0}-decays and $B^+ \to D^0h^+$ and background from wrongly reconstructed D^0 for $\bar{B}^0 \to D^{*+}\pi^-$. b) Model for background from D^{*0}-decays and $B^+ \to D^0h^+$ and background from wrongly reconstructed D^0 and background from other uncategorised decays in $\bar{B}^0 \to D^{*+}K^-$. Both are in the $D^0 \to K^-\pi^+$ channel.

The individual PDFs were fixed to the shape parameters found in the isolated fits. The joint PDF then was used to fit a data sample without the signal, allowing the total normalisation of the background and other bachelor hadron component to float,



(A)

FIGURE 5.21: Model for $\bar{B}^0 \to D^{*+} l^- \bar{\nu}$ background, background with a true D^0 and a wrongly assigned slow pion as well as other uncategorised background for $\bar{B}^0 \to D^{*+}\pi^-$in the $D^0 \to K^-\pi^+$ channel.



(A) (B)

FIGURE 5.22: Background model of the $\bar{B} \to D^{*+}\pi^-$ mode (a), and of the $\bar{B} \to D^{*+}K^-$ mode (b) in the $D^0 \to K^-\pi^+$ channel, evaluated using a MC sample equivalent to fives times the Belle data set.

while the individual PDF fractions were fixed to the true fractions in the MC. The resulting fit was found to describe the MC very well within less than one standard deviation, see Fig. 5.22.

The combined background and signal model, M, for $\bar{B}^0 \to D^{*+}\pi^-$is

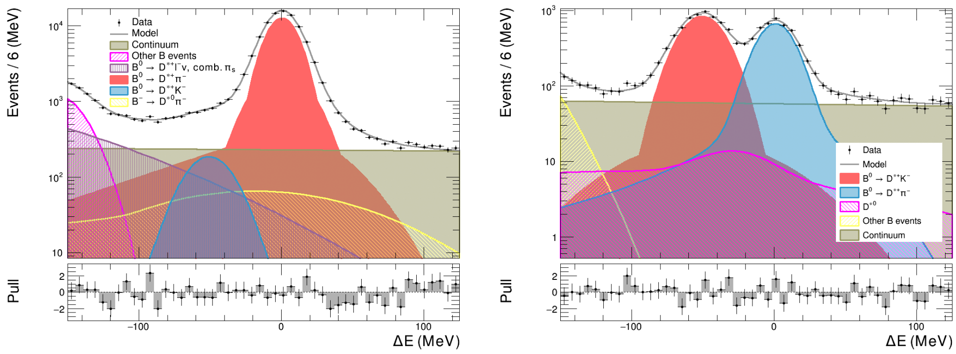
$$\begin{aligned} M^{\bar{B}^0\to D^{*+}\pi^-}_{D^0\to K^-\pi^+} &= N_{\bar{\mathrm{q}}\mathrm{q}1}\mathrm{Ch}^1_{\mathrm{bkg1}} + N_{\mathrm{bkg}}\left[r_{\mathrm{bkg2}}\mathrm{G}_{\mathrm{bkg2}} + r_{\mathrm{bkg3}}\mathrm{G}_{\mathrm{bkg3}} + (1 - r_{\mathrm{bkg2}} - r_{\mathrm{bkg3}})\mathrm{Cb}_{\mathrm{bkg4}}\right] \\ &+ \mathrm{F}_{\bar{B}^0\to D^{*+}\pi^-}(N_\pi, \mu_\pi, \beta_\pi) \\ &+ \mathrm{F}_{\bar{B}^0\to D^{*+}K^-}(N_K, \mu_K, \beta_K) \end{aligned} , \tag{5.11}$$

and for $\bar{B}^0 \to D^{*+}K^-$

$$
\begin{aligned}
M^{\bar{B}^0 \to D^{*+}K^-}_{D^0 \to K^-\pi^+} = {} & N_{\bar{\mathrm{q}}\mathrm{q}2}\mathrm{Ch}^1_{\mathrm{bkg}5} + N_{\mathrm{bkg}}\left[r_{\mathrm{bkg}6}\mathrm{G}_{\mathrm{bkg}6} + (1 - r_{\mathrm{bkg}6})\left(f_1\mathrm{G}_{\mathrm{bkg}7} + (1 - f_1)\mathrm{G}_{\mathrm{bkg}8}\right)\right] \\
& + \mathrm{F}_{\bar{B}^0 \to D^{*+}\pi^-}(N_\pi, \mu_\pi, \beta_\pi) \\
& + \mathrm{F}_{\bar{B}^0 \to D^{*+}K^-}(N_K, \mu_K, \beta_K)
\end{aligned}
, \qquad (5.12)
$$

where N_i are yields, r_i are fixed ratios in the PDF describing the background from B-events, normalised to $1 = r_{\mathrm{bkg}6} + (1 - r_{\mathrm{bkg}6}) = r_{\mathrm{bkg}2} + r_{\mathrm{bkg}3} + (1 - r_{\mathrm{bkg}2} - r_{\mathrm{bkg}3})$. Fractions of PDFs within a background component are described by f_i, free signal means are μ_i and calibration factors are denoted β_i. The PDFs are rank one Chebishev polynomials, Ch^1_i, Gaussian functions, G_i, and Crystal Ball functions, Cb_i. The floated parameters are the signal component means, calibration factors and all yields. All other PDF parameters are fixed and their values are given in Tables 5.13 and 5.12. The signal means and calibration factors are shared and fit simultaneously for both models.

To check the consistency of the models, the signal is added to the sample and a non-simultaneous fit is performed, floating the signal mean, a calibration factor for the signal width and a yield for the combined background PDF, as well as for the signal, depicted in Fig. 5.23. The resulting fits work well and the extracted signal yields reproduce the truth matched signal expectation within the statistical uncertainty.



(A) (B)

FIGURE 5.23: Signal and background model overlaid with MC data for the $\bar{B} \to D^{*+}\pi^-$ mode (a), and the $\bar{B} \to D^{*+}K^-$ mode (b) in the $D^0 \to K^-\pi^+$ channel.

5.5.3 Background models in the $D^0 \to K^-2\pi^+\pi^-$ channel

The background models for the $D^0 \to K^-2\pi^+\pi^-$ channel are fitted analogous way as in $D^0 \to K^-\pi^+$, the results are listed in Table 5.14 and Table 5.15.

TABLE 5.14: Parameters of all PDFs of $\bar{B}^0 \to D^{*+}\pi^-$in the $D^0 \to K^-2\pi^+\pi^-$ channel.

$\mathbf{\bar{B}^0 \to D^{*+}\pi^-}$	PDF: $F_{\bar{B}^0\to D^{*+}\pi^-}$
$\sigma_1 = \beta_{\pi^-} \cdot 17.80 \pm 0.65$ MeV	$\mu_1 \in [-2, -2]$ MeV
$\sigma_2 = \beta_{\pi^-} \cdot 10.05 \pm 0.12$ MeV	$\mu_2 = \mu_1 = \mu_3$
$\sigma_{\text{Crystal Ball}} = \beta_{\pi^-} \cdot 37.14 \pm 0.63$ MeV	
$\alpha_\pi = 0.57 \pm 0.02$	$n_\pi = 99.99 \pm 75.53$
$f_1 = 0.227 \pm 0.021$	$f_2 = 0.103 \pm 0.004$
$\mathbf{q\bar{q}}$	PDF: Ch_1
$a_0 = -0.099 \pm 0.011$	yield free
$\mathbf{\bar{B}^0 \to D^{*+}\rho^-}$	PDF: G_1
$\sigma = 18.10 \pm 2.23$ MeV	$\mu = -161.43 \pm 5.86$ MeV
$r_1 = 0.145 \pm 0.001$	
$\mathbf{\bar{B}^0 \to D^{*+}l^-\nu}$**, other**	PDF: G_2
$\sigma = 78.94 \pm 3.67$ MeV	$\mu = -221.63 \pm 13.08$ MeV
$r_2 = 0.195 \pm 0.001$	
$\mathbf{B^+ \to D^0h^+}$	PDF: G_3
$\sigma = 107.38 \pm 8.59$ MeV	$\mu = -30.19 \pm 6.17$ MeV
bkg. fraction $r_3 = 0.032 \pm 0.001$	
combinatorial $\boldsymbol{\pi_s}$	PDF: G_4 and G_5
$\sigma_1 = 50.24 \pm 71.34$ MeV	$\mu_1 = -200.47 \pm 117.11$ MeV
$\sigma_2 = 51.54 \pm 3.32$ MeV	$\mu_2 = -24.57 \pm 5.80$ MeV
$f_1 = 0.204 \pm 0.062$	$r_4 = 0.046 \pm 0.001$
Fake $\mathbf{D^0,\ D^{*0}}$	PDF: G_6 and G_7
$\sigma_1 = 39.16 \pm 8.87$ MeV	$\mu_1 = 5.09 \pm 5.56$ MeV
$\sigma_2 = 187.98 \pm 27.63$ MeV	$\mu_2 = -144.19 \pm 40.95$ MeV
$f_2 = 0.089 \pm 0.038$	$1 - r_1 - r_2 - r_3 - r_4 = 0.539 \pm 0.001$
$\mathbf{\bar{B}^0 \to D^{*+}K^-}$	PDF: $F_{\bar{B}^0\to D^{*+}K^-}$
all parameters as in $\bar{B}^0 \to D^{*+}K^-$	yield free

The $D^0 \to K^-2\pi^+\pi^-$ channel has two more tracks in the final state and hence more possibilities to end up with a incorrectly reconstructed D^0-meson. These make up a large fraction of the background, see Fig. 5.13, as compared to the $D^0 \to K^-\pi^+$ channel where this background is less abundant.

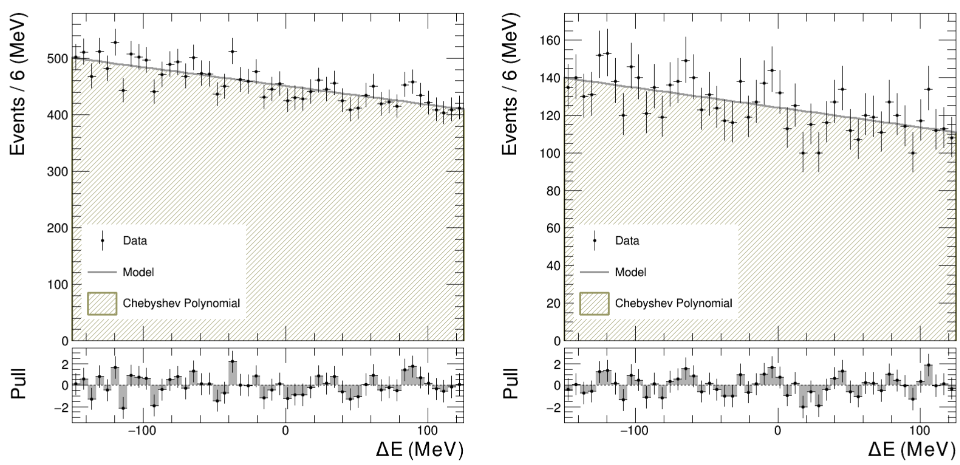
The background from continuum is parametrised similarly to the $D^0 \to K^-\pi^+$ channel with a Chebishev polynomial of rank one, depicted in Fig. 5.24.

Background from decays of $\bar{B}^0 \to D^{*+}\rho^-$ are described with a Gaussian shape, depicted in Fig. 5.25. The PDF is extracted together with the fixed PDF for continuum background, to allow the fit without restricting the range.

A peaking background arises from $\bar{B}^0 \to D^{*+}\pi^-$misidentified as $\bar{B}^0 \to D^{*+}K^-$and vice versa. They will be parametrised with the same model, so that the $\bar{B}^0 \to D^{*+}\pi^-$ peak in the $\bar{B}^0 \to D^{*+}\pi^-$ channel has the same PDF as in the $\bar{B}^0 \to D^{*+}K^-$ channel and vice versa, see Fig. 5.26. The models for $\bar{B}^0 \to D^{*+}K^-$ in $\bar{B}^0 \to D^{*+}\pi^-$ and vice versa, shown here are placeholders to fit the isolated background; In the simultaneous fit the PDF extracted from the isolated signal will be used for $\bar{B}^0 \to D^{*+}K^-$ and $\bar{B}^0 \to D^{*+}\pi^-$.

TABLE 5.15: Parameters of all PDFs of $\bar{B}^0 \to D^{*+}K^-$in the $D^0 \to K^-2\pi^+\pi^-$ channel.

$\mathbf{\bar{B}^0 \to D^{*+}K^-}$	PDF: $F_{\bar{B}^0\to D^{*+}K^-}$
$\sigma_1 = \beta_K \cdot 13.76 \pm 0.18$ MeV	$\mu_1 \in [-51, -48]$ MeV
$\sigma_2 = \beta_K \cdot 49.21 \pm 2.63$ MeV	$\mu_2 = \mu_1$
$\alpha_K = 10.85 \pm 0.61$	$n_K = 0.97 \pm 64.34$
$f_1 = 0.911 \pm 0.017$	
$\mathbf{q\bar{q}}$	PDF: Ch_2
$a_0 = -0.115 \pm 0.021$	yield free
$\mathbf{\bar{B}^0 \to D^{*+}\rho^-, \bar{B}^0 \to D^{*+}l\nu}$	PDF: G_8
$\sigma = 105.60 \pm 197.41$ MeV	$\mu = -791.10 \pm 466.10$ MeV
$r_8 = 0.129 \pm 0.009$	
$\mathbf{D^{*0}, B^+ \to D^0h^+}$	PDF: G_9 and G_{10}
$\sigma_1 = 155.54 \pm 126.22$ MeV	$\mu_1 = -121.06 \pm 157.94$ MeV
$\sigma_2 = 36.29 \pm 126.22$ MeV	$\mu_2 = -37.83 \pm 157.94$ MeV
$f_3 = 0.806 \pm 0.259$	$r_9 = 0.184 \pm 0.011$
Fake $\mathbf{D^0}$, fake $\boldsymbol{\pi_s}$, other	PDF: G_{11}
$\sigma = 95.29 \pm 5.10$ MeV	$\mu = -59.89 \pm 5.28$ MeV
$1 - r_8 - r_9 = 0.685 \pm 0.027$	
$\mathbf{\bar{B}^0 \to D^{*+}\pi^-}$	PDF: $F_{\bar{B}^0\to D^{*+}\pi^-}$
all parameters as in $\bar{B}^0 \to D^{*+}\pi^-$	yield free

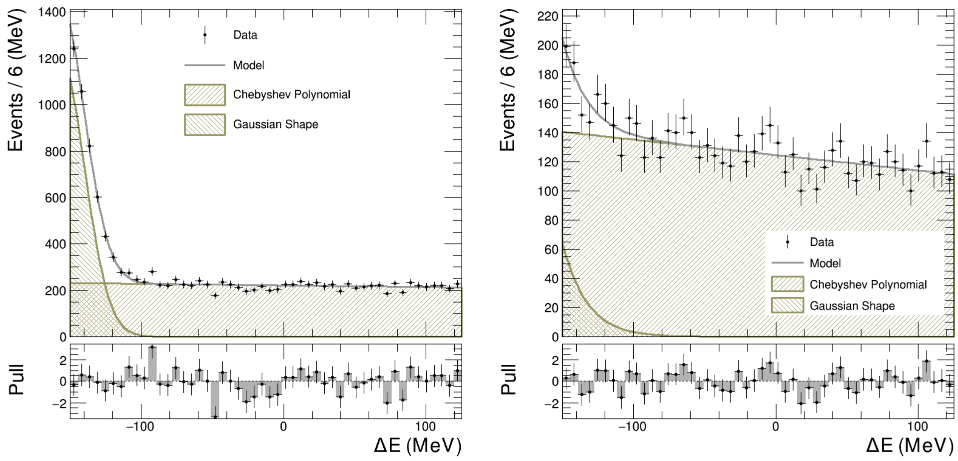


(A) (B)

FIGURE 5.24: Background model for the continuum background for $\bar{B}^0 \to D^{*+}\pi^-$without (a) and kaon mode (b). This is the $D^0 \to K^-2\pi^+\pi^-$ channel, evaluated using a data sample equivalent of fives times the Belle data set.

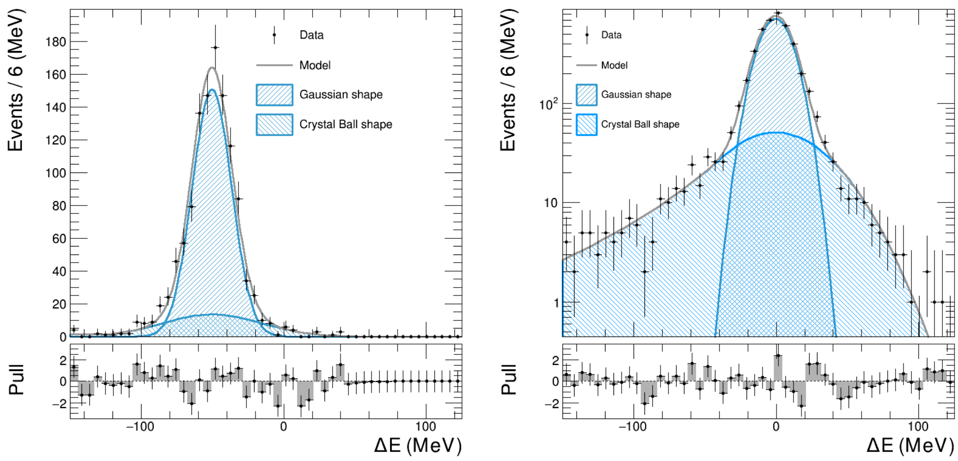
Background originating from a true D^0-meson that was combined with an incorrect combinatorial slow pion can peak in ΔM if the slow pion comes from another D^*-decay. This background is parametrised in isolation for $\bar{B}^0 \to D^{*+}\pi^-$with a double Gaussian, as depicted in Fig. 5.27a.

In $\bar{B}^0 \to D^{*+}K^-$they combined together with incorrectly reconstructed D^0-mesons and all uncategorised decays, shown in Fig. 5.29a. A single Gaussian was chosen to describe the distribution. In $\bar{B}^0 \to D^{*+}\pi^-$background from wrongly reconstructed D^0 were joined together with D^0 from D^{*0}, see Fig. 5.29b.



(A) (B)

FIGURE 5.25: Model for $\bar{B}^0 \to D^{*+}\rho^-$ background in $\bar{B}^0 \to D^{*+}\pi^-$(a) and $\bar{B}^0 \to D^{*+}K^-$(b). This is the $D^0 \to K^-2\pi^+\pi^-$ channel, evaluated using a data sample equivalent of fives times the Belle data set.

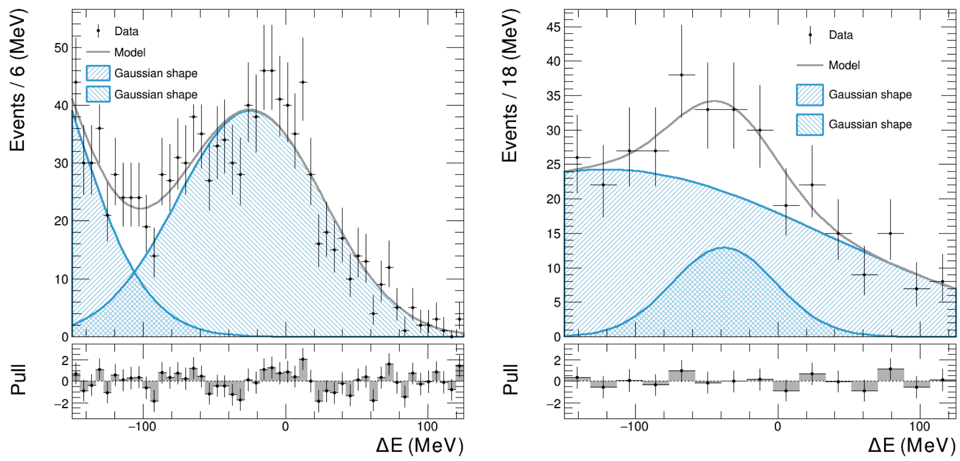


(A) (B)

FIGURE 5.26: Model for the $\bar{B} \to D^{*+}K^-$ contribution in $\bar{B}^0 \to D^{*+}\pi^-$without (a) and the $\bar{B} \to D^{*+}\pi^-$ contribution in $\bar{B}^0 \to D^{*+}K^-$(b). This is the $D^0 \to K^-2\pi^+\pi^-$ channel, evaluated using a data sample equivalent of fives times the Belle data set. Both of these will be fit with a calibration factor for the the width that is simultaneously fit in both hadron channels.

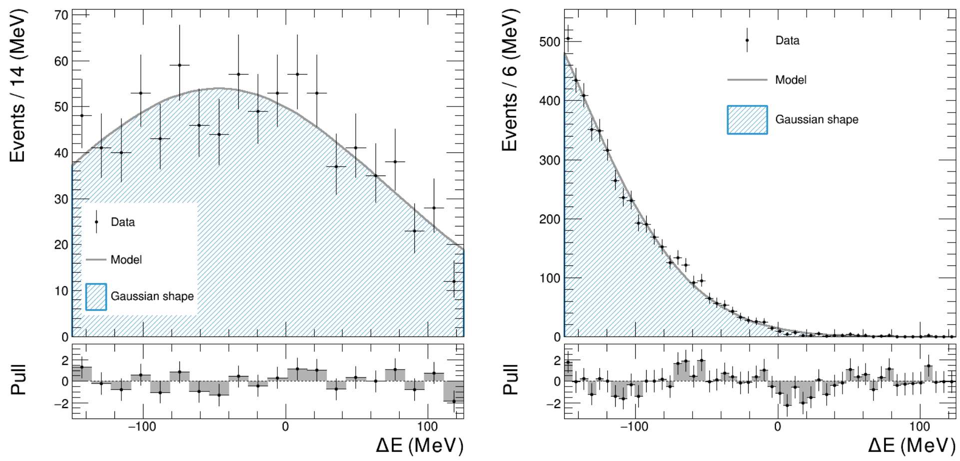
Background from $B^+ \to D^0\pi^+$ were fitted with a single Gaussian depicted in Fig. 5.28a. In $\bar{B}^0 \to D^{*+}K^-$these were combined with D^{*0} background, depicted in Fig. 5.27b.

$\bar{B}^0 \to D^{*+}\pi^-$has another background from semi-leptonic decays such as $\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu$, where the muon mimics the bachelor pion. These are described together with all uncategorised background, by a Gaussian that has its mean far out the signal window towards negative values of ΔE, as displayed in Fig. 5.28b. These are far less abundant in $\bar{B}^0 \to D^{*+}K^-$and were added to the $\bar{B}^0 \to D^{*+}\rho^-$ background PDF component, since they have a similar exponential shape that does not fit in any of the other PDFs but is not abundant enough to justify its own PDF.



(A) (B)

FIGURE 5.27: a) Background model for combinatorial slow pion background in $\bar{B}^0 \to D^{*+}\pi^-$. b) Background model for decays involving D^{*0} and $B^+ \to D^0\pi^+$ background in $\bar{B}^0 \to D^{*+}K^-$.

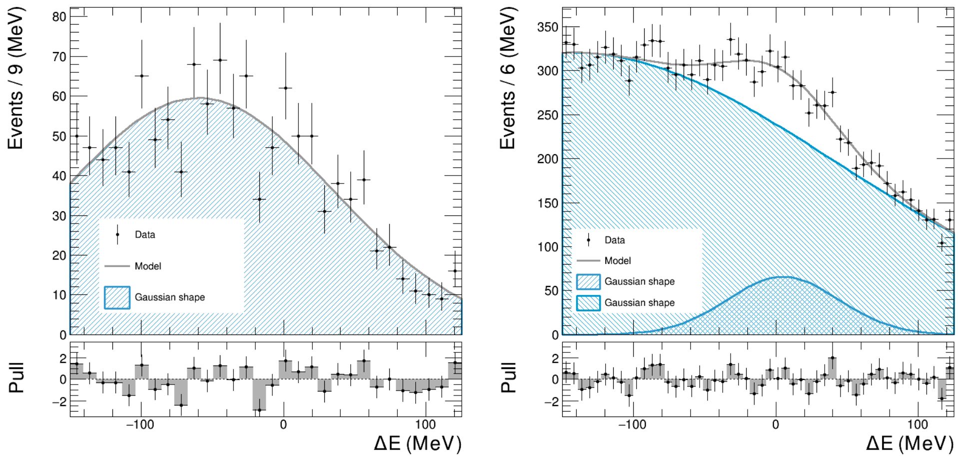


(A) (B)

FIGURE 5.28: a) Background model for background from $B^+ \to D^0\pi^+$ in $\bar{B}^0 \to D^{*+}\pi^-$. b) Background model for semi-leptonic and uncategorised decays in $\bar{B}^0 \to D^{*+}\pi^-$.

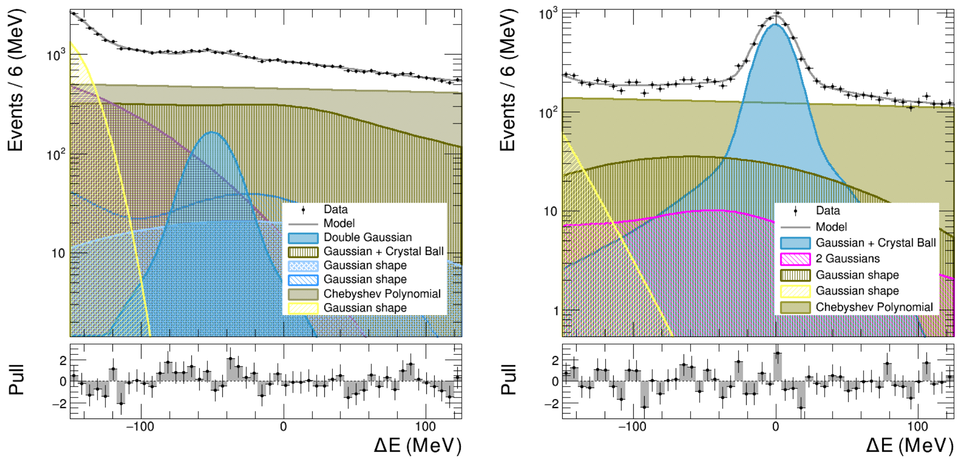
The combined background and signal model, M, for $\bar{B}^0 \to D^{*+}\pi^-$is

$$
\begin{aligned}
M^{\bar{B}^0\to D^{*+}\pi^-}_{D^0\to K^-2\pi^+\pi^-} = &\, N_{\bar{q}q1}\mathrm{Ch}^1_1 + N_{\mathrm{bkg}}[r_1\mathrm{G}_1 + r_2\mathrm{G}_2 + r_3\mathrm{G}_3 + r_4(f_1\mathrm{G}_4 + (1-f_1)\mathrm{G}_5) \\
&+ (1 - r_1 - r_2 - r_3 - r_4)(f_2\mathrm{G}_6 + (1-f_2)\mathrm{G}_7)] \\
&+ F_{\bar{B}^0\to D^{*+}\pi^-}(N_\pi, \mu_\pi, \beta_\pi) \\
&+ F_{\bar{B}^0\to D^{*+}K^-}(N_K, \mu_K, \beta_K)
\end{aligned}
, \qquad (5.13)
$$



(A) (B)

FIGURE 5.29: a) Model for background from incorrectly reconstructed D^0-mesons, combinatorial slow pions and uncategorised background in $\bar{B}^0 \to D^{*+}K^-$. b) Background model for the fake D^0-meson background and background from D^{*0}-decays in $\bar{B}^0 \to D^{*+}\pi^-$.



(A) (B)

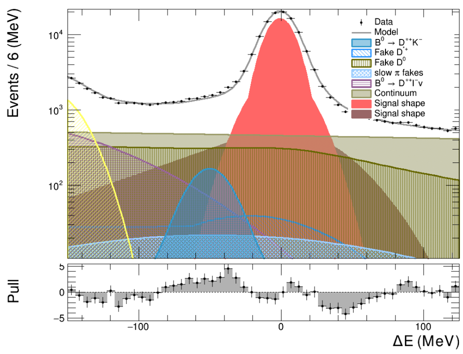
FIGURE 5.30: Background model of the $\bar{B} \to D^{*+}\pi^-$ mode (a), and the $\bar{B} \to D^{*+}K^-$ mode (b), in the $D^0 \to K^-2\pi^+\pi^-$ channel, evaluated using a data sample equivalent of fives times the Belle data set. The PDFs of $\bar{B} \to D^{*+}K^-$ in $\bar{B} \to D^{*+}\pi^-$ and $\bar{B} \to D^{*+}\pi^-$ in $\bar{B} \to D^{*+}K^-$ will be added a calibration factor and replaced by the respective signal PDF in a simultaneous fit.

and for $\bar{B}^0 \to D^{*+}K^-$

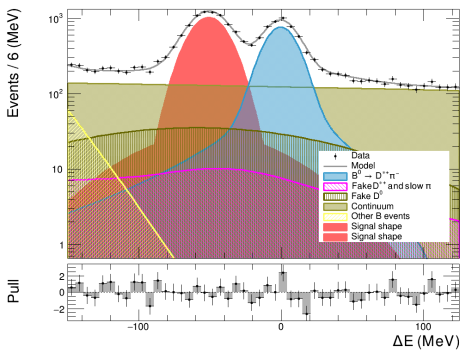
$$\begin{aligned} M_{D^0\to K^-2\pi^+\pi^-}^{\bar{B}^0\to D^{*+}K^-} = & N_{\bar{q}q2}\mathrm{Ch}_2^1 + N_{\mathrm{bkg}}[r_8\mathrm{G}_8 + r_9\left(f_3\mathrm{G}_9 + (1-f_3)\mathrm{G}_{10}\right) \\ & + (1-r_8-r_9)\mathrm{G}_{11}] \\ & + F_{\bar{B}^0\to D^{*+}\pi^-}(N_\pi, \mu_\pi, \beta_\pi) \\ & + F_{\bar{B}^0\to D^{*+}K^-}(N_K, \mu_K, \beta_K) \end{aligned} , \tag{5.14}$$

where N_i are yields, r_i fixed ratios in the PDF describing the background from B-events, normalised to $1 = r_8 + r_9 + (1 - r_8 - r_9) = r_1 + r_2 + r_3 + r_4 + (1 - r_1 - r_2 - r_3 - r_4)$. Fractions of PDFs within a background component are described by f_i, μ_i are free signal means and β_i are calibration factors. The PDFs are rank one Chebishev polynomials, Ch_i^1, Gaussian functions, G_i, and Crystal Ball functions, Cb_i. The floated parameters are the signal component means, calibration factors and all yields. All other PDF parameters are fixed and their values may be found in Table 5.15 and Table 5.14. The signal means and calibration factors are shared and fit simultaneously for both models. The signal PDFs, F_i, are the same as in $D^0 \to K^-\pi^+$given in Eq. 5.8 and Eq. 5.9.

The individual PDFs were combined fixing their fractions to the value obtained from MC, except for continuum background. The fit was performed, by allowing the combined background normalisation and the continuum background normalisation float, the results are depicted in Fig. 5.30. With the combined PDF the background was fitted and the yield was found to be within one standard deviation of the truth matched yield. Performing the fit with floating total signal yields, a yield for all earlier mentioned background and a separate continuum yield, as well as a calibration factor for the width of the signal PDFs, it was found that these describe the data well. However, a small disagreement around the steep slopes of the signal PDF are observed, similarly to the $D^0 \to K^-\pi^+$ channel, these lead to biases, which will be treated as systematic uncertainties. The results using MC samples equivalent of fives times the Belle data set are depicted in Fig. 5.31.



(A)

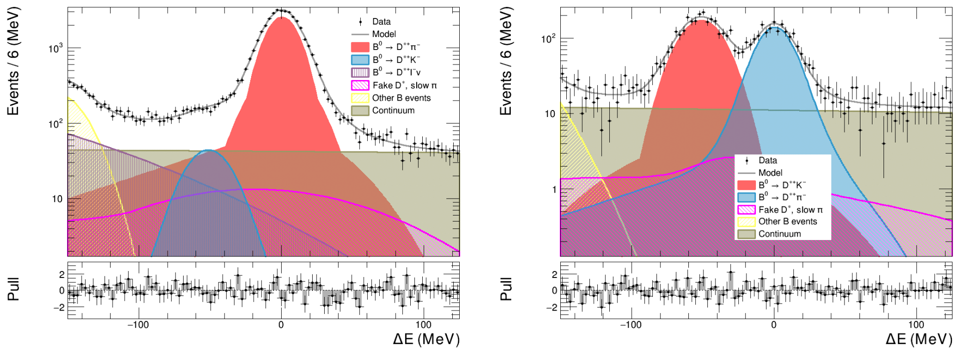


(B)

FIGURE 5.31: Full model including signal in $\bar{B}^0 \to D^{*+}\pi^-$(a), and (b) in the $D^0 \to K^-2\pi^+\pi^-$ channel.

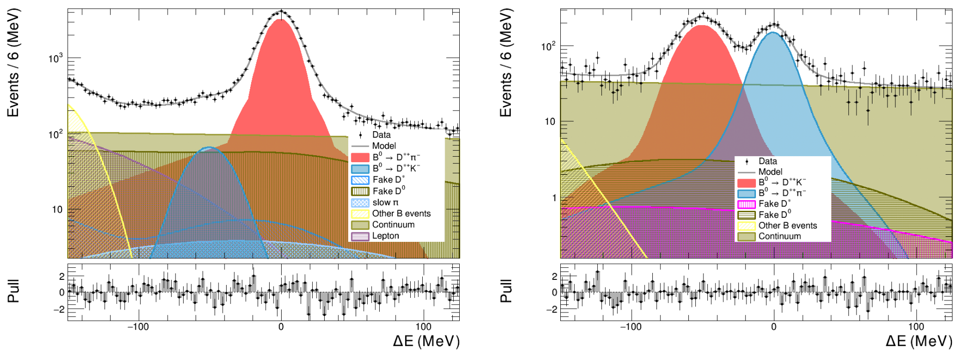
5.5.4 Fit results using a single stream of MC

The extracted PDFs were used to perform a simultaneous fit of the PDF of $\bar{B}^0 \to D^{*+}\pi^-$ and $\bar{B}^0 \to D^{*+}K^-$ in both data samples to a data set equivalent to the size of the Belle dataset. The fit was performed with all parameters of the PDFs fixed, except for the widths and central values of the $\bar{B}^0 \to D^{*+}h^-$ components. The central values were allowed to float in a window of ± 2 MeV around the expected value. The results are depicted in Fig. 5.32 and Fig. 5.33. The correlation matrices of these test



(A) (B)

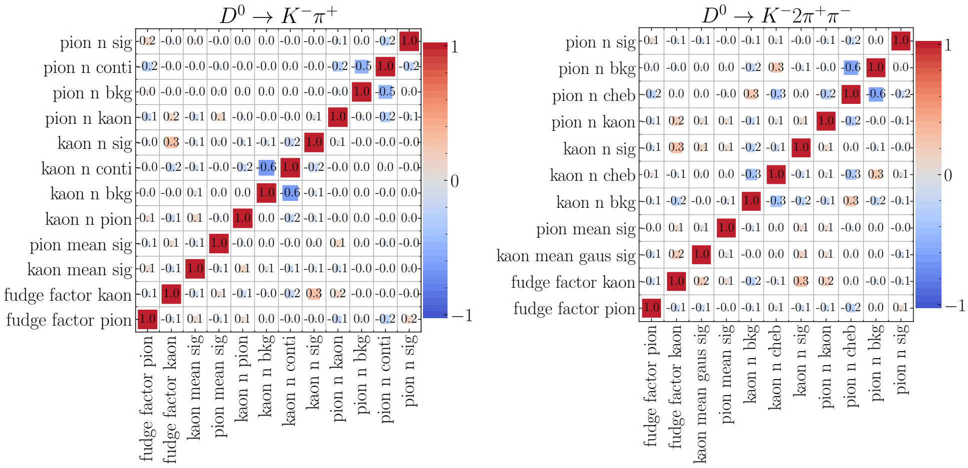
FIGURE 5.32: Projection of the fit results for the simultaneous fit on MC. The results for the $D^0 \to K^-\pi^+$ channel are shown for the $\bar{B}^0 \to D^{*+}\pi^-$ (a), and $\bar{B}^0 \to D^{*+}\pi^-$ (b). These results are evaluated on stream 1. The PDFs for the hadron peaks are the simultaneously extracted in both hadron modes.



(A) (B)

FIGURE 5.33: Projection of the fit results for the simultaneous fit on MC. The results for the $D^0 \to K^-2\pi^+\pi^-$ channel are shown for $\bar{B}^0 \to D^{*+}\pi^-$ (a), and $\bar{B}^0 \to D^{*+}\pi^-$ (b). These results are evaluated on stream 1. The PDFs for the hadron peaks are the simultaneously extracted in both hadron modes.

fits are depicted in Fig. 5.34. The largest correlations found are between the yields of the continuum background and the yields of background from B-events. This is expected as they are both part of the total background.



(A) (B)

FIGURE 5.34: Correlation matrix of the floating fit parameters in the $D^0 \to K^-\pi^+$ channel (a), and the $D^0 \to K^-2\pi^+\pi^-$ channel (b).

5.6 Reconstruction efficiency

The efficiency is evaluated using MC equivalent to five times the Belle data set. The total number of decays into the respective final states on MC particle level before detector simulation is taken as the denominator for the efficiency, ϵ, such that

$$\epsilon = \frac{N_{\text{obs}}}{N_{\text{total}}}, \tag{5.15}$$

where N_{obs} is the true number of truth matched reconstructed candidates after all selection criteria and N_{total} is the total number of decays at generator level, before reconstruction. The results are listed in Table 5.16. The simulated signal decay modes, in particular the D^0-modes, and their branching ratios into the respective final state can be found in Table 5.4.

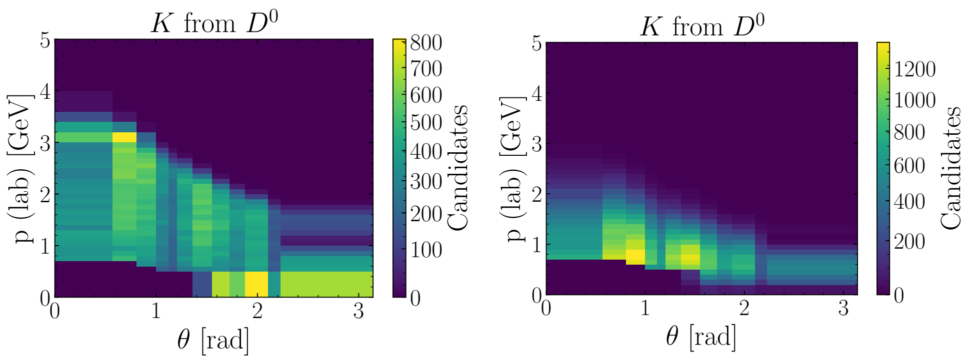
TABLE 5.16: Average number of decays in MC and extracted efficiencies. The efficiency errors are propagated from the Gaussian yield uncertainties.

	$\mathbf{D^0 \to K^-\pi^+}$	$\mathbf{D^0 \to K^-2\pi^+\pi^-}$
N_{total} $\bar{B}^0 \to D^{*+}\pi^-$	266419	518125
N_{total} $\bar{B}^0 \to D^{*+}K^-$	20204	39502
N_{obs} $\bar{B}^0 \to D^{*+}\pi^-$	87044	92487
N_{obs} $\bar{B}^0 \to D^{*+}K^-$	5725	5920
$\epsilon_{\bar{B}^0 \to D^{*+}\pi^-}$	0.3267 ± 0.0012	0.1785 ± 0.0006
$\epsilon_{\bar{B}^0 \to D^{*+}K^-}$	0.2833 ± 0.0042	0.1498 ± 0.0020

5.6.1 Corrections to the efficiency

The overall reconstruction efficiency is determined with MC but is different for data and MC due to mis-modelling for tracks with momenta $p_{\text{lab}} < 200$ MeV. For this analysis only the K/π-identification and slow-π-tracking mis-modelling need to be corrected because no lepton-identification selection criteria was used.

***K*/*π*-identification** The experiment provides measurements of correction factors for data-MC differences, binned in angles and momentum. However, the previously obtained corrections [57] are incomplete and do not cover high polar angle, low momentum regions. The regions that are not available in the official corrections will be corrected by efficiencies for pions determined in Appendix B. Events with kaons that do have correction factors will be removed from the analysis. For $D^0 \to K^-\pi^+$ this affects 6.1 % of the candidates and 14.4 % for $D^0 \to K^-2\pi^+\pi^-$. The bachelor kaon tracks are not affected as their momentum is high enough to always fall into bins that have good measurements. The phase space distribution of the kaon tracks that survive the selection are depicted in Fig. 5.35.



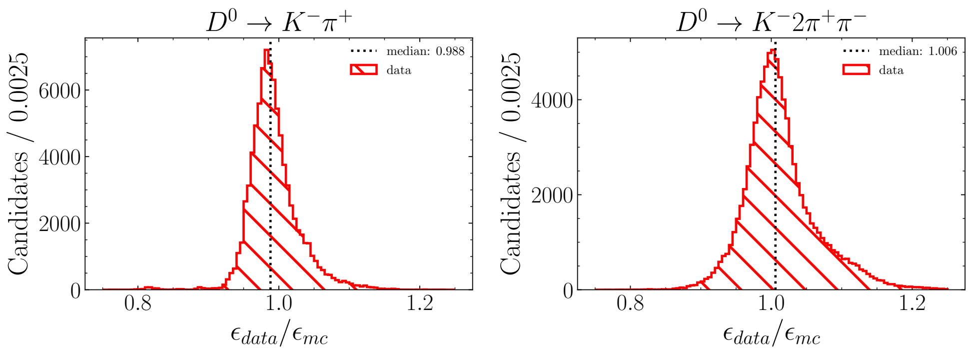
(A) (B)

FIGURE 5.35: Distribution of the kaon tracks in bins of θ and momentum that have correction factor and are used in this analysis for a) $D^0 \to K^-\pi^+$ and b) $D^0 \to K^-2\pi^+\pi^-$.

The final efficiency will be corrected for data-MC disagreement, such that

$$\epsilon' = \epsilon \cdot \frac{\epsilon_{\text{data}}}{\epsilon_{\text{MC}}}, \tag{5.16}$$

where ϵ is the efficiency and ϵ' the corrected efficiency. The distributions of the correction factors are depicted in Fig. 5.36. The efficiency will be corrected by the median, μ, of these distributions. The numerical values are $\mu_{D^0\to K^-\pi^+} = 0.988$ and $\mu_{D^0\to K^-2\pi^+\pi^-} = 1.006$. The individual K/π-efficiencies can be found in Fig. 5.37.



