Dissertation

Measurement of CP Parameters in $B \to D^0_{CP\pm} K$ Decays With the BABAR Detector

Till Moritz Karbach

June 3, 2009

Vom Fachbereich Physik der Technischen Universität Dortmund zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation.

1. Gutachter: Prof. Dr. B. Spaan
2. Gutachter: Priv.-Doz. Dr. R. Klingenberg
Tag der mündlichen Prüfung: 10. Juli 2009

All work presented in this document is the author's own work, unless stated otherwise. The document is composed in majestic plural.

This work is supported by the Bundesministerium für Bildung und Forschung and Deutsche Forschungsgemeinschaft (Germany).

Abstract

We study the $B^- \to D^0\pi^-$ and $B^- \to D^0K^-$ decays and their charge conjugates, with D^0 reconstructed in the *CP*-odd eigenstates (D^0_{CP-}) $K^0_S\pi^0$, $K^0_S\phi$, $K^0_S\omega$, in the *CP*-even eigenstates (D^0_{CP+}) K^+K^-, $\pi^+\pi^-$ and in the (non-*CP*) flavor eigenstates (D^0, $\overline{D}^0$) $K^-\pi^+$, $K^-\pi^+\pi^0$, and $K^-\pi^+\pi^+\pi^-$. We use a data sample of 467×10^6 $\Upsilon(4S)$ decays collected with the *BABAR* detector at the PEP-II e^+e^- storage ring, representing the final $\Upsilon(4S)$ dataset of the *B*-factory. We measure the ratios of the branching fractions $R_{CP\pm}$ and the direct *CP* asymmetries $A_{CP\pm}$,

$$R_{CP\pm} = 2\,\frac{\mathcal{B}(B^-\to D^0_{CP\pm}K^-)+\mathcal{B}(B^+\to D^0_{CP\pm}K^+)}{\mathcal{B}(B^-\to D^0K^-)+\mathcal{B}(B^+\to \overline{D}^0K^+)},$$

$$A_{CP\pm} = \frac{\mathcal{B}(B^-\to D^0_{CP\pm}K^-)-\mathcal{B}(B^+\to D^0_{CP\pm}K^+)}{\mathcal{B}(B^-\to D^0_{CP\pm}K^-)+\mathcal{B}(B^+\to D^0_{CP\pm}K^+)},$$

to be

$$\begin{aligned} R_{CP+} &= 1.128 \pm 0.081 \pm 0.037, \\ R_{CP-} &= 1.041 \pm 0.069 \pm 0.030, \\ A_{CP+} &= +0.197 \pm 0.060 \pm 0.031, \\ A_{CP-} &= -0.096 \pm 0.064 \pm 0.009, \end{aligned}$$

where the first error is statistical and the second one systematic. The statistical significance of A_{CP+} being non-zero is 3.4 standard deviations, constituting evidence for direct *CP* violation in charged B decays. We also express the results in terms of the so called Cartesian coordinates x_+, x_-, and r^2: $x_+ = -0.059\pm0.034(\text{stat})\pm0.014(\text{syst})$, $x_- = 0.102\pm0.036(\text{stat})\pm0.015(\text{syst})$, $r^2 = 0.084 \pm 0.054(\text{stat}) \pm 0.024(\text{syst})$. These results will help to better constrain the phase parameter $\gamma = \arg(-V_{ud}V^*_{ub}/V_{cd}V^*_{cb})$ of the Cabibbo-Kobayashi-Maskawa quark mixing matrix.

Contents

1 Introduction

The Standard Model of particle physics describes the interactions of the most fundamental particles. It is grounded on the U(1)×SU(2)$_L$ symmetry of the unified electro-weak interaction, and on the SU(3) quark flavor symmetry of the strong interaction. The current experimental data suggest nature is invariant under the transformations of these symmetry groups, which is reflected by the according invariance of the Standard Model Lagrangian. However, nature is not invariant under certain discrete symmetry transformations. This work explores the violation of *CP* symmetry, where the *CP* operator is a junction of the parity operator P, which reflects all spacial dimensions through the origin, and the charge conjugation operator, C. In the Standard Model, *CP* violation is a consequence of the weak interaction. Testing the *CP* violation mechanism is a sensitive test of the Standard Model in general.