(A) (B)

FIGURE 5.36: Distribution of the efficiency corrections to for each $\bar{B} \to D^{*+}h^-$ candidate a) $D^0 \to K^-\pi^+$ and b) $D^0 \to K^-2\pi^+\pi^-$.

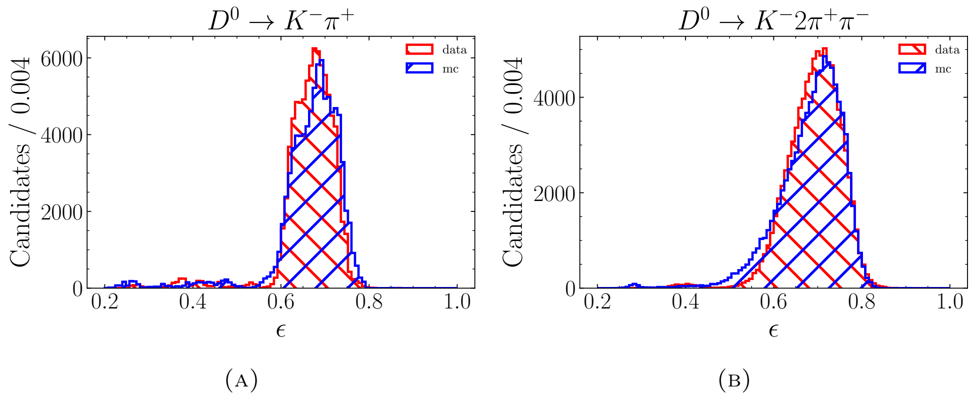


FIGURE 5.37: Distribution of the K/π-ID efficiencies in data and MC in a) $D^0 \to K^-\pi^+$ and b) $D^0 \to K^-2\pi^+\pi^-$.

Slow-π-identification The tracking of low momentum pions is especially challenging and the experiment provides measurements of mis-modelling in bins of lab frame track momentum. The correction is following the same procedure as described for K/π-identification using Eq. 5.16. The same correction factor of $\epsilon_{\text{data}}/\epsilon_{\text{MC}} = 0.993$ is applied for all channels, see Fig. 5.38 for the distributions.

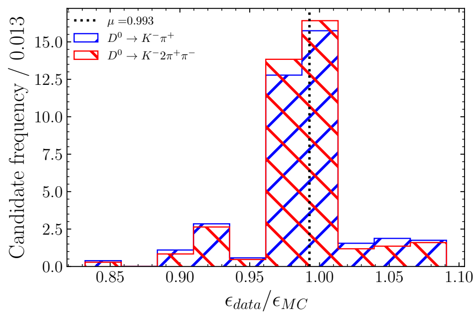


FIGURE 5.38: Distribution of the slow-π correction factors.

5.7 Branching ratio

The branching ratios are calculated from the signal yield, N_{meas}, as

$$\mathcal{B}(B^0 \to D^{*+}h^-) = \frac{N_{\text{meas}}}{N_{B^0} \cdot \mathcal{B}(D^{*+}) \cdot \mathcal{B}(D^0) \cdot \epsilon'_h}, \tag{5.17}$$

where ϵ'_h is the corrected efficiency, $\mathcal{B}(D^0 \to K^-\pi^+) = 0.03950 \pm 0.00031$, $\mathcal{B}(D^0 \to K^-2\pi^+\pi^-) = 0.0823 \pm 0.0014$ and $\mathcal{B}(D^{*+} \to D^0\pi^+ = 0.667 \pm 0.005$ the D^0 and D^{*+} branching ratios from 2020 PDG and N_{B^0} the total number of neutral B-mesons in the data set, which is given by

$$N_{B^0} = N_{B\bar{B}} \cdot f_{B^0} = (750.0 \pm 13.8) \cdot 10^6, \tag{5.18}$$

where the fraction of $\Upsilon(4S)_{b\bar{b}} \to B^0\bar{B}^0$ decays is taken from 2020 PDG as $f_{B^0} = 0.486 \pm 0.006$ and the total number of $B\bar{B}$-pairs was measured by the experiments performance group as $N_{B\bar{B}} = (771.581 \pm 10.566) \cdot 10^6$.

5.8 Systematic uncertainty evaluation

Systematic uncertainties arise from various sources when combining measurements to perform a complex measurement like the extraction of physics parameters from detector hits. The uncertainties that have to be considered are due to detector performance and quantities of the branching ratio, Eq. 5.17. These are are:

- limited MC sample size,
- tracking,
- correction to the K/π-identification efficiency,
- correction to slow-π-identification efficiency,
- number of B^0-mesons,
- branching ratio of the D^0 and D^*-decay,
- fixed PDF parameters in the fits,
- biases due to the choice of PDFs.

Ratios of measurement may lead to some of the systematic errors cancelling. This section will discuss the various sources of uncertainty in more detail and how they are incorporated into the analysis.

Systematic uncertainties of the reconstruction efficiency. The reconstruction efficiency contains all efficiencies related to K/π-ID, tracking and detector acceptance. Detector effects such as the reconstruction efficiency are treated as statistical and arise from the limited sample size given by MC.

Systematic uncertainties of tracking. Systematic uncertainties originating from track reconstruction are described by a single value per track for tracks with $p_{\text{lab}} > 200$ MeV. The total uncertainty is given by the number of tracks in the final state as a correlated sum such that

$$\sigma_{\text{tracking}} = n_{\text{tracks}} \cdot 0.35\%, \tag{5.19}$$

where n_{tracks} is the number of tracks, omitting the slow-π, which are evaluated separately. The number of 0.35% per track was provided by the experiment performance group.

Branching ratios and B^0-meson counting The uncertainties of the branching ratios $\mathcal{B}(\Upsilon(4S)_{b\bar{b}} \to B^0\bar{B}^0)$, $D^{*+} \to D^0\pi^+$, $D^0 \to K^-\pi^+$ and $D^0 \to K^-2\pi^+\pi^-$ are taken from 2020 PDG. Furthermore, there is an uncertainty on the number of total B-meson pairs in the data set, measured by the experiment performance group. Its uncertainties are propagated using Eq. 5.18. The resulting uncertainties are listed in Table. 5.17.

Systematic uncertainties from K/π and slow-π selection Systematic uncertainties arise from statistical and systematic errors to on the slow-π tracking and particle-ID. The evaluation is done in bins in momentum and θ for K/π-identification and in bins of momenta for slow-π-tracking. Consequently, the tracks falling into each bin have to

TABLE 5.17: Uncertainties of branching ratios of various decay channels in percent.

Source	Value
$\mathcal{B}(\Upsilon(4S)_{b\bar{b}} \to B^0\bar{B}^0)$	$\sigma_{\mathcal{B}(\Upsilon(4S)_{b\bar{b}} \to B^0\bar{B}^0)} = 1.23\%$
Number of $B\bar{B}$ pairs	$\sigma_{N(B\bar{B})} = 1.37\%$
Number of B^0-mesons	$\sigma_{N(B^0)} = 1.84\%$
$\mathcal{B}(D^{*+} \to D^0\pi^+)$	$\sigma_{\mathcal{B}(D^{*+} \to D^0\pi^+)} = 0.74\%$
$\mathcal{B}(D^0 \to K^-\pi^+)$	$\sigma_{\mathcal{B}(D^0 \to K^-\pi^+)} = 0.78\%$
$\mathcal{B}(D^0 \to K^-2\pi^+\pi^-)$	$\sigma_{\mathcal{B}(D^0 \to K^-2\pi^+\pi^-)} = 1.7\%$

be counted and their uncertainties combined. The total uncertainty for each of these categories is calculated as squared sums for statistical and uncorrelated errors

$$\sigma_{\text{uncorr}} = \frac{1}{N_{\text{particles}}} \sum_{\text{bin}=j} (\sigma^j_{\text{uncorr}} \cdot N^j_{\text{bin}})^2, \tag{5.20}$$

and linear sums for correlated errors such that

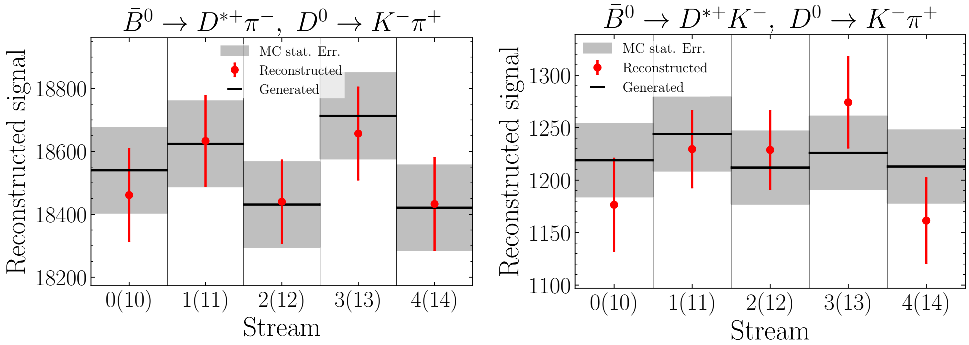
$$\sigma_{\text{corr}} = \frac{1}{N_{\text{particles}}} \sum_{\text{bin}=j} \sigma^j_{\text{corr}} \cdot N^j_{\text{bin}}, \tag{5.21}$$

where N_{bin} refers to the total number of particles in a bin.

Systematic uncertainties from the fit procedure Systematic uncertainties due to the fit procedure may arise due to fixed parameters in the PDFs and require a detailed study of how the fit behaves for different input samples. These effects are studied in the following sections. First, the fit is tested on five data samples each of the size of the Belle data set. This is followed by a toy MC method that resamples the existing MC and repeat large numbers of fits. The results then are used to check if the generated signal yields and branching ratios can be measured and if biases are present. Furthermore, the toy MC method is used to study the effect of changing fixed ratios of the PDF yields and parameters of the PDFs to obtain systematic uncertainties related to these.

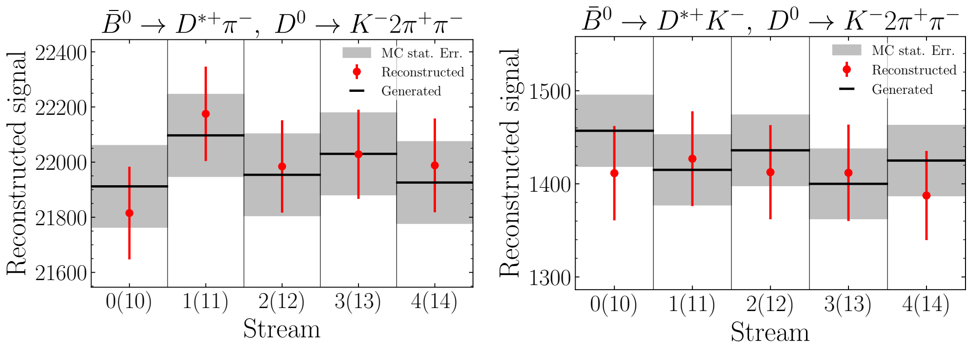
5.8.1 Stream tests

To understand the quality of the fit procedure, the fit is repeated first on the individual MC streams and the generated and measured signal yields are compared, see Fig. 5.39 and Fig. 5.40 for the extracted yield in the respective D^0-meson channel. The resulting pulls are depicted in Fig. 5.41. It was found that the fit reproduces the signal yield reliably within one standard deviation in both D^0 channels.



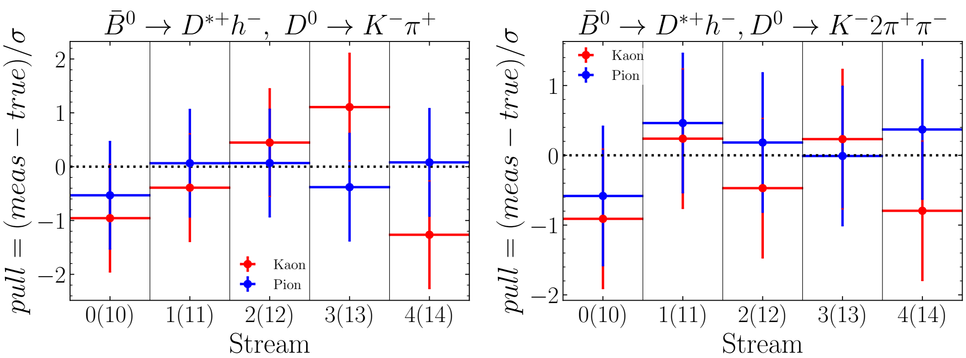
(A) (B)

FIGURE 5.39: Evaluation of the fit quality by performing a stream test, fitting each MC stream individually in the $D^0 \to K^-\pi^+$ channel.



(A) (B)

FIGURE 5.40: Evaluation of the fit quality by performing a stream test, fitting each MC stream individually in the $D^0 \to K^-2\pi^+\pi^-$ channel.

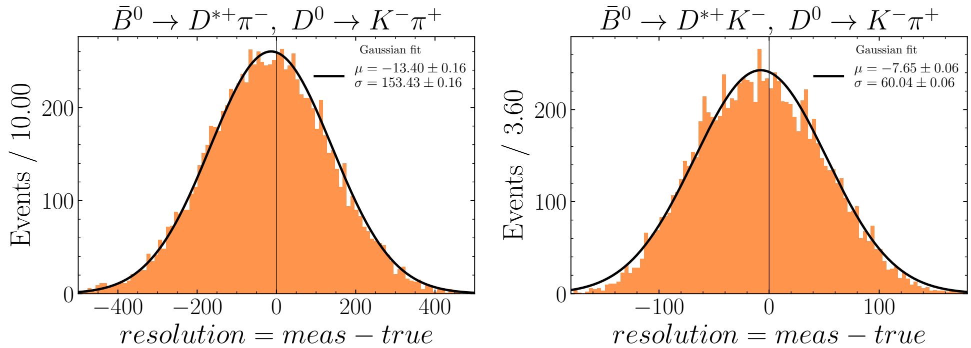


(A) (B)

FIGURE 5.41: Evaluation of the fit pull distributions for each stream.

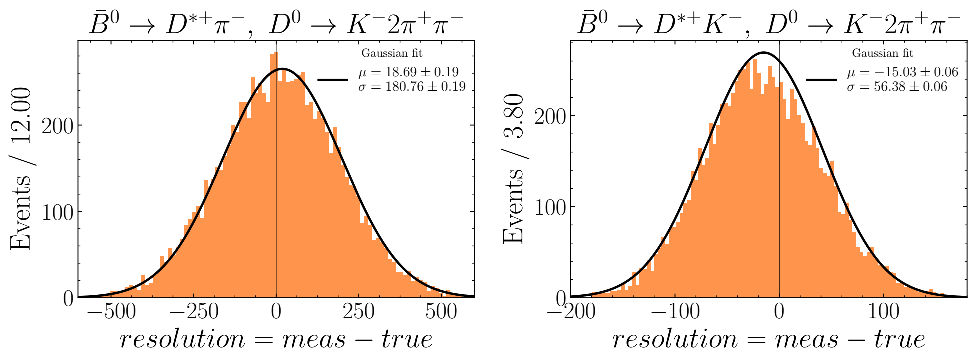
5.8.2 Systematic uncertainties due to the signal extraction method

Testing of individual streams of MC can not give an exact quantitative evaluation of eventual present biases and resolutions, a statistical resampling, toy MC, method was used to evaluate them. For this procedure a random stream was chosen and the data were histogramed in 100 bins. The individual values in each of the bins where then replaced by a random number following a Poissonian distribution using the previous value as central value. The histograms then were randomised and fit with the nominal PDFs 5000 times. The resulting distributions for the resolution and pulls are depicted in Fig. 5.42 and Fig. 5.46 for the $D^0 \to K^-\pi^+$ channel, and Fig. 5.43 and Fig. 5.47 for the $D^0 \to K^-2\pi^+\pi^-$ channel. In some of the pull distributions a spike at zero can be seen, these are due to outliers in the uncertainty distributions. The resolutions and pulls reproduce the expected number of particles and hence the fits are found to work well. Furthermore, the calibration factors are agreeable with an almost Gaussian distributions around unity, see Fig. 5.44 and Fig. 5.45, which also points towards well working fits. Fit biases and systematic effects arising from fixing PDF parameters are studied using the same toy MC method, see Sec. 5.8.4.



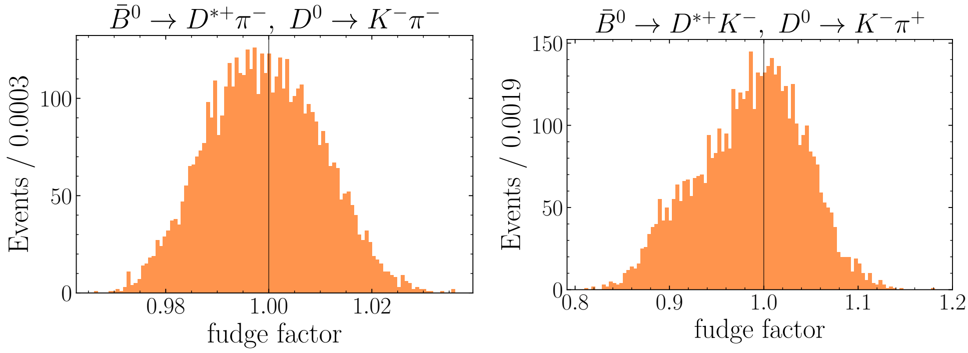
(A) (B)

FIGURE 5.42: Distributions of the resolution in 10000 toy MC experiments for $\bar{B}^0 \to D^{*+}\pi^-$(a) and $\bar{B}^0 \to D^{*+}K^-$(b) in the $D^0 \to K^-\pi^+$ channel.



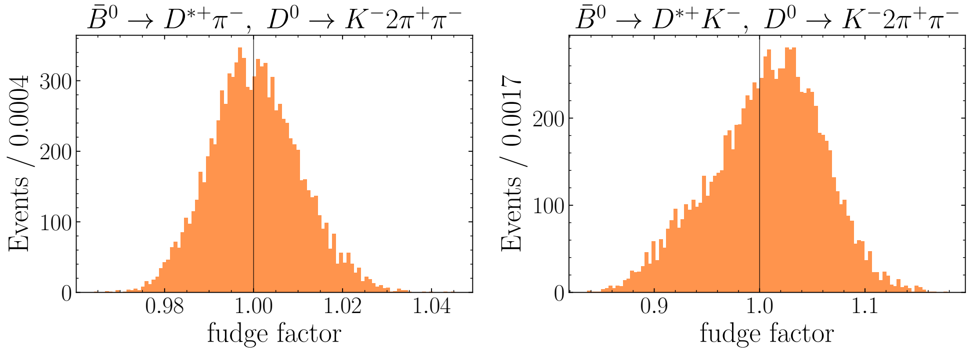
(A) (B)

FIGURE 5.43: Distributions of the signal yield resolution in 5000 toy MC experiments for $\bar{B}^0 \to D^{*+}\pi^-$(a) and $\bar{B}^0 \to D^{*+}K^-$(b) in the $D^0 \to K^-2\pi^+\pi^-$ channel.



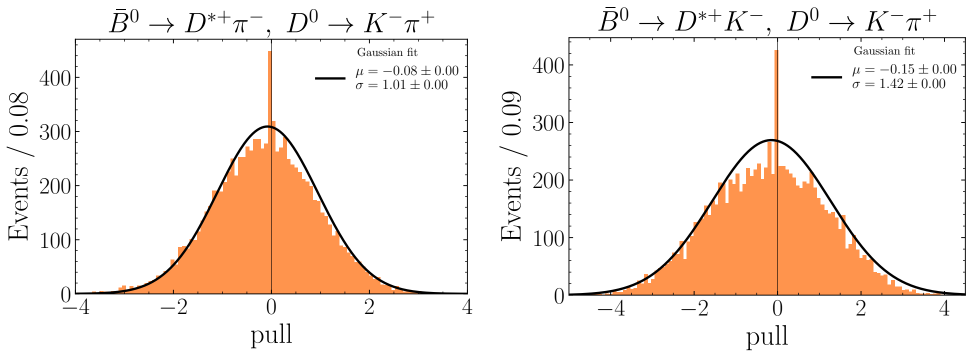
(A) (B)

FIGURE 5.44: Distributions of the calibration factors, defined in Eq. 5.10, in 5000 toy MC experiments for $\bar{B}^0 \to D^{*+}\pi^-$(a) and $\bar{B}^0 \to D^{*+}K^-$(b) in the $D^0 \to K^-\pi^+$ channel.



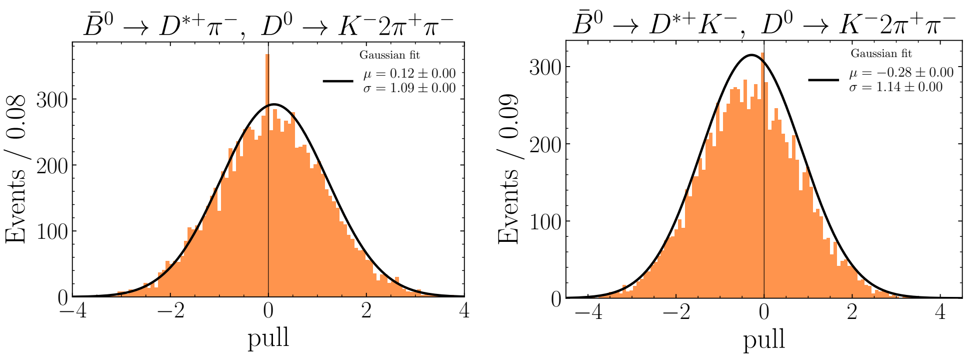
(A) (B)

FIGURE 5.45: Distributions of the calibration factors, defined in Eq. 5.10, in 5000 toy MC experiments for $\bar{B}^0 \to D^{*+}\pi^-$(a) and $\bar{B}^0 \to D^{*+}K^-$(b) in the $D^0 \to K^-2\pi^+\pi^-$ channel.



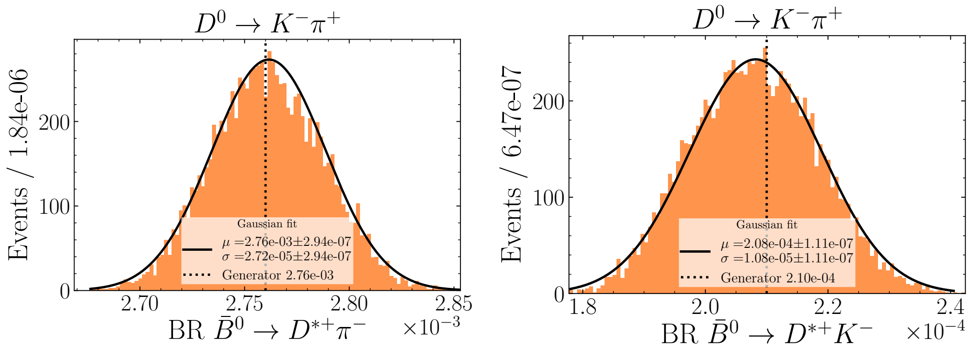
(A) (B)

FIGURE 5.46: Pull distributions of the signal yield in 5000 toy MC experiments for $\bar{B}^0 \to D^{*+}\pi^-$(a) and $\bar{B}^0 \to D^{*+}K^-$(b) in the $D^0 \to K^-\pi^+$ channel.



(a) (b)

Figure 5.47: Pull distributions of the signal yield in 5000 toy MC experiments for $\bar{B}^0 \to D^{*+}\pi^-$(a) and $\bar{B}^0 \to D^{*+}K^-$(b) in the $D^0 \to K^-2\pi^+\pi^-$ channel.



(A) (B)

FIGURE 5.48: Distributions of the branching ratio measured in 5000 toy MC experiments for $\bar{B}^0 \to D^{*+}\pi^-$(a), and $\bar{B}^0 \to D^{*+}K^-$(b) in the $D^0 \to K^-\pi^-$ channel.

5.8.3 Branching ratio closure test

In order to ensure self consistency of the method the branching ratio is measured on MC and compared to the generated value. The branching ratios are calculated using the previously obtained yields from the toy study. For each toy experiment the branching ratio is calculated as

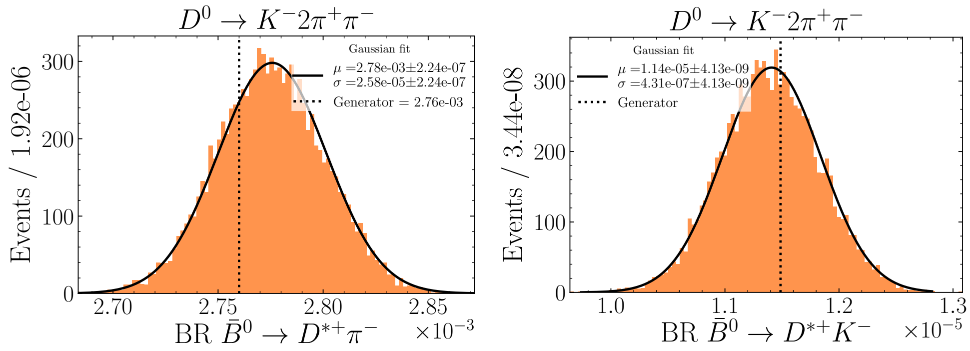
$$\mathcal{B}(\bar{B}^0 \to D^{*+}h^-) = \frac{N_{\text{meas}}}{N_{B^0_{\text{generated}}} \cdot \mathcal{B}(D^{*+}) \cdot \mathcal{B}(D^0) \cdot \epsilon_h}. \tag{5.22}$$

The results for the $D^0 \to K^-\pi^+$ channel are depicted in Fig. 5.48 and for the $D^0 \to K^-2\pi^+\pi^-$ channel in Fig. 5.49, as well as Fig. 5.52 and Fig. 5.53 for the combined case. The generator truth is calculated from the number of B-mesons and the generator branching ratios as in Table 5.4. In both cases the generated values are reproduced. However, for $D^0 \to K^-2\pi^+\pi^-$ a small deviation can be observed. As the total number of MC-truth matched candidates was reproduced by the fit, this may be due to a small fluctuation in the number of B-mesons decaying into $\bar{B}^0 \to D^{*+}\pi^-$in the MC.

The ratio of the branching ratios, defined as

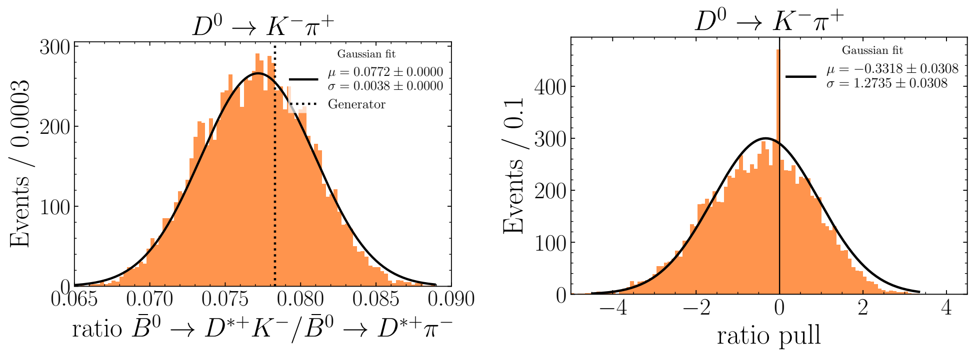
$$\mathcal{R}_{K/\pi} = \frac{\mathcal{B}(\bar{B}^0 \to D^{*+}K^-)}{\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-)}, \tag{5.23}$$

was calculated using the efficiency corrected yields. The results on MC are shown in Fig. 5.50 and Fig. 5.51.



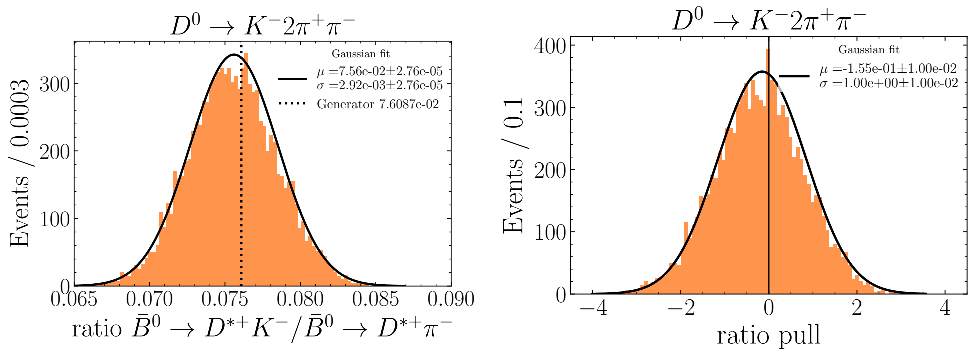
(A) (B)

FIGURE 5.49: Distributions of the branching ratio measured in 5000 toy MC experiments for $\bar{B}^0 \to D^{*+}\pi^-$(a), and $\bar{B}^0 \to D^{*+}K^-$(b), in the $D^0 \to K^-2\pi^+\pi^-$ channel.



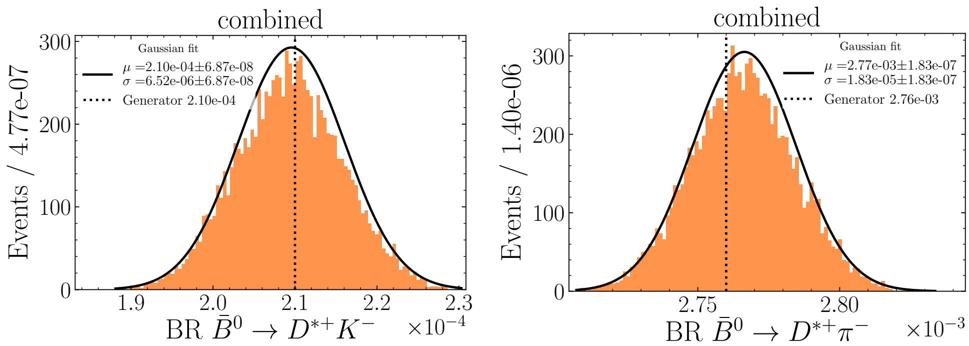
(A) (B)

FIGURE 5.50: Distributions of the ratio, $\mathcal{R}_{K/\pi} = \mathcal{B}(\bar{B}^0 \to D^{*+}K^-)/\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-)$, measured in 5000 toy MC experiments for $\bar{B}^0 \to D^{*+}\pi^-$(a) and $\bar{B}^0 \to D^{*+}K^-$(b) in the $D^0 \to K^-\pi^-$ channel.



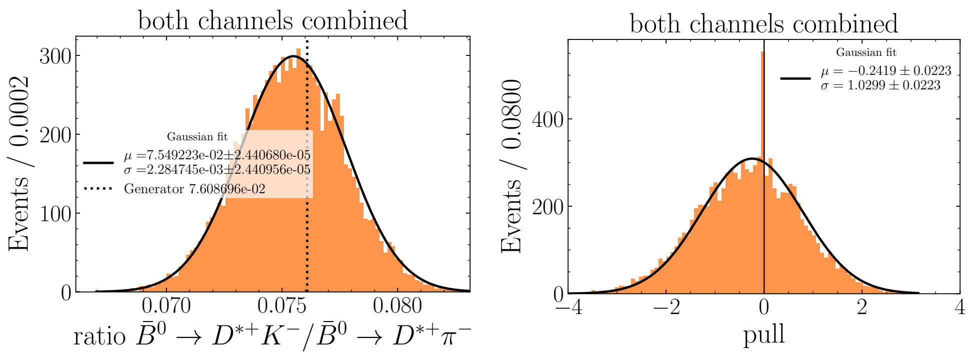
(A) (B)

FIGURE 5.51: Distributions of the ratio, $\mathcal{R}_{K/\pi} = \mathcal{B}(\bar{B}^0 \to D^{*+}K^-)/\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-)$, (a) and the pull distributions (b), measured in 5000 toy MC experiments in the $D^0 \to K^-2\pi^+\pi^-$ channel.



(A) (B)

FIGURE 5.52: Combined branching ratio measurements for $\bar{B}^0 \to D^{*+}K^-$ (a), and $\bar{B}^0 \to D^{*+}\pi^-$ (b).

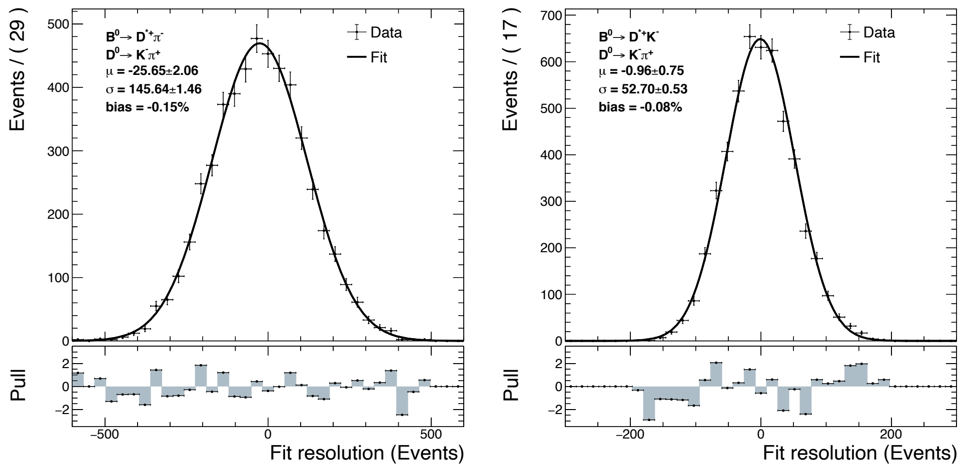


(A) (B)

FIGURE 5.53: Distributions of the ratio, $\mathcal{R}_{K/\pi} = \mathcal{B}(\bar{B}^0 \to D^{*+}K^-)/\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-)$, when both D^0-meson channels are combined (a) and the pull distributions (b).

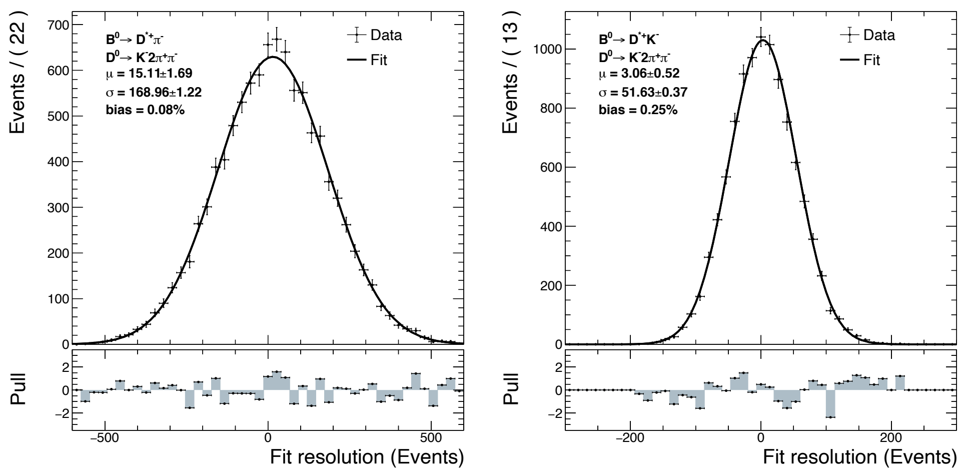
5.8.4 Fit biases

The distributions found for the resolution, defined as the distance between measured and generated signal yield, of the fits in the toy studies have shown small biases, see Fig. 5.42 and Fig. 5.43. In order to obtain more precise estimations the fits to the signal yield resolution distributions are repeated with unbinned fits, depicted in Fig. 5.54 and Fig. 5.55. The bias is calculated as deviation of the mean divided by the expected number in percent. The bias will not be corrected but used as a systematic uncertainty for the ΔE-signal yield results.



(A) (B)

FIGURE 5.54: Distributions of the signal yield resolutions found by the toy studies (a) for $\bar{B}^0 \to D^{*+}\pi^-$and (b) for $\bar{B}^0 \to D^{*+}K^-$in $D^0 \to K^-\pi^+$. The resolutions are fit with unbinned Gaussian fit to determine the bias, defined as the distance between zero and the mean of the Gaussian PDF.



(A) (B)

FIGURE 5.55: Distributions of the signal yield resolutions found by the toy studies (a) for $\bar{B}^0 \to D^{*+}\pi^-$and (b) for $\bar{B}^0 \to D^{*+}K^-$in $D^0 \to K^-2\pi^+\pi^-$. The resolutions are fit with unbinned Gaussian fit to determine the bias, defined as the distance between zero and the mean of the Gaussian PDF.

5.8.5 Systematic uncertainties related to PDF parameters

For systematic uncertainties arising from fixed background ratios in the PDFs, the fits are repeated in toy MC studies varying the fixed ratios according to the uncertainties on their branching ratios as given in the PDG. The effect is found to be almost negligible, $< 0.1\%$ on the branching ratio. Furthermore, PDFs where widths are fixed were varied by a Gaussian distribution with 15% variation around the central value found from MC. The impact was found to be minimal, i.e. $< 0.1\%$, and was fixed to 0.1%.

5.9 Systematic uncertainties of the ratio $R_{K/\pi}$

Systematic uncertainties on the ratio $R_{K/\pi} = \mathcal{B}(\bar{B}^0 \to D^{*+}K^-)/\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-)$ arise from the previously discussed uncertainties of the branching ratios. For corrections, ϵ_h, to the efficiencies and their uncertainties, σ_{ϵ_h}, this means that when calculating $R_{K/\pi}$ from the number of measured decays N,

$$R_{K/\pi} = \frac{N_{\text{meas}}(\bar{B}^0 \to D^{*+}K^-) \cdot (\epsilon_\pi \pm \sigma_{\epsilon_\pi})}{N_{\text{meas}}(\bar{B}^0 \to D^{*+}\pi^-) \cdot (\epsilon_K \pm \sigma_{\epsilon_K})}. \tag{5.24}$$

One can now assume that many sources of systematic uncertainty are correlated, as the underlying physics process i.e. the D^{*+} side of the decay is similar, such that

$$\sigma_{\epsilon_K} = \rho \cdot \sigma_{\epsilon_\pi}, \tag{5.25}$$

where $\rho \in [-1, 1]$ is the correlation coefficient. This requires the K/π-identification related uncertainties for tracks from the D^0-decay and those of the bachelor hadron to be separately treated. The assumed correlation coefficients for all sources of systematic uncertainty are depicted in Fig. 5.56 and a detailed summary of the numerical values can be found in the summary tables of Sec. 5.10.

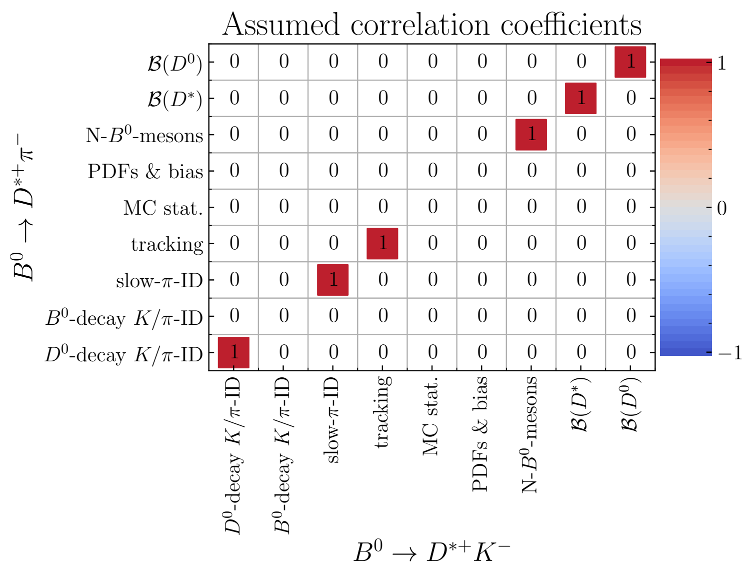


FIGURE 5.56: Assumed correlation coefficients for the sources of systematic uncertainty of $R_{K/\pi}$.

TABLE 5.18: Expected statistical and systematic uncertainties. The total error is taken from the squared sum assuming no correlations between the errors. The entries marked with † are used in the ratio, for the particle ID these are based on taking only the bachelor hadron into account. All other uncertainties cancel in the ratio.

	$\boldsymbol{D^0 \to K^-\pi^+}$	
type	$\bar{B} \to D^{*+}\pi^-$	$\bar{B} \to D^{*+}K^-$
π-ID stat.	0.78%(0.72%†)	0.54%
π-ID sys.	0.60%(0.44%†)	0.27%
K-ID stat.	0.76%	1.05%(0.72%†)
K-ID sys.	0.53%	1.15%(0.61%†)
K-ID run dep. sys.	0.30%	0.30%
π_{slow} stat.	0.79%	0.79%
π_{slow} sys.	0.01%	0.01%
π_{slow} corr.	1.33%	1.33%
3 tracks tracking sys.	1.05%	1.05%
MC stat.	0.39%†	1.4%†
fixed yields bkg. PDF	0.10%†	0.10%†
fixed shapes bkg. PDF	0.10%†	0.10%†
fit bias	0.15%†	0.15%†
N-B^0-mesons	1.84%	1.84%
$\mathcal{B}(D^{*+} \to D^0\pi^+)$	0.74%	0.74%
$\mathcal{B}(D^0 \to K^-\pi^+)$	0.78%	0.78%
total (Br.)	3.20%	3.60%
total (ratio)	1.93%	1.93%
fit stat. err.	0.84%	4.00%

5.10 Systematic uncertainty summary

Systematic uncertainties are evaluated using the methods described above. There are two cases to consider for branching ratios and ratios of branching ratios. For the branching ratios the individual contributions are listed and the squared sum is taken as the total systematic uncertainty. The D^0-channels are combined as an average taking correlations into account using a covariance based approach. For the ratio $R_{K/\pi}$ all correlations are accounted for with the same covariance based approach as discussed in Sec.5.9. As a result, all correlated quantities cancel and only a few uncertainties contribute. The results are listed in Table 5.18 and Table 5.19 and the combined results may be found in Table 5.20, including the correlation coefficients used to combine the D^0-channels. For the track dependent systematics $N_{\mathrm{tracks}}(D^0 \to K\pi)/N_{\mathrm{tracks}}(D^0 \to K3\pi) = 3/5$ is taken as the correlation coefficient, excluding slow-π, which are treated as fully correlated.

It was found that for this analysis the systematic uncertainties are the dominant source of uncertainty. This means that even with a larger data sample as may be provided by Belle II in the future the precision of this measurement may not be improved unless the systematic uncertainties are reduced. This is feasible as the uncertainties related to low momentum slow-π tracking can greatly be improved with a larger dataset.

Table 5.19: Expected statistical and systematic uncertainties. The total error is taken from the squared sum assuming no correlations between the errors.

	$D^0 \to K^-2\pi^+\pi^-$	
type	$\bar{B} \to D^{*+}\pi^-$	$\bar{B} \to D^{*+}K^-$
π-ID stat.	0.95%(0.65%†)	0.20%
π-ID sys.	0.52%(0.46%†)	0.20%
K-ID stat.	0.72%	1.03%(0.72%†)
K-ID sys.	0.57%	0.62%(0.62%†)
K-ID run dep. sys.	0.30%	0.30%
π_{slow} stat.	0.79%	0.79%
π_{slow} sys.	0.01%	0.01%
π_{slow} corr.	1.33%	1.33%
5 tracks tracking sys.	1.75%	1.75%
MC stat.	0.35%†	1.39%†
fixed yields bkg. PDF	0.10%†	0.10%†
fixed shapes bkg. PDF	0.10%†	0.10%†
fit bias	0.08%†	0.74%†
N-B^0-mesons	1.84%	1.84%
$\mathcal{B}(D^{*+} \to D^0\pi^+)$	0.74%	0.74%
$\mathcal{B}(D^0 \to K^-2\pi^+\pi^-)$	1.70%	1.70%
total (Br.)	3.81%	4.05%
total (ratio)	1.89%	1.89%
fit stat. err. (Br.)	0.78%	3.70%

Table 5.20: Expected statistical and systematic uncertainties for the combined, averaged results. The individual categories are combined, using correlation coefficients as given.

	Combined		
Type	$\bar{B} \to D^{*+}\pi^-$	$\bar{B} \to D^{*+}K^-$	Corr. coeff.
π-ID stat.	0.77%(0.58%†)	0.34%	3/5
π-ID sys.	0.50%(0.41%†)	0.21%	3/5
K-ID stat.	0.66%	0.93%(0.64%†)	3/5
K-ID sys.	0.49%	0.80%(0.55%†)	3/5
K-ID run dep. sys.	0.27%	0.27%	3/5
π_{slow} stat.	0.79%	0.79%	1
π_{slow} sys.	0.01%	0.01%	1
π_{slow} corr.	1.33%	1.33%	1
tracking sys.	1.26%	1.26%	3/5
MC stat.	0.26%†	0.99%†	0
fixed yields bkg. PDF	0.07%†	0.07%†	0
fixed shapes bkg. PDF	0.07%†	0.07%†	0
fit bias	0.09%†	0.37%†	0
N-B^0-mesons	1.60%	1.60%	1
$\mathcal{B}(D^{*+} \to D^0\pi^+)$	0.74%	0.74%	1
$\mathcal{B}(D^0)$	0.94%	0.94%	0
total (Br.)	3.25%	3.42%	
total (ratio)	1.50%	1.50%	
fit stat. err. (Br.)	0.57%	2.74%	

5.11 Cross checks

This analysis is optimised purely on MC and therefore it is important to cross check that the modelling of the shapes of all selection variables is correct. To achieve this data and MC are compared in regions outside of the signal region. The plots shown in this section are based on the full Belle data sets and five streams of generic MC weighted such that they match the expectation of a single stream of MC. The overall normalisation is floated in the signal fits and therefore small disagreements of the normalisation will have no impact.

The first control region discussed is in M_{bc} and allows to check background from charm continuum and combinatorial B^0. This allows i.e. for a check of the modelling of background where a true D^+ is combined a random hadron.

The second control region is a region outside of the ΔM selection which allows for a check of the micromodeling of background sources such as $\bar{B}^0 \to D^{*0}\pi^0$.

After that strong requirements on the kaon-ID are used to isolate signal like background and signal that passes the selection.

There is no side band that would allow for a high purity check of B-meson background. To reduce background from continuum, all cross checks are done with and without continuum suppression applied, see Appendix A for detail on the method. This improves the isolation of background from B-mesons.

5.11.1 Corrections weights

For the plots shown in the side bands, all correction weights are applied to the MC. The weights were already discussed in Sec. 5.6.1. However, an additional run dependent factor to account for total normalisation differences between MC and data, as well as a factor of 1/5 to scale fives stream down to the size of the Belle data set, were applied.

5.11.2 Data control region in M_{bc}

To evaluate the accuracy of the MC, a side band region in $M_{\mathrm{bc}} < 5.27$ GeV was analysed. This is done without selections on the D^0 mass and ΔM, with particle-ID cuts applied and without continuum suppression, Fig. 5.57 and Fig. 5.59, and with continuum suppression, Fig. 5.58 and Fig. 5.60, for the two D^0 channels respectively. Overall there is a small constant offset in $\bar{B}^0 \to D^{*+}\pi^-$, most likely related to the size of the charm component in the continuum background; see for example Fig. 5.57a and Fig. 5.57b for the ΔE and M_{bc} distributions in the $D^0 \to K^-\pi^+$ channel and Fig. 5.59a and Fig. 5.59b for the $D^0 \to K^-2\pi^+\pi^-$ channel. This hypothesis is further strengthened when looking into the D^0 mass peaks, Fig. 5.57c and Fig. 5.59c for example or the ΔM peak, Fig. 5.57e and Fig. 5.59e respectively. The distributions of the continuum classification variable is depicted in Fig. 5.57f and Fig. 5.59f and the shape agrees well between data and MC for values larger then 0.1. The most important variable for the classification, the Fox-Wolfram moment R_2, is depicted in Fig. 5.57g and Fig. 5.59g. The shape agrees between data and MC and the continuum suppression does not introduce any distortion. However, in $D^0 \to K^-2\pi^+\pi^-$ one can see a larger fraction of B-like events, which is probably due to a slightly larger amount of fake D^0. When comparing the variables without and with continuum suppression, Fig. 5.57 and Fig. 5.58 as well as Fig. 5.59 and Fig. 5.60 this becomes more obvious. For $\bar{B}^0 \to D^{*+}K^-$ the same observations can be made.

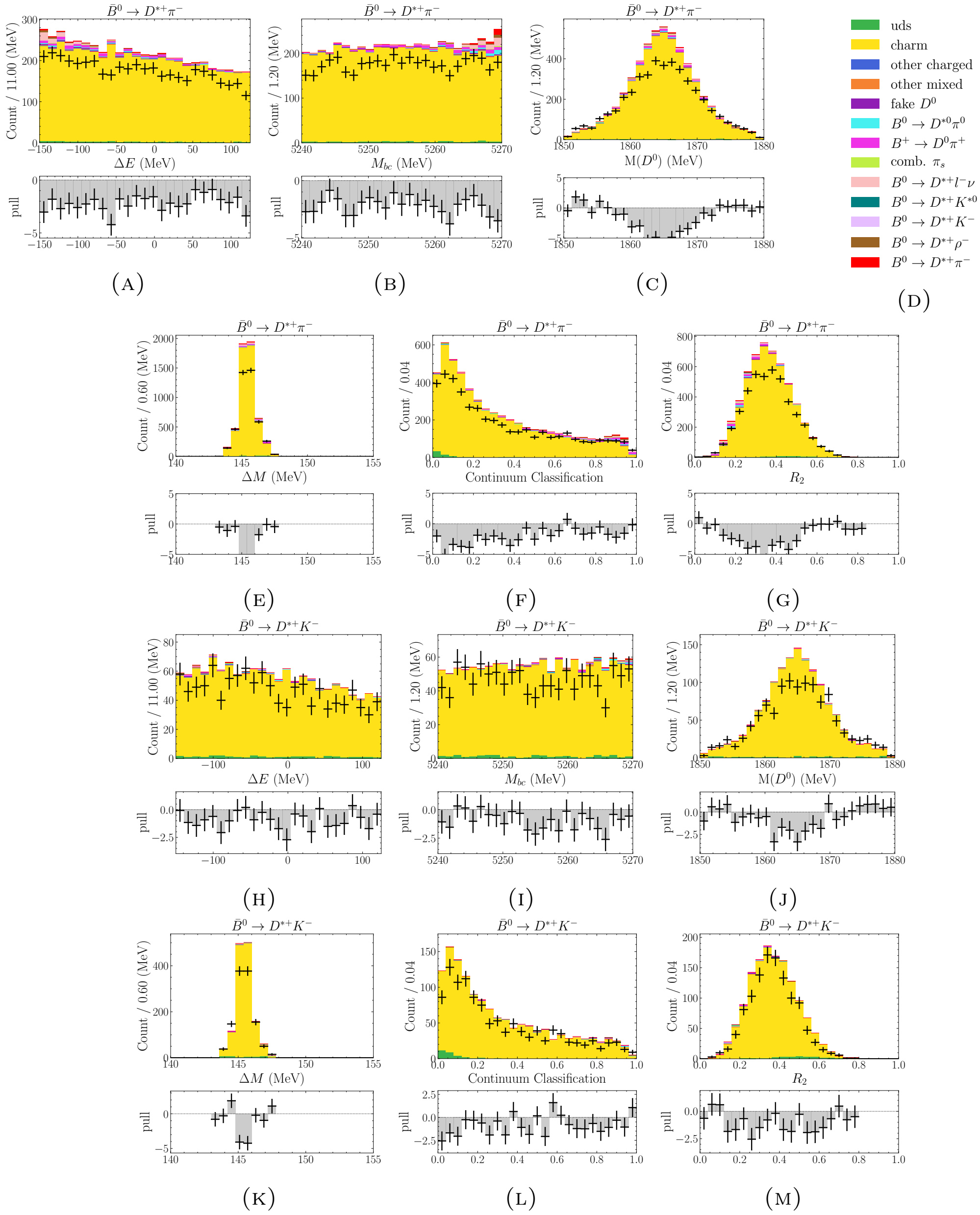


FIGURE 5.57: Variables in the side band $M_{bc} < 5270$ MeV for the $D^0 \to K^-\pi^+$ channel. All plots are made without continuum suppression applied.

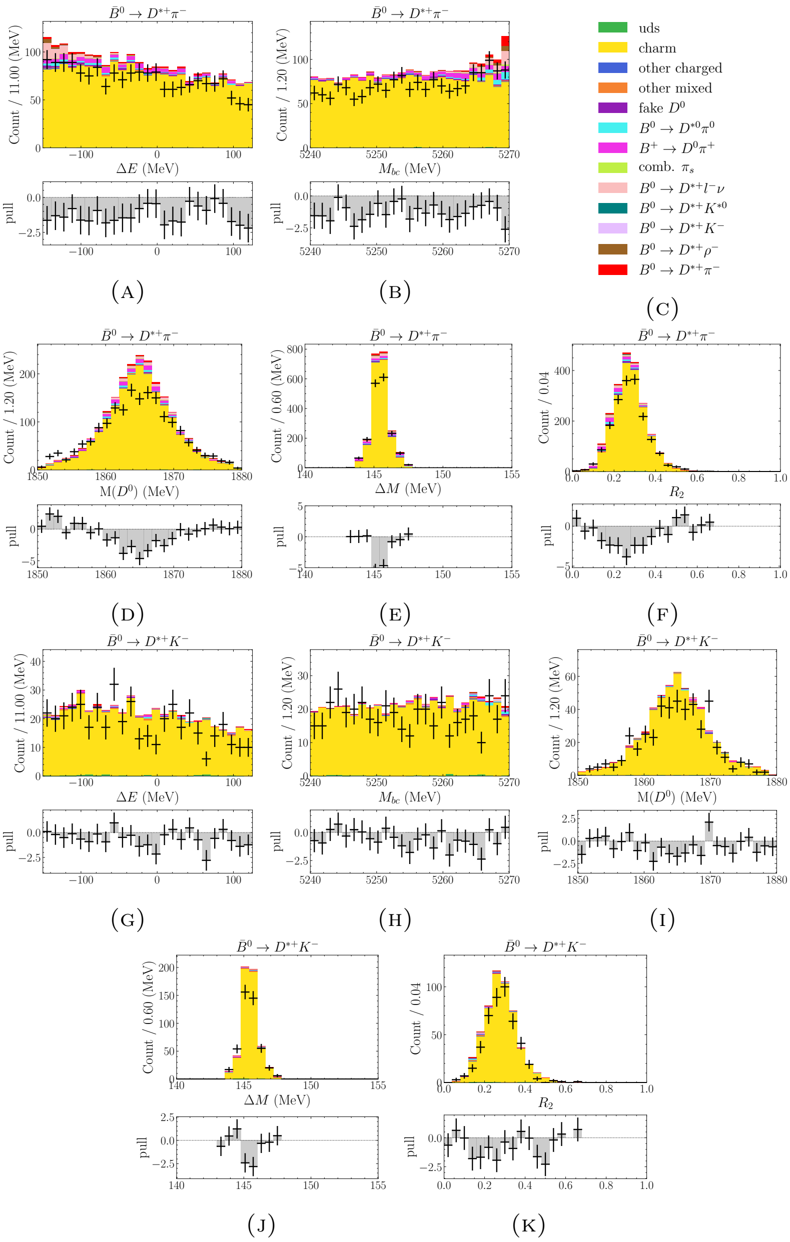


FIGURE 5.58: Variables in the side band $M_{bc} < 5270$ MeV for the $D^0 \to K^-\pi^+$ channel with continuum suppression applied. All plots are made with continuum suppression applied.

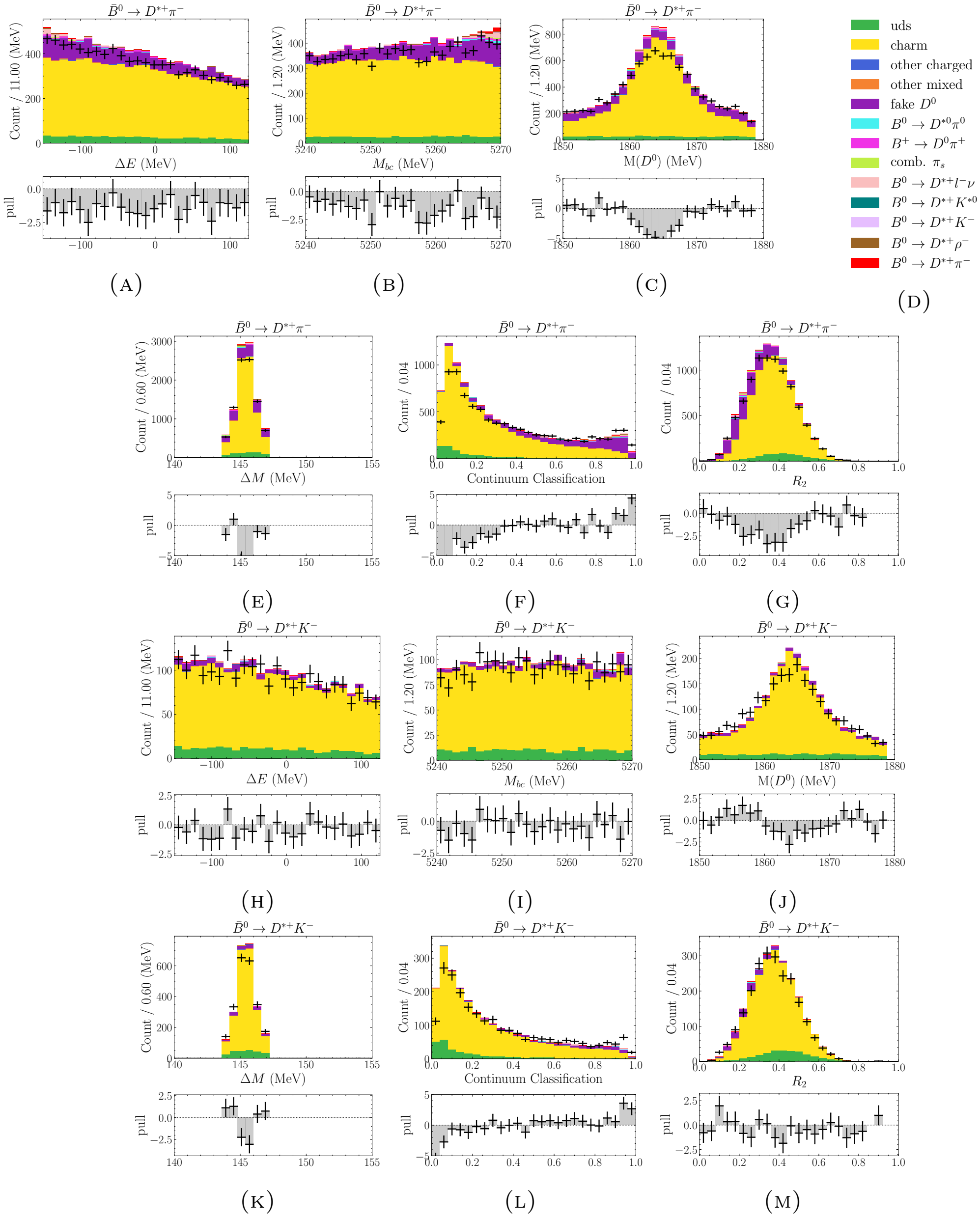


FIGURE 5.59: Variables in the side band $M_{bc} < 5270$ MeV for the $D^0 \to K^-2\pi^+\pi^-$ channel. All plots are made without continuum suppression applied.

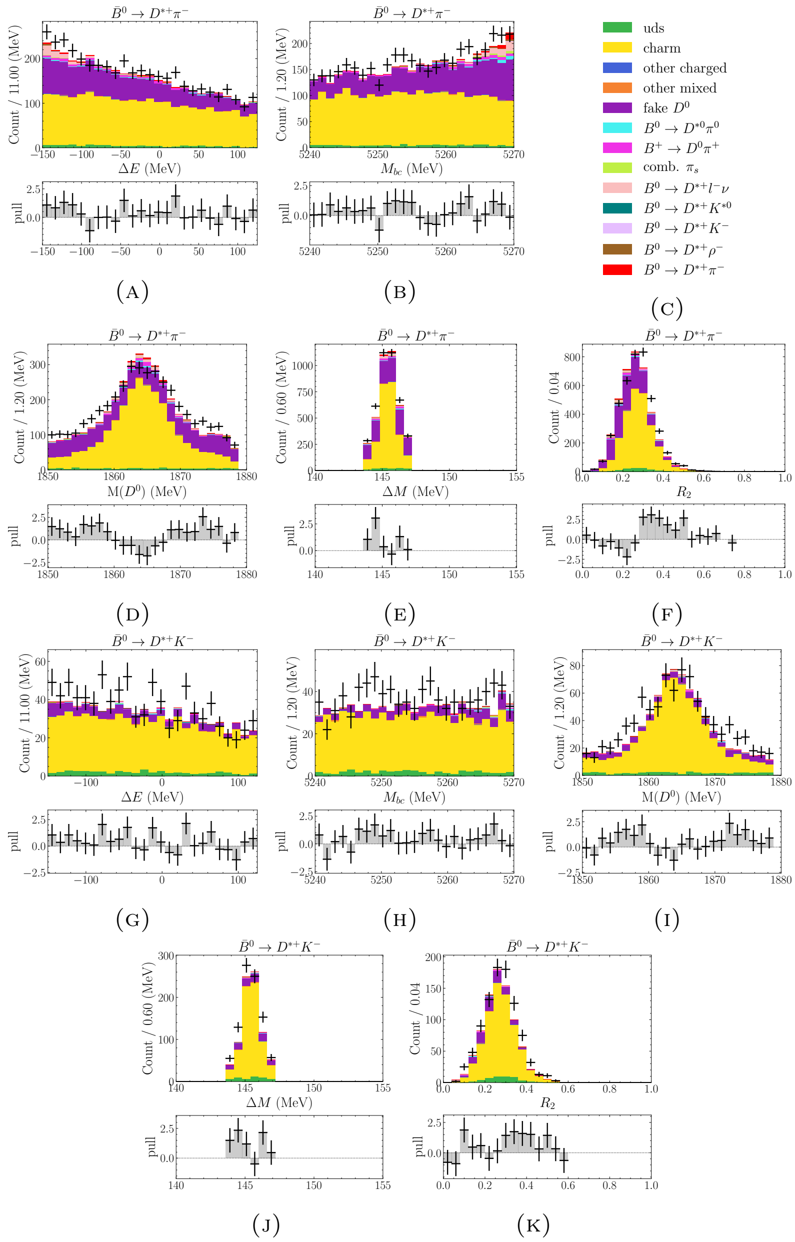


FIGURE 5.60: Variables in the side band $M_{bc} < 5270$ MeV for the $D^0 \to K^-2\pi^+\pi^-$ channel with continuum suppression applied. All plots are made with continuum suppression applied.

5.11.3 Data control region in ΔM

Another useful control region is available outside of the signal region selected in ΔM as $143\,\mathrm{MeV} < \Delta M \cup 148\,\mathrm{MeV} < \Delta M$. Other selections applied are the K/π-ID selections and D^0-mass selection. This side band allows to check data and MC agreement for background components that do not involve correctly reconstructed D^{*+} candidates. The largest non-continuum background visible in this control region is background from neutral D^{*0} for the $D^0 \to K^-\pi^+$ channel, see Fig. 5.61 without continuum suppression Fig. 5.62 with continuum suppression. In the $D^0 \to K^-2\pi^+\pi^-$ channel incorrectly reconstructed D^0 are the largest component in this control region, see Fig. 5.63 and Fig. 5.64. A small fraction of signal still passes the control region selections as seen in Fig. 5.61a for pions in $D^0 \to K^-\pi^+$. For kaons Fig. 5.61h the same happens and additionally a small peak from the pion component that passes the selection is visible. One can see this behaviour in M_{bc}, ΔM and ΔE exactly where the pion signal would be expected, hence it was concluded that this is due to pions passing the kaon selection. This conclusion is further strengthened by the fact that this happens in both D^0 modes. Taking this effect into account the agreement between data and MC is good. The same observation can be made when looking at the figures for $D^0 \to K^-2\pi^+\pi^-$, however here the relative abundance of the background is much larger such that this effect is less obvious. Overall good agreement for the $\bar{B}^0 \to D^{*+}\pi^-$ channel can be seen and a smaller constant offset for the $\bar{B}^0 \to D^{*+}K^-$ channel can be attributed to efficiency differences of the K/π-ID selection. Comparing the figures with and without continuum suppression, Fig. 5.61 with Fig. 5.62 and Fig. 5.63 with Fig. 5.64, a similar effect as seen in the M_{bc} side band becomes visible, that is that in the $D^0 \to K^-2\pi^+\pi^-$ channel the fake-D^0 background is either larger or the continuum suppression efficiency is smaller in data as compared to MC.

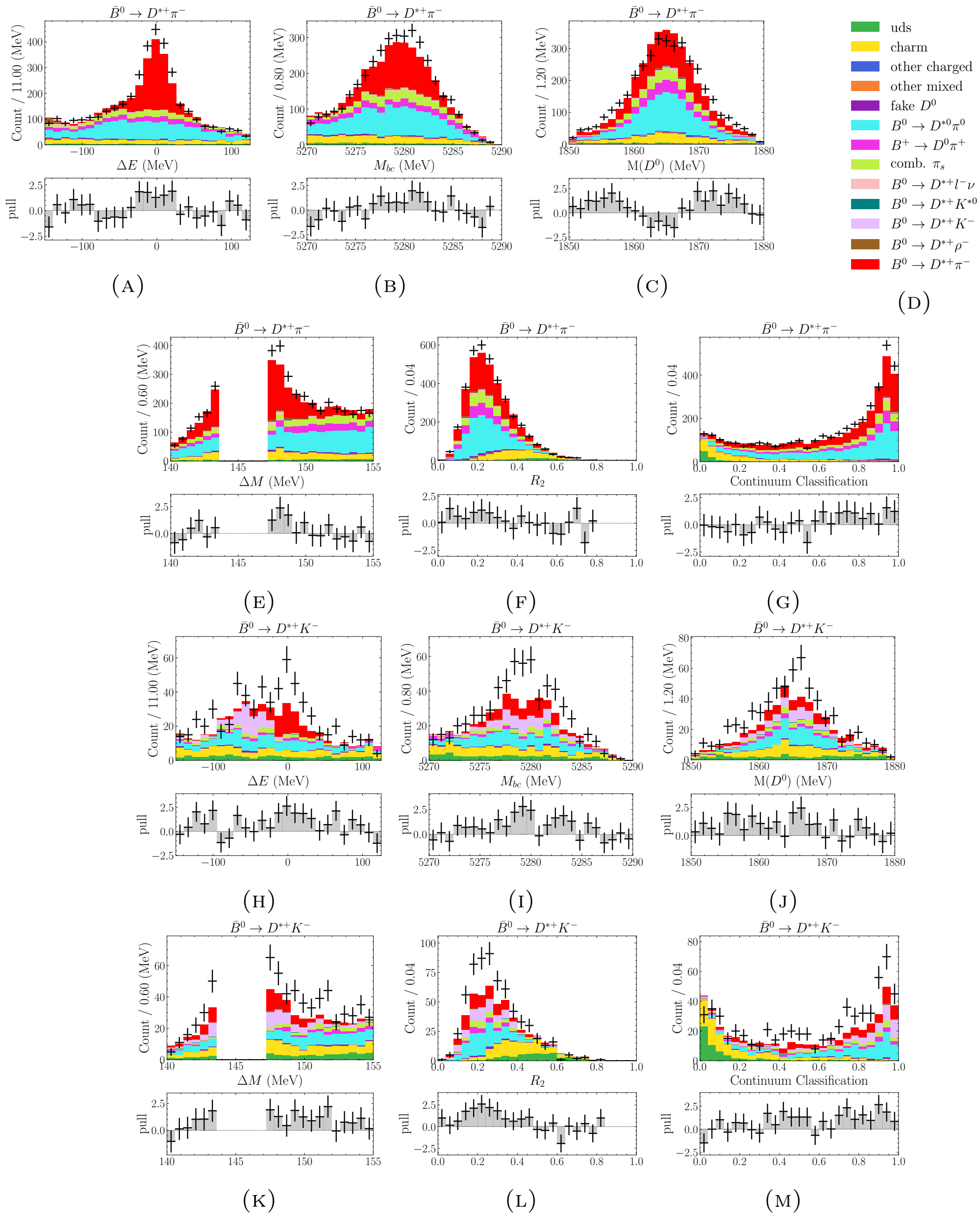


FIGURE 5.61: Variables in the control region $143\,\mathrm{MeV} < \Delta M \cup 148\,\mathrm{MeV} < \Delta M$ in the $D^0 \to K^-\pi^+$ channel. All plots are made without continuum suppression applied.

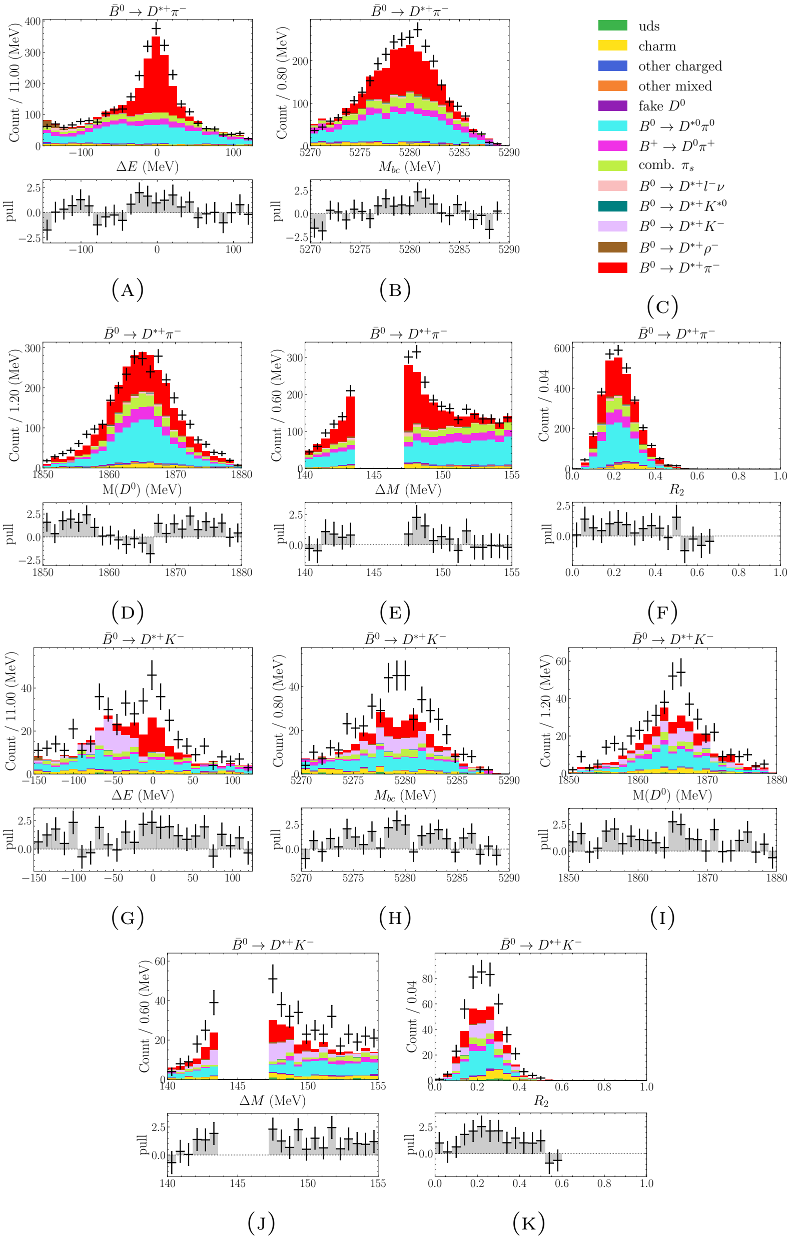


FIGURE 5.62: Variables in the control region 143 MeV $< \Delta M \cup$ 148 MeV $< \Delta M$ for the $D^0 \to K^-\pi^+$ channel with continuum suppression applied.

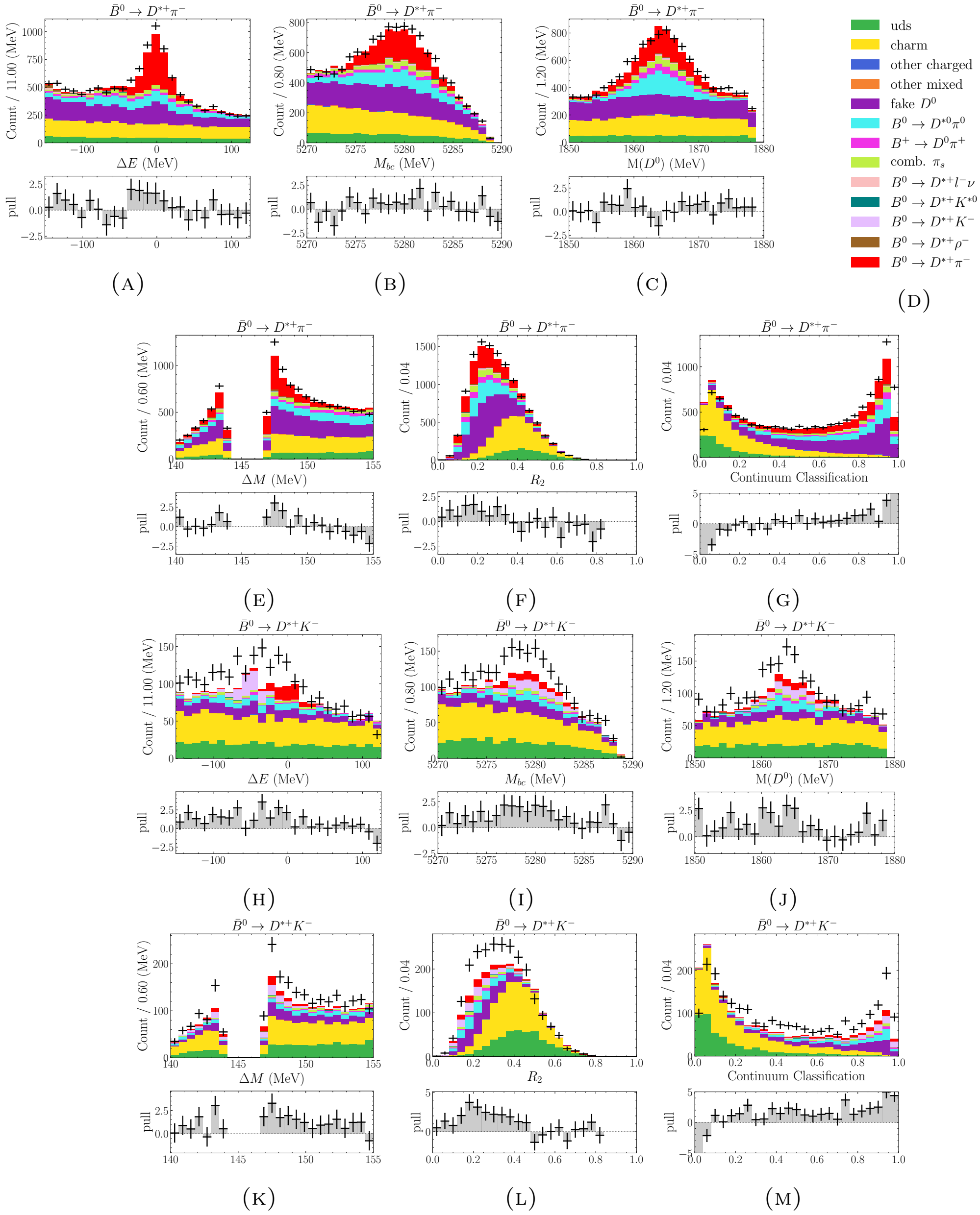


FIGURE 5.63: Variables in the control region $143\,\text{MeV} < \Delta M \cup 148\,\text{MeV} > \Delta M$ in the $D^0 \to K^-2\pi^+\pi^-$ channel. All plots are made without continuum suppression applied.

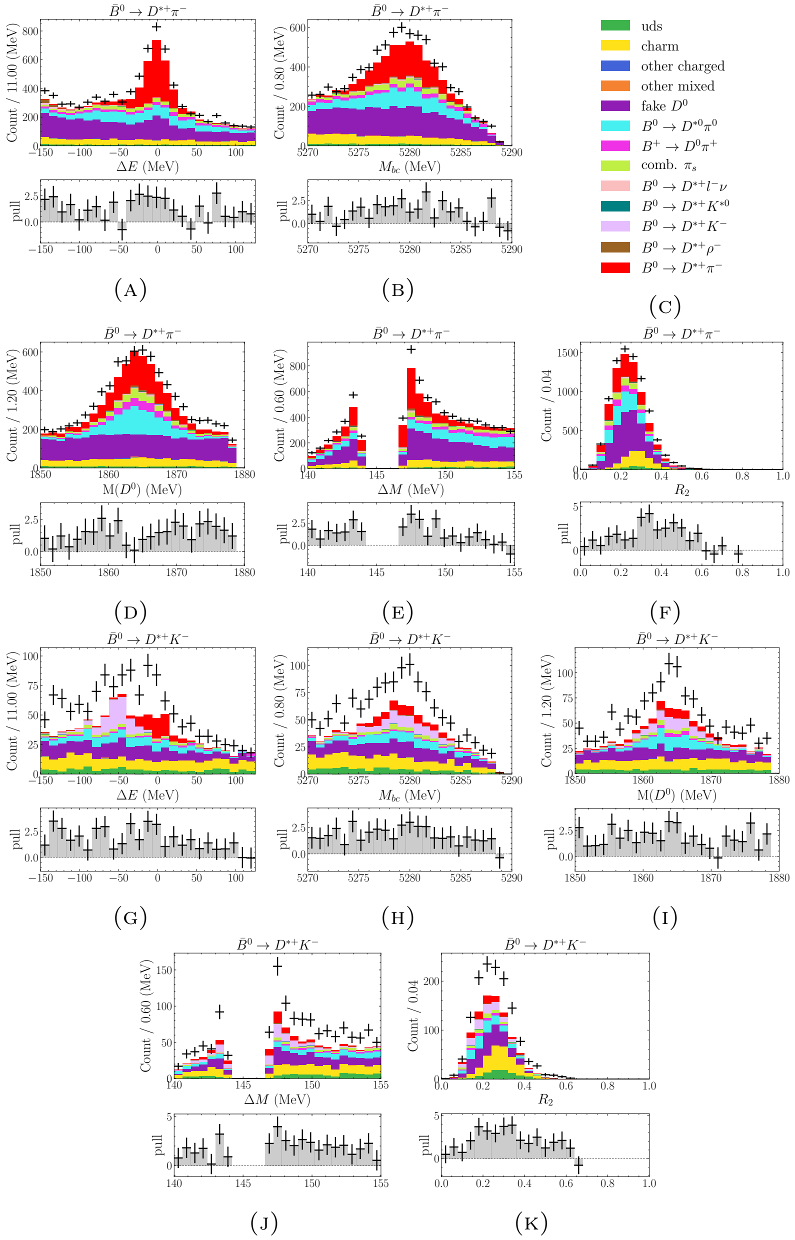
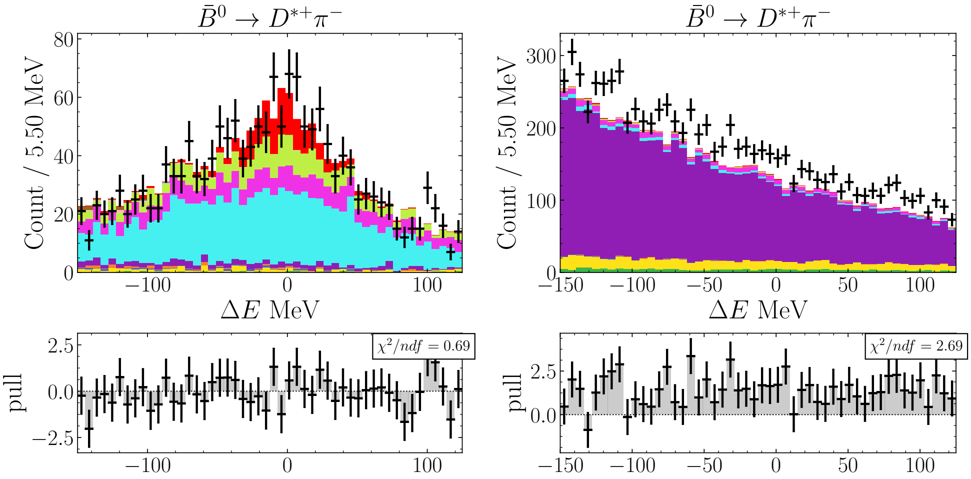


FIGURE 5.64: Variables in the control region $143\,\text{MeV} < \Delta M \cup 148\,\text{MeV} < \Delta M$ for the $D^0 \to K^-2\pi^+\pi^-$ channel with continuum suppression applied.

5.11.4 Continuum depleted charm background

Two background types need special consideration as shape differences may impact the performance of the signal extraction fit. These are D^{*0} in $D^0 \to K^-\pi^+$ in the channel and fake-D^0 in the $D^0 \to K^-2\pi^+\pi^-$ channel and both a situated between the $\bar{B}^0 \to D^{*+}\pi^-$ and $\bar{B}^0 \to D^{*+}K^-$ signal peaks in δE. Since no isolating control region is available these will be isolated in the ΔM control region with an additional strong selection on the continuum suppression to remove continuum backgrounds. For $D^0 \to K^-\pi^+$ we require a classification, clf, of $clf > 0.6$, depicted in Fig. 5.65a, and for $D^0 \to K^-2\pi^+\pi^-$, depicted in Fig. 5.65b, we require a classification $clf > 0.85$. No significant shape mismatched can be observed, apart from a constant offset in the $D^0 \to K^-2\pi^+\pi^-$ channel.

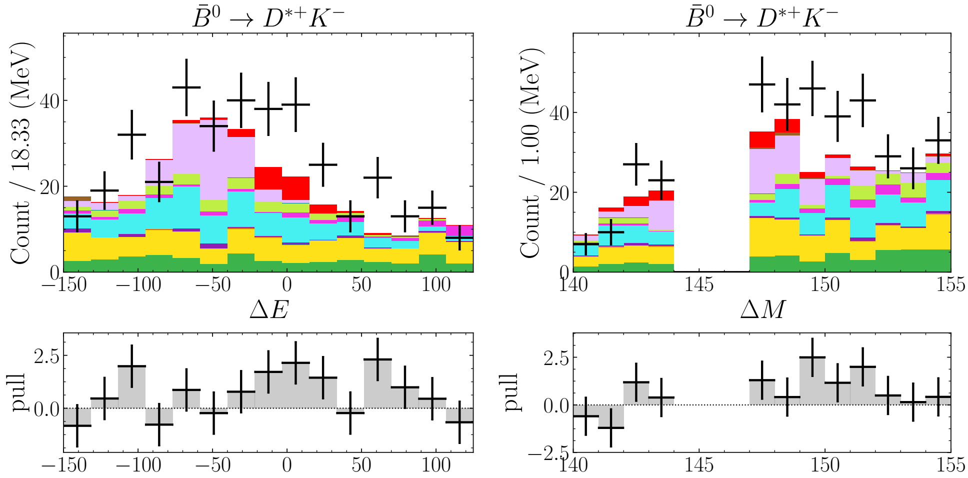


(A) (B)

FIGURE 5.65: Enriched backgrounds of D^{*0} (light blue) in the $D^0 \to K^-\pi^+$ in the channel and fake-D^0 (purple) in the $D^0 \to K^-2\pi^+\pi^-$ channel.

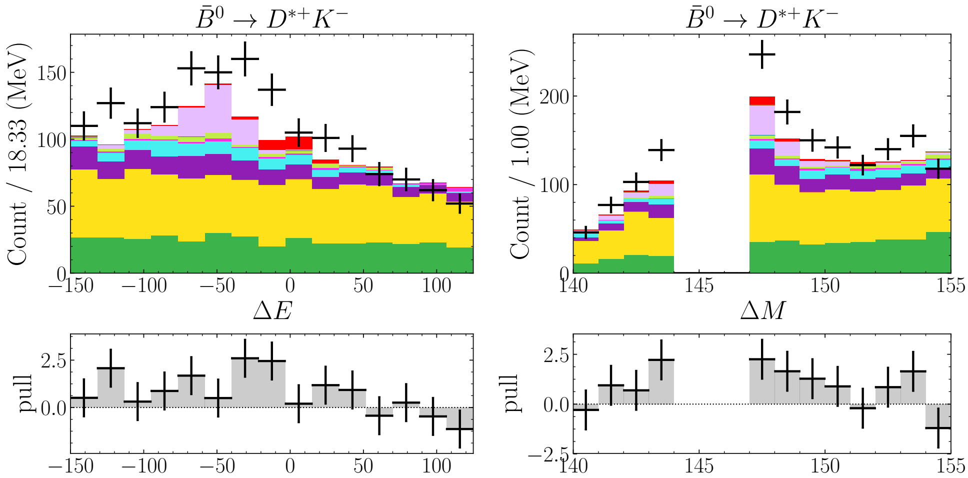
5.11.5 Peaking background at $\Delta E = 0$ in the ΔM control region

To further investigate the origin of the peaking background in $\bar{B}^0 \to D^{*+}K^-$, Fig 5.64g, the selection on the kaon ID was increased to $K/\pi > 0.9$. The resulting distributions for ΔM and ΔE are depicted in Fig. 5.66 for $D^0 \to K^-\pi^+$ and Fig. 5.67 for $D^0 \to K^-2\pi^+\pi^-$. For $D^0 \to K^-\pi^+$ the result is a clear reduction of the peak at $\Delta E = 0$, when comparing to Fig. 5.61h. The remaining structure still peaks in $\Delta E = 0$ but it is not clear if this component also peaks in the signal region of ΔM. For $D^0 \to K^-2\pi^+\pi^-$ the peak at $\Delta = 0$ is also clearly smaller when comparing to Fig. 5.63h. The remaining structure is $\bar{B}^0 \to D^{*+}K^-$, as it peaks in both ΔM and ΔE in the signal region. To further strengthen this hypothesis the K/π-ID variable of the bachelor hadron is depicted in Fig. 5.68.