The couplings of the quark fields to the weak interaction are described by the 3×3 mixing matrix V of the quark mass. It is called *CKM* matrix [1], named after Nicola Cabibbo, who introduced the mixing idea in the u, d, s, c quark sector, and after Makoto Kobayashi and Toshihide Maskawa, who in 1973 extended it to a then unknown third quark family, t, b. Their extension introduces a free phase parameter, yielding a mechanism to explain *CP* violation in the Standard Model. For this work, Kobayashi and Maskawa received the Noble Price[1] in 2008.

A powerful strategy to test the mechanism of *CP* violation is to measure the four independent parameters of the unitary CKM matrix. These parameters are related to the angles α, β, γ of an unitarity triangle in the complex plane. Out of the angles, the angle γ is the least precisely known.

The goal of this work is to help in constraining the allowed parameter region of γ. It follows the GLW method, which was suggested by Gronau, London, and Wyler in 1991 [2, 3], their initials naming the method. The method exploits the interference between the $B^\pm \to D^0 K^\pm$ and $B^\pm \to \overline{D}^0 K^\pm$ decay amplitudes, where the D^0 and $\overline{D}^0$-mesons decay to common *CP* eigenstates. The charge averaged ratios

$$R_{CP\pm} = 2\,\frac{\mathcal{B}(B^- \to D^0_{CP\pm} K^-) + \mathcal{B}(B^+ \to D^0_{CP\pm} K^+)}{\mathcal{B}(B^- \to D^0 K^-) + \mathcal{B}(B^+ \to \overline{D}^0 K^+)} \tag{1.1}$$

and the direct *CP* asymmetries

$$A_{CP\pm} = \frac{\mathcal{B}(B^- \to D^0_{CP\pm} K^-) - \mathcal{B}(B^+ \to D^0_{CP\pm} K^+)}{\mathcal{B}(B^- \to D^0_{CP\pm} K^-) + \mathcal{B}(B^+ \to D^0_{CP\pm} K^+)} \tag{1.2}$$

are related to the angle γ. We will extract $R_{CP\pm}$ and $A_{CP\pm}$ by means of a three-dimensional simultaneous extended maximum likelihood fit to data collected by the *BABAR* Experiment. The *BABAR* Experiment is hosted by the SLAC National Laboratory in California, USA, and constitutes one of the two B-Factories in the world, the Belle Experiment at the KEK laboratory in Japan being the second one. The data taking phase of *BABAR* ended in April 2008, concluding a very successful running period of almost ten years. The final dataset consists of 467 million $B\overline{B}$ pairs.

[1] Together with Yoichiro Nambu.

There have been three subsequent GLW measurements using *BABAR* data. The author contributed to the more recent two of them, while the first GLW measurement was carried out by different authors. The first measurement was based on 236×10^6 $B\overline{B}$ pairs and was published in 2006 [4]. In 2007, a partial result on 382×10^6 $B\overline{B}$ pairs was obtained by an analysis group consisting of the author and a fellow *BABAR* collaborator. The analysis strategy of the first measurement was re-implemented: the author contributed the final event selection and the final fitter part, while the collaborator developed the pre-selection part. The partial result was published in 2008 [5]. The third analysis is the main analysis presented in this document. It was performed on the final dataset of 467 million $B\overline{B}$ pairs, the author now being the leading analyst. The author introduced additional D^0 final states to the analysis, as well as a new strategy for the final selection and a new and more powerful fit strategy.

This document is organized as follows. We first give an overview of the scientific background of *CP* violation and the physics of *B*-Factories (Chapter 2). Then we introduce the GLW method as well as the analysis technique (Chapter 3). Chapter 4 explains the main features of the *BABAR* Experiment, before Chapter 5 presents the 2007 partial result. The following Chapters 6, 7, and 8 define the event selection and compare the measured data to the expectation from simulation. Chapters 9 and 10 introduce and validate the final fit strategy. A non-reducible kind of peaking background is discussed in Chapter 11, and systematic uncertainties are discussed in Chapter 12. Finally Chapter 13 presents the result on the parameters $A_{CP\pm}$ and $R_{CP\pm}$ and discusses the implications on the angle γ.

2 Scientific Background

2.1 *CP* Violation in the Standard Model

One distinguishes three kinds of *CP* violation (see, for instance, Reference [6]). The most intuitive one is called *direct CP* violation. It implies, that the decay rate Γ of a particle P decaying into a final state f is different from the rates observed for the *CP*-conjugate process:

$$\Gamma(P \to f) \neq \Gamma(\bar{P} \to \bar{f}). \tag{2.1}$$

The second type is called *CP violation in oscillation*. Here, the rate at which a neutral particle P^0 oscillates into its anti-particle $\bar{P}^0$ is different from the *CP*-conjugate process, $\Gamma(P^0 \to \bar{P}^0) \neq \Gamma(\bar{P}^0 \to P^0)$. The third type is called *CP violation in interference between oscillation and decay*. Recalling the decay rate is the absolute square of the decay amplitude, $\Gamma = |A|^2$, the third type refers to the cases where the rate $|A(P^0 \to \bar{P}^0 \to f) + A(P^0 \to f)|^2$ is different from its *CP*-conjugate. For this work, only effects of direct *CP* violation are important, whose formalism we briefly review in the following. Introducing the amplitudes $A(P \to f) \equiv A_f$ and $A(\bar{P} \to \bar{f}) \equiv \bar{A}_{\bar{f}}$, Eq. 2.1 is equivalent to the asymmetry $|A_f|^2 - |\bar{A}_{\bar{f}}|^2$ being non-zero. One can rewrite any amplitude in terms of its magnitude and its phases, $A = |A|e^{i(\delta+\Phi)}$, where δ refers to a *strong* phase, and Φ refers to a *weak* phase. They differ in their behavior under *CP* transformation. The strong phase is invariant, $e^{i\delta} \to e^{i\delta}$, while the weak phase gains a sign, $e^{i\Phi} \to e^{-i\Phi}$. Now A_f and $\bar{A}_{\bar{f}}$ become

$$\begin{aligned} A_f &= |A|e^{i(\delta+\Phi)}, \\ \bar{A}_{\bar{f}} &= |A|e^{i(\delta-\Phi)}. \end{aligned} \tag{2.2}$$

From this it is apparent, that one will only see direct *CP* violation if there is interference between different contributing amplitudes,

$$A_f \equiv \sum_n |A_n| e^{i(\delta_n+\Phi_n)}. \tag{2.3}$$

Then, the asymmetry becomes

$$|A_f|^2 - |\bar{A}_{\bar{f}}|^2 = -2 \sum_{n,m} |A_n||A_m| \sin(\delta_n - \delta_m) \sin(\Phi_n - \Phi_m), \tag{2.4}$$

which is only non-zero in case of more than two interfering amplitudes with both different strong phases and different weak phases. Only then direct *CP* violation can occur. While there are many sources of strong phases in the Standard Model, there is only one source of weak phases: the quark mixing mechanism.