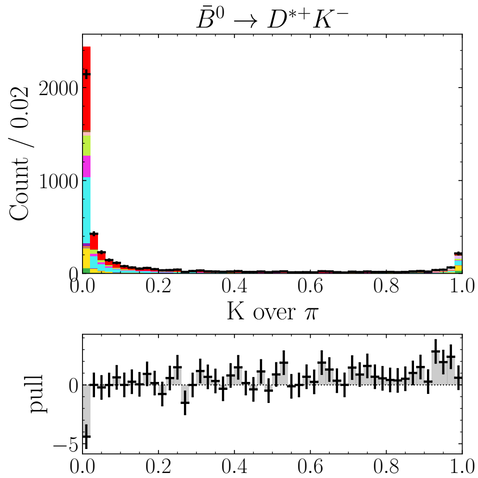
(A) (B)

FIGURE 5.66: control region in ΔM with tight selection of $K/\pi > 0.9$ in the $D^0 \to K^-\pi^+$ channel. All plots are made without continuum suppression applied.

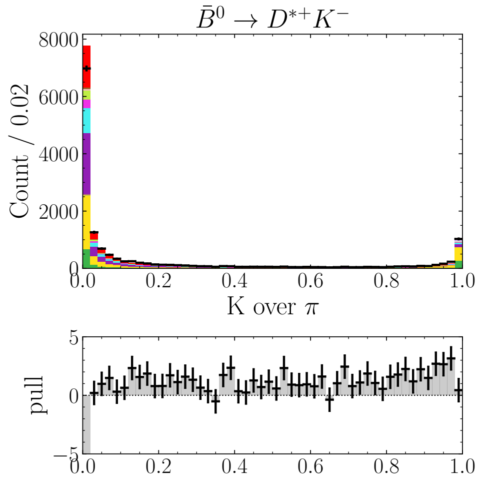


(A) (B)

FIGURE 5.67: control region in ΔM with a tight selection of $K/\pi > 0.9$ in the $D^0 \to K^-2\pi^+\pi^-$ channel. All plots are made without continuum suppression applied.



(A)

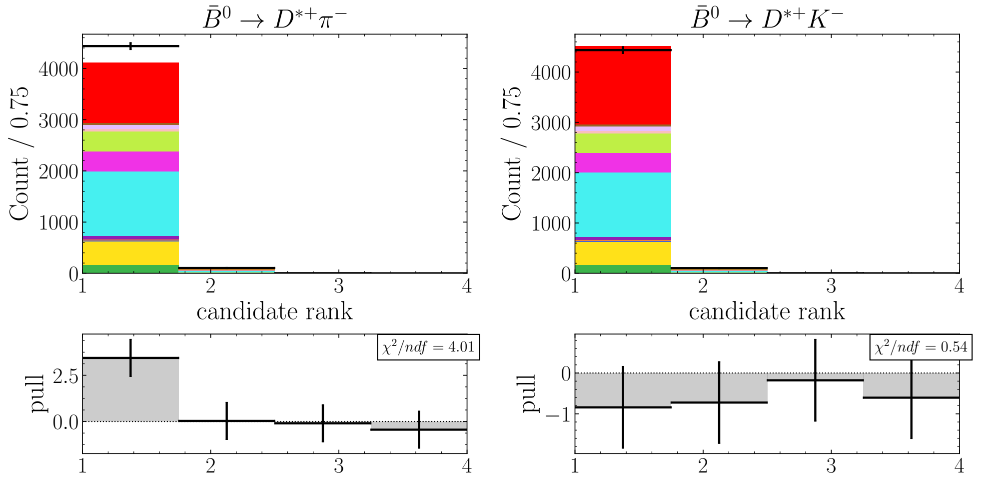


(B)

FIGURE 5.68: Distributions of the KID variable of the bachelor hadron within the ΔM control region for (a) $D^0 \to K^-\pi^+$and (b) $D^0 \to K^-2\pi^+\pi^-$.

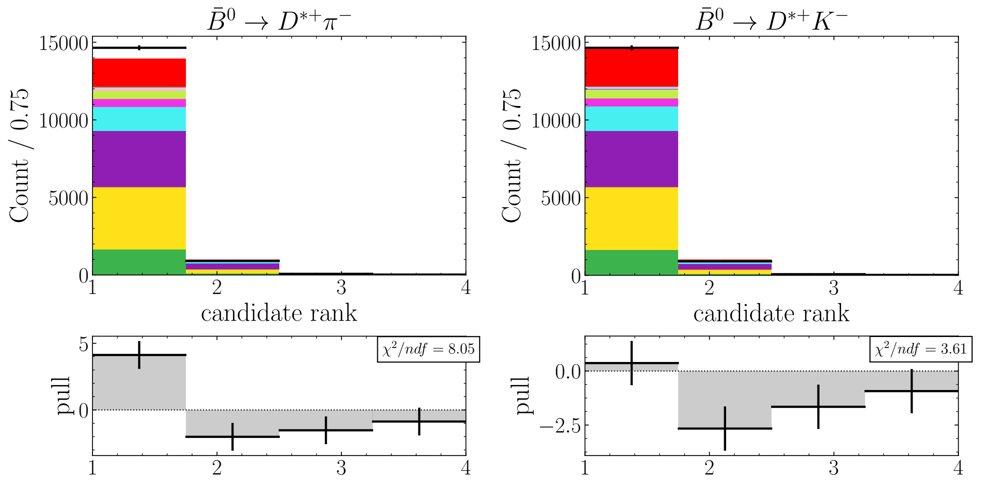
5.11.6 Candidate selection in the ΔM control region

A very important cross check is that the candidate selection works as expected on data. In order to verify this we check the distributions of the ranking variable in the ΔM control region to ensure that B-meson like backgrounds have the same relative abundance in MC and data and that the relative number of events with more then one candidate is the same. This is the case, apart from a slightly more abundant number of events with a single candidate, see Fig. 5.69 and Fig. 5.70. This is probably due to a larger number of signal events passing the ΔM selection due to the peak being slightly wider in data, and is hence not a problem.



(A) (B)

FIGURE 5.69: Variable ranking of (a) $\bar{B}^0 \to D^{*+}\pi^-$and (b) $\bar{B}^0 \to D^{*+}K^-$in the ΔM control region for the $D^0 \to K^-\pi^+$ in the channel.



(A) (B)

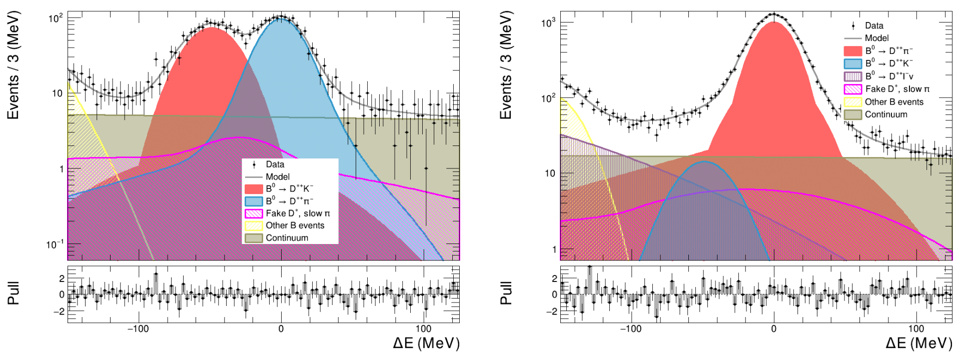
FIGURE 5.70: Variable ranking of (a) $\bar{B}^0 \to D^{*+}\pi^-$and (b) $\bar{B}^0 \to D^{*+}K^-$in the ΔM control region for the $D^0 \to K^-2\pi^+\pi^-$ in the channel.

5.12 Results

The final fit results to data are depicted in Fig. 5.71 for $D^0 \to K^-\pi^+$ and Fig. 5.72 $D^0 \to K^-2\pi^+\pi^-$. The fits work well with no pulls visible. The branching ratios a re calculated as given in Eq. 5.17.

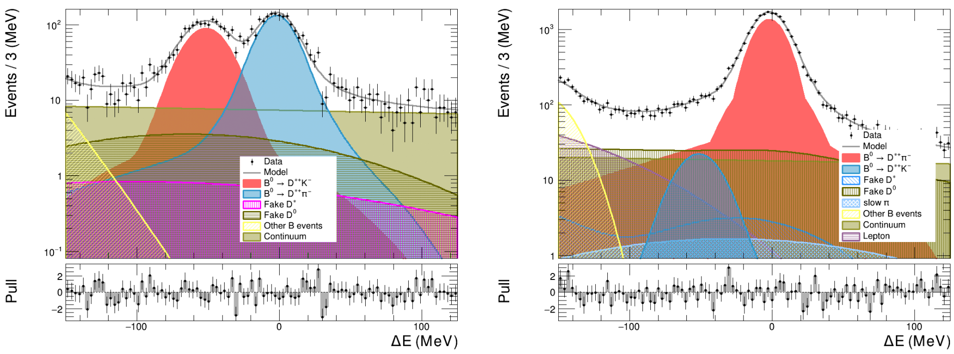
The systematic uncertainties are evaluated in Sec. 5.10. The results and a comparison to theoretical prediction and prior measurements is displayed in Fig. 5.73 for the branching ratios and Fig. 5.74 for their ratio. The numerical values are listed in Table 5.21 for the branching ratios and Table 5.22 for the ratio.

For $\mathcal{B}(\bar{B}^0 \to D^{*+}K^-)$ a very similar central value as in a previous Belle measurement on a 10.6 fb^{-1} dataset in 2001 was obtained. An improvement of a factor of 3.5 was achieved, due to a reduction of the statistical uncertainty from 20% $\to$ 2.74% and a reduction of the systematic uncertainties 5% $\to$ 3.42%. The value is compared to two theory models, taking uncertainties from experiment and theory into account the distance is 1.1σ, for Huber et al. NNLO and 2.7σ for the Bordone et al. [19] NNLO prediction. The statistical uncertainty is 0.57% and the systematic uncertainty 3.25%. The same evaluation is made for $\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-)$. The distance to the theoretical value is 1.8σ for Huber et al. [15] at NNLO. For the ratio $\mathcal{R}_{K/\pi} = \mathcal{B}(\bar{B}^0 \to D^{*+}K^-)/\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-)$ a discrepancy of 2.7σ to the theory prediction is found. The total experimental uncertainty was reduced to 3.2% from 5.7% (BaBar) and 5.5% (LHCb). The BaBar measurement is performed without separating bachelor kaons from pions in a single fit in ΔE, therefore a larger uncertainty is expected.



(A) (B)

FIGURE 5.71: Final fit results to data of $\bar{B}^0 \to D^{*+}K^-$ (a) and $\bar{B}^0 \to D^{*+}\pi^-$ (b) in the $D^0 \to K^-\pi^+$ channel.



(A) (B)

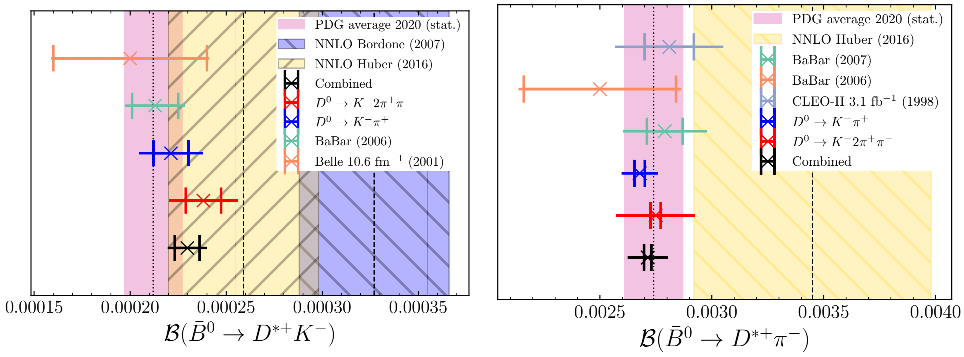
FIGURE 5.72: Final fit results to data of $\bar{B}^0 \to D^{*+}K^-$ (a) and $\bar{B}^0 \to D^{*+}2\pi^-\pi^+$ (b) in the $D^0 \to K^-\pi^+$ channel.

TABLE 5.21: Results of the branching ratio measurement. The first uncertainty is statistical, the second is systematic. For the discrepancy, the first value is with respect to Huber et al. [15] values of $\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-) = (3.45 \pm 0.53) \times 10^{-3}$ and $\mathcal{B}(\bar{B}^0 \to D^{*+}K^-) = (2.59 \pm 0.39) \times 10^{-4}$, the second value is with respect to the Bordone et al. [19] prediction of $\mathcal{B}(\bar{B}^0 \to D^{*+}K^-) = (3.27 \pm 0.39) \times 10^{-4}$. Experimental and theoretical uncertainties are accounted for in the calculation of the discrepancy.

$D^0 \to K^-\pi^+$	Result	Discrepancy
$\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-)$	$(2.604 \pm 0.023 \pm 0.076) \times 10^{-3}$	1.8σ
$\mathcal{B}(\bar{B}^0 \to D^{*+}K^-)$	$(2.152 \pm 0.088 \pm 0.078) \times 10^{-4}$	$1.1\sigma(2.7\sigma)$
$D^0 \to K^-2\pi^+\pi^-$		
$\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-)$	$(2.749 \pm 0.023 \pm 0.011) \times 10^{-3}$	1.4σ
$\mathcal{B}(\bar{B}^0 \to D^{*+}K^-)$	$(2.381 \pm 0.092 \pm 0.097) \times 10^{-4}$	$0.5\sigma(2.2\sigma)$
Combined		
$\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-)$	$(2.677 \pm 0.016 \pm 0.087) \times 10^{-3}$	1.6σ
$\mathcal{B}(\bar{B}^0 \to D^{*+}K^-)$	$(2.267 \pm 0.064 \pm 0.078) \times 10^{-4}$	$1.8\sigma(2.5\sigma)$

TABLE 5.22: Results of the ratio $\mathcal{R}_{K/\pi} = \mathcal{B}(\bar{B}^0 \to D^{*+}K^-)/\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-)$ measurement. The first uncertainty is statistical the second is systematic. Note that Huber et al. [15] predict the exact same value and uncertainty of $\mathcal{R}_{K/\pi} = (0.75 \pm 0.02) \times 10^{-1}$ at NLO and NNLO as Beneke et al. [3] at NLO. Note that the distance measured in standard deviation grows when combining the values, this is due to correlations in the uncertainties.

Channel	Result	Discrepancy
$D^0 \to K^-\pi^+$	$\mathcal{R}_{K/\pi} = (8.254 \pm 0.350 \pm 0.147) \times 10^{-2}$	1.8σ
$D^0 \to K^-2\pi^+\pi^-$	$\mathcal{R}_{K/\pi} = (8.527 \pm 0.336 \pm 0.150) \times 10^{-2}$	2.5σ
Combined	$\mathcal{R}_{K/\pi} = (8.390 \pm 0.243 \pm 0.115) \times 10^{-2}$	2.7σ



(A) (B)

FIGURE 5.73: Final results of the branching ratio measurements. The inner uncertainty is the statistical uncertainty and the outer is the combined uncertainty $\sigma_{\text{total}} = \sqrt{\sigma^2_{\text{stat.}} + \sigma^2_{\text{sys.}}}$.

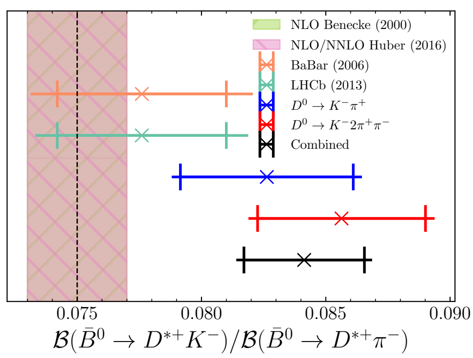


FIGURE 5.74: Final results of the branching ratio measurements. The inner uncertainty is the statistical uncertainty and the outer is the combined uncertainty $\sigma_{\text{total}} = \sqrt{\sigma^2_{\text{stat.}} + \sigma^2_{\text{sys.}}}$.

Chapter 6

Phenomenological interpretation

In this chapter the phenomenological implications of the measured values of $\mathcal{B}(\bar{B}^0 \to D^{*+}K^-)$ and $\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-)$ will be discussed. External inputs from momentum transfer dependent measurements of $d\Gamma(\bar{B}_d \to D^{*+}l^-\bar{\nu})/dq^2$ are used to cancel form factors and measure the model parameter, a_1, which encodes fundamental assumptions made for the prediction of branching ratios of hadronic B-decays. QCD factorisation predicts $a_1 = 1$ in the most naive picture. When taking higher order corrections into account the prediction expects a quasi universal value of $a_1 = 1.05$, independent of the decay topology, as discussed in Sec. 2.2.6. Other groups find discrepancy between these values and the available measurement of the order of 5% [11], see Sec. 2.2 for a detailed discussion of the theoretical background. This work is performed in close collaboration with the authors of Ref. [58], who kindly provided the the semi-leptonic inputs.

6.1 Measuring a_1

Table 6.1: Input parameters for the a_1 calculation. Values not listed here are taken from Ref. [58].

Description	Parameter	Value
Lifetimes	τ_{B_d}	(1.540 ± 0.004) ps
CKM matrix element	$\lvert V_{ud}\rvert$	0.97420 ± 0.00021
	$\lvert V_{us}\rvert$	0.2243 ± 0.0005
Decay constants	f_K	(0.1302 ± 0.0014) GeV
	f_ρ	(0.216 ± 0.006) GeV
	f_{K^*}	(0.211 ± 0.007) GeV
	f_{a^+}	(0.238 ± 0.01) GeV
	f_D	(0.2119 ± 0.0011) GeV
	X_h	1 ± 0.0007
Branching ratios	$\mathcal{B}(\bar{B} \to D^{*+}\rho^-)$	$(6.8 \pm 0.9) \times 10^{-3}$
	$\mathcal{B}(\bar{B} \to D^{*+}K^{*-})$	$(3.3 \pm 0.6) \times 10^{-4}$
	$\mathcal{B}(\bar{B} \to D^{*+}a^-)$	$(1.30 \pm 0.27) \times 10^{-2}$
Invariant masses	$M(\pi^+)$	$(0.13957061 \pm 0.00000024)$ GeV
	$M(K^+)$	(0.493677 ± 0.000016) GeV
	$M(\rho^+)$	(0.77526 ± 0.00025) GeV
	$M(K^{*+})$	(0.89176 ± 0.00025) GeV
	$M(a^+)$	(1.230 ± 0.040) GeV

The parameter a_1 is a complex number and hence only the absolute value can be measured. Starting from Eq. 2.6 one finds

$$|a_1(D^*h)| = \left[\left(6\pi^2\tau_B|V_{uq}|f_h^2X_h\right)^{-1}\frac{\Gamma(\bar{B}_d \to D^{*+}h^-)}{d\Gamma(\bar{B}_d \to D^{*+}l^-\bar{\nu})/dq^2|_{q^2=m_h}}\right]^{\frac{1}{2}}. \tag{6.1}$$

Once a value for $d\Gamma(\bar{B}_d \to D^{*+}l^-\bar{\nu})/dq^2$ is found, the prediction is straightforward. The inputs for the parameters not measured in this thesis are taken from the PDG, as listed in Table 6.1. The parameter $X_h = 1$ for vector mesons and $X_h = 1 \pm m_h^2/m_B^2$ for non-vector mesons is used. The values for $d\Gamma(\bar{B}_d \to D^{*+}l^-\bar{\nu})/dq^2$ are directly extracted from a Belle measurement [59] using the $D^{*+} \to D^0\pi^+$ and $D^0 \to K^-\pi^+$channel.

The full differential decay rate of semi-leptonic decays in the massless lepton limit is given by [17]

$$\begin{aligned}
&\frac{d\Gamma(B^0 \to D^{*-}\ell^+\nu_\ell)}{dwd\cos\theta_\ell d\cos\theta_\mathrm{v} d\chi} = \\
&\frac{\eta_\mathrm{EW}^2 3m_B m_{D^*}^2}{4(4\pi)^2}G_F^2|V_{cb}|^2\sqrt{w^2-1}(1-2wr+r^2) \\
&\{(1-\cos\theta_\ell)^2\sin^2\theta_\mathrm{v}H_+^2(w) + (1+\cos\theta_\ell)^2\sin^2\theta_\mathrm{v}H_-^2(w) \\
&+4\sin^2\theta_\ell\cos^2\theta_\mathrm{v}H_0^2 - 2\sin^2\theta_\ell\sin^2\theta_\mathrm{v}\cos 2\chi H_+(w)H_-(w) \\
&-4\sin\theta_\ell(1-\cos\theta_\ell)\sin\theta_\mathrm{v}\cos\theta_\mathrm{v}\cos\chi H_+(w)H_0(w) \\
&+4\sin\theta_\ell(1+\cos\theta_\ell)\sin\theta_\mathrm{v}\cos\theta_\mathrm{v}\cos\chi H_-(w) - H_0(w)\},
\end{aligned} \tag{6.2}$$

where

$$w = \frac{P_B \cdot P_{D^*}}{m_B m_{D^*}} = \frac{m_B^2 + m_{D^*}^2 - q^2}{2m_B m_{D^*}}, \tag{6.3}$$

is the hadronic recoil, and θ_V, θ_l are helicity angles and χ an acoplanar angle as depicted in Fig. 6.1, G_F is the Fermi Constant and $r = m_{D^*}/m_{B^0}$. For the electro-weak

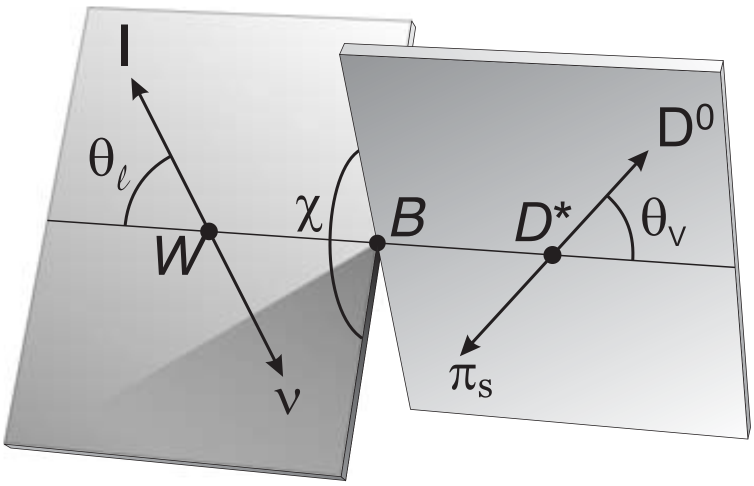


FIGURE 6.1: Definition of the decay planes of $\bar{B}_d \to D^{*+}l^-\bar{\nu}$ decays [59].

corrections the value $\eta_\mathrm{EW} = 1.0066$ [60] is used. The helicity amplitudes $H(\omega)_{\pm,0}$ depend on a parametrisation choice. There are two different parametrisations commonly used in the filed, Caprini-Lellouch-Nauberg (CLN) [17] and Boyd-Grinstein and Lebed (BGL) [61].

In CLN the helicity amplitudes are parametrised as

$$H_i(w) = m_B \frac{R^*(1-r^2)(w+1)}{2\sqrt{1-2wr+r^2}} h_{A_1}(w) \left| \tilde{H}_i(w) \right|, \tag{6.4}$$

with

$$\tilde{H}_\pm(w) = \frac{\sqrt{1-2wr+r^2}(1 \mp \sqrt{w-1/w+1}R_1(w)))}{2\sqrt{1-2wr+r^2}}, \tag{6.5}$$

$$\tilde{H}_0(w) = 1 + \frac{(w-1)(1-R_2(w))}{(1-r)}, \tag{6.6}$$

and

$$R^* = \frac{2\sqrt{m_B m_{D^*}}}{m_B + m_{D^*}}. \tag{6.7}$$

Furthermore, the following relations are needed

$$\begin{aligned} h_{A_1}(w) &= h_{A_1}(1)\left[1 - 8\rho^2 z + (53\rho^2 - 15)z^2 \right. \\ &\quad \left. -(231\rho^2 - 91)z^3\right], \\ R_1(w) &= R_1(1) - 0.12(w-1) + 0.05(w-1)^2, \\ R_2(w) &= R_2(1) - 0.11(w-1) - 0.06(w-1)^2, \end{aligned} \tag{6.8}$$

where $z = (\sqrt{w+1} - \sqrt{2})/(\sqrt{w+1} + \sqrt{2})$. There are four free parameters ρ^2, $R_1(1)$ and $R_2(1)$ and the normalisation, $h_{A_1}(1)$. These are found by a fit to data.

In BGL the parametrisation is chosen as

$$\begin{aligned} H_0(w) &= \mathcal{F}_1(w)/\sqrt{q^2}\ , \\ H_\pm(w) &= f(w) \mp m_B m_{D^*}\sqrt{w^2-1}g(w)\ . \end{aligned} \tag{6.9}$$

The functions $f(w)$, $\mathcal{F}_1(w)$ and $g(w)$ can be written as power series in z, such that

$$\begin{aligned} f(z) &= \frac{1}{P_{1+}(z)\phi_f(z)} \sum_{n=0}^{i} a_n^f z^n\ , \\ \mathcal{F}_1(z) &= \frac{1}{P_{1+}(z)\phi_{\mathcal{F}_1}(z)} \sum_{n=0}^{j} a_n^{\mathcal{F}_1} z^n\ , \\ g(z) &= \frac{1}{P_{1-}(z)\phi_g(z)} \sum_{n=0}^{k} a_n^g z^n\ , \end{aligned} \tag{6.10}$$

where P_i are Blaschke Factors and ϕ_i are functions to account for B_c resonances. The bounds i, j, k describe a cutoff that is done to reduce the dimensionality of the fit. In the following the notation BGL(i,j,k) will be used to indicate the number of terms that are used. Where indicated as lattice QCD (LQCD) the values $F(1) = 0.906 \pm 0.013$ [62] is taken.

6.1.1 Testing $SU(3)$ symmetry

One can also test whether a_1 is a universal factor, independent of the quarks contained in the hadronic branching ratio of the enumerator. A value of $a_1(K)/a(\pi) = 1$ would mean that $SU(3)$-symmetry holds, as suggested in Ref. [3]. The test is done by measuring the ratios of a_1 for different particles species, i.e. for K and π:

$$\frac{a_1^2(K^-)}{a_1^2(\pi^-)} = \frac{\Lambda_K \Gamma(\bar{B}_d \to D^{*+}K^-)/d\Gamma(\bar{B}_d \to D^{*+}l^-\bar{\nu})/dq^2|_{q^2=m_K}}{\Lambda_\pi \Gamma(\bar{B}_d \to D^{*+}\pi^-)/d\Gamma(\bar{B}_d \to D^{*+}l^-\bar{\nu})/dq^2|_{q^2=m_\pi}}, \tag{6.11}$$

where

$$\Lambda_h = \left(|V_{uq}| f_h^2 X_h\right)^{-1}. \tag{6.12}$$

6.1.2 Systematic uncertainties

Systematic uncertainties of a_1 For $a_1(\pi^-)$ and $a_1(K^-)$ all measurements are done at Belle and it can be assumed that most of the systematic uncertainties of $\mathcal{B}(\bar{B}^0 \to D^{*+}h^-)$ and $\mathcal{B}(\bar{B}^0 \to D^{*+}l^-\bar{\nu})|_{q^2=m(h)}$ are fully correlated and hence cancel in the ratio. This enables us to cancel systematic uncertainties arising from the D^{*+} reconstruction and experiment related quantities, such as tracking and B-meson counting, as explained in Sec. 5.9 for $R_{K/\pi}$. The numerical values for the semi-leptonic systematic uncertainty are given in Ref.[59] and the method combining them is described Ref.[58]. The assumed correlation coefficients between these systematic uncertainties and the uncertainties on the hadronic branching ratio, Table 5.18, in are shown in Fig. 6.2. For all other calculations where the enumerator branching ratio is not

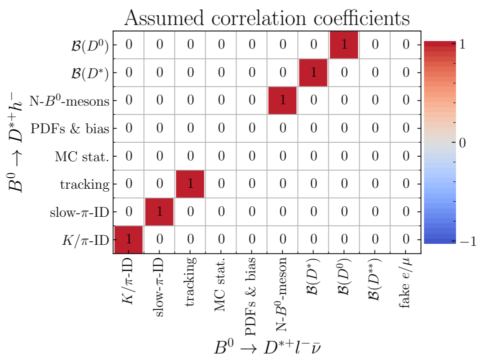


FIGURE 6.2: Assumed correlation coefficients between the two decay channels.

measured at Belle the full systematic uncertainty without these correlations is used.

Systematic uncertainties of the ratio $a_1(K)/a_1(\pi)$ For the ratio $a_1(K)/a_1(\pi)$ the systematic uncertainties are cancelled for correlated parameters as in 5.9. Additionally correlations for the semi-leptonic decay rate at the points $q^2 = m_\pi$ and $q^2 = m_K$ are found from the toy MC distributions done for Ref. [58]. The found values are depicted in Fig. 6.3.

6.1.3 Results

The semi-leptonic differential decay rates is depicted for selected parametrisations in Fig. 6.4[1]. The corresponding results for a_1 are depicted in Fig. 6.5 and the numerical

[1]Kindly provided by Daniel Ferlewicz, using data measured by Eiasha Waheed.

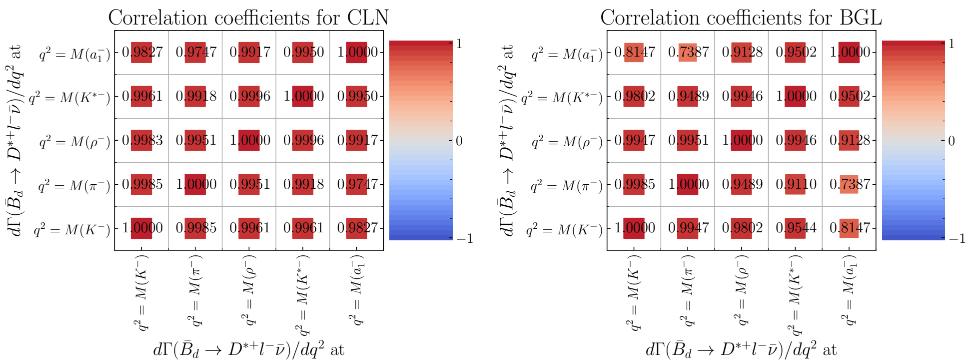


FIGURE 6.3: Correlation coefficients of $d\Gamma(\bar{B}_d \to D^{*+}l^-\bar{\nu})/dq^2$ for different points of q^2.

values can be found in Table 6.2. For $a_1(\rho^-)$, $a_1(K^{*-})$ and $a_1(a_1^-)$ the hadronic branching ratios were taken from the PDG. Thus, there may be discrepancies in the parameters used to calculate the branching ratios of the numerator, which may cause a deviation from unity. These could be different due to values of $\mathcal{B}(D^{*+} \to D^0\pi^+)$, $\mathcal{B}(B^0 \to K^-\pi^+)$, the decay constants f_h, or $|V_{uq}|$, so we will discuss in detail only the values found using the branching ratios measured in this thesis in the following. The deviation is calculated as the distance between the central values of measurement and theory

$$d = x_{\text{meas}} - x_{\text{theory}} \tag{6.13}$$

and the uncertainties are propagated. Values of $a_1(\bar{B} \to D^{*+}\pi^-) = 0.904 \pm 0.020$ and $a_1(\bar{B} \to D^{*+}K^-) = 0.926 \pm 0.026$ is found for BGL(2, 2, 2). When comparing to values found from BaBar data [11], of $a_1(\bar{B} \to D^{*+}\pi^-) = 0.98 \pm 0.04$ and $a_1(\bar{B} \to D^{*+}K^-) = 0.96 \pm 0.05$, this corresponds to an improvement of the total uncertainty for pions of 4.0% → 2.2% and for kaons of 5.2% → 2.8% and a shift of the central values towards lower values.

For $a_1(K)/a_1(\pi)$ the results are shown in Fig. 6.6 and the numerical values are listed in Table 6.3. The value of $a_1(K^-)^2/a_1(\pi^-)^2 = 1.047 \pm 0.054$ was found, consistent with $SU(3)$-symmetry. Furthermore, the ratio was calculated for different particle species. Their branching ratios are measured by different experiments and therefore it can not be ensured that the input parameters are consistent. Hence these may only be taken as a reference. However, also these mostly agree with $SU(3)$-symmetry.

TABLE 6.2: Results for a_1. The deviations take both, total experimental and theoretical uncertainties into account.

Parametrisation	Result	Deviation (σ)	Deviation (%)
$\pi^- CLN, LQCD$	$a_1 = 0.900 \pm 0.016$	8.4σ	$(-16.0 \pm 1.8)\%$
$\pi^- BGL(2,2,2), LQCD$	$a_1 = 0.910 \pm 0.019$	6.9σ	$(-15.0 \pm 2.1)\%$
$K^- CLN, LQCD$	$a_1 = 0.922 \pm 0.024$	5.5σ	$(-13.9 \pm 2.5)\%$
$K^- BGL(2,2,2), LQCD$	$a_1 = 0.931 \pm 0.026$	4.9σ	$(-13.1 \pm 2.6)\%$
$\rho^- CLN, LQCD$	$a_1 = 0.832 \pm 0.061$	3.5σ	$(-21.1 \pm 6.0)\%$
$\rho^- BGL(2,2,2), LQCD$	$a_1 = 0.836 \pm 0.062$	3.4σ	$(-20.7 \pm 6.0)\%$
$K^{*-} CLN, LQCD$	$a_1 = 0.803 \pm 0.079$	3.1σ	$(-23.8 \pm 7.6)\%$
$K^{*-} BGL(2,2,2), LQCD$	$a_1 = 0.807 \pm 0.079$	3.0σ	$(-23.5 \pm 7.6)\%$
$a_1^- CLN, LQCD$	$a_1 = 0.980 \pm 0.111$	0.7σ	$(-7.1 \pm 10.6)\%$
$a_1^- BGL(2,2,2), LQCD$	$a_1 = 0.983 \pm 0.111$	0.6σ	$(-6.7 \pm 10.7)\%$

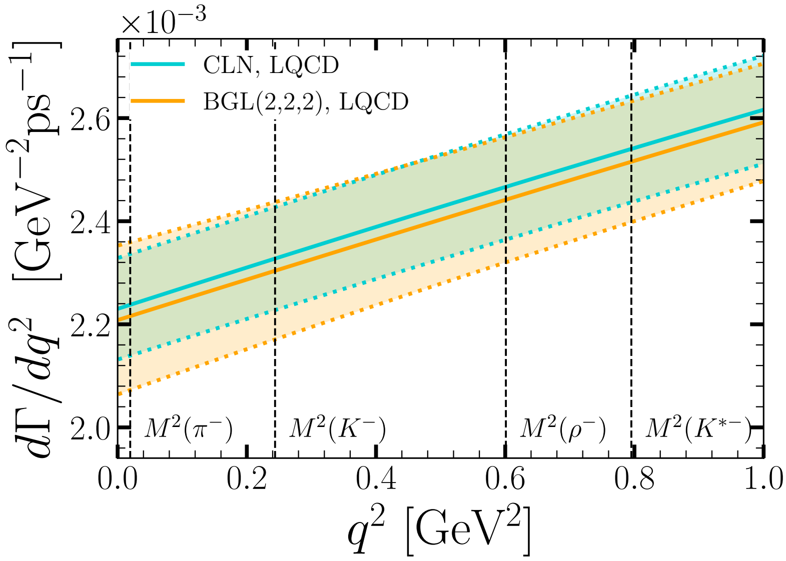


FIGURE 6.4: Fit results of the semi-leptonic decay rates $d\Gamma(\bar{B}_d \to D^{*+}l^-\bar{\nu})/dq^2$ as a function of momentum transfer to the D^* .

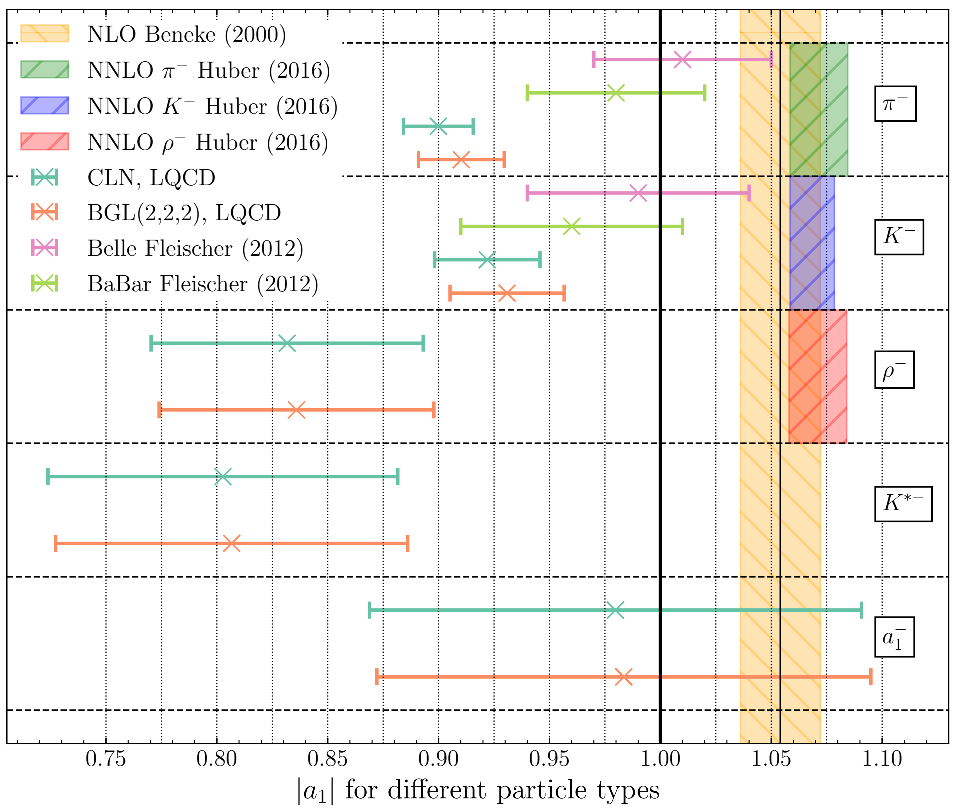


FIGURE 6.5: Results for a_1. The theory predictions are taken from Ref. [3, 15].

6.2 Conclusion

This is the first measurement of $a_1(\pi^-)$ and $a_1(K^-)$ performed by a single experiment. The treatment of systematic uncertainties is more uniform than in any previous

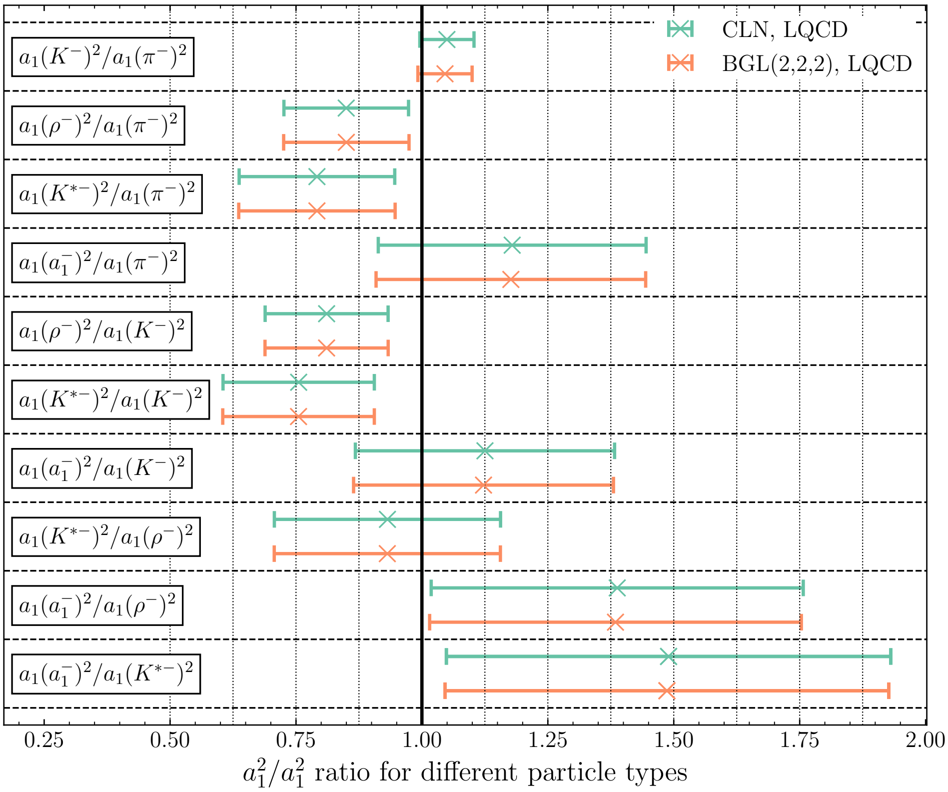


FIGURE 6.6: Results for ratio of a_1^2 for different particle species. Deviations from unity would suggest $SU(3)$ symmetry breaking effects.

TABLE 6.3: Results for the double ratio $a_1(p_1)^2/a_1(p_2)^2$.