The Standard Model explains the fact fermions have mass, and quarks in particular, through the Yukawa-Coupling to the Higgs-field. The ground state of the Higgs-field has a different symmetry than the Lagrangian. Through this mechanism of spontaneous symmetry breaking the fermion fields acquire mass terms in the Lagrangian. The situation is complicated by the fact that the mass eigenstates of the down-type quarks, d, s, b, aren't eigenstates to the

electro-weak interaction at the same time. The mass eigenstates of the up-type quarks, u, c, t, however, are. This has two important consequences. At first it allows an up-type quark of one generation to decay into a down-type quark of another generation. For instance, a c-quark is allowed to decay into a d-quark, $c \rightarrow W^+ d$. At second, the complex unitary $\times$ 3 matrix, which relates the mass eigenstates $(d, s, b)^T$ to the electro-weak eigenstates $(d', s', b')^T$,

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.5}$$

contains a free phase parameter: the weak phase. This mass mixing matrix V is called *CKM* matrix [1]. Its elements V_{ij} govern the weak decays of up-type quarks to down-type quarks and vice versa:

$$A(u \rightarrow dW^+) \sim V_{ud} \ , \qquad A(\bar{u} \rightarrow \bar{d}W^-) \sim V_{ud}^* \ , \tag{2.6}$$

where u and d stand for generic up- and down-type quarks. The fact that the complex conjugate of the matrix element enters the amplitude for decays of the anti-quarks is crucial for *CP* violation: here we have different amplitudes for particles and anti-particles. The CKM matrix was introduced to diagonalize the mass matrix of the down-type quarks, hence V_{CKM} is unitary, $V_{\mathrm{CKM}} V_{\mathrm{CKM}}^\dagger = 1$. This yields nine unitarity conditions which the V_{ij} need to fulfill, one of which has gained special interest:

$$V_{ud} V_{ub}^* + V_{cd} V_{cb}^* + V_{td} V_{tb}^* = 0. \tag{2.7}$$

It is a sum of complex numbers, which equals zero, and can be depicted as a triangle in the complex plane (Figure 2.1). The enclosed area of this triangle is a measure of the *CP* violation in the Standard Model: the larger the area, the more *CP* violation is predicted. A lot of effort has been put into measuring the dimensions of this triangle, including this work, to test the unitarity of the CKM matrix. A violation of unitarity is considered a strong argument for new physical processes beyond the scope of the Standard Model.

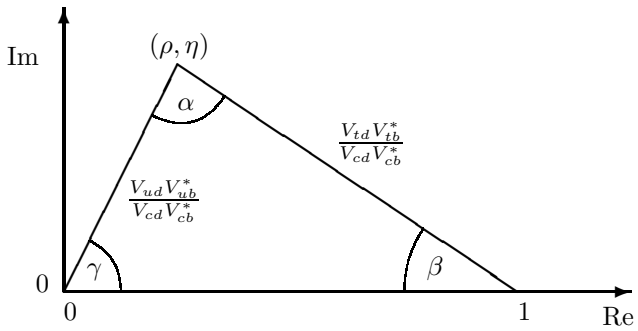

Figure 2.1: Unitarity triangle of the CKM matrix, corresponding to Eq. 2.7. The Wolfenstein parameters ρ and η are defined in Eq. 2.8.

The CKM matrix has four independent parameters in total. Starting with the 18 parameters of a general 3×3 complex matrix, on subtracts the above unitarity conditions (9), and 5 relative phases which can be absorbed by re-phasing the six quark fields, without loss of generality. The remaining four parameters can be thought of as three Euler angles and one complex phase. There are several parameterizations of the CKM matrix, approximating it through four independent parameters. One of the more popular approaches is that of

Wolfenstein [7], where the matrix elements are expressed in terms of powers of $\lambda \simeq \sin\theta_c$, where θ_c is the Cabibbo angle:

$$V_{\mathrm{CKM}} = \begin{pmatrix} 1-\frac{\lambda^2}{2} & \lambda & \lambda^3 A(\rho - i\eta) \\ -\lambda & 1-\frac{\lambda^2}{2} & \lambda^2 A \\ \lambda^3 A(1-\rho-i\eta) & -\lambda^2 A & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{2.8}$$

In the Wolfenstein parameterization, the only two CKM matrix elements which have a complex phase are V_{ub} and V_{td}. The present values [8] of the Wolfenstein parameters are $\lambda = 0.2272 \pm 0.0010$, $A = 0.818^{+0.007}_{-0.017}$, $\rho = 0.221^{+0.064}_{-0.028}$, $\eta = 0.340^{+0.017}_{-0.045}$. Using these values one can obtain the magnitudes of the $|V_{ij}|$, rounded to one significant digit:

$$|V_{\mathrm{CKM}}| = \begin{pmatrix} 1 & 0.2 & 0.004 \\ 0.2 & 1 & 0.04 \\ 0.008 & 0.04 & 1 \end{pmatrix} \tag{2.9}$$

It has become common to call amplitudes involving $|V_{ij}|$ smaller than one *CKM suppressed*. To emphasize the fact $|V_{ub}|$ and $|V_{td}|$ are especially tiny, one often speaks of *doubly Cabibbo suppressed* amplitudes when referring to amplitudes containing these matrix elements.

A promising way to determine the dimensions of the unitarity triangle is to use the decays of B-mesons. These particles are formed out of a b-quark and a light anti-quark, $\bar{u}$ or $\bar{d}$.[1] They have about five times the proton mass, $m(B^{\pm}) = 5279.15 \pm 0.31\,\mathrm{MeV}/c^2$ [8], providing enough phase space to allow for a large variety of final states. Their lifetime is relatively long, leading to a long flight length of $c\tau = 491.1\,\mu\mathrm{m}$, which is well accessible with modern particle detectors. The primary decay chain involves the $b \to c$ amplitude, often producing D-mesons in the final state: $\mathcal{B}(B^+ \to \overline{D}^0 X) = 79 \pm 4\,\%$ [8]. By combining carefully selected amplitudes of B decays one can construct observables which are sensitive to one of the three angles of the unitarity triangle:

$$\beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right),\ \alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right),\ \gamma = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right). \tag{2.10}$$

To measure the angle β, the decay channel $B^0 \to J/\psi K_S^0$ is most suitable. It is dominated by a color suppressed[2] $b \to c$ tree diagram, in which, following Eq. 2.8, the Standard Model doesn't predict a weak phase. But without a weak phase there is no direct *CP* violation in this mode. Instead the measurement is based on *CP* violation in the interference between oscillation and decay. The B^0–$\overline{B^0}$ oscillations are described by a factor q/p, which in the Standard Model is, to a good approximation, $q/p = V_{tb}^* V_{td}/V_{tb}V_{td}^* = e^{-2i\beta}$. In the analysis of $B^0 \to J/\psi K_S^0$ decays one therefore extracts the quantity $\sin(2\beta)$, so one can determine β up to discrete ambiguities. The $B^0 \to J/\psi K_S^0$ decays have a relatively high branching ratio of $\approx 10^{-3}$, and they are referred to as the *golden mode*. As a consequence, this analysis has become a precision measurement at present B-Factories. *BABAR* determines $\sin(2\beta)$ to be $\sin(2\beta) = 0.714 \pm 0.032(\mathrm{stat}) \pm 0.018(\mathrm{syst})$ [9].

The angle α is determined through the decay channels $B^0 \to \pi^+\pi^-$ and $B^0 \to \rho\rho$, involving the $b \to u$ amplitudes. However, these channels only allow for a measurement of an effective angle, α_{eff}. This is because the main tree-level amplitude is polluted by a sizable contribution of loop diagrams (so called *penguins*), which involve a different weak phase. The difference

[1] In general, mesons consist of a quark and an anti-quark. Whenever we refer to a meson's quark content we imply one of the quarks denotes an anti-quark.

[2] Color suppression is a effect of Quantum Chromo Dynamics. Real particles need to be color neutral. In internal tree diagrams the color of the produced quarks is constrained to compensate that of the initial quarks. Considering three colors, the factor is $\approx (1/3)^2 \approx 0.1$.