Parametrisation	Ratio	Result
CLN, LQCD	$a_1(K^-)^2/a_1(\pi^-)^2$	1.050 ± 0.054
BGL(2,2,2), LQCD	$a_1(K^-)^2/a_1(\pi^-)^2$	1.046 ± 0.054
CLN, LQCD	$a_1(\rho^-)^2/a_1(\pi^-)^2$	0.850 ± 0.124
BGL(2,2,2), LQCD	$a_1(\rho^-)^2/a_1(\pi^-)^2$	0.850 ± 0.124
CLN, LQCD	$a_1(K^{*-})^2/a_1(\pi^-)^2$	0.792 ± 0.154
BGL(2,2,2), LQCD	$a_1(K^{*-})^2/a_1(\pi^-)^2$	0.792 ± 0.155
CLN, LQCD	$a_1(a_1^-)^2/a_1(\pi^-)^2$	1.179 ± 0.266
BGL(2,2,2), LQCD	$a_1(a_1^-)^2/a_1(\pi^-)^2$	1.177 ± 0.268
CLN, LQCD	$a_1(\rho^-)^2/a_1(K^-)^2$	0.811 ± 0.122
BGL(2,2,2), LQCD	$a_1(\rho^-)^2/a_1(K^-)^2$	0.811 ± 0.122
CLN, LQCD	$a_1(K^{*-})^2/a_1(K^-)^2$	0.755 ± 0.150
BGL(2,2,2), LQCD	$a_1(K^{*-})^2/a_1(K^-)^2$	0.755 ± 0.150
CLN, LQCD	$a_1(a_1^-)^2/a_1(K^-)^2$	1.125 ± 0.257
BGL(2,2,2), LQCD	$a_1(a_1^-)^2/a_1(K^-)^2$	1.122 ± 0.258
CLN, LQCD	$a_1(K^{*-})^2/a_1(\rho^-)^2$	0.932 ± 0.225
BGL(2,2,2), LQCD	$a_1(K^{*-})^2/a_1(\rho^-)^2$	0.931 ± 0.225
CLN, LQCD	$a_1(a_1^-)^2/a_1(\rho^-)^2$	1.388 ± 0.369
BGL(2,2,2), LQCD	$a_1(a_1^-)^2/a_1(\rho^-)^2$	1.384 ± 0.369
CLN, LQCD	$a_1(a_1^-)^2/a_1(K^{*-})^2$	1.490 ± 0.441
BGL(2,2,2), LQCD	$a_1(a_1^-)^2/a_1(K^{*-})^2$	1.486 ± 0.440

measurements. For the measurement of $|a_1|$ it can be stated that the results are independent of the parametrisation choice, although BGL$(2, 2, 2)$ was chosen for further discussions as it is the most general and model independent choice. It is remarkable that independently of the choice of form-factor parametrisations, all values found are multiple standard deviations smaller than the theory prediction. Higher order corrections to the theory suggest even larger values and thus higher discrepancies [15]. For $a_1(\pi^-)$ and $a_1(K^-)$ this can mean three things, either large, $13\% - 16\%$ non-factorisable contribution to the matrix elements, new physics contributions to the Wilson Coefficients or both.

Recent theoretical analyses of non-factorisable contributions in $B^0 \to J/\psi K_S$ decays suggest contributions of the size $O(10^{-3})$ [63], which is also in clear disagreement with the result obtained above.

In the ratios of a_1 for different particle types one can see that all of these are consistent with unity within one standard deviation. This makes a strong case for the statement that a_1 is indeed a universal quantity and $SU(3)$-symmetry holds in hadronic B-decays.

Chapter 7

Implementation and development of a global particle vertex fitting algorithm for Belle II

In order to fully exploit the upgraded Belle II detector and to improve the signal to noise ratio measurements for all physics analysis the software has to evolve as well. The development of an algorithm in this context is described in this chapter. The algorithm was initially developed for the BaBar experiment [64]. It was ported to the Belle II framework, by modifying and building upon the initial algorithm whose source code was kindly provided by its developer, Wouter Hulsbergben. However, the entire codebase, apart from the inheritance scheme was replaced by more efficient code and new features were added.

7.1 Particle decay vertex fitting

Particle decay vertex fitting techniques are widely used in particle and nuclear physics. Beyond the suppression of background, applications range from the improvement of particle momentum resolution (under the assumption they originate from some vertex point), to the determination of the presence of intermediate particles, and the precision determination of decay vertex positions. One can, for example, combine the measurements of two charged pion tracks originating from the decay of a K^0_S to extract the decay vertex position, flight length, and four-vector and their uncertainties. By performing a kinematic fit, one obtains an improvement of the pion track momenta and can use the χ^2 of the fit result to suppress background.

In order to construct more complex decay topologies, one usually combines cascades of these fits starting with long lived stable particles such as electrons, muons, pions and photons, forming intermediate resonances and finally the full decay of interest. For example, in the decay $B^0 \to J/\psi K^0_\mathrm{S}$, where $J/\psi \to \mu^+\mu^-$ and $K^0_\mathrm{S} \to \pi^+\pi^-$, one would first fit the J/ψ and K^0_S candidates and then use these to construct the B^0 candidates, as depicted in Fig. 7.1a and Fig. 7.1b. However, this only works well if the final state particles are charged and leave traces in the tracking detectors. Neutral particles can not be tracked in Belle II; they are only measured by their energy deposition in the calorimeter. Single layer crystal calorimeters do not offer directional information on where the particle associated to an energy deposition originated from. That means that the decay vertex cannot be extracted from a fit that exploits only the calorimeter information. In order to obtain the momentum vector, it could be assumed that the particle originates from the primary interaction point and travels directly into the cluster's center of gravity. This can introduce a large bias on the momentum direction. Consider, for example, the decay $B^0 \to J/\psi K^0_\mathrm{S}$, where the kaon

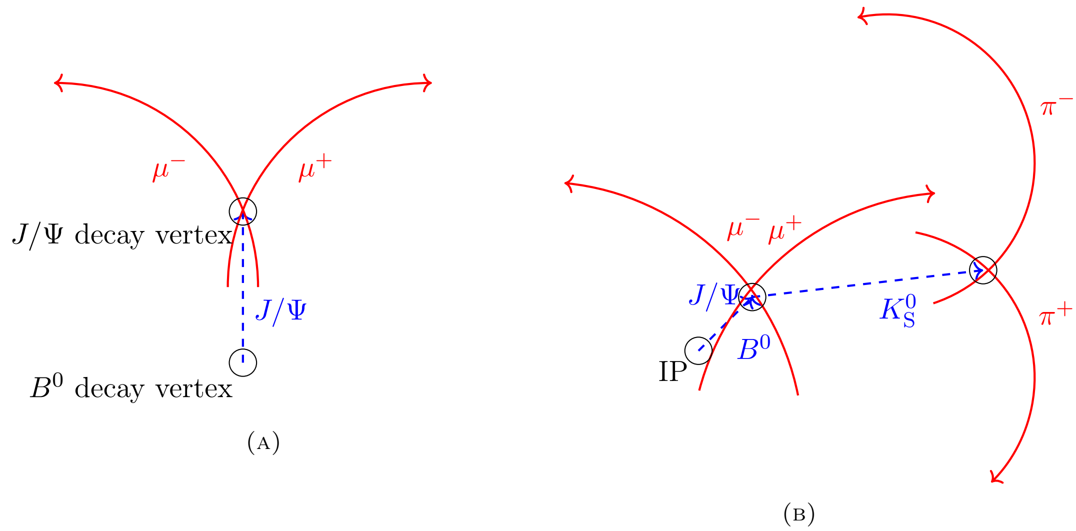


FIGURE 7.1: a) Depiction of a $J/\psi \to \mu^+\mu^-$ decay. The red lines show the track helix approximations obtained from the tracking detectors, the blue dashed lines show the decaying particle momentum vectors found by the fit. Since the decay length of the J/ψ is too short to be seen in the detector, its decay vertex is taken to be the one of the B^0. b) Depiction of a $B^0 \to J/\psi(\to \mu^+\mu^-)K_S^0(\to \pi^+\pi^-)$ decay. Measured track helices do not necessarily overlap in three dimensions. The depicted length ratios are not to scale.

instead decays to neutral final states, $K_S^0 \to \pi^0\pi^0$, and the pions decay, $\pi^0 \to \gamma\gamma$, as displayed in Fig. 7.2. Weakly decaying intermediate particles such as K_S^0 mesons have flight lengths of up to tens of centimetres at Belle II. Thus for a neutral particle in such a particle decay chain, the assumption that it originates from the primary e^+e^- collision is not sufficient. Furthermore, decays of particles into an intermediate state and a charged particle, i.e. $D^{*+} \to D^0\pi^+$, form a decay vertex that is only constraint by one measurement and thus susceptible to measurement errors.

The method presented overcomes these issues by globally fitting the entire decay tree in a single fit, taking into account all intermediate particles, extracting all involved particle's four-momenta, vertex positions, flight lengths and their covariance matrices, using a Kalman Filter as described in Ref. [64]. The software environment of Belle II is used together with the C++ template library EIGEN [65] for matrix operations, which provides fast execution times for the fit algorithm. Furthermore physics applications of the fitter with Belle II Monte Carlo samples and studies of performance characteristics are presented.Fig. 7.3 and 7.4 show an event display depicting simulated particles traversing the Belle II detector. In the decay $\Upsilon(4S) \to \bar{B}^0B^0$, one meson decays as $\bar{B}^0 \to D\omega\pi\pi$, and the other as $B^0 \to K_S J/\psi$, with $J/\psi \to \mu^+\mu^-$ and $K_S^0 \to \pi^0\pi^0$ with $\pi^0 \to \gamma\gamma$. Fig. 7.3 depicts the full detector geometry and Fig. 7.4 shows a close-up of the inner vertex detectors. In this example it is shown that the decay vertex of the K_S^0 can be highly displaced.

7.2 Extended Kalman Filter

Vertex fitting can be formulated as a least squares minimization problem. The computational challenge in finding a solution lies in matrix inversions, which naively scale as $\mathcal{O}(n^3)$, where n is the dimension of the matrix. In a naive approach, this is equal to the number of parameters extracted in the fit. An extended Kalman Filter is an

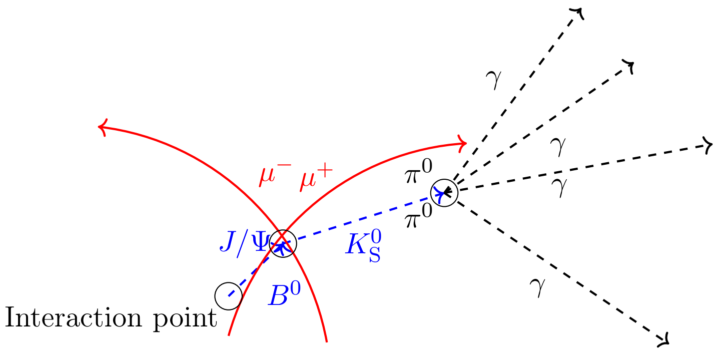


FIGURE 7.2: Depiction of a the decay $B^0 \to K^0_S J/\psi$. The red lines show the track helix approximations obtained by the tracking detectors, the blue dashed lines show the composite particle momentum vectors found by the fit. The dashed black lines depict the photon momenta found by the fit. Note that these can only be determined by the fit as the directional information of the calorimeter is not sufficient. The initial guess is that they point from the interaction point towards the calorimeter cluster. The decay lengths of the J/ψ and π^0 are too short to be seen in the detector therefore the vertex positions are taken from the particle above them in the hierarchy.

iterative approach to find the least squares estimator by defining a series of constraints (knowledge of parameters from measurements and symmetries) on a hypothesis (a set of particle parameters in this case). The hypothesis has different states during the stages of the filtering process. The filtering process is an iterative algorithm applying the constraints sequentially and updating the state with respect to the constraints. The sequence is repeated until convergence is reached or divergence is observed. In the case of convergence, the last state describes the best hypothesis for the parameters that can be found. An Extended Kalman Filter in the gain matrix formulation [66, 67, 64] is used in this algorithm. The state vector $\boldsymbol{x}$ holds the particle parameters to be optimized by the fit. The parameters depend on the type of particle. The most general parametrization takes the form

$$\boldsymbol{x} = \{x_1, y_1, z_1, \theta_1, p_{x1}, p_{y1}, p_{z1}, E_1, ..., x_n, y_n, z_n, \theta_n, p_{xn}, p_{yn}, p_{zn}, E_n\}, \tag{7.1}$$

where vertex coordinates of the i-th particle are denoted as $\{x_i, y_i, z_i\}$, its decay length is denoted as θ_i and the four momentum is $\{p_{x,i}, p_{y,i}, p_{z,i}, E_i\}$, and n is the number of particles in the fitted topology. This implies that $n \approx 8 \times$ number of particles. However, for final state particles the decay vertex is not of interest, so that only the momenta have to be parametrised. The problem is split into k constraints given by measurements and other knowledge of parameters (see Sec. 7.3 for the definitions of the measurement $\boldsymbol{m}$ and hypothesis $\boldsymbol{h}$ for each of the different constraints). The minimized χ^2 can be expressed as a weighted sum of k sets of equations, which are the constraints. For a given iteration, α, of the Kalman Filter, one can write

$$\chi^2_\alpha = \sum_k \chi^2_k, \tag{7.2}$$

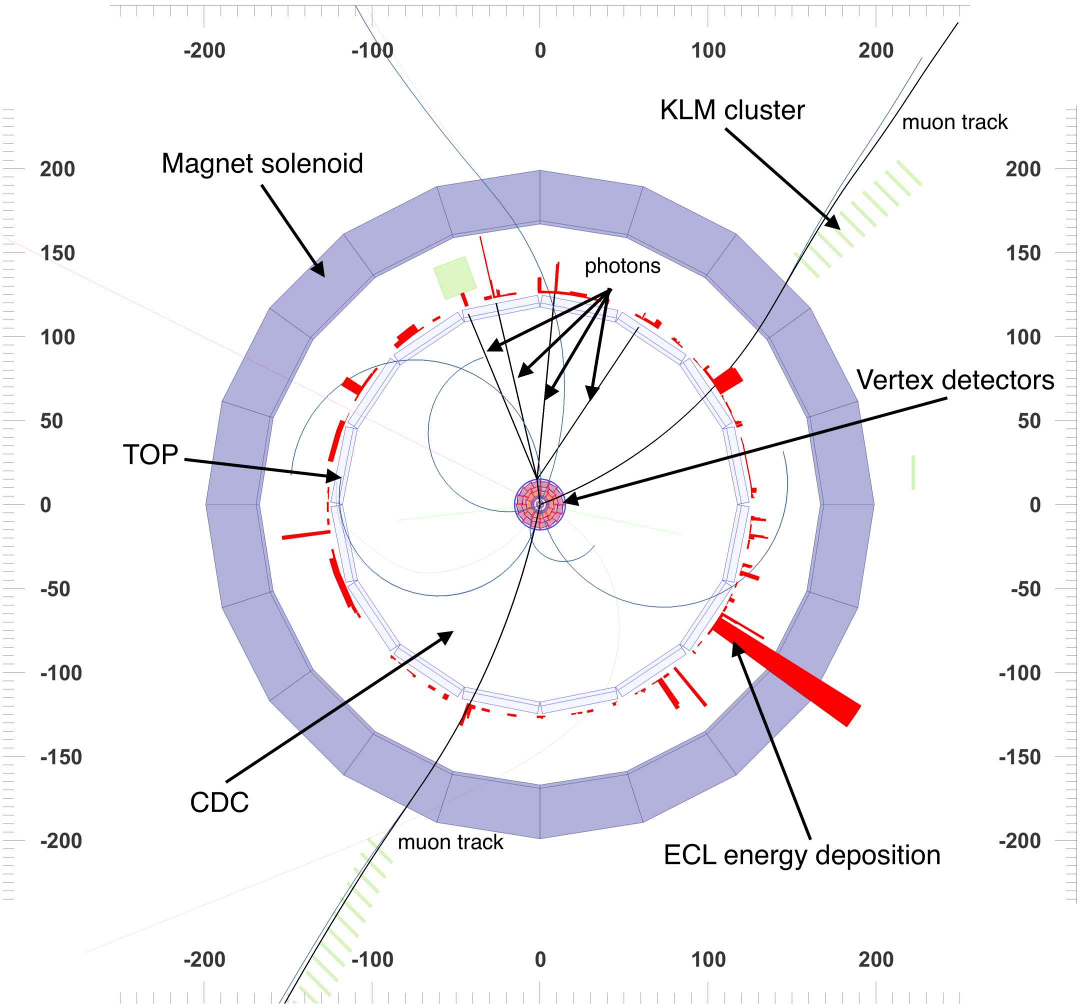


FIGURE 7.3: Event display projected onto the $x-y$ plane depicting the process of $\bar{e}^+e^- \to \Upsilon(4S) \to \bar{B}^0 B^0$, with $\bar{B}^0 \to D\omega\pi\pi$ and $B^0 \to K_S J/\psi$, where $J/\psi \to \mu^+\mu^-$ and $K_S^0 \to \pi^0\pi^0$ with $\pi^0 \to \gamma\gamma$. All particles of the signal B^0 decay are indicated by black lines. The rest of the event correspond to the decay of the $\bar{B}^0$. All distances are measured in centimetres. For a close-up of the vertex detector see Fig. 7.4.

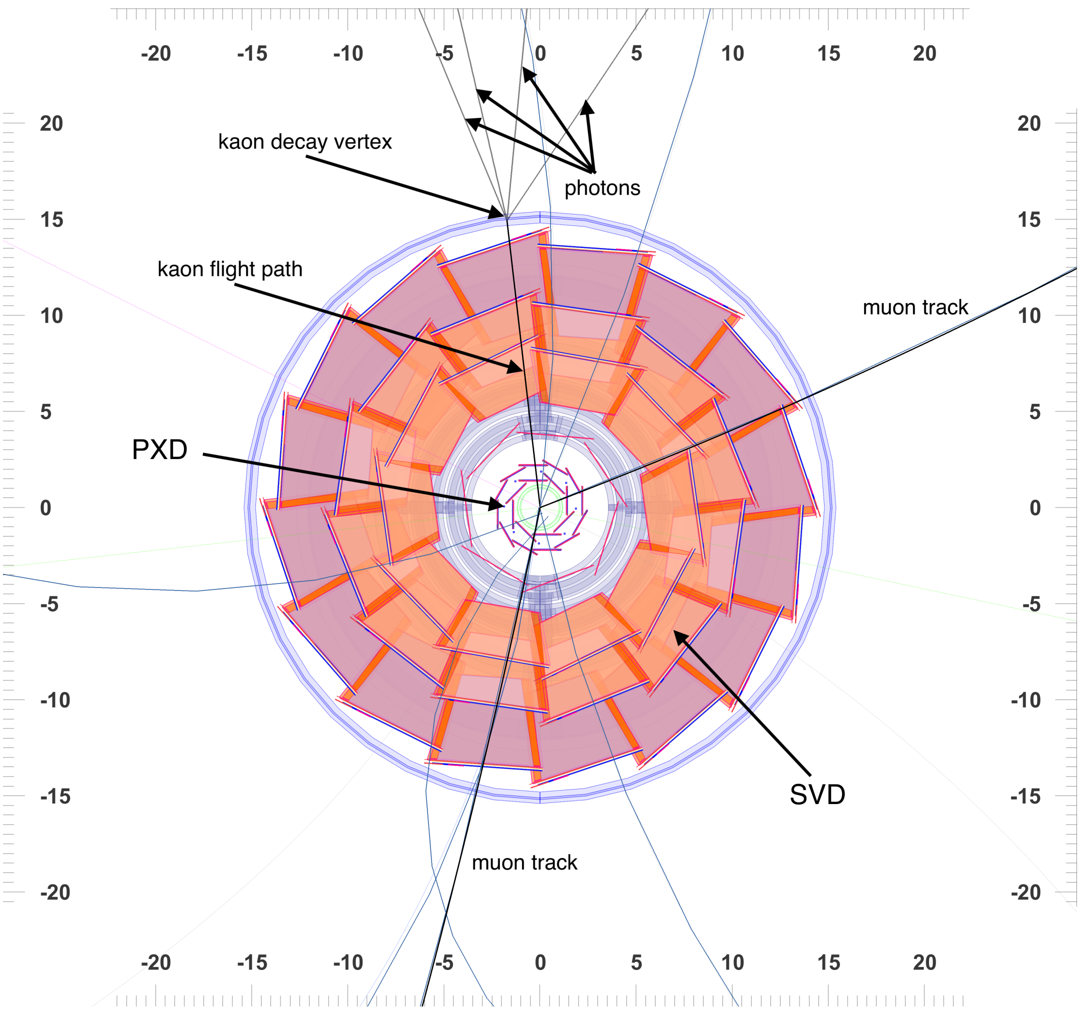


FIGURE 7.4: Close-up of Fig. 7.3. In this event, the K^0_S travels until the edge of the vertex detectors before decaying.

and

$$\begin{aligned}\chi_k^2 =&(\boldsymbol{x_k} - \boldsymbol{x}_{k-1})^T \boldsymbol{C}_{k-1}^{-1} (\boldsymbol{x_k} - \boldsymbol{x}_{k-1}) + \\ &(\boldsymbol{m}_k - \boldsymbol{h}_k(\boldsymbol{x_k}))^T \boldsymbol{V}_k^{-1} (\boldsymbol{m}_k - \boldsymbol{h}_k(\boldsymbol{x}_k)),\end{aligned} \tag{7.3}$$

using the measurement covariance matrix $\boldsymbol{V}_k^{-1}$ transported to the $\boldsymbol{m} - \boldsymbol{h}$ system as a weight, and the covariance matrix $\boldsymbol{C}_{k-1}^{-1}$ of the hypothesis, such that the hypothesis of each constraint k depends on the outcome of the previous constraint. Minimizing Eq. 7.3 yields a rule to find a new state $\boldsymbol{x}_k^\alpha$ for the current iteration and constraint, that is

$$\boldsymbol{x}_k^\alpha = \boldsymbol{x}_{k-1}^\alpha - \boldsymbol{K}_k^\alpha \boldsymbol{r}_k^\alpha, \tag{7.4}$$

where $\boldsymbol{K}$ is a gain matrix, which will be defined below. However, we first define the residual $\boldsymbol{r}$ as the distance between measurement and hypothesis

$$\boldsymbol{r}_k^\alpha = \boldsymbol{m}_k - \boldsymbol{h}_k(\boldsymbol{x}_{k-1}^\alpha) \ . \tag{7.5}$$

The hypothesis $\boldsymbol{h}_k(\boldsymbol{x}^\alpha)$ can be linearised around a reference state of the previous iteration $\boldsymbol{x}^{\alpha-1}$ using the Jacobian

$$\boldsymbol{H}_k^{\alpha-1} := \frac{\partial \boldsymbol{r}_k(\boldsymbol{x}^{\alpha-1})}{\partial \boldsymbol{x}^{\alpha-1}} = -\frac{\partial \boldsymbol{h}_k(\boldsymbol{x}^{\alpha-1})}{\partial \boldsymbol{x}^{\alpha-1}}, \tag{7.6}$$

such that

$$\boldsymbol{h}_k(\boldsymbol{x}_{k-1}^\alpha) = \boldsymbol{h}_k(\boldsymbol{x}^{\alpha-1}) - \boldsymbol{H}_k^{\alpha-1} \cdot (\boldsymbol{x}_{k-1}^\alpha - \boldsymbol{x}^{\alpha-1}). \tag{7.7}$$

Note that $\boldsymbol{x}^{\alpha-1}$ is always taken after the last constraint in iteration $\alpha - 1$ was filtered and the minus sign is coming from generalising $\boldsymbol{H}$ as the derivative of the residual. Equation 7.5 thus becomes

$$\boldsymbol{r}_k^\alpha = \boldsymbol{m}_k - \boldsymbol{h}_k(\boldsymbol{x}^{\alpha-1}) + \boldsymbol{H}_k^{\alpha-1} \cdot (\boldsymbol{x}_{k-1}^\alpha - \boldsymbol{x}^{\alpha-1}). \tag{7.8}$$

The gain matrix is calculated for every constraint in every iteration and is defined as

$$\boldsymbol{K}_k = \boldsymbol{C}_{k-1} (\boldsymbol{H}_k^{\alpha-1})^T \boldsymbol{R}_k^{-1}. \tag{7.9}$$

The only matrix to be inverted is of the dimension of the constraint and not the dimension of the state-space. $\boldsymbol{R}_k$ is defined below. The current state's covariance matrix is obtained via propagation of uncertainties

$$\boldsymbol{C}_k = \boldsymbol{C}_{k-1} - \boldsymbol{C}_{k-1} \boldsymbol{H}_k^{\alpha-1} \boldsymbol{R}_k^{-1} \boldsymbol{H}_k^{\alpha-1} \ \boldsymbol{C}_{k-1}^T \tag{7.10}$$

There are other formulations of Eq. 7.10 that are numerically more stable as discussed in [64]. The covariance matrix of the residual system can be found to be

$$\boldsymbol{R}_k = \boldsymbol{V}_k + \boldsymbol{H}_k^{\alpha-1} \boldsymbol{C}_{k-1} (\boldsymbol{H}_k^{\alpha-1})^T. \tag{7.11}$$

For constraints where no measurement is available, for example the kinematic or geometric constraint, it can be shown that Eq.7.11 is valid if $\boldsymbol{V}_k = 0$ [64].
Linearizing the hypothesis around a reference state, see Eq. 7.7, is a major improvement to the stability of the fit. The authors of Ref. [64] pointed us towards this improvement and more details can be found in Ref. [68]. In this formulation of the LSE problem, the update of the covariance matrix, Eq. 7.10, is the computational bottleneck, as it is a dense $m \times m$ matrix with the dimension of the state vector m and must be calculated and filled k times for each iteration. One can define a fit as converged when the change of the χ^2 is less than 2% and the resulting covariance

matrix is invertible, which is tested by checking its rank. A Kalman Filter is sensitive to the initial state vector and its associated covariance matrix chosen. One can choose values of the measurement in the respective constraint multiplied by a factor 1000 as the initial value of the hypothesised covariance matrix. For the initial vertex positions one can use the points of closest approach of track-helices or those of the “parent particles", if they are available, zero otherwise.

7.3 Parametrising and constraining the decay chain

To parametrize the decay chain, one can use a set of parameters that describe the properties of the particles. And perform a number of reductions on these properties to reduce the dimensionality of the problem. For final state particles, one can only save the momenta, as they do not have decay vertices. For their production vertices the decay vertices of their respective composite particles, referred to as the parents, are used. The energy of each final state particle is calculated using the momenta and its nominal mass hypotheses, taking the mass values provided by the particle data group (PDG) [50]. Intermediate particles are classified in two categories: particles that decay dominantly via the strong force, and those that decay via the weak force. Strongly decaying particles with a lifetime of less than 10^{-14} s, which corresponds to boosted flight lengths of less than 1 μm, are treated as if they instantly decay (the detector has a vertex position resolution in the x-y plane for charged particles of $20 - 30$ μm). The hypothesised quantities are their energy and momenta, while the production and decay vertices coincide and are taken from their parent particle's decay vertex. For weakly decaying particles, we additionally measure a decay vertex and a flight length, defined as the distance between the production and decay vertices in three dimensions.

7.3.1 Parametrising the constraints

Constraints are defined by Eq. 7.7, note that Lagrange multipliers could be eliminated, which allows for this convenient definition of constraints [64]. The resulting Jacobians take the form of $m \times n$ matrices where m is the dimension of the state vector and n is the dimension of the respective constraint. Thus, only few of its elements are non-zero. For example, for a three dimensional point constraint k, the hypothesis of particle number j with $\boldsymbol{h}_j = \{x_j, y_j, z_j\}$ and $\boldsymbol{x}$ as in Eq. 7.1, only the j-th diagonal block is non-zero

$$-\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}} = \boldsymbol{H} = \begin{pmatrix} 0 & -\mathbb{1}_3 & 0 \end{pmatrix} . \tag{7.12}$$

The blocks filled with zero correspond to the parameters of particle $\boldsymbol{x}_i \neq \boldsymbol{x}_j$. The columns filled with zeros will be omitted throughout this section, for brevity.
In the following the definitions of constraints that have been implemented in the Belle II software, based on the specific geometry of Belle II are listed.

7.3.1.1 Reconstructed track

A track can be parametrised with a five parameter helix. In Belle II it was chosen to use a perigee-parametrised helix, such that the helix is defined at the perigee, the point of closest approach of the helix to the origin of the coordinate system. The corresponding transformations to transport a helix to that point are discussed in Ref. [39]. A description of the parameters can be found in Table 3.2, and a depiction of the helix is in Fig. 3.10. One can parametrise tracks such that the model's dependence on Cartesian parameters is expressed as

$$\boldsymbol{h}_{\text{track}}(\boldsymbol{x}) = \begin{pmatrix} d_0 \\ \phi_0 \\ \omega \\ z_0 \\ \tan\lambda \end{pmatrix} = \begin{pmatrix} A(1+U)^{-1} \\ \text{atan2}(p_y, p_x) - \text{atan2}(\omega \cdot \Delta_{\parallel}, 1 + \omega \cdot \Delta_{\perp}) \\ a \cdot q/p_t \\ z + l \cdot \tan\lambda \\ p_z/p_t \end{pmatrix} . \tag{7.13}$$

Where atan2 refers to the ϕ domain corrected inverse tangent function. The same parametrisation for the hypothesis and the measurement is used. The measurement quantities are labelled with the index m. The residuals of iteration α then become

$$\boldsymbol{r}^{\alpha}_{track}(\boldsymbol{x}) = \begin{pmatrix} d_{0,m} - d_0 \\ \phi_{0,m} - \phi_0 \\ \omega_m - \omega \\ z_{0,m} - z_0 \\ \tan\lambda_m - \tan\lambda \end{pmatrix} + \boldsymbol{H}^{\alpha-1} \cdot (\boldsymbol{x}^{\alpha}_{k-1} - \boldsymbol{x}^{\alpha-1}) . \tag{7.14}$$

The Jacobian block $\boldsymbol{A} := -\partial\boldsymbol{h}/\partial\boldsymbol{x}$ is defined as the derivatives with respect to the vertex position, and $\boldsymbol{B} := -\partial\boldsymbol{h}/\partial\boldsymbol{p}$ as the derivatives with respect to momentum. The positions of these blocks in the Jacobian depend on the topology fitted and the particle represented by the track. The state vector is ordered hierarchically. This means that the decay vertex parameters come before its momentum, followed by the child particle's parameters. The full Jacobian $\boldsymbol{H}$ then takes the following form

$$\boldsymbol{H} = \begin{pmatrix} \cdots & & \cdots & & \cdots \\ \cdots & & \cdots & & \cdots \\ \cdots & \boldsymbol{A} & \cdots & \boldsymbol{B} & \cdots \\ \cdots & & \cdots & & \cdots \\ \cdots & & \cdots & & \cdots \end{pmatrix} . \tag{7.15}$$

For the non-zero elements of the Jacobian blocks, denoted by $-\partial d_0/\partial x = \boldsymbol{A}_{d_0,x}$, we derive the spatial components as

$$\begin{aligned} \boldsymbol{A}_{d_0,x} &= -\frac{p_{y0}}{p_{t0}}, & \boldsymbol{A}_{d_0,y} &= \frac{p_{x0}}{p_{t0}}, \\ \boldsymbol{A}_{\phi_0,x} &= -\frac{a \cdot q \cdot p_{x0}}{p_{t0}^2}, & \boldsymbol{A}_{\phi_0,y} &= -\frac{a \cdot q \cdot p_{y0}}{p_{t0}^2}, \\ \boldsymbol{A}_{z_0,x} &= \frac{p_x \cdot p_{x0}}{p_{t0}^2}, & \boldsymbol{A}_{z_0,y} &= \frac{p_x \cdot p_{y0}}{p_{t0}^2}, & \boldsymbol{A}_{z_0,z} &= -1, \end{aligned} \tag{7.16}$$

and for the momenta

$$\begin{aligned} \boldsymbol{B}_{d_0,p_x} = & \frac{y((aq)^2 r + 2aqp_y x + 2p_y^2\beta))}{p_{t0}p_t^2\beta^2} \\ & + \frac{p_x(2p_y x\beta + aq(y^2(-2+\beta) + x^2\beta))}{p_{t0}p_t^2\beta^2}, \\ \boldsymbol{B}_{d_0,p_y} = & -\frac{2p_x^2 x\beta + 2p_x y(p_y - aqx + p_y p_{t0}/p_t)}{p_{t0}p_t^2\beta^2} \\ & - \frac{aq(aqrx - p_y(x^2(-2+\beta) + y^2\beta))}{p_{t0}p_t^2\beta^2}, \end{aligned} \tag{7.17}$$

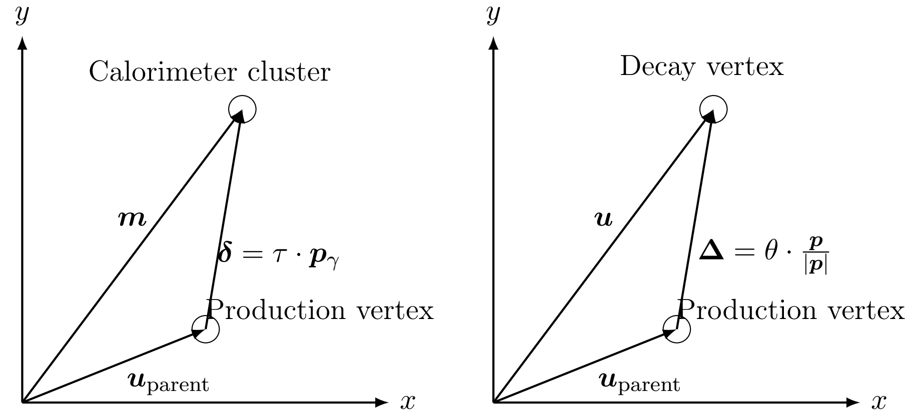


FIGURE 7.5: (a) The photon constraint, Eq. 7.21, reduced to two dimensions for simplicity. The vector $\boldsymbol{\delta}$ is defined as pointing from the photon's production vertex to the measured calorimeter cluster, indicated with the photons parent's coordinate vector $\boldsymbol{u}$ and measurement vector $\boldsymbol{m}$. (b) Geometric constraint, Eq. 7.31. The vector $\boldsymbol{\Delta}$ is defined pointing from the particles decay and production vertex, indicated with the particle's and its parent's coordinate vector $\boldsymbol{u}$.

and

$$\begin{aligned}
\boldsymbol{B}_{\omega,p_x} &= \frac{aqp_x}{p_t^3},\ \boldsymbol{B}_{\omega,p_y} = \frac{aqp_y}{p_t^3},\\
\boldsymbol{B}_{z_0,p_x} &= -\frac{p_z(p_x^2 x - p_y(aqr + p_y x) + 2p_x p_y y)}{p_{t0}^2 p_t^2},\\
\boldsymbol{B}_{z_0,p_y} &= -\frac{p_z(p_x(aqr + 2p_y x) - p_x^2 y + p_y^2 y)}{p_{t0}^2 p_t^2},\\
\boldsymbol{B}_{z_0,p_y} &= -(aq)^{-1}\text{atan2}\left(aq(p_y y - p_x x),\ p_x^2 + p_y p_{y0} - aqp_x y\right),\\
\boldsymbol{B}_{\tan\lambda,p_x} &= \frac{p_z p_x}{p_t^3},\ \boldsymbol{B}_{\tan\lambda,p_y} = \frac{p_z p_y}{p_t^3},\ \boldsymbol{B}_{\tan\lambda,p_z} = -p_t^{-1}.
\end{aligned} \tag{7.18}$$

7.3.1.2 Reconstructed photon

For photons the position of the calorimeter cluster and its energy are measured and from that one can infer the vertex parameters. The geometry, depicted in Fig. 7.5, gives

$$0 = \boldsymbol{u}_{\text{parent}} + \boldsymbol{\delta} - \boldsymbol{m}, \tag{7.19}$$

substituting $\boldsymbol{\delta} = \tau \cdot \boldsymbol{p}$ and inserting the energy relation, one gets

$$\boldsymbol{h}_{\text{photon}}(\boldsymbol{x}) = \begin{pmatrix} u_x + \tau \cdot p_x \\ u_y + \tau \cdot p_y \\ u_z + \tau \cdot p_z \\ \sqrt{p_x^2 + p_y^2 + p_z^2} \end{pmatrix} \quad \text{and} \quad \boldsymbol{m}_{\text{photon}}(\boldsymbol{x}) = \begin{pmatrix} m_x \\ m_y \\ m_z \\ E_m \end{pmatrix}, \tag{7.20}$$

where $\{u_x, u_y, u_z\}$ are the production vertex coordinates, $\{p_x, p_y, p_z\}$ are the parameters of the momentum vector pointing from the production vertex to the calorimeter cluster, $\{m_x, m_y, m_z, E_m\}$ are the position and measured energy of the corresponding ECL cluster. The parameter τ is a scalar with the units of time, it can be eliminated when writing down the residual to reduce the dimensionality of the equation system and avoid a trivial local minimum of $\boldsymbol{r}_\gamma$ at $\tau = 0$ when taking $\{u_x, u_y, u_z\} = 0$ as the starting point of the first iteration. Since the geometry of the detector is cylindrical, we can not simply eliminate any of the dimensions as this could introduce a pole in the residual equations. Therefore we sort the momenta and eliminate the dimension with the highest momentum such that we form a 3-dimensional equation system

$$\boldsymbol{r}'^{\alpha}_{\text{photon}}(\boldsymbol{x}) = \begin{pmatrix} (m_i - u_i) - (m_k - u_k)\frac{p_i}{p_k} \\ (m_j - u_j) - (m_k - u_k)\frac{p_j}{p_k} \\ E_m - \sqrt{p_i^2 + p_j^2 + p_k^2} \end{pmatrix} + \boldsymbol{H}^{\alpha-1} \cdot (\boldsymbol{x}^{\alpha}_{k-1} - \boldsymbol{x}^{\alpha-1}) \ , \tag{7.21}$$

where the indices i, j, k indicate the dimensions by order of increasing momentum $p_k \geq p_i \geq p_j$. We define $\boldsymbol{A}_{i,u_k} := -\partial h'_i/\partial u_k$ and $\boldsymbol{B}_{i,p_k} := -\partial h'_i/\partial p_k$ with the hypothesis of the reduced system r'. Thus, the non-zero entries are

$$\begin{aligned}
\boldsymbol{A}_{0,u_k} &= \frac{p_i}{p_k}, & \boldsymbol{A}_{0,u_i} &= -1, \\
\boldsymbol{A}_{1,u_k} &= \frac{p_j}{p_k}, & \boldsymbol{A}_{1,u_j} &= -1, \\
\boldsymbol{B}_{0,p_k} &= p_k^{-2}, & \boldsymbol{B}_{0,p_i} &= \frac{u_k - m_k}{p_k}, \\
\boldsymbol{B}_{1,p_k} &= p_k^{-2}, & \boldsymbol{B}_{1,p_j} &= \frac{u_k - m_k}{p_k}, \\
\boldsymbol{B}_{2,p_k} &= -\frac{p_k}{|\boldsymbol{p}|}, & \boldsymbol{B}_{2,p_i} &= -\frac{p_i}{|\boldsymbol{p}|}, & \boldsymbol{B}_{2,p_j} &= -\frac{p_j}{|\boldsymbol{p}|}.
\end{aligned} \tag{7.22}$$

The full Jacobian then takes the form

$$\boldsymbol{H} = \begin{pmatrix} \cdots & & \cdots & & \cdots \\ \cdots & \boldsymbol{A} & \cdots & \boldsymbol{B} & \cdots \\ \cdots & & \cdots & & \cdots \end{pmatrix} . \tag{7.23}$$

The covariance matrix of the measurement must be transformed into the reduced system. For that

$$\boldsymbol{V'} = \boldsymbol{F}\boldsymbol{V}\boldsymbol{F}^T, \tag{7.24}$$

with the transport matrix $\boldsymbol{F} = \partial r'/\partial m$, which depends on the sorting of the momenta such that the non-zero entries are

$$\begin{aligned}
\boldsymbol{F}_{0,m_k} &= -\frac{p_i}{p_k}, & \boldsymbol{F}_{0,m_i} &= 1, \\
\boldsymbol{F}_{1,m_k} &= -\frac{p_j}{p_k}, & \boldsymbol{F}_{1,m_j} &= 1, \\
\boldsymbol{F}_{2,E_k} &= 1,
\end{aligned} \tag{7.25}$$

is used. This constraint is not parametrized in p_t in order to keep the derivatives in Eq. 7.22 as computationally simple as possible.

7.3.1.3 Reconstructed K_L^0

K_L^0 are treated in the same way as photons, except that the nominal mass provided by the PDG is used in the energy calculation. The KLM detector is used for the cluster position measurement instead of the calorimeter. It cannot provide a precise energy measurement. Instead, it extrapolates the energy deposited by a particle as $E = c \cdot n$, where n is the number of hit cells in the cluster and c is a constant with the units

GeV. This approach makes the energy measurement for K_L^0 much less resolved than for photons.

7.3.1.4 Kinematic constraint

The kinematic constraint enforces four-momentum conservation, meaning it fits the four-momentum of the parent as the sum of the child momenta

$$\boldsymbol{r}^{\alpha}(\boldsymbol{x}) = \boldsymbol{p}_{\text{particle}} - \sum_i \boldsymbol{p}_{i,\text{child}} + \boldsymbol{H}^{\alpha-1} \cdot (\boldsymbol{x}^{\alpha}_{k-1} - \boldsymbol{x}^{\alpha-1}) \ . \tag{7.26}$$

The Jacobian for a particle with *two* children, which in this example are taken to be stable particles, can be defined as

$$\boldsymbol{H} = \begin{pmatrix} \cdots & & \cdots & & \cdots & & \cdots \\ \cdots & & \cdots & & \cdots & & \cdots \\ \cdots & \boldsymbol{A} & \cdots & \boldsymbol{B}_1 & \cdots & \boldsymbol{B}_2 & \cdots \\ \cdots & & \cdots & & \cdots & & \cdots \end{pmatrix}, \tag{7.27}$$

with the blocks $\boldsymbol{A}, \boldsymbol{B}$ as

$$\boldsymbol{A} = -\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{p}_{\text{particle}}} = \mathbb{1}_4, \tag{7.28}$$

and

$$\boldsymbol{B}_i = -\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{p}_{\text{child},i}} = -1 \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ p_{x,i}/E_i & p_{y,i}/E_i & p_{z,i}/E_i & 0 \end{pmatrix}, \tag{7.29}$$

Note that the energy row of $\boldsymbol{B}_i$ depends on how the particle is parametrized. Composite particles, for example, are parametrized with an energy variable in the state vector, resulting in $\boldsymbol{B} = -\mathbb{1}_4$, while for stable particles Eq. 7.29 is used.