$\alpha - \alpha_{\text{eff}}$ can be estimated through an isospin analysis of the flavor SU(2) related modes. Alternatively one can benefit from an angular analysis of the $\rho\rho$ vector-vector final state to disentangle the tree-level from the penguin contributions. *BABAR* measures α to be in the range $[83.3°, 105.8°]$ at the 68 % confidence level [10, 11].

The third angle, γ, is the most difficult to measure. At present the most relevant decay channels are $B^{\pm} \to D^{(*)0} K^{(*)\pm}$ and $B^{\pm} \to \bar{D}^{(*)0} K^{(*)\pm}$. In fact, the goal of this work is to constrain γ using $B^- \to D^0 K^-$ and $B^- \to \overline{D}^0 K^-$ decays and their charge conjugates. The methods to measure γ are based on the interference between a color-favored $b \to c$ amplitude, and a color and doubly Cabibbo suppressed $b \to u$ transition, carrying the weak phase. Apart from the low branching fractions due to the small value of $|V_{ub}|$, the main difficulty is that the interfering amplitudes have very different magnitudes. Thus the *CP* violating interference term is small: $|B \to \overline{D}^0 K / B^- \to D^0 K^-| \simeq 0.1$. As for the near future, other channels may become equally important. These are, most prominently, the $B_s^0 \to D_s^{\mp} K^{\pm}$ decays, which will be accessible at the LHCb experiment [12]. Here, the amplitude depends on $\gamma + \phi_s$, where the angle ϕ_s originates from B_s^0–$\bar{B}_s^0$ mixing and can be measured directly in $B_s^0 \to J/\psi\,\phi$ decays. Also the $B^0 \to \pi^+\pi^-$ and $B_s^0 \to K^+K^-$ decays will be well accessible at LHCb: through an analysis of the time-dependent *CP* asymmetries one gets access to γ when both channels are combined.

Figure 2.2 summarizes the current experimental status by constraining the apex the the unitarity triangle in the complex plane using all measurements available to date. The difficulties associated with the measurement of the angle γ are illustrated by the large error band associated to it.

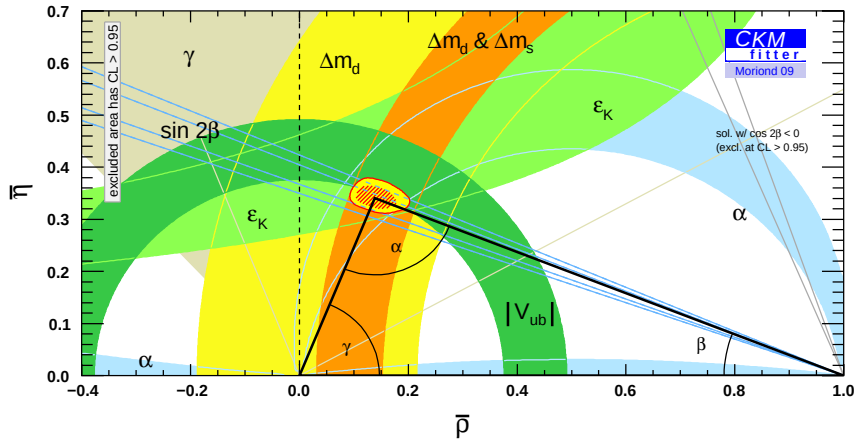

Figure 2.2: Results of a global fit to all available measurements constraining the CKM unitarity triangle (CKM*fitter* group, 2009 [13]). There is very good consistency with the Standard Model prediction: the triangle is closing within current accuracy.

2.2 B-Factory

To perform measurements of the three angles of the unitarity triangle, one needs $\mathcal{O}(10^8)$ B decays, because the relevant branching ratios are on the order of 10^{-6}. Such large samples have been collected in the recent past by the so called *B-Factories*. These are particle

accelerators producing strong electron and positron beams, at an e^+e^- center-of-mass energy of $\sqrt{s} = 10.58$ GeV. This equals the invariant mass of the $\Upsilon(4S)$ resonance, which is a bound state of two b-quarks. The $\Upsilon(4S)$ decays in more than 96 % of the time into a $B\overline{B}$ pair, allowing for an efficient production of B-mesons:

$$e^+e^- \to \Upsilon(4S) \to B\overline{B}. \tag{2.11}$$

It is an interesting fact that the $\Upsilon(4S)$ invariant mass equals almost exactly the mass to two B-mesons. The excess energy is only 21 MeV, resulting in a B-momentum of 328 MeV/c. This provides an helpful kinematical constraint, because B-mesons are almost at rest in the center-of-mass system. However, to be able to measure significant flight lengths of the B-mesons, one needs relativistic momenta. Else one couldn't determine the life time differences between the B^0 and the $\overline{B^0}$, which is crucial to measure B^0–$\overline{B^0}$ oscillations needed for instance to measure the CKM angle β. The solution to this problem is to use asymmetric beam energies, resulting in a sufficient Lorentz boost of the e^+e^- center-of-mass system.

Because of the kinematical constraint provided by the initial $\Upsilon(4S)$ state, the energy of each B-meson in the e^+e^- center-of-mass frame (CM) must be equal to $\sqrt{s}/2$. This can be exploited by defining two variables, m_{ES} (*energy-substituted mass*) and ΔE (*energy difference*):

$$m_{\mathrm{ES}} \equiv \sqrt{\left(\frac{1}{2}s + \mathbf{p}_0 \cdot \mathbf{p}_B\right)^2 / E_0^2 - \mathbf{p}_B^2}, \tag{2.12}$$

$$\Delta E \equiv E_B^* - E_{\mathrm{beam}}^*. \tag{2.13}$$

The momenta $\mathbf{p}_i$, $\mathbf{p}_B$, and the energy E_0 are measured in the laboratory frame and the subscripts 0 and B refer to the e^+e^- system and the reconstructed B-meson, respectively; E_B^* is the reconstructed CM energy of the B candidate. $E_{\mathrm{beam}}^* = \sqrt{s}/2$ is the reconstructed CM beam energy. For correctly reconstructed B decays m_{ES} peaks at the B mass, with a resolution typically smaller than 3 MeV/c^2, and ΔE peaks at 0, with a resolution on the order of 20 MeV. Both the ΔE and m_{ES} variables are widely used at the B-Factories.

Only a fraction of the e^+e^- collisions actually produces an $\Upsilon(4S)$, all other collisions produce either quark/anti-quark pairs or lepton/anti-lepton pairs. The cross sections for all of these processes are given in Table 2.1. The $e^+e^- \to q\overline{q}$ cross sections are of comparable size to the $b\overline{b}$ cross section. When analyzing B-decays, one therefore expects a *continuum* background.

Table 2.1: e^+e^- production cross-sections at $\sqrt{s} = m(\Upsilon(4S))$ within the experimental acceptance of the *BABAR* detector.

$e^+e^- \to$	$b\overline{b}$	$c\overline{c}$	$s\overline{s}$	$u\overline{u}$	$d\overline{d}$	$\tau^+\tau^-$	$\mu^+\mu^-$	e^+e^-
σ [nb]	1.05	1.30	0.35	1.39	0.35	0.94	1.16	~ 40

The event rate in an experiment detecting events of colliding particles is proportional to the interaction cross section. The relating factor is called the *luminosity* $\mathcal{L}$: $\mathrm{d}N/\mathrm{d}t = \mathcal{L} \cdot \sigma$. The luminosity is proportional to the beam currents. As a consequence, high luminosity B-Factories aim for strong beam currents of 2-3 A. The collected data samples are measured in terms of the *integrated luminosity*, $L = \int \mathcal{L}\mathrm{d}t$. Thus the expected number of events of a certain type contained in the sample is $N = L \cdot \sigma$.