7.3.1.5 Geometric constraint

The geometric constraint fits the decay length parameter θ for composite particles, see Fig. 7.5. Accounting for the geometry we have

$$0 = \boldsymbol{u}_{\text{parent}} + \boldsymbol{\Delta} - \boldsymbol{u} \ . \tag{7.30}$$

Instead of directly extracting a flight vector $\boldsymbol{\Delta}$, the unit vector of the momentum is used as it is well constrained by the previously filtered kinematic constraints, substituting $\boldsymbol{\Delta} = \theta \cdot \boldsymbol{p}/|\boldsymbol{p}|$, allows for a more accurate estimation of θ, the decay length parameter in our model. In contrast with [64], the decay length rather than decay time is chosen, because it makes the fit more linear. The residual is defined as

$$\boldsymbol{r}^{\alpha}(\boldsymbol{x}) = \boldsymbol{u}_{\text{parent}} + \theta \cdot \frac{\boldsymbol{p}}{|\boldsymbol{p}|} - \boldsymbol{u} + \boldsymbol{H}^{\alpha-1} \cdot (\boldsymbol{x}^{\alpha}_{k-1} - \boldsymbol{x}^{\alpha-1}) \ . \tag{7.31}$$

using

$$\boldsymbol{A} = -\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{u}_{\text{parent}}} = \mathbb{1}_3, \ \boldsymbol{B} = -\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{u}} = -\mathbb{1}_3, \ \boldsymbol{C} = -\frac{\partial \boldsymbol{h}}{\partial \theta} = \frac{1}{|\boldsymbol{p}|} \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix}, \tag{7.32}$$

and

$$\boldsymbol{D} = -\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{p}} = \frac{\theta}{|\boldsymbol{p}|^3} \begin{pmatrix} (p_y^2 + p_z^2) & -p_x p_y & -p_x p_z \\ -p_y p_x & (p_x^2 + p_z^2) & -p_y p_z \\ -p_z p_x & -p_z p_y & (p_x^2 + p_y^2) \end{pmatrix}, \tag{7.33}$$

such that

$$\boldsymbol{H} = \begin{pmatrix} \dots & & \dots & & \dots & & \dots & & \dots \\ \dots & \boldsymbol{A} & \dots & \boldsymbol{B} & \dots & \boldsymbol{C} & \dots & \boldsymbol{D} & \dots \\ \dots & & \dots & & \dots & & \dots & & .... \end{pmatrix}. \tag{7.34}$$

7.3.1.6 Mass constraint

The mass constraint requires a particle four-vector to be consistent with its nominal mass. The particle is treated as a measurement with infinite precision and use the mass value provided by PDG such that

$$r^\alpha(\boldsymbol{x}) = m_{\text{PDG}}^2 - E^2 + |\boldsymbol{p}|^2 + \boldsymbol{H}^{\alpha-1} \cdot (\boldsymbol{x}_{k-1}^\alpha - \boldsymbol{x}^{\alpha-1}), \tag{7.35}$$

with

$$\boldsymbol{H} = \begin{pmatrix} \dots & 2p_x & 2p_y & 2p_z & -2E & \dots \end{pmatrix}. \tag{7.36}$$

7.3.1.7 Beam spot constraint

B-mesons have a short lifetime and decay very close the beam spot, therefore it is useful to constrain the reconstructed particles to a volume within that region. The constraint is implemented by considering the initial e^+e^- collision as an abstract parent particle with the parameters and uncertainties of the beam spot, thus constraining the *production* vertex position of its children to an area within the beam spot's three dimensional uncertainty region using

$$\boldsymbol{r}^\alpha(\boldsymbol{x}) = \boldsymbol{s} - \boldsymbol{h}_{\text{production}} + \boldsymbol{H}^{\alpha-1} \cdot (\boldsymbol{x}_{k-1}^\alpha - \boldsymbol{x}^{\alpha-1}), \tag{7.37}$$

with the Jacobian

$$\boldsymbol{H} = \begin{pmatrix} \dots & & \dots \\ \dots & -\mathbb{1}_3 & \dots \\ \dots & & \dots \end{pmatrix}. \tag{7.38}$$

Here $\boldsymbol{s}$ denotes the beam spot position vector and $\boldsymbol{h}_{\text{production}}$ is the fitted production vertex of the particle. The beam spot is typically determined by averaging the x, y and z positions of the e^+e^- collisions over many interactions. SuperKEKB's nano-beam scheme will give x, y and z collision vertex distributions of $\sigma_x^* = 11\ \mu$m, $\sigma_y^* = 0.062\ \mu$m, and $\sigma_z^* = 500\ \mu$m respectively. This will imply powerful *production* vertex constraints for physics analyses.

7.3.1.8 Custom origin constraint

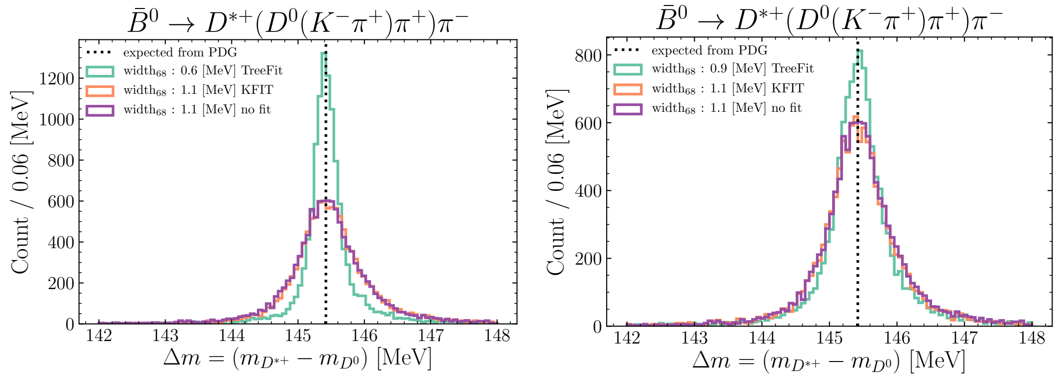
Similar to the beam spot constraint, one can construct another geometric constraint by defining a custom vertex position and an associated uncertainty to be the origin of the decay chain. This can be very useful when the decay chain contains particles that can not be detected, for example, neutrinos or long lived dark matter particles. In these decay chains it is not useful to fit the full chain, as the missing particle's four momentum would lead to an incorrect assumption for the kinematic constraint of the parent particle. However, knowing that the particle originates from a B-meson decay, one can define a geometric constraint corresponding to the volume where B-mesons decay on average. A beam energy constraint is not necessarily needed, as one can mass constrain the $\Upsilon(4S)$ particle, since all four-vectors in $e^+e^- \rightarrow \Upsilon(4\mathrm{S})$ are well known.

7.4 Applications of the fitter in Belle II

In this section use cases of the algorithm within the Belle II experiment are presented. Of special interest are decay chains containing one and two π^0-mesons, as well as D^{*+}-mesons. There are numerous channels where the phase space is large enough to add a π^0 to a vertex with two charged tracks, which makes this a very common structure in decay trees and therefore an interesting target to fit in a wide spectrum of analyses. Many decay channels of the B-meson contain D^{*+}-mesons, hence improving the reconstruction of these is a major goal. Monte Carlo samples from the Belle II experiment were used and background rejection and signal resolution improvements, as well as the extraction of various other parameters and their resolutions were analysed. Background can arise from non-signal processes, and from incorrect combinations of particles used in building the tree. The selection criteria used for all analyses in this section are listed in Tab. 7.1. Implicit charge conjugation is used, meaning that charged particles appearing in the decay chains imply the inclusion of opposite-sign charges. Because some distributions are not Gaussian, the inter quantile width containing 68% of the distribution as a measure to compare them is used. It is defined as $\mathrm{width}_{68} = \mathrm{q}(84\%) - \mathrm{q}(16\%)$.

7.4.1 Comparing the global fit to staged fits - the effect of the geometric constraint

The described method is compared to the standard staged fit method, KFIT [69], that was employed by the Belle experiment. KFIT performs stage wise fits, fitting the children particles of each vertex in a separate fit, so that only the measurements of the children particles are accounted for. This is done in a global fit to the 4-vectors and vertex position in cartesian space. In the case of the decay $D^{*+} \rightarrow D^0\pi^+$ in the decay chain $\bar{B}^0 \rightarrow D^{*+}(D^0(K^-\pi^+)\pi^+)\pi^-$, the first decay vertex is fitted using $D^0 \rightarrow K^-\pi^+$ the resulting D^0-meson is then fitted together with a pion track to form a D^{*+}-meson. This means that the decay vertex of the D^{*+}-meson is only constrained by the measurement of the π^+-track and the direction of the D^0-meson. The π^+-tracks of D^{*+}-meson decays have in the environment of the Belle II experiment usually low transversal momenta and thus are more likely to have large measurement uncertainties. In the global fit the final state particles are constrained via their respective constraints, tracks in this case. The four-vectors of intermediate particles are built using kinematic constraints. These, together with the decay vertex of the D^0-meson and the decay vertex of the D^{*+} are optimised at the same time, via a geometric constraint. As a result the resolution on the correlated quantity $\Delta m = m(D^{*+}) - m(D^0)$ can be improved by about 30% to a width of 0.57 MeV, depicted in Fig. 7.6a. If this constraint is lifted the improvement is diminished, see Fig. 7.6b. However, the individual resolutions on the masses of the D^{*+}-meson and D^0-meson are not improved. The improvement can be observed in semi-leptonic decay chains such as $\bar{B}^0 \rightarrow D^{*+}(D^0(K^-\pi^+)\pi^+)\mu^-\bar{\nu}_\mu$ displayed in Fig. 7.7. However, in this case the resulting kinematic parameters of the B^0-meson will be strongly biased due to the missing four-vector of the neutrino.



(A) (B)

FIGURE 7.6: Resolution comparison of the variable Δm for TreeFit a) with and b) without geometrically constraint D^0-meson and B^0-meson to KFIT.

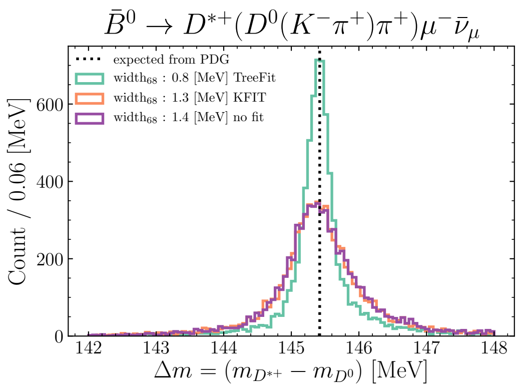


FIGURE 7.7: Comparison of the Δm-resolution in a semi-leptonic decay chain for different fitters. The resolution is worse than in a kinematically fully constrained decay chain, Fig. 7.6a, but the relative improvement compared to KFIT is similar.

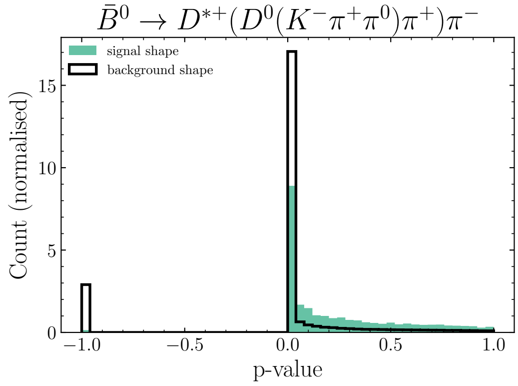


FIGURE 7.8: P-value distributions of fits to $\bar{B}^0 \to D^{*+}(D^0(K^-\pi^+\pi^0)\pi^+)\pi^-$. Failed fits were assigned a p-value of -1. The shape of the distribution is non-flat due to the use of a mass constraint on the π^0. The preselection criteria are listed in Table 7.1.

7.4.2 Fitting decay chains containing neutral particles

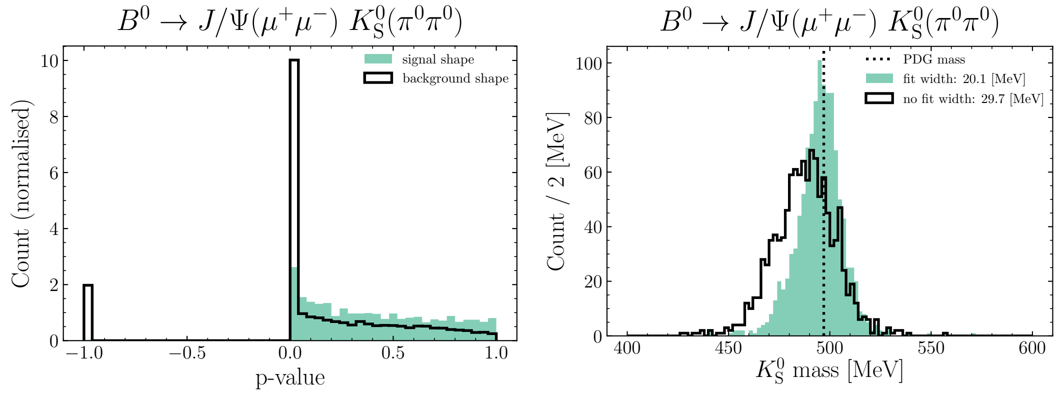
A high rate mode with a single π^0, is $\bar{B}^0 \to D^{*+}(D^0(K^-\pi^+\pi^0(\gamma\gamma))\pi^+)\pi^-$. This channel is of interest as it is the highest branching ratio D^0 decay mode. in this example the fitter is used to suppress background. If combinations for which the fit failed - i.e. those with a p-value ≤ 0 - are rejected, about 15% of the background while rejecting only about 1% signal, as displayed in Fig. 7.8. If a fit fails, a p-value of -1 is assigned for display purposes only. Note that since some of the constraints are non-linear and the uncertainties are not exactly Gaussian in some cases, p-values can have non-flat shapes.

7.4.3 Fitting decay vertices of neutral particles

Next study the decay chain $B^0 \to K^0_S(\pi^0(\gamma\gamma)\pi^0(\gamma\gamma))J/\psi(\mu^+\mu^-)$ is analysed. By performing the fit a large amount of the background can be removed by requiring that it passes the fit, see Fig. 7.9a. The only well defined vertex in this chain is given by the $J/\psi \to \mu^+\mu^-$ decay. The other vertices are very uncertain due to the absence of charged tracks. The best assumption that can be made for the production vertex position of the four photons is that they originate from the interaction point, if they are fitted in a single stage fit oblivious of the J/ψ. Performing a fit with mass constrained π^0-mesons improves the extracted mass of the K^0_S, so that after the fit, it is centred around the true value, as depicted in Fig. 7.9b. It is then possible to further reject background outside the nominal mass window and improve the signal purity when analysing this channel.

7.4.4 Using a beamspot constraint to improve the decay vertex resolution of B-mesons

A beam constraint enforces the production vertex of the B^0-meson to lie within the beamspots volume, see Sec. 7.3.1.7. It can be used to improve the decay vertex resolution of B-mesons and all children particles in the decay chain, depicted in Fig. 7.10a, Fig. 7.10b and Fig. 7.11. The resolution for B-mesons was improved by 15% to a value of about 33 μm in the example of the decay chain $\bar{B}^0 \to D^{*+}(D^0(K^-\pi^+)\pi^+)\pi^-$. In the no fit case $0, 0, 0$ is taken to be the vertex and no error is assigned. In terms of the defined resolution this can be better than the fit using KFIT in the D^{*+} case, because



(A) (B)

FIGURE 7.9: a) P-value of the fits to $B^0 \to J/\psi K^0_S$. b) Fitted mass of the K^0_S (green) and the mass before the fit (black). The mass distribution is centred around the true value after performing the fits with a π^0 mass constraint. Qualitatively this is the same result as obtained in Ref. [64]. The bias was removed and the width reduced by 9.6 MeV.

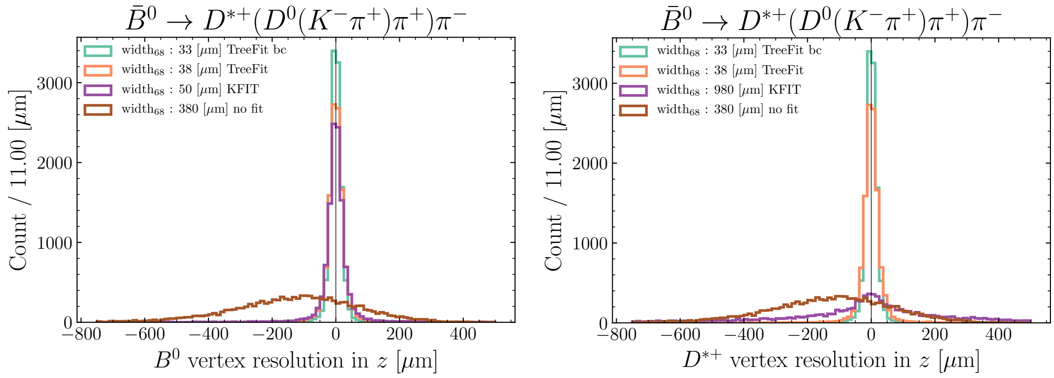
the pion involved in the decay has a low momentum, typically of around 150 MeV. These have measurement with a very large uncertainties and a fit involving only the D^0 and the pion then is dominated by this uncertainty. The very long tails for KFIT are due to very forward or very backward tracks. When fitting the entire decay chain with TreeFit, the tracks with better measurements help to constrain the vertex via geometric constraints. The efficiency of both algorithms is comparable.

7.4.5 Extracting the decay length of D^0-mesons from D^{*+} decays using a geometric constraint

The geometric constraint, see 7.3.1.5, constrains production and decay vertices of long lived particles in the decay chain. This allows for the extraction of flight lengths and thus lifetimes of intermediate particles such as D^0-mesons. This study is performed on $\bar{B}^0 \to D^{*+}(D^0(K^-\pi^+)\pi^+)\pi^-$ decays and extract the decay length of the D^0-meson. The results are depicted in Fig. 7.12. Since D^{*+}-mesons decay almost instantly, the three dimensional distance between the B^0 and the D^0 decay vertices is used. To improve the resolution on the B^0 vertex a beamspot constraint is applied.

7.5 Performance

Vertex fitting is computationally the most expensive operation during physics analysis. This is due to the large number of possible combinations that arise when reconstructing complex decay topologies. Physics analyses in the Belle II experiment are performed on a global grid of computing clusters [70] occupying thousands of CPUs simultaneously. Reducing the execution time of these analyses is a major effort of the collaboration. To achieve this the EIGEN library for matrix operations is used. EIGEN ensures memory locality is respected and operations are vectorized when possible. Modern CPUs are much faster in performing operations than retrieving memory. Allocating memory used locally in time and space allows modern CPUs to guess and prefetch the memory that is used next, such that no clock cycles are wasted [71]. The algorithm based on EIGEN was profiled using the Valgrind [72]



(A) (B)

FIGURE 7.10: a) Resolution of the z coordinate of B-mesons with and without applying a beam constraint. b) Resolution of the z coordinate of D^{*+}-mesons with and without applying a beam constraint. The fits using a beam constraint are indicated as TreeFit bc. Over- and underflow bins are not depicted.

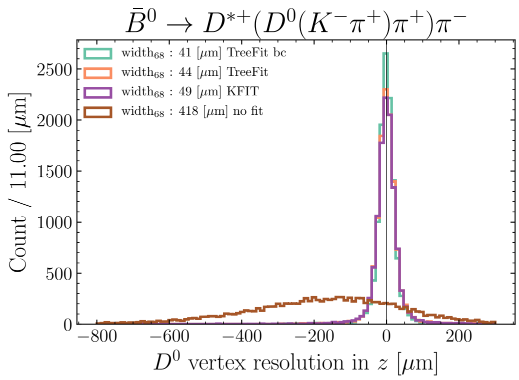
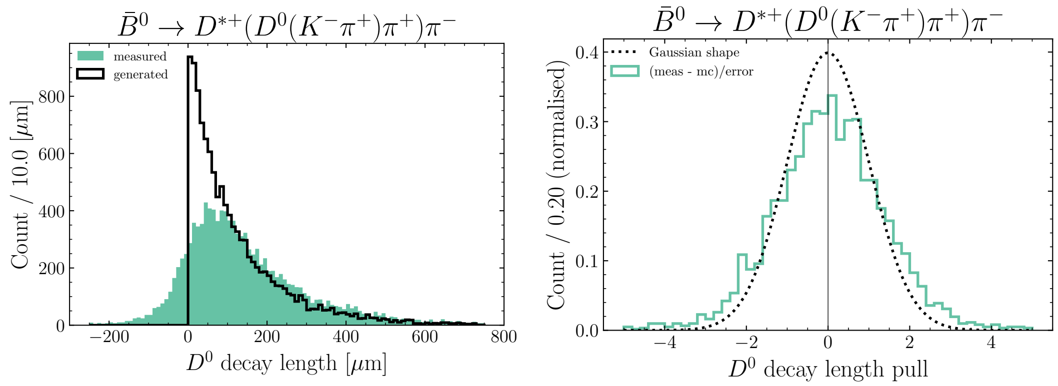


FIGURE 7.11: Resolution of the z coordinate of D-mesons with and without applying a beam constraint. The fits using a beam constraint are indicated as TreeFit bc.

TABLE 7.1: Selection criteria used in the analyses in this section. All particles use the same criteria if appearing in the decay chains. The invariant mass is obtained by summing the particle's daughter four-momenta before performing the fit. The particle ID for K^+/π^+ is defined as the likelihood ratio $\mathcal{LR}(K^+(\pi^+)) = \mathcal{L}(K^+(\pi^+))/(\mathcal{L}(\pi^+) + \mathcal{L}(K^+))$ and the beam energy constrained mass m_{bc} [13].

particle	pre selection applied
γ	$E_\gamma > 0.075$ GeV
π^0	0.145 GeV $> m(\gamma\gamma) > 0.125$ GeV
π^+	$\mathcal{LR}(\pi^+) > 0.5$
K^+	$\mathcal{LR}(K^+) > 0.5$
D^0	1.9 GeV $> m(K^-\pi^+\pi^0) > 1.7$ GeV
D^*	$\Delta m = m(D^*) - m(D^0) < 0.155$ GeV
B^0	$m_{\text{bc}} = \sqrt{E^2_{\text{beam}}/4 - p^2} > 5.27$ GeV

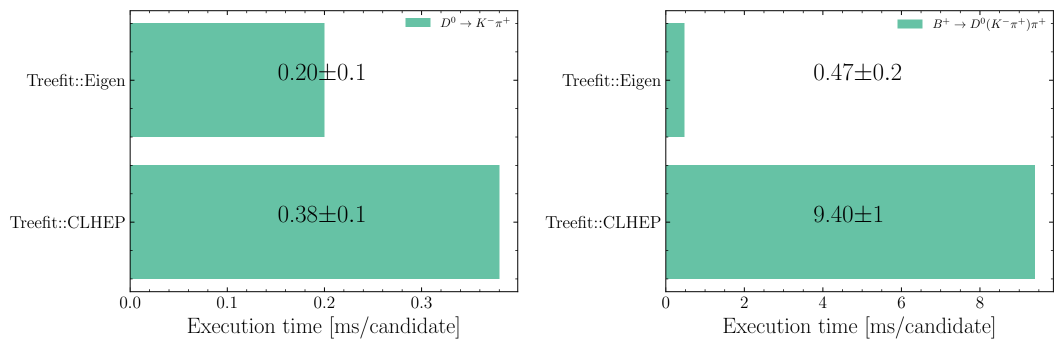


(A) (B)

FIGURE 7.12: a) The extracted decay length (green) and the generated decay length (black). The negative tail of the extracted decay length is due to the detectors resolution function. b) Pull (measured − generated/uncertainty) of the decay length distribution. The shape of a Gaussian with a mean of zero and a RMS of one is plotted for reference. The shape of the pull matches with a Gaussian, apart from a small mismatch in the central region. This mismatch can be attributed to an underestimation of the tracking parameter uncertainties, where a similar mismatch was observed.

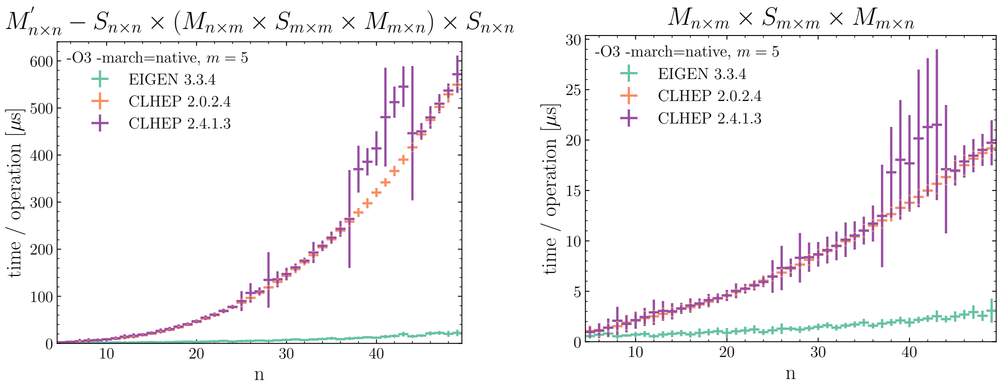
module Callgrind. It was found that one iteration of the algorithm spends 69% of the time on the implementation of Eq. 7.10, 27% on Eq. 7.9 and 4% on other operations, when fitting the decay chain $B^+ \to D^0(K^-\pi^+)\pi^+$. The performance using EIGEN was compared to a previous implementation using the CLHEP library [73], a C++ library widely used for algebraic operations in particle physics. By using EIGEN the execution time was speed up by a factor two in simple cases and more than an order of magnitude in more complex topologies, see Fig. 7.13. These numbers were obtained using EIGEN-3.3.4 and CLHEP-2.2.0.4 and compiled with clang-4.0 using the $-O3$ flag for full optimisation including vectorisation.

In order to asses the scaling with increasing matrix sizes, a stand-alone program to directly compare the libraries was compiled, to investigate n times m matrices, M, and symmetrical matrices, S. For the latter symmetric matrix types in both



(A) (B)

FIGURE 7.13: a) Average execution time per candidate fitting two tracks. b) Average execution time per candidate fitting a topology with an intermediate decay. The errors are estimations from repeating the study 10 times.



(A) (B)

FIGURE 7.14: a) Execution time scaling of Eq. 7.10. For increasing matrix sizes n. b) Execution time scaling for the inner matrix product of Eq. 7.10. The notation S indicates a symmetrical matrix, M a non symmetrical, the dimension m was fixed to $m = 5$, which is equivalent to Eq. 7.10 being calculated for a track constraint.

libraries were used. Optimized matrix operation implementations provided by the libraries are used where possible. One matrix dimension is fixed to the largest occurring number in the algorithm, which is $m = 5$ for track constraints. Then the other dimension n is increased, which is equivalent to adding more particles to the decay chain. The computation of Eq. 7.10 is repeated 500 times and the median is taken as a data point. The process is repeated 100 times, each time with randomised initial matrices and the width of the distribution is taken as the uncertainty. Much stronger exponential scaling of the operations involving the CLHEP library, depicted in Fig. 7.14, was observed. Since the version of CLHEP used in the algorithm is outdated, it wa additionally compared to a more recent version of the library. No change in performance was found for the operations investigated. The compiler flags $-O3$ and $-march = native$ were used when building our program to enable all possible optimisations the hardware allows. It was found that omitting these flags drastically worsens the performance of EIGEN, while CLHEP has to be compiled prior to usage and therefore is oblivious to these options. The cache-miss ratewas measured to be close to zero, which persists if an upper limit for the size of the matrices in EIGEN is pecified. From this it was concluded that our matrices are mostly heap allocated and the main performance gain is driven by vectorisation.

7.6 Conclusion

An improved implementation of a global vertex fitting tool, based on Ref. [64], tailored for the environment of the Belle II experiment was implemented in the context of this thesis. It can be used for various purposes, such as the extraction of particle production and decay vertices, decay lengths, particle four-momenta and rejection of backgrounds, as well as the extraction of the respective uncertainties. The global fitting technique is particularly powerful in fitting and reducing background in modes that contain neutral particles and can significantly improve the resolution on Δm an important variable for the reconstruction of D^*-mesons, which will be very important for the Belle II physics analysis program and will directly benefit analyses of CP-violation and the CKM unitarity angle γ, such as Ref. [74]. The execution time of the algorithm was reduced by an order of magnitude, by restricting the maximum possible size of dynamic matrices and enabling vectorised matrix operations.

Chapter 8

Conclusion

The SM of particle physics, even though incomplete, provides explanations for a wide range of phenomena in great detail. Experimentally it is the most well-tested theory in physics. As the experiments on the high energy frontier at the LHC, still have not found any evidence for physics beyond the SM, high precision tests of the SM become even more interesting. The Belle experiment and its successor the Belle II experiment provide excellent experimental conditions to perform these types of tests as a large data sets of B-mesons are available in a clean environment and the detector is well understood.

This thesis presents new measurements of the branching ratios $\mathcal{B}(\bar{B}^0 \to D^{*+}K^-)$ and $\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-)$, as well as their ratio $R_{K/\pi}$, using $(771 \pm 10) \times 10^6$ B-meson pairs recorded by the Belle experiment. The branching ratios were $\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-) = (2.677 \pm 0.016 \pm 0.087) \times 10^{-3}$ and $\mathcal{B}(\bar{B}^0 \to D^{*+}K^-) = (2.267 \pm 0.064 \pm 0.078) \times 10^{-4}$, and the ratio was found to be $\mathcal{R}_{K/\pi} = (8.413 \pm 0.243 \pm 0.126) \times 10^{-2}$. Compared to the predictions this suggests that either $SU(3)$-symmetry breaking contributions exists and/or the contribution of the exchange topology to $\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-)$ is underestimated in the predictions. The absolute uncertainties of 4.0% and 3.0% make these the most precise measurements of branching ratios at the Belle experiment. The ratio, $R_{K/\pi}$, was measured in a way that cancels the systematic uncertainty almost completely, allowing for a measurement of the ratio with 3.2% total uncertainty. This has never been measured before and provides the highest precision test of $SU(3)$ symmetry in B-decays to date.

The measured branching ratio have been used together with momentum dependent semi-leptonic decay rates in collaboration with the authors of Ref. [58] to measure $a_1(\bar{B} \to D^{*+}\pi^-) = 0.904 \pm 0.020$ and $a_1(\bar{B} \to D^{*+}K^-) = 0.926 \pm 0.026$, a parameter encoding fundamental assumptions about effects of the strong force in B-decays. This is the first measurement of this kind pertformed at a single experiment and it is shown that there must be either large non-factorisable contributions to the matrix elements of B-decays and/or new physics involved, to explain a deviation $15\% - 17\%$ from the factorisation assumption. Furthermore, $SU(3)$-symmetry was tested by measuring ratio $a_1(K^-)^2/a_1(\pi^-)^2 = 1.047 \pm 0.054$. This is agreeable with unity, so no evidence of a dependency on the involved quark flavour and thus for $SU(3)$ breaking effects was found.

In preparation for measurements of this kind, and many others at Belle II, a novel global vertex fitting algorithm was implemented and extended for the need of the Belle II experiment. The execution speed as compared to existing implementations could be improved by an order of magnitude, which will drastically speed up many analyses and fundamentally improve physics performance of measurements. It allows to improve the mass invariant resolution for D^{*+}-meson reconstruction by $\approx 30\%$

and therefore improves background suppression by the same amount. It allows to uniformly treat charged and neutral final states and outperforms previous methods. As a results analyses of direct CP-violation and the CKM unitarity angle γ, such as Ref. [74], will benefit greatly from this new tool. It was used for the measurement of the K/π-ID corrections in Appendix B. Furthermore, the algorithm has become the standard vertex fitting tool of the Belle II collaboration and will be used by many analyses in the future. This work is published in Ref. [75].

Looking towards the future, the Belle II experiment will collect 50 ab^{-1} and perform many high precision analyses of B-decays. The work in this thesis has tested the theoretical framework used to calculate the theoretical predictions with high precision and found these to be insufficient. Thus, the stage was set for improvements on the theory side, which in turn are the basis of phenomenological interpretations of future measurements. In particular CP-violation measurements stand to benefit from reductions of the hadronic uncertainties on the matrix elements. As consequence the ground work for a better understanding of the physics hiding in the Belle II data set was laid.

Appendix A

Continuum suppression

The reaction $e^+e^- \rightarrow q\bar{q}$, where $q \in \{u, d, s, c\}$, is ≈ 4 times more likely to occur than $e^+e^- \rightarrow B\bar{B}$, due to the small cross-section of this process, see Fig. 3.1. In this section the possibility of using a multivariate method to reduce these backgrounds was explored. Furthermore, the toy MC method described in Sec. 5.8.2 was used to evaluate the signal resolution, defined as the distance between the found and the MC-truth matched value, with the reduced background. The improvement was found to be too small to justify the employment of such a complex method in the signal region. However, it was used in the control regions to isolate B background.

A.1 Continuum suppression

In order to separate out $B\bar{B}$-events from continuum background, a Boosted Decision Tree (BDT) from the XGBOOST python library as described in Ref. [76] was used. This classifier takes a number of input variables $\vec{x}$ and maps these to an output variable

$$f(\vec{x}) \rightarrow y, y \in [0, 1], \tag{A.1}$$

describing how likely an event is to have originated from continuum ($q\bar{q}$), or $B\bar{B}$. The larger the value the higher the chance that the event is a $B\bar{B}$-event. The output of the classification after applying all other preselection is depicted in Fig. A.4 for the $D^0 \rightarrow K^-\pi^+$ mode and Fig. A.5 for the $D^0 \rightarrow K^-2\pi^+\pi^-$ mode. The input variables are described in turn below.

TABLE A.1: Hyper parameters chosen for the BDT training.

Parameter	Value
Learning rate	0.01
Depth of each tree	3
Number of trees	150
Random state	False

A.1.1 Input variables

Thrust Thrust is a concept to define the direction of particle jets [77]. The thrust-axis is the axis in the 3d-plane that maximise the projection off all momenta for a B-candidate or the rest of event onto it. The cosine of the thrust axis relative to the z-axis is a good discriminant as continuum events usually are more boosted towards small angles relative to the z-axis, while B-events are spherical and do not have a preferred direction.

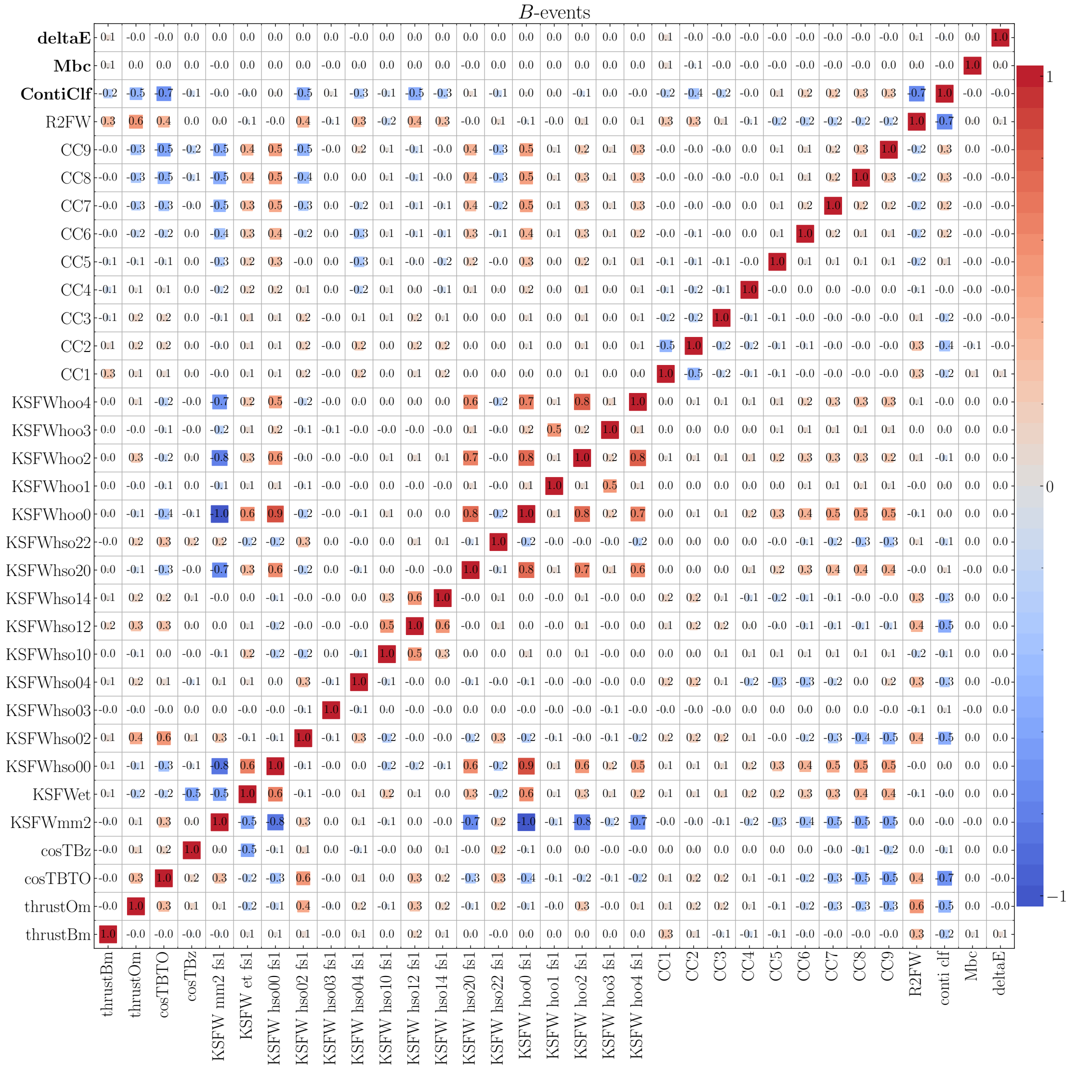


FIGURE A.1: Correlation matrix for B-events between the variables used in the training of the classifier. The bold printed variables were not fed to the classifier. The names "Pion" and "Kaon" describe the corresponding MC-truth variables for the respective components.

Fox Wolfram Moments The Fox Wolfram Moments [78], H_l, in the following abbreviated as “KSFW_ij” where an additional label indicates “et" for transversal energy, “mm2" for missing mass squared, “Hso" for particle j in the signal and particle i in the rest of event and “Hoo" if both particles are in the rest of event. The moments are calculated as

$$H_l = \sum_{i,j}^{N} |\boldsymbol{p}_i|\,|\boldsymbol{p}_j|\, P_l \cos\theta_{ij},\; R_l = \frac{H_l}{H_0}, \tag{A.2}$$

where P_l is the l-th order Legendre Polynomial and θ_{ij} is the angle between i-th and j-th particle momentum $\boldsymbol{p}$. The ratio of the second to the 0-th order polynomial R_2 is a very good discrimination variable.

Cleo-Cones Cleo-Cones, first employed by the CLEO experiment [79], are defined as the scalar momentum flow around the thrust axis in concentric cones of 10° size, such

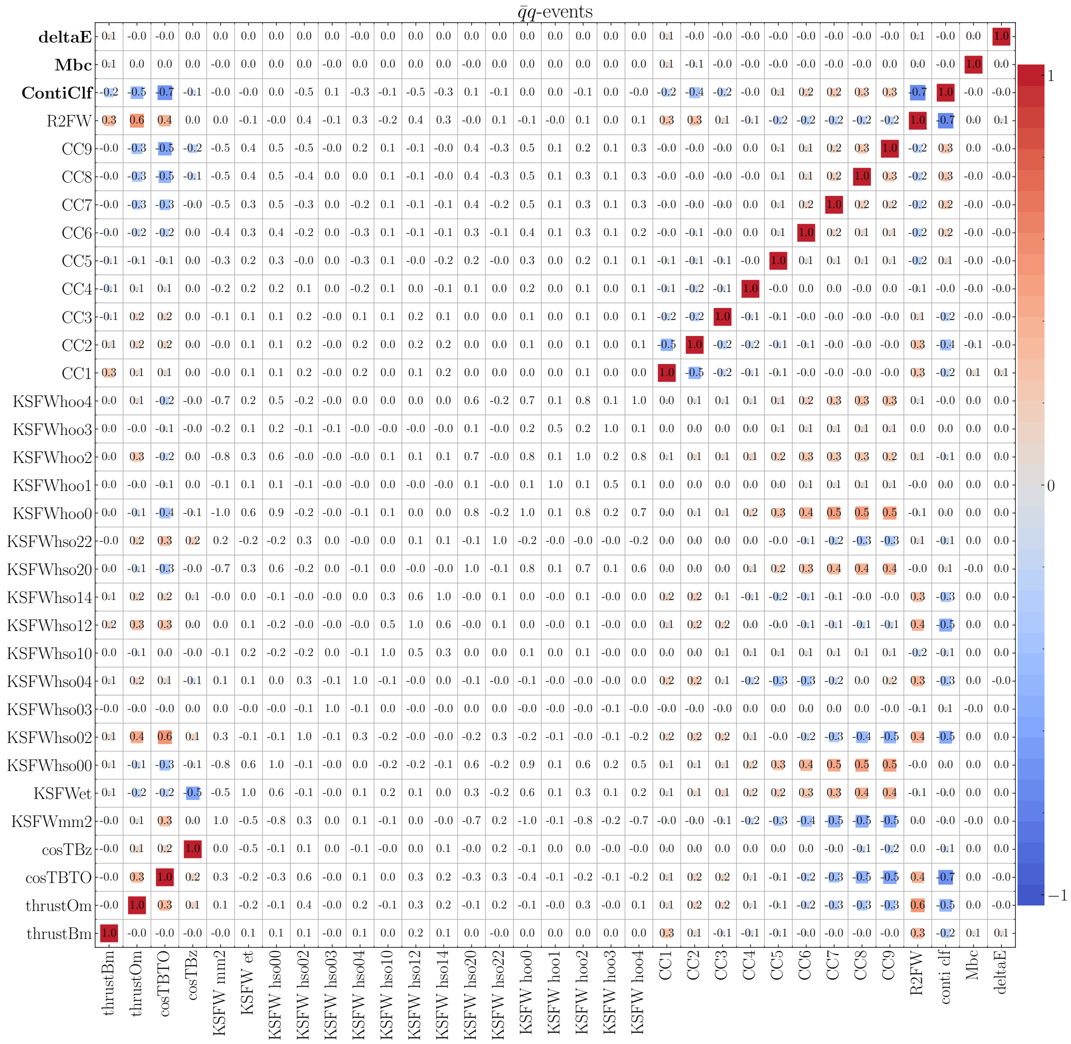


FIGURE A.2: Correlation matrix for $\bar{q}q$-events between the variables used in the training of the classifier. The bold printed variables were not fed to the classifier. The names "Pion" and "Kaon" describe the corresponding MC-truth variables for the respective components.

that CC_2 is the momentum flow in a cone of 20° opening angle minus the momentum flow in a cone of 10° opening angle around the thrust axis.

A.1.2 Continuum suppression inputs in the M_{bc} control region

The variables that are used as inputs for the continuum classification give a direct insight into the event shape modelling. Hence, verifying the modelling of these provides confidence in the overall modelling. Furthermore, it has to be checked in order to justify its usage for the continuum suppression. This is done in a data control region with $M_{bc} < 5270$ MeV. All other signal selection criteria are applied. The continuum suppression variables do not change with the direction of the hadrons K/π-ID selection and therefore only the pion mode is shown in Fig. A.6 and Fig. A.7 for the $D^0 \to K^-\pi^+$ channel and Fig. A.8 and Fig. A.9 for the $D^0 \to K^-2\pi^+\pi^-$ channel. Overall good agreement of the shapes is observed with a small disagreement of the total normalisation $D^0 \to K^-\pi^+$, most likely due to the charm continuum component as discussed for the M_{bc} control region.