3 Outline of the Analysis

3.1 GLW Method

Many of the proposals on how to measure the CKM angle γ involve decays of B-mesons, both neutral and charged ones, into a neutral D-meson and a kaon: the D-meson consists of a charm quark (c) and a light quark (u, d), the kaon contains a strange quark (s) and a light quark. Analyzing the decays of neutral B-mesons, however, comes with additional complications. For instance, the self-tagging $B \to D^0 K^{*0}$ final states, in which the charge of the kaon from the $K^{*0} \to K^+\pi^-$ decay indicates the flavor of the initial B-meson, are color-suppressed; in non-self-tagging final states one faces B^0–$\overline{B^0}$ oscillations. The decays of charged B-mesons are more promising: $B^\pm \to \overline{D}^{(*)0} K^{(*)\pm}$ and $B^\pm \to D^{(*)0} K^{(*)\pm}$. They proceed via two amplitudes, $b \to c\bar{u}s$ and the doubly Cabibbo suppressed $b \to u\bar{c}s$, the latter containing the weak phase. These amplitudes interfere when the D-meson is being reconstructed in a final state common to both D and $\overline{D}$, giving access to the angle γ. Figure 3.1 shows Feynman diagrams of the two interfering B decays.

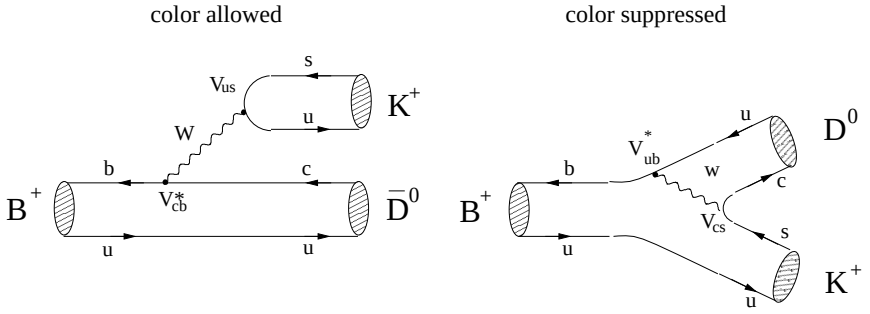

Figure 3.1: Feynman diagrams of the $B \to \overline{D}^0 K$ and $B \to D^0 K$ decays. There are actually two more tree level amplitudes contributing. They are smaller compared to those depicted and contain the same weak phase.

The GLW method was suggested by Gronau, London, and Wyler in 1991 [2, 3]. It exploits the interference between the non-excited $B^\pm \to D^0 K^\pm$ and $B^\pm \to \overline{D}^0 K^\pm$ decay amplitudes, where the common final state to the D^0 and $\overline{D}^0$-mesons is chosen to be a *CP* eigenstate. A great advantage of the GLW method is that it is theoretically very clean. This is because the $B \to D^0 K$ decays are governed by tree level decays rather than penguin loop diagrams. Such loop diagrams could potentially be affected by physics beyond the Standard Model, because new processes are expected to enter with new particles in the loop. This could introduce new weak phases, too, which would dilute the measurement of γ. In addition, the GLW method extracts relevant hadronic parameters directly from the data, so there is no need to depend on quantum-chromo-dynamical calculations introducing hadronic uncertainties.

At the heart of the original GLW method lie two equations of amplitudes, which can be represented by triangles in the complex plane (Figure 3.2). It is possible to relate the sides

to each other, and to identify 2γ as the angle between two of the sides. This enables one to determine the weak phase γ through a measurement of branching ratios. Due to geometrical reasons, γ can only be determined up to an eight-fold ambiguity: the sign of γ remains unknown, one can't distinguish γ from $\gamma + \pi$, and there is a possibility to confuse γ with the relative strong phase of the final states. In the neutral D system, the two *CP* eigenstates are defined as

$$\left|D^0_{CP\pm}\right\rangle = \frac{1}{\sqrt{2}}\left(\left|D^0\right\rangle