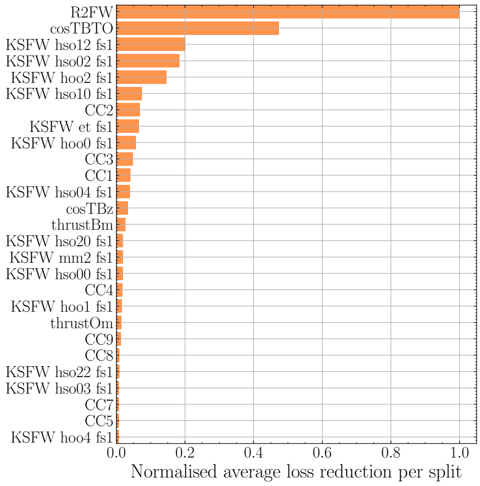
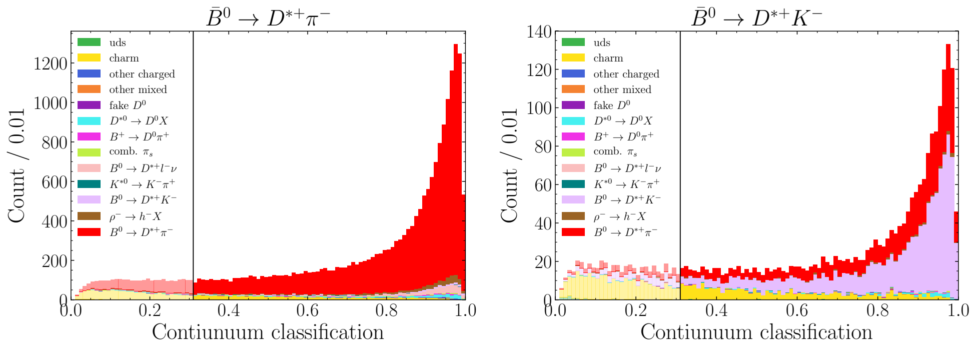


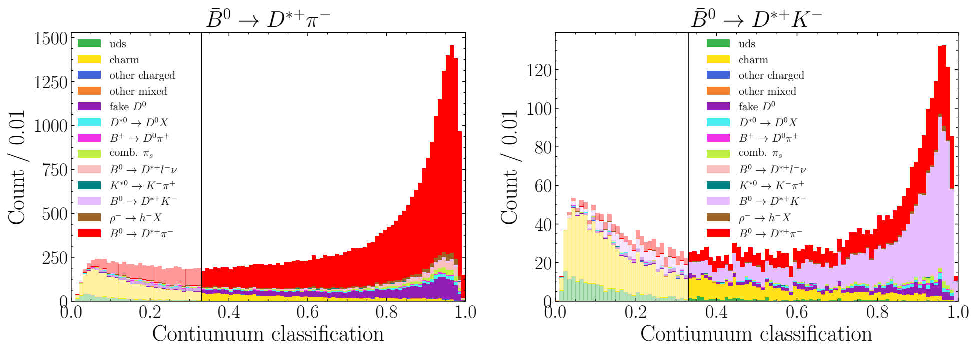
FIGURE A.3: Feature importance ranked by the average loss reduction of a selection in the respective feature.



(A)

(B)

FIGURE A.4: Output of the continuum classification by component for the $D^0 \to K^-\pi^+$ mode. The line indicates the optimal selection threshold.



(A)

(B)

FIGURE A.5: Output of the continuum classification by component for the $D^0 \to K^-2\pi^+\pi^-$ mode. The line indicates the optimal selection threshold.

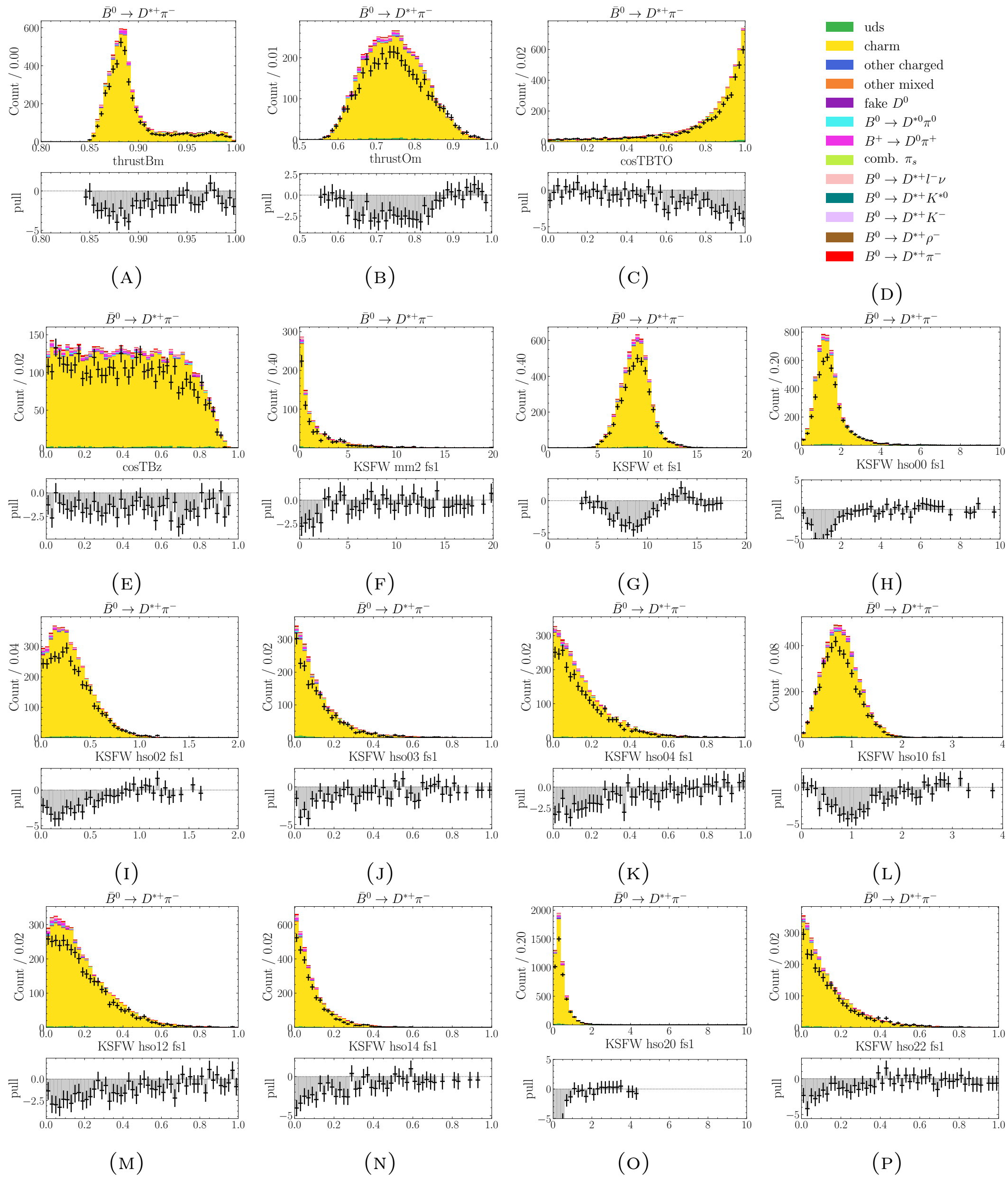


FIGURE A.6: Variables used for the continuum suppression classifier in the control region $M_{\text{bc}} < 5270$ MeV for the $D^0 \to K^-\pi^+$ channel. KSFW-moments are abbreviated with KSFW and Cleo Cones are abbreviated as CC followed by the number.

A.1.3 Training of the classifier

The classifier was trained with continuum events, before splitting into kaon and pion data set. Training variables are all Cleo-Cones and KSFW-moments calculated from the final state daughter particles of the B-meson counted as signal. The full list of input variables and their correlations can be found in Fig. A.1 for B-events and Fig. A.2 for $\bar{q}q$-events. The variables "deltaE" and "Mbc" were excluded from the training and are shown to ensure they are minimally biased. The importance of the features, ranked by the average loss reduction a selection on a feature introduces when it is used to split the trees, is depicted in Fig. A.3. The best variable, as measured by the average reduction of the loss function when using the variable, is the ratio of

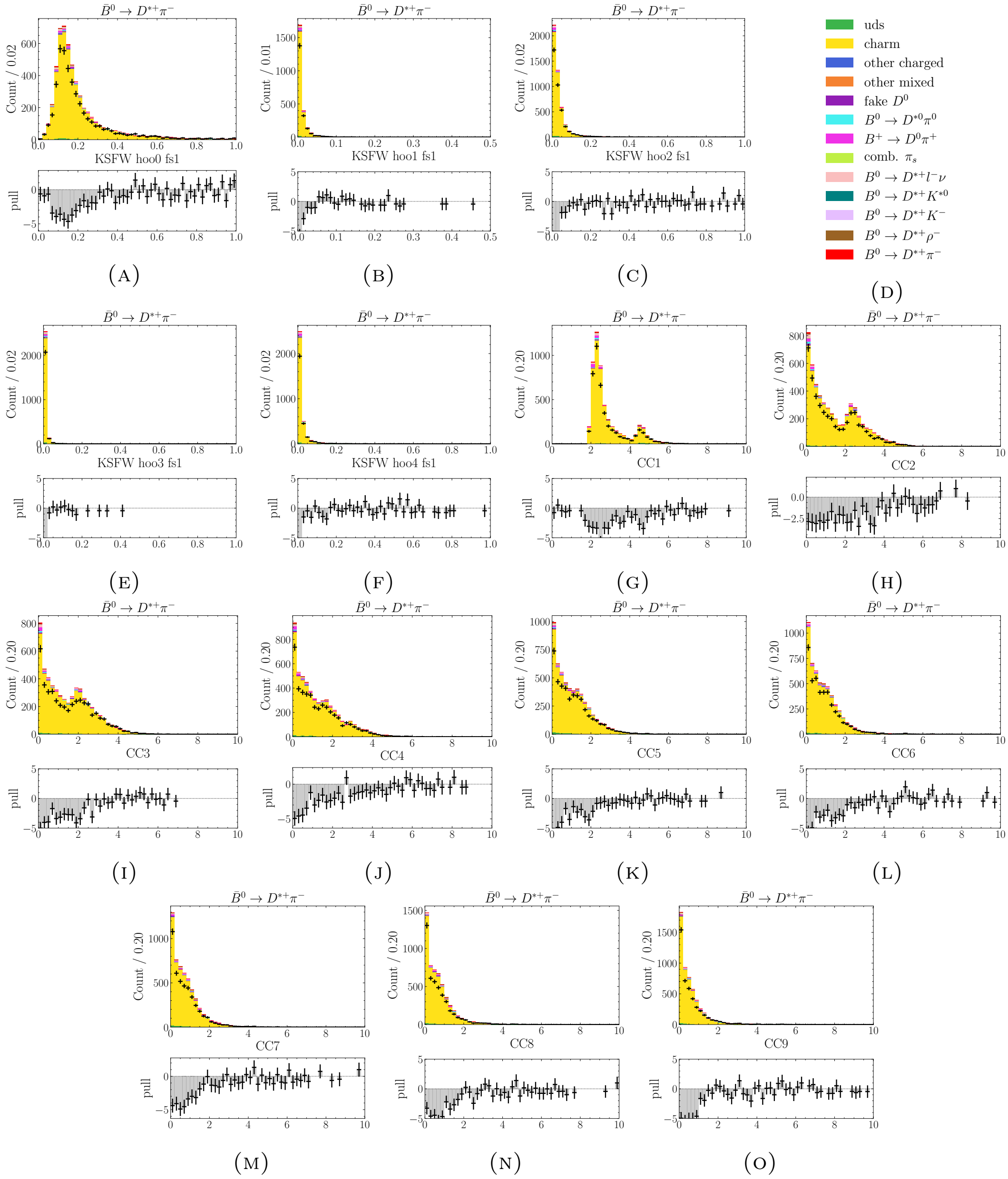


FIGURE A.7: Variables used for the continuum suppression classifier in the control region $M_{\mathrm{bc}} < 5270$ MeV for the $D^0 \to K^-\pi^+$ channel. KSFW-moments are abbreviated with KSFW and Cleo Cones are abbreviated as CC followed by the number.

the first and second Fox-Wolfram-Moments, R_2. The hyper parameters chosen for the training were optimised using a grid search method and are listed in Tab. A.1. The parameter optimisation suggested to use deeper trees with a tree depth of ≈ 8, however, a smaller value was chosen to avoid correlations of the classifier output and the signal variable “deltaE” in the signal window. To optimise the threshold point in the classifier output we define a figure of merit as

$$\mathrm{FOM} = \frac{S}{\sqrt{S+B}}, \tag{A.3}$$

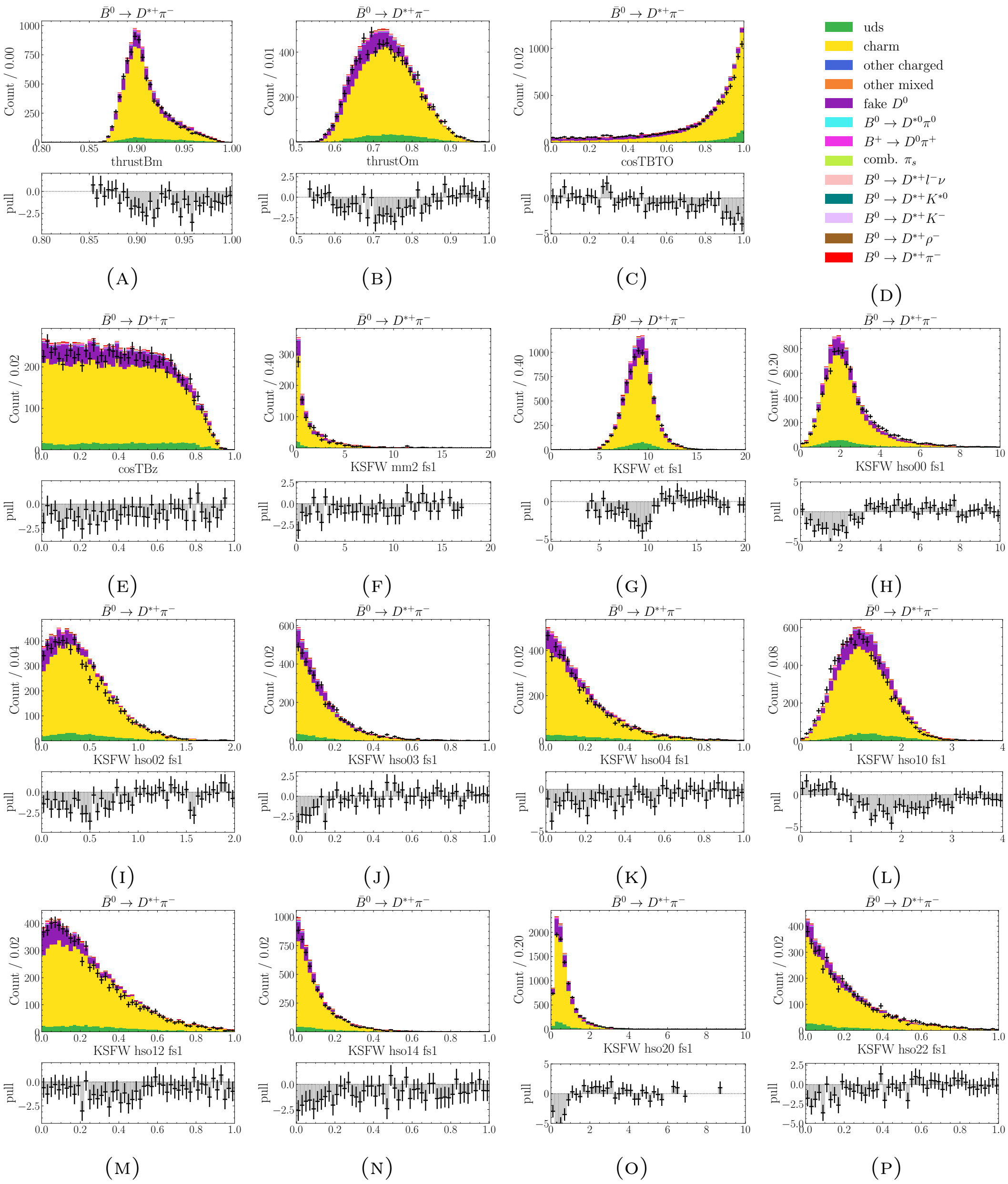


FIGURE A.8: Variables used for the continuum suppression classifier in the control region $M_{\text{bc}} < 5270$ MeV for the $D^0 \to K^-\pi^+$ channel. KSFW-moments are abbreviated with KSFW and Cleo Cones are abbreviated as CC followed by the number.

where S refers to the number of $\bar{B}B$ events and B to the number of continuum events. We take the maximum of this expression as the optimal threshold point. Due to the dependency on the signal fraction we choose to optimise the continuum suppression for the two D^0 decay modes separately, see Fig. A.10 and Fig. A.11. In principle a classifier can introduce bias and sculpt the signal variable. To ensure the absence of these effects the efficiency for the different components is plotted in bins of the signal variable, Fig. A.12. The efficiency of the individual components, defined as the number of true events of a component that survive the optimised selection on the classifier output, normalised to the total number of events in that component, was found to be constant within the statistical uncertainties. Furthermore, analysing

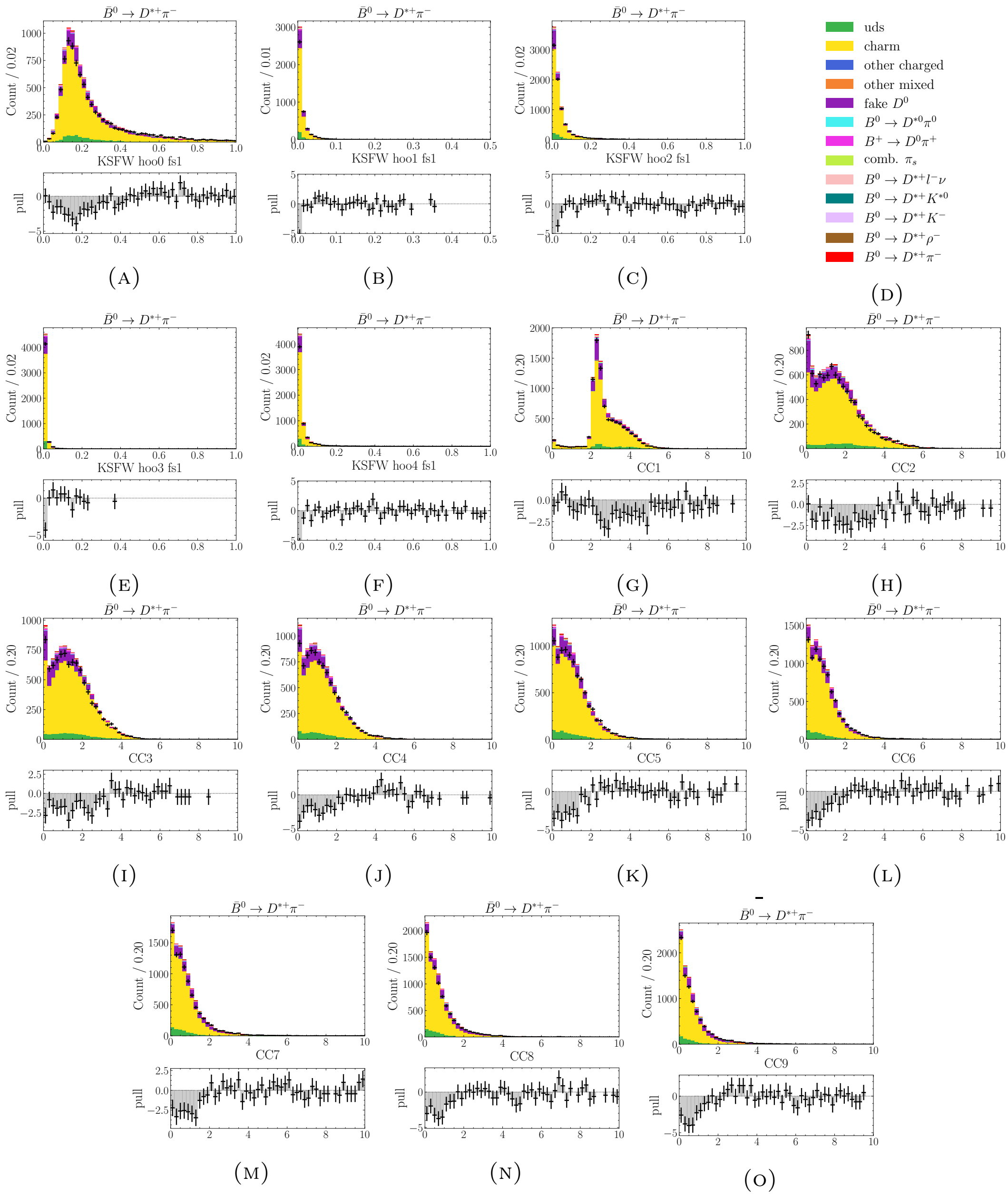
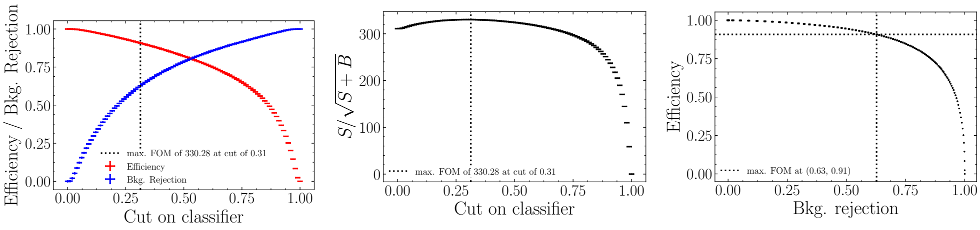


FIGURE A.9: Variables used for the continuum suppression classifier in the control region $M_{\mathrm{bc}} < 5270$ MeV for the $D^0 \to K^-2\pi^+\pi^-$ channel. KSFW-moments are abbreviated with KSFW and Cleo Cones are abbreviated as CC followed by the number.

the two dimensional histograms of the signal variable and the classification output, depicted in Fig. A.13 for the $D^0 \to K^-\pi^+$ channel and Fig. A.14 for the $D^0 \to K^-2\pi^+\pi^-$ channel has not shown any non-linear correlations or sculpturing.

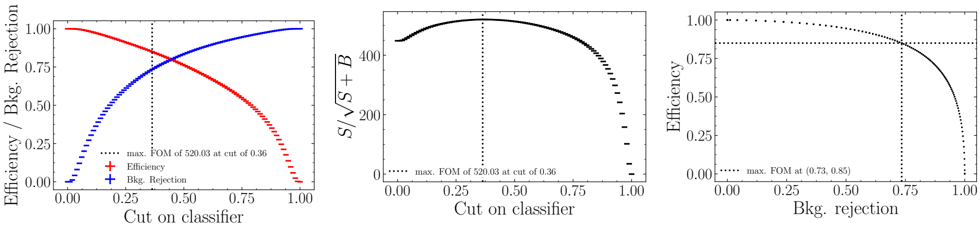
A.1.4 Performance of the continuum suppression

The performance of the continuum suppression algorithm is measured by the change of the signal resolution obtained by the fits with continuum suppression applied. Overall a small improvement for the kaon channel was found, see Fig. A.15 and Fig. A.16. The total effect on the ratio, Fig. A.17, is $\approx 3\%$ and too small to justify applying



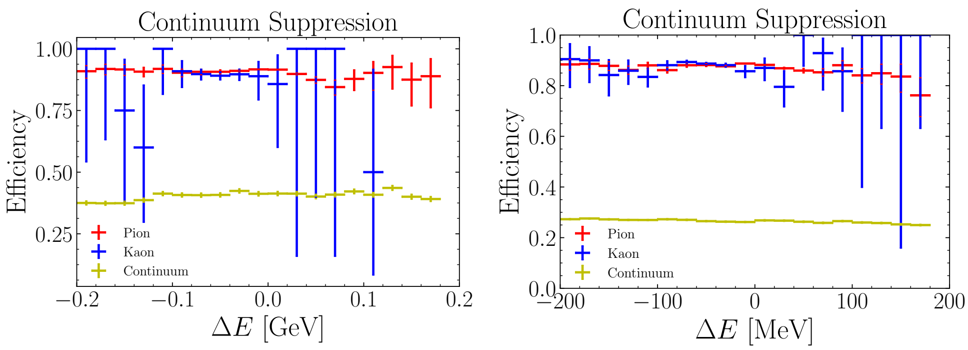
(A) (B) (C)

FIGURE A.10: Receiver operating characteristics of the continuum suppression classifier in the $D^0 \to K^-\pi^+$ mode. Efficiency and signal refer to $\bar{B}B$ events.



(A) (B) (C)

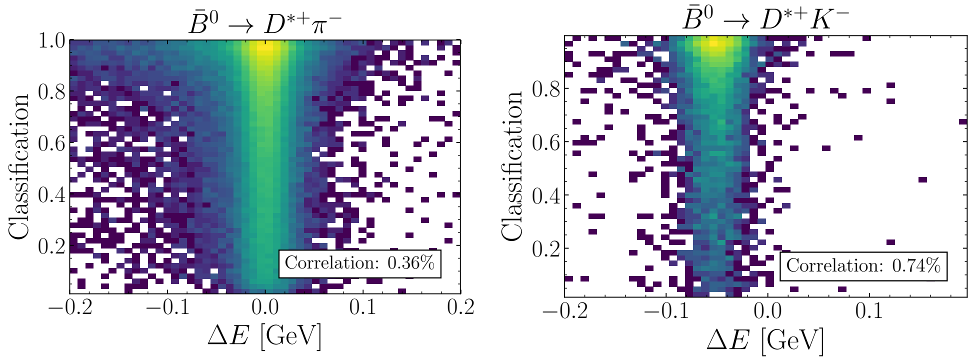
FIGURE A.11: Receiver operating characteristics of the continuum suppression classifier in the $D^0 \to K^-2\pi^+\pi^-$ mode. Efficiency and signal refer to $\bar{B}B$ events.



(A) (B)

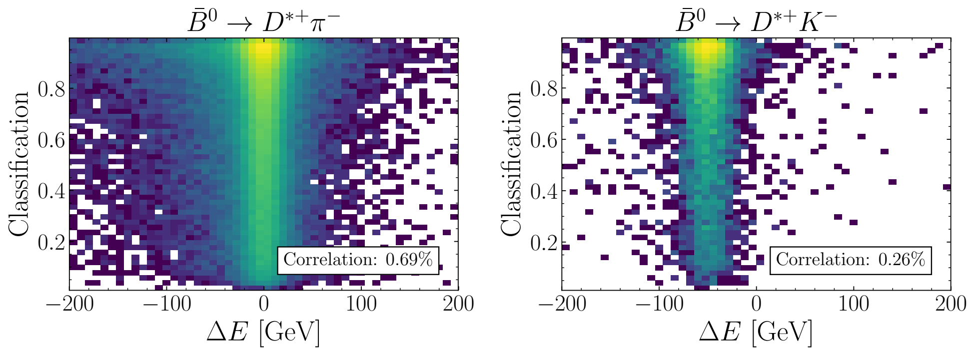
FIGURE A.12: Efficiency of the continuum classification by component in bins or the signal variable for the $D^0 \to K^-\pi^+$ mode (a) and the $D^0 \to K^-2\pi^+\pi^-$ mode (b). Efficiencies are defined as the number of true events of a component that survive the optimised selection on the classifier output, normalised to the total number of events in that component.

the continuum suppression and risk eventual additional biases that would have to be accounted for with a systematic uncertainty. The potential reason for the weak effect can be attributed to the well defined signal peaks and general small signal over background ratio. The main driver of the fit uncertainty is due to B-like backgrounds and a suppression of continuum backgrounds does not improve the resolution significantly, as its relative abundance is already very small. Therefore the decision was made not to employ continuum suppression for the signal fits in this analysis. However, the continuum suppression will be used as a tool to isolate B-like backgrounds when



(A) (B)

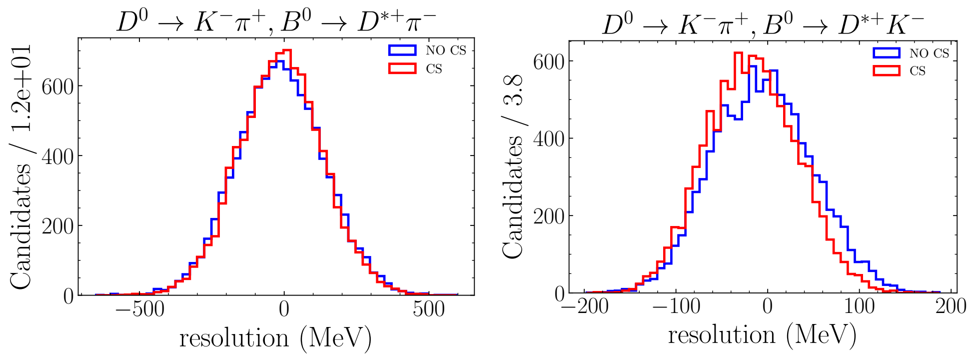
FIGURE A.13: Correlation check of the signal variable with the classifier output for the $D^0 \to K^-\pi^+$ mode.



(A) (B)

FIGURE A.14: Correlation check of the signal variable with the classifier output for the $D^0 \to K^-2\pi^+\pi^-$ mode.

comparing data and MC in Sec. 5.11.



(A) (B)

FIGURE A.15: Distributions of the resolutions, defined as the truth matched yield minus the signal yield determined by the fit, found by the toy studies with and without continuum suppression (a) for the pion mode and (b) for the kaon mode in $D^0 \to K^-\pi^+$.

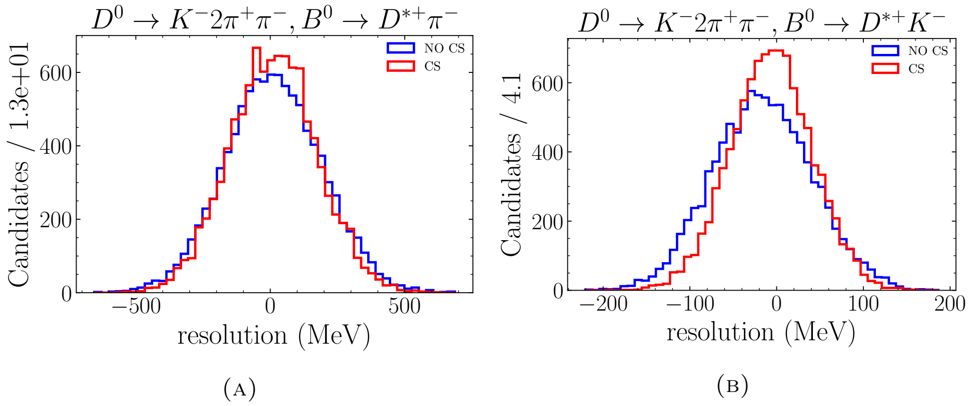


FIGURE A.16: Distributions of the resolutions, defined as the truth matched minus the fit signal yield, found by the toy studies with and without continuum suppression (a) for the pion mode and (b) for the kaon mode in $D^0 \to K^-2\pi^+\pi^-$.

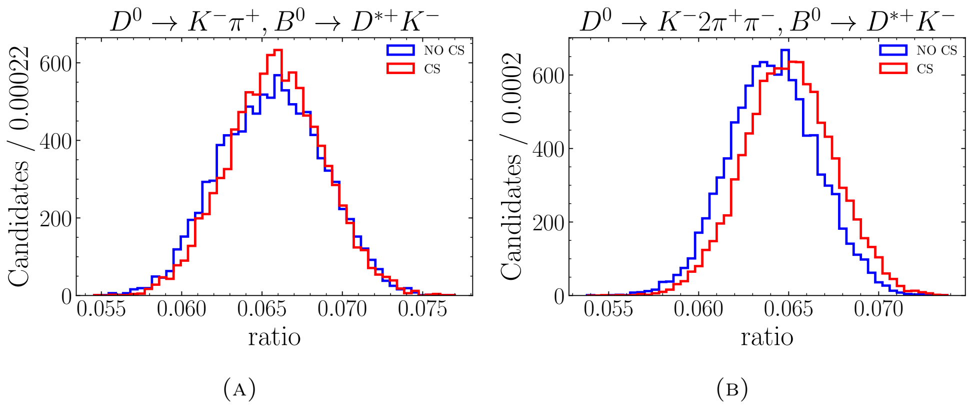


FIGURE A.17: Distributions of the resolutions, defined as the truth matched minus the fit result, of the ratio found by the toy studies with and without continuum suppression (a) for $D^0 \to K^-\pi^+$ and (b) $D^0 \to K^-2\pi^+\pi^-$.

Appendix B

Measuring low momentum pion efficiencies using a K_S^0 sample

Previous studies of pion identification at Belle have utilised an inclusive D^* sample [57] to study pion and kaon identification performance over a broad phase-space. However, these studies have been unable to measure the low momentum regions that are populated in multi-body and inclusive decays. This work presents a complimentary study of pion identification at low momentum via $K_S^0 \to \pi^+\pi^-$ decays.

This study uses the full on-resonance data-set corresponding to 702.6 fb^{-1}. Due to the changes made to the detector when upgrading the SVD, the sample is analysed separately for both SVD configurations. To prepare the analysis and study the performance in Monte Carlo we use the standard $\bar{B}B$ and continuum samples of the last official MC campaign (`mixed, charged, uds, charm`). Three streams are used for SVD1; two for SVD2, to balance the total number of events in SVD1 and SVD2.

B.1 Procedure

This analysis is performed in `basf2` via `b2bii` using *release-04-02-01*. To reconstruct the sample we search for tracks originating near the interaction point, requiring $|dr| <$ 2 cm and $|dz| <$ 5 cm. Oppositely charged tracks are combined to reconstruct K_S^0 candidates if the invariant mass of the pair is between $0.4 - 0.6$ GeV. A vertex fit is performed using `treeFitter` [75], where we require $P(\chi^2) > 0.001$. Finally the mass selection of the updated candidate is tightened to $0.45 < M < 0.55$ GeV and we require the cosine of the angle between the momentum and vertex vector of the K_S^0 to be above 0.998 to reject the majority of combinatorial background. We define the same θ bins as in Ref. [57]. For momentum bins we set $\{0, 0.2, 0.3, 0.4, 0.5, ..., 1.5\}$ GeV, as the bin boundaries.
We only measure for one selection on the likelihood ratio $L(K/\pi) > 0.6$ and assume $L(K/\pi) > 0.6$ is used for kaon identification and $L(K/\pi) < 0.6$ for pion selection. In each bin positive and negative tracks are combined with candidates where both tracks are found in a single bin excluded to avoid double counting K_S^0 candidates. The signal is modelled as a triple Gaussian where the parameters are first extracted by fitting the truth-matched MC distribution in the full region with no likelihood selection applied. The background is modelled as a second degree Chebyshev polynomial with free parameters for each bin. When fitting either side of the likelihood selection the mean of the Gaussians is allowed to float separately and the widths share a common fudge-factor to allow for differences between MC and Data for each bin and selection region. The ratios between the Gaussian PDFs are fixed to the values found from MC previously. Example fits are depicted in Fig. B.1. The efficiency, ϵ, is measured

directly using the simultaneous fit functions

$$\begin{aligned} F_{L(K/\pi)>0.6} &= N_{\text{bkg1}} \text{Ch}^2_{\text{bkg1}}(a_0, a_1) \\ &\quad + \epsilon N_{\text{tot}}(r_1 \text{G}_1(\mu, \beta \cdot \sigma_1) + r_2 \text{G}_2(\mu, \beta \cdot \sigma_2) + r_3 \text{G}_3(\mu, \beta \cdot \sigma_3)), \end{aligned} \tag{B.1}$$

and

$$\begin{aligned} F_{L(K/\pi)<0.6} &= N_{\text{bkg2}} \text{Ch}^2_{\text{bkg2}}(a_0, a_1) \\ &\quad + [1 - \epsilon] N_{\text{tot}}(r_1 \text{G}_1(\mu, \beta \cdot \sigma_1) + r_2 \text{G}_2(\mu, \beta \cdot \sigma_2) + r_3 \text{G}_3(\mu, \beta \cdot \sigma_3)), \end{aligned} \tag{B.2}$$

where N_i are yields, r_i fixed ratios, Ch^2 second degree Chebyshev Polynomials with free parameters and G_i are Gaussian functions with free shared mean μ, fixed widths σ_i and a shared "fudge-factor" β. For MC, the truth matched efficiency is used and the fit is only done to determine small biases introduced by the fitting procedure, which will be added as an uncorrelated systematic uncertainty. The biases are added in quadrature to systematic uncertainties associated to the PDF parameters. These are evaluated using a toy MC method, where the fixed ratios, r_i in Eq. B.1 and Eq. B.2, and widths, σ_i, of the Gaussians are randomly drawn from Gaussian PDFs. The widths of these are defined by the uncertainties of the ratios. In total 200 toy experiments are performed for each bin and it is required that the fits must have converged without errors. Systematic uncertainties are propagated to the data-MC correction factor, $r_{\text{data}-\text{MC}} = \epsilon_{\text{data}}/\epsilon_{MC}$, assuming full correlations.

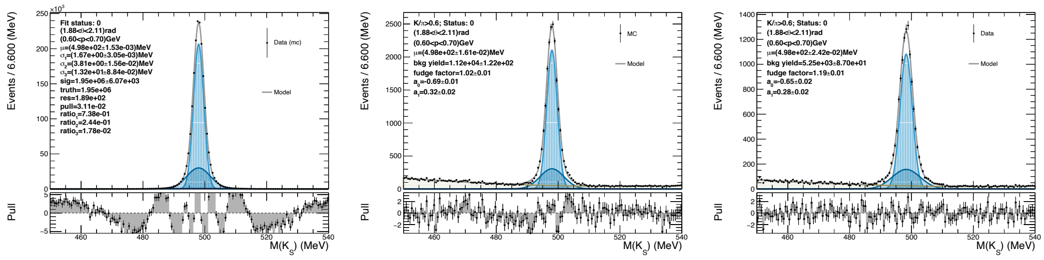


FIGURE B.1: Extracted PDF (a) and fit results for MC (b) and data (c) in the bin $p \in [0.6, 0.7]$ GeV and $\theta \in [1.88, 2.11]$ rad for $L(K/\pi) > 0.6$ and the SVD2 detector configuration.

B.2 Results

In general we observe good fits to data. In some bins the the fit functions fail to perfectly model the signal peak as can be seen in the pull distributions. However, we do not add more degrees of freedom to the signal model, because it is impossible to optimise a single model for all bins in an automated way. Furthermore, we explore re-parametrisations with additional width terms for specific bins and find that the changes are of negligible size. We compare our results to the existing tables where available and find good agreement, see Fig. B.2, Fig. B.3, Fig. B.4 Fig. B.5, Fig. B.6 and Fig. B.7. For the analysis the new weights are used in bins where the D^*-samples provided by the performance group do not have measurements. This work was done in collaboration with Marcel Hohmann.

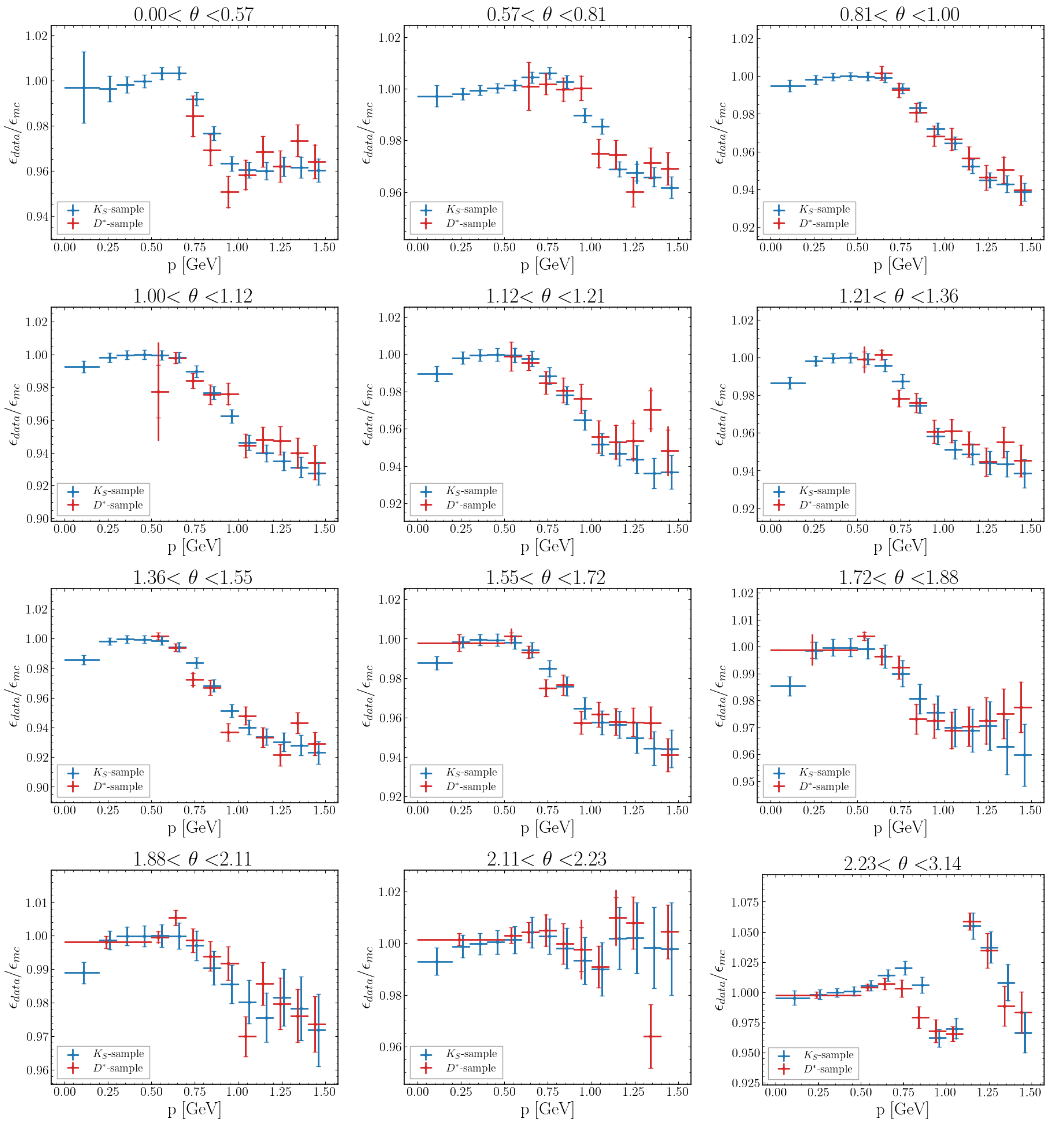


FIGURE B.2: Ratio of pion ID efficiencies in data over MC for SVD1.

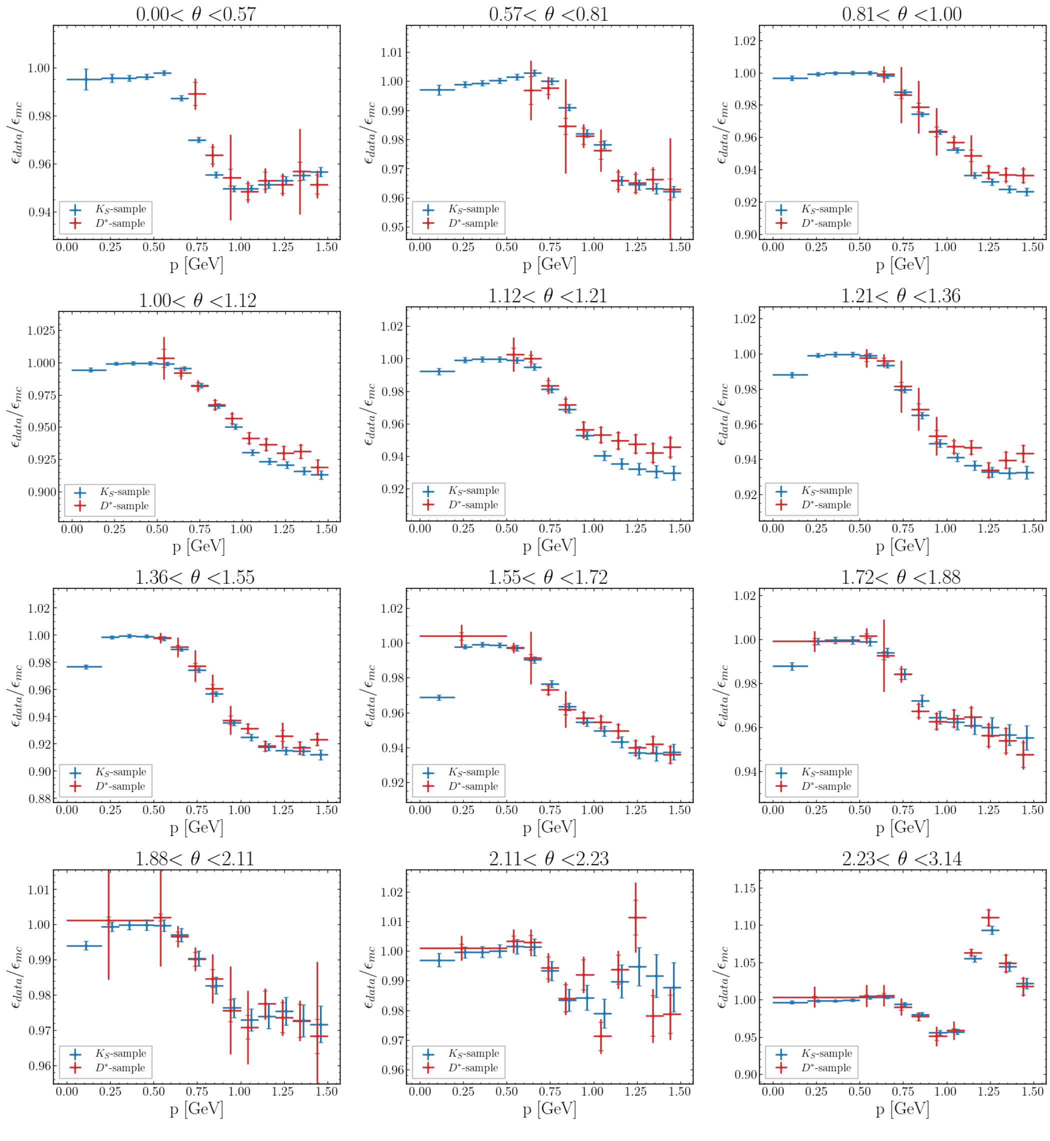


FIGURE B.3: Ratio of pion ID efficiencies in data over MC for SVD2.

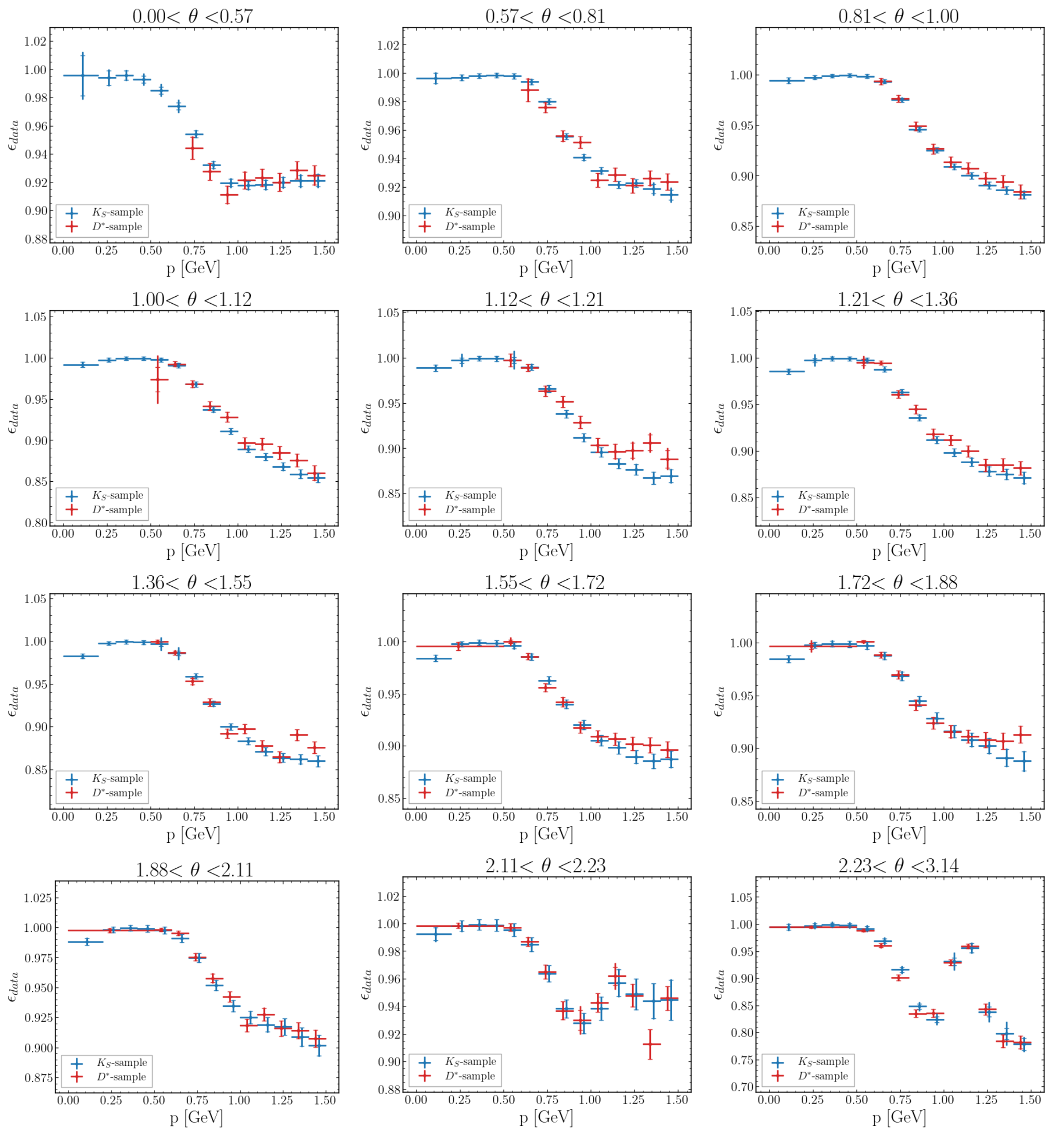


FIGURE B.4: Pion ID efficiencies in data for SVD1.

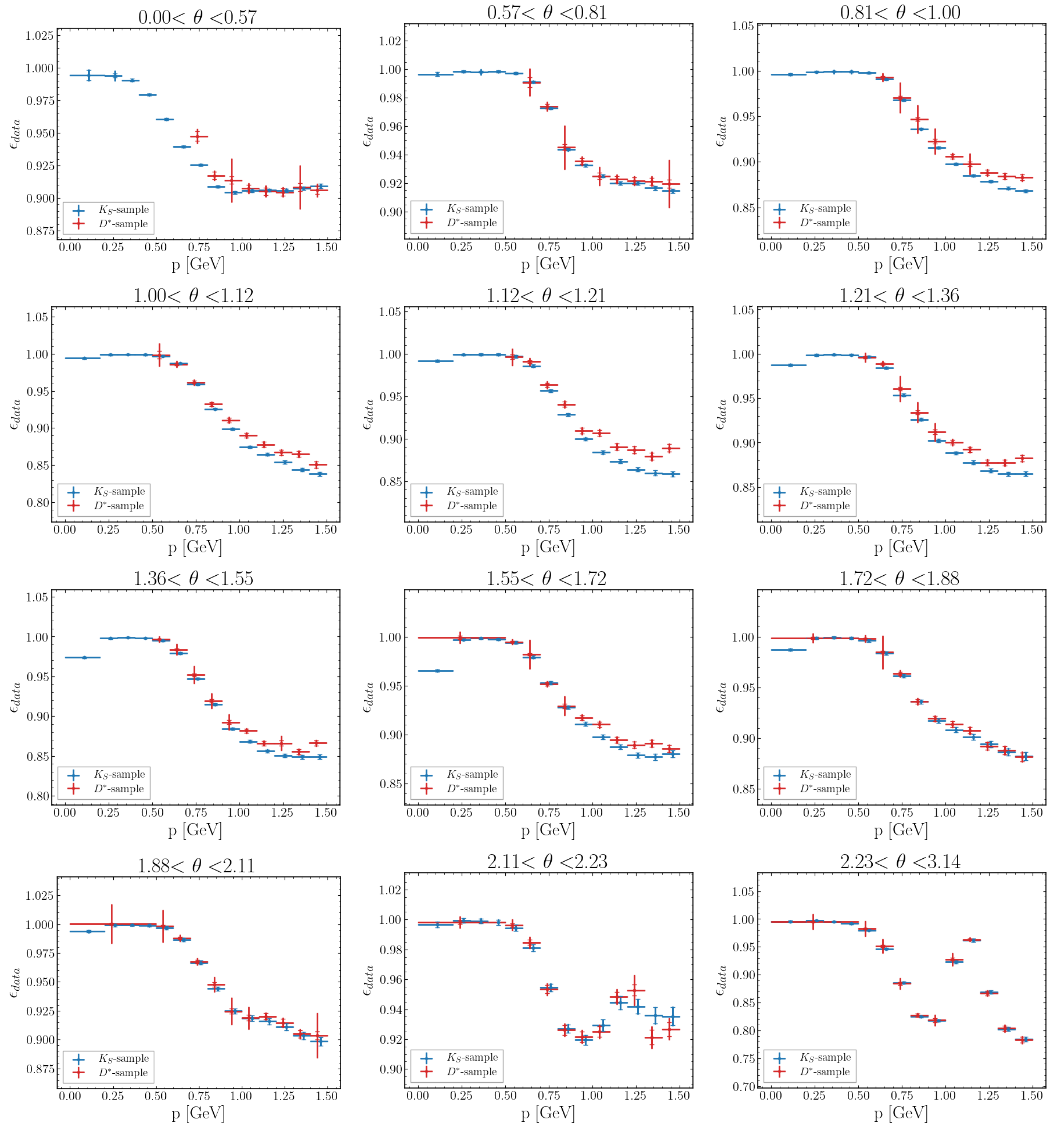


FIGURE B.5: Pion ID efficiencies in data for SVD2.

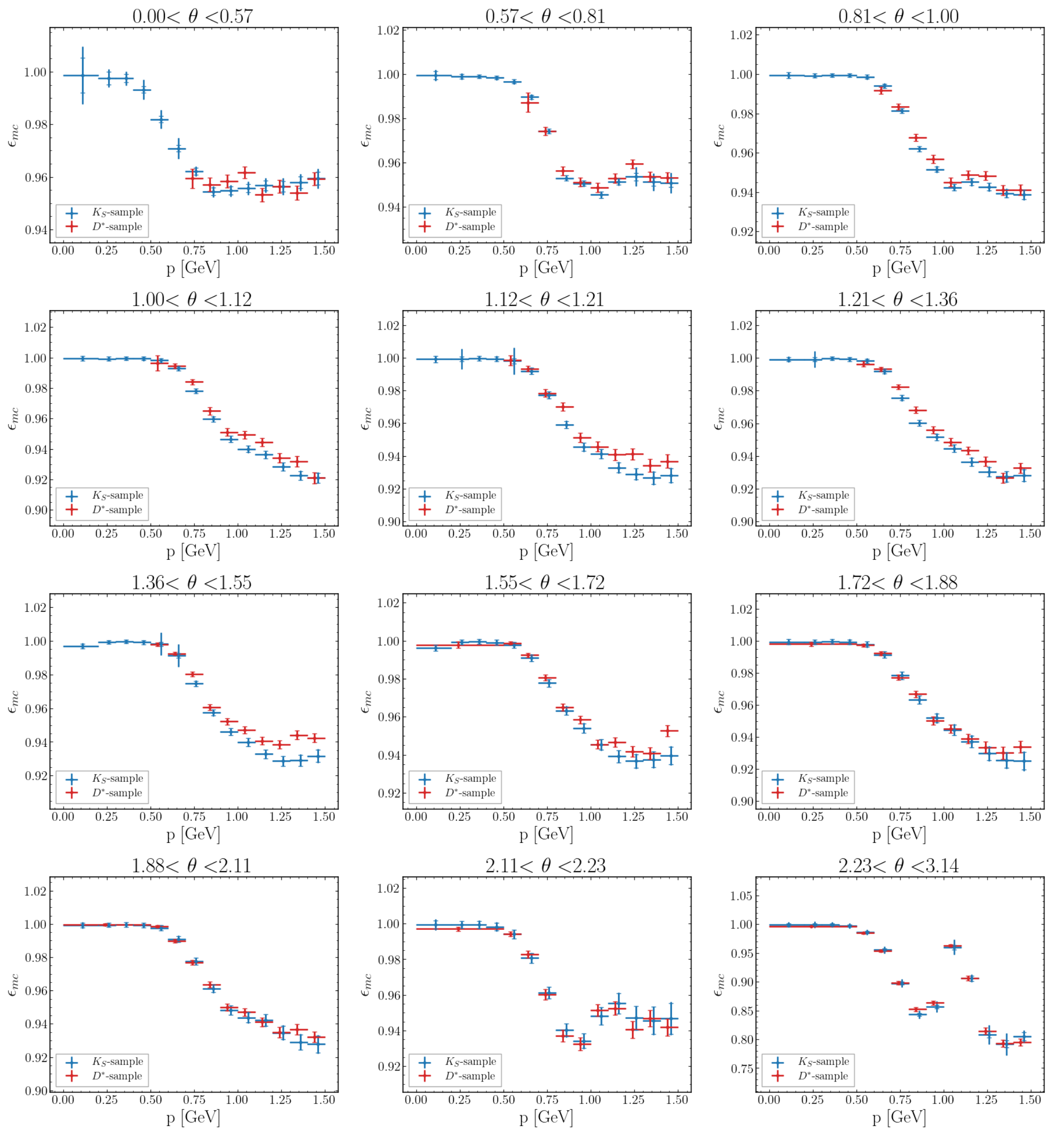


FIGURE B.6: Pion ID efficiencies in MC for SVD1.

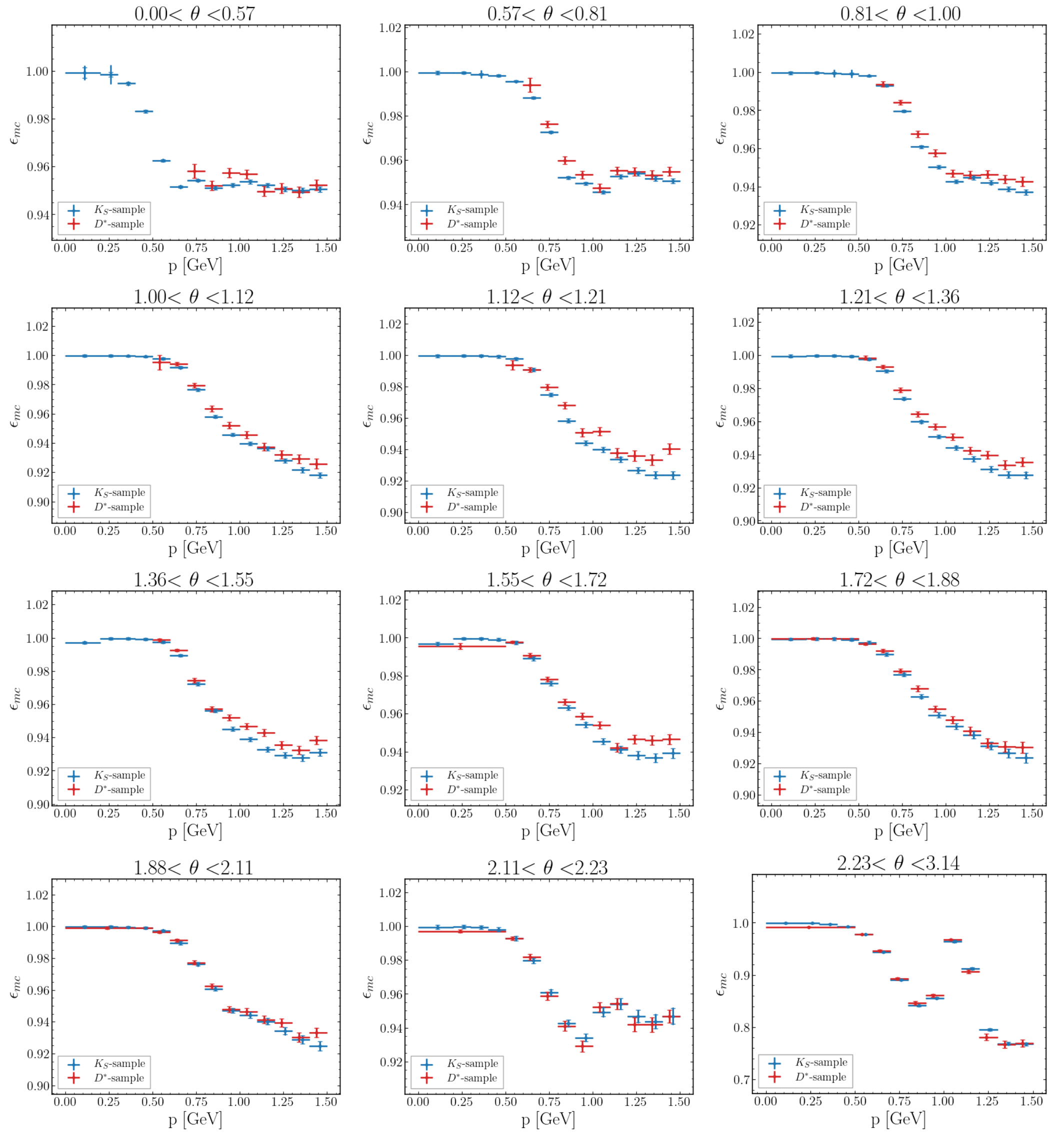


FIGURE B.7: Pion ID efficiencies in MC for SVD2.

Appendix C

Probability density functions

The PDFs used in this thesis are defined in detail here. All parameters shown in the equations can be extracted in the fits.

Gaussian PDF The Gaussian PDF for a single parameter is defined as

$$f(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma}\right). \tag{C.1}$$

Crystal Ball PDF The Crystal Ball function [80] is a modified Gaussian that is used to describe processes with radiative losses. It is defined as

$$f(x|\mu,\sigma,\alpha,n) = N \cdot \begin{cases} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma}\right) & \text{if } \frac{(x-\mu)}{\sigma} > -\alpha \\ A \cdot (B - \frac{(x-\mu)}{\sigma})^{-n} & \text{if } \frac{(x-\mu)}{\sigma} \leq -\alpha \end{cases} \tag{C.2}$$

using

$$\begin{aligned} A &= \left(\frac{n}{|\alpha|}\right)^n \exp\left(-\frac{1}{2}\alpha^2\right), \\ B &= \frac{n}{|\alpha|} - |\alpha|, \end{aligned} \tag{C.3}$$

and

$$\begin{aligned} N &= \frac{1}{\sigma(C+D)}, \\ C &= \frac{n}{|\alpha|}\frac{1}{n-1} \exp\left(-\frac{1}{2}\alpha^2\right), \\ D &= \sqrt{\frac{\pi}{2}}\left(1 + \operatorname{erf}\left(\frac{|\alpha|}{\sqrt{2}}\right)\right). \end{aligned} \tag{C.4}$$

Chebyshev Polynomials Chebyshev Polynomials [81] are defined via a recursive relation for a random variable $u \in [-1,1]$, $u \in \mathbb{R}$, such that

$$T_{n+1}(u) = 2uT_n(u) - T_{n-1}(u), \tag{C.5}$$

with $T_0 = 1$ and $T_1 = u$. The polynomial PDF then is build from these terms by adding coefficients

$$f(u|C_n) = 1 + \sum_{n>0} c_n T_n(u), \tag{C.6}$$

where $c \in [-1,1]$. The transformation from any standard random variable $x \to u$ is given by

$$f(x) = \frac{x - \frac{1}{2}(x_{max} + x_{min})}{\frac{1}{2}(x_{max} - x_{min})}. \tag{C.7}$$

Empirical ΔM background function When reconstructing D^{*+}-mesons the combinatorial backgrounds have a shape that was found to follow an exponential function, such that

$$D(x|\Delta M_0, a, b, c) = \left[1 - \exp\left(\frac{(x - \Delta M_0)}{c}\right)\right]\left(\frac{x}{\Delta M_0}\right)^a + b\left(\frac{x}{\Delta M_0} - 1\right). \quad \text{(C.8)}$$

Bibliography

[1] E. Dogen. *Shobogenzo, Die Schatzkammer des wahren Dharma-Auges.* Ed. by Ritsunen G. Linnebach and Gudo Wafu Nishijima-Roshi. Werner Kristkeitz Verlag, 2013. ISBN: 978-3-921508-90-9.

[2] A. Manohar and M. Wise. *Heavy Quark Physics.* Cambridge, 2000. ISBN: 0-521-64241-8.

[3] M. Beneke et al. "QCD factorization for exclusive, non-leptonic B meson decays: General arguments and the case of heavy-light final states". In: *arXiv:hep-ph/0006124v1* (2000).

[4] F. Englert and R. Brout. "Broken symmetry and the mass of gauge vector mesons". In: *Phys. Rev. Lett.* 13 (9 Aug. 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[5] P. W. Higgs. "Broken symmetries and the masses of gauge bosons". In: *Phys. Rev. Lett.* 13 (16 Oct. 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[6] ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Physics Letters B* 716.1 (2012), pp. 1–29. DOI: https://doi.org/10.1016/j.physletb.2012.08.020. URL: http://www.sciencedirect.com/science/article/pii/S037026931200857X.

[7] CMS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Physics Letters B* 716.1 (2012), pp. 30–61. DOI: https://doi.org/10.1016/j.physletb.2012.08.021. URL: http://www.sciencedirect.com/science/article/pii/S0370269312008581.

[8] R. Davis, D. S. Harmer, and K. C. Hoffman. "Search for neutrinos from the sun". In: *Phys. Rev. Lett.* 20 (21 May 1968), pp. 1205–1209. DOI: 10.1103/PhysRevLett.20.1205. URL: https://link.aps.org/doi/10.1103/PhysRevLett.20.1205.

[9] Belle Collaboration. "Observation of Large CP-violation in the neutral B meson system". In: *Phys. Rev. Lett.* 87 (9 Aug. 2001), p. 091802. DOI: 10.1103/PhysRevLett.87.091802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.87.091802.

[10] BABAR Collaboration. "Observation of CP-violation in the B^0 meson system". In: *Phys. Rev. Lett.* 87 (9 Aug. 2001), p. 091801. DOI: 10.1103/PhysRevLett.87.091801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.87.091801.

[11] R. Fleischer, N. Serra, and N. Tuning. "Tests of factorization and $SU(3)$ relations in B decays into heavy-light final states". In: *Phys. Rev.* D83 (2011), p. 014017. DOI: 10.1103/PhysRevD.83.014017. arXiv: 1012.2784 [hep-ph].

[12] A. Pich and E. de Rafael. "Strong CP-violation in an effective chiral lagrangian approach". In: *Nuclear Physics B* 367.2 (1991), pp. 313–333. DOI: https://doi.org/10.1016/0550-3213(91)90019-T. URL: http://www.sciencedirect.com/science/article/pii/055032139190019T.

[13] E. Kou et al. "The Belle II physics book". In: *Progress of Theoretical and Experimental Physics* 2019.12 (Dec. 2019). DOI: 10.1093/ptep/ptz106. URL: http://dx.doi.org/10.1093/ptep/ptz106.

[14] G. Buchalla, A. J. Buras, and M. E. Lautenbacher. "Weak decays beyond leading logarithms". In: *Rev. Mod. Phys.* 68 (4 Oct. 1996), pp. 1125–1244. DOI: 10.1103/RevModPhys.68.1125. URL: https://link.aps.org/doi/10.1103/RevModPhys.68.1125.

[15] T. Huber, S. Kränkl, and X.-Q. Li. "Two-body non-leptonic heavy-to-heavy decays at NNLO in QCD factorization". In: *JHEP* 09 (2016), p. 112. DOI: 10.1007/JHEP09(2016)112. arXiv: 1606.02888 [hep-ph].

[16] D. Bigi and P. Gambino. "Revisiting $B \to D\ell\nu$". In: *Phys. Rev. D* 94.9 (2016), p. 094008. DOI: 10.1103/PhysRevD.94.094008. arXiv: 1606.08030 [hep-ph].

[17] I. Caprini, L. Lellouch, and M. Neubert. "Dispersive bounds on the shape of $\bar{B} \to D^{(*)}$ lepton anti-neutrino form-factors". In: *Nucl. Phys. B* 530 (1998), pp. 153–181. DOI: 10.1016/S0550-3213(98)00350-2. arXiv: hep-ph/9712417.

[18] M. Bordone et al. "Heavy-Quark expansion for $\bar{B}_s \to D_s^{(*)}$ form factors and unitarity bounds beyond the $SU(3)_F$ limit". In: *Eur. Phys. J. C* 80.4 (2020), p. 347. DOI: 10.1140/epjc/s10052-020-7850-9. arXiv: 1912.09335 [hep-ph].

[19] M. Bordone et al. "A puzzle in $\bar{B}^0_{(s)} \to D^{(*)+}_{(s)}\{\pi^-, K^-\}$ decays and extraction of the f_s/f_d fragmentation fraction". In: *The European Physical Journal C* 80.10 (Oct. 2020). DOI: 10.1140/epjc/s10052-020-08512-8. URL: http://dx.doi.org/10.1140/epjc/s10052-020-08512-8.

[20] S. Iguro and T. Kitahara. "Implications for new physics from novel puzzle in $\bar{B}^0_{(s)} \to D^{(*)+}_{(s)}\{\pi^-, K^-\}$ decays". In: *Physical Review D* 102.7 (Oct. 2020). DOI: 10.1103/physrevd.102.071701. URL: http://dx.doi.org/10.1103/PhysRevD.102.071701.

[21] G. Aad et al. "Search for new resonances in mass distributions of jet pairs using 139 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Journal of High Energy Physics* 2020.3 (Mar. 2020). DOI: 10.1007/jhep03(2020)145. URL: http://dx.doi.org/10.1007/JHEP03(2020)145.

[22] M. Aaboud et al. "Search for vector-boson resonances decaying to a top quark and bottom quark in the lepton plus jets final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Physics Letters B* 788 (Jan. 2019), pp. 347–370. DOI: 10.1016/j.physletb.2018.11.032. URL: http://dx.doi.org/10.1016/j.physletb.2018.11.032.

[23] A. M. Sirunyan et al. "Search for narrow and broad dijet resonances in proton-proton collisions at $\sqrt{s} = 13$ TeV and constraints on dark matter mediators and other new particles". In: *Journal of High Energy Physics* 2018.8 (Aug. 2018). DOI: 10.1007/jhep08(2018)130. URL: http://dx.doi.org/10.1007/JHEP08(2018)130.

[24] A. M. Sirunyan et al. "Search for resonant $t\bar{t}$ production in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Journal of High Energy Physics* 2019.4 (Apr. 2019). DOI: 10.1007/jhep04(2019)031. URL: http://dx.doi.org/10.1007/JHEP04(2019)031.

[25] A. M. Sirunyan et al. "Search for high mass dijet resonances with a new background prediction method in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Journal of High Energy Physics* 2020.5 (May 2020). DOI: 10.1007/jhep05(2020)033. URL: http://dx.doi.org/10.1007/JHEP05(2020)033.

[26] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev. D* 98 (3 Aug. 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[27] T. Browder and K. Honscheid. "B mesons". In: *Progress in Particle and Nuclear Physics* 35 (1995), pp. 81–219. DOI: https://doi.org/10.1016/0146-6410(95)00042-H. URL: http://www.sciencedirect.com/science/article/pii/014664109500042H.

[28] W. Herr and B. Muratori. *The concept of luminosity.* 2006. DOI: 10.5170/CERN-2006-002.361. URL: https://cds.cern.ch/record/941318.

[29] K. Akai et al. "RF systems for the KEK B factory". In: *Nucl. Instrum. Meth. A* 499 (2003), pp. 45–65. DOI: 10.1016/S0168-9002(02)01773-4.

[30] A. Abashian et al. "The Belle detector". In: *Nucl. Instrum. Meth. A* 479 (2002), pp. 117–232. DOI: 10.1016/S0168-9002(01)02013-7.

[31] The Belle II collaboration. "Belle II Technical Design Report". In: *arxiv.org* (Nov. 2010). eprint: 1011.0352.

[32] M. Gelb, T. Keck, et al. "B2BII: Data conversion from Belle to Belle II". In: *Computing and Software for Big Science* 2.1 (Nov. 2018). DOI: 10.1007/s41781-018-0016-x. URL: http://dx.doi.org/10.1007/s41781-018-0016-x.

[33] D. J. Lange. "The EvtGen particle decay simulation package". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 462.1 (2001). BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines, pp. 152–155. DOI: https://doi.org/10.1016/S0168-9002(01)00089-4. URL: http://www.sciencedirect.com/science/article/pii/S0168900201000894.

[34] T. Sjöstrand, S. Mrenna, and P. Skands. "PYTHIA 6.4 physics and manual". In: *Journal of High Energy Physics* 2006.05 (May 2006), pp. 026–026. DOI: 10.1088/1126-6708/2006/05/026. URL: http://dx.doi.org/10.1088/1126-6708/2006/05/026.

[35] E. Barberio, B. van Eijk, and Z. Was. "Photos a universal Monte Carlo for QED radiative corrections in decays". In: *Computer Physics Communications* 66.1 (1991), pp. 115–128. DOI: https://doi.org/10.1016/0010-4655(91)90012-A. URL: http://www.sciencedirect.com/science/article/pii/001046559190012A.

[36] R. Brun et al. *GEANT3*. Sept. 1987.

[37] S. Agostinelli et al. "Geant4 a simulation toolkit". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8. URL: http://www.sciencedirect.com/science/article/pii/S0168900203013688.

[38] J. Brodzicka, T. Browder, et al. *Physics achievements from the Belle experiment.* 2012. arXiv: `1212.5342 [hep-ex]`.

[39] V. Karimaki. "Effective circle fitting for particle trajectories". In: *Nucl. Instrum. Meth.* A305 (1991), pp. 187–191. DOI: `10.1016/0168-9002(91)90533-V`.

[40] N. Dawe. URL: `https://github.com/ndawe/tikz-track`.

[41] F. James. *Statistical methods in experimental physics.* 2006. ISBN: 978-981-270-527-3.

[42] D. Bortoletto et al. "Inclusive and exclusive decays of B mesons to final states including charm and charmonium mesons". In: *Phys. Rev.* D45 (1992), pp. 21–35. DOI: `10.1103/PhysRevD.45.21`.

[43] C. Bebek et al. "Exclusive decays and masses of the B mesons". In: *Phys. Rev.* D36 (1987), p. 1289. DOI: `10.1103/PhysRevD.36.1289`.

[44] R. Giles et al. "Two-body decays of B mesons". In: *Phys. Rev.* D30 (1984), p. 2279. DOI: `10.1103/PhysRevD.30.2279`.

[45] M. S. Alam et al. "Exclusive hadronic B decays to charm and charmonium final states". In: *Phys. Rev.* D50 (1994), pp. 43–68. DOI: `10.1103/PhysRevD.50.43`. arXiv: `hep-ph/9403295 [hep-ph]`.

[46] R. Akers et al. "Observation of exclusive decays of B mesons at LEP". In: *Phys. Lett.* B337 (1994), pp. 196–206. DOI: `10.1016/0370-2693(94)91469-9`.

[47] H. Albrecht et al. "Reconstruction of B mesons". In: *Phys. Lett.* B185 (1987), p. 218. DOI: `10.1016/0370-2693(87)91558-9`.

[48] H. Albrecht et al. "Determination of the branching ratio for the decay $B^0 \to D^{*-}\pi^+$". In: *Phys. Lett.* B182 (1986), pp. 95–100. DOI: `10.1016/0370-2693(86)91084-1`.

[49] H. Albrecht et al. "Exclusive hadronic decays of B mesons". In: *Z. Phys.* C48 (1990), pp. 543–552. DOI: `10.1007/BF01614687`.

[50] C. Patrignani et al. "Review of Particle Physics". In: *Chin. Phys.* C40.10 (2016), p. 100001. DOI: `10.1088/1674-1137/40/10/100001`.

[51] B. Aubert et al. "Measurement of the absolute branching fractions $B \to D\pi, D^*\pi$ with a missing mass method". In: *Phys. Rev.* D74 (2006), p. 111102. DOI: `10.1103/PhysRevD.74.111102`. arXiv: `hep-ex/0609033 [hep-ex]`.

[52] B. Aubert et al. In: *Phys. Rev. Lett.* 87 (2001), p. 111801. DOI: `10.1103/PhysRevLett.87.111801`. arXiv: `hep-ex/0104051 [hep-ex]`.

[53] G. Brandenburg et al. "A new measurement of $\bar{B} \to D^*\pi$ branching fractions". In: *Phys. Rev. Lett.* 80 (1998), pp. 2762–2766. DOI: `10.1103/PhysRevLett.80.2762`. arXiv: `hep-ex/9706019 [hep-ex]`.

[54] B. Aubert et al. "Measurement of branching fractions and resonance contributions for $B^0 \to \bar{D}^0 K^+\pi^-$ and search for $B^0 \to D^0 K^+\pi^-$ decays". In: *Phys. Rev. Lett.* 96 (2006), p. 011803. DOI: `10.1103/PhysRevLett.96.011803`. arXiv: `hep-ex/0509036 [hep-ex]`.

[55] K. Abe et al. "Observation of Cabibbo suppressed $\bar{B} \to D^{(*)}K^-$ decays at BELLE". In: *Phys. Rev. Lett.* 87 (2001), p. 111801. DOI: `10.1103/PhysRevLett.87.111801`. arXiv: `hep-ex/0104051 [hep-ex]`.

[56] R. Aaij et al. "Study of $B^0 \to D^{*-}\pi^+\pi^-\pi^+$ and $B^0 \to D^{*-}K^+\pi^-\pi^+$ decays". In: *Phys. Rev.* D87.9 (2013), p. 092001. DOI: `10.1103/PhysRevD.87.092001`. arXiv: `1303.6861 [hep-ex]`.

[57] S. Nishida. “Study of kaon and pion identification using an inclusive D^* sample”. In: *Belle Internal* (Jan. 2005). Belle Note 779.

[58] D. Ferlewicz, P. Urquijo, and E. Waheed. “Revisiting fits to $B^0 \to D^{*-}l^+\nu_l$ to measure $|V_{cb}|$ with novel methods and preliminary LQCD data at non-zero recoil”. In: (Aug. 2020). arXiv: `2008.09341 [hep-ph]`.

[59] E. Waheed et al. “Measurement of the CKM matrix element $|V_{cb}|$ from $B^0 \to D^{*-}l^+\nu$ at Belle”. In: *Physical Review D* 100.5 (Sept. 2019). DOI: `10.1103/physrevd.100.052007`. URL: `http://dx.doi.org/10.1103/PhysRevD.100.052007`.

[60] A. Sirlin. “Large m_W, m_Z behaviour of the $O(\alpha)$ corrections to semileptonic processes mediated by W”. In: *Nuclear Physics B* 196.1 (1982), pp. 83–92. DOI: `https://doi.org/10.1016/0550-3213(82)90303-0`. URL: `http://www.sciencedirect.com/science/article/pii/0550321382903030`.

[61] C. Glenn Boyd, B. Grinstein, and R. F. Lebed. “Precision corrections to dispersive bounds on form factors”. In: *Physical Review D* 56.11 (Dec. 1997), pp. 6895–6911. DOI: `10.1103/physrevd.56.6895`. URL: `http://dx.doi.org/10.1103/PhysRevD.56.6895`.

[62] J. A. Bailey et al. “Update of $|V_{cb}|$ from the $\bar{B} \to D^* l\bar{\nu}$ form factor at zero recoil with three-flavor lattice QCD”. In: *Phys. Rev. D* 89 (11 June 2014), p. 114504. DOI: `10.1103/PhysRevD.89.114504`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.89.114504`.

[63] H. Li and S. Mishima. “Penguin pollution in the $B^0 \to J/\psi K_S$ decay”. In: *Journal of High Energy Physics* 2007.03 (Mar. 2007), pp. 009–009. DOI: `10.1088/1126-6708/2007/03/009`. URL: `http://dx.doi.org/10.1088/1126-6708/2007/03/009`.

[64] W. D. Hulsbergen. “Decay chain fitting with a Kalman filter”. In: *Nucl. Instrum. Meth.* A552 (2005), pp. 566–575. DOI: `10.1016/j.nima.2005.06.078`. arXiv: `physics/0503191 [physics]`.

[65] G. Guennebaud, J. Benoît, et al. *Eigen v3*. URL: `http://eigen.tuxfamily.org`.

[66] R. E. Kalman. “A new approach to linear filtering and prediction problems”. In: *Transactions of the ASME–Journal of Basic Engineering* 82.Series D (1960), pp. 35–45.

[67] R. Frühwirth. “Application of Kalman filtering to track and vertex fitting”. In: *Nucl. Instrum. Methods Phys. Res., A* 262.HEPHY-PUB-503 (June 1987), 444. 19 p. URL: `https://cds.cern.ch/record/178627`.

[68] R. Frühwirth and R. K. Bock. “Data analysis techniques for high-energy physics experiments”. In: *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.* 11 (2000). Ed. by H. Grote, D. Notz, and M. Regler, p. 252.

[69] P. Avery. *Applied Fitting Theory I, IV, VI.* 1999.

[70] T. Hara. “Computing at the Belle II experiment”. In: *Journal of Physics: Conference Series* 664.1 (Dec. 2015), p. 012002. DOI: `10.1088/1742-6596/664/1/012002`. URL: `https://doi.org/10.1088%5C%2F1742-6596%5C%2F664%5C%2F1%5C%2F012002`.

[71] M. Kowarschik and C. Weiß. “An overview of cache optimization techniques and cache-aware numerical algorithms”. In: *Algorithms for Memory Hierarchies — Advanced Lectures, volume 2625 of Lecture Notes in Computer Science.* Springer, 2003, pp. 213–232.

[72] N. Nethercote and J. Seward. "Valgrind: A framework for heavyweight dynamic binary instrumentation". In: *SIGPLAN Not.* 42.6 (June 2007), pp. 89–100. DOI: 10.1145/1273442.1250746. URL: http://doi.acm.org/10.1145/1273442.1250746.

[73] L. Lonnblad. "CLHEP: A project for designing a C++ class library for high-energy physics". In: *Comput. Phys. Commun.* 84 (1994), pp. 307–316. DOI: 10.1016/0010-4655(94)90217-8.

[74] Belle collaboration. "Evidence for direct CP violation in the decay $B^{\pm} \to D^{(*)}K^{\pm}, D \to K^0_{rmS}\pi^+\pi^-$ and measurement of the CKM phase ϕ_3". In: *Phys. Rev. D* 81 (11 June 2010), p. 112002. DOI: 10.1103/PhysRevD.81.112002. URL: https://link.aps.org/doi/10.1103/PhysRevD.81.112002.

[75] J.-F. Krohn, F. Tenchini, P. Urquijo, et al. "Global decay chain vertex fitting at Belle II". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 976 (2020), p. 164269. DOI: https://doi.org/10.1016/j.nima.2020.164269. URL: http://www.sciencedirect.com/science/article/pii/S0168900220306653.

[76] T. Chen and C. Guestrin. "XGBoost: A scalable tree boosting system". In: KDD '16 (2016), pp. 785–794. DOI: 10.1145/2939672.2939785. URL: http://doi.acm.org/10.1145/2939672.2939785.

[77] E. Farhi. "Quantum chromodynamics test for jets". In: *Phys. Rev. Lett.* 39 (25 Dec. 1977), pp. 1587–1588. DOI: 10.1103/PhysRevLett.39.1587. URL: https://link.aps.org/doi/10.1103/PhysRevLett.39.1587.

[78] G. C. Fox and S. Wolfram. "Observables for the analysis of event shapes in e^+e^- annihilation and other processes". In: *Phys. Rev. Lett.* 41 (23 Dec. 1978), pp. 1581–1585. DOI: 10.1103/PhysRevLett.41.1581. URL: https://link.aps.org/doi/10.1103/PhysRevLett.41.1581.

[79] CLEO II Collaboration. "Search for exclusive charmless hadronic B decays". In: *Phys. Rev. D* 53 (3 Feb. 1996), pp. 1039–1050. DOI: 10.1103/PhysRevD.53.1039. URL: https://link.aps.org/doi/10.1103/PhysRevD.53.1039.

[80] J. E. Gaiser. "Charmonium spectroscopy from radiative decays of the J/Ψ and Ψ'". In: *SLAC-R-255* (1982).

[81] W. Verkerke and D. P. Kirkby. "The RooFit toolkit for data modeling". In: *eConf* C0303241 (2003). Ed. by L. Lyons and Muge Karagoz, MOLT007. arXiv: physics/0306116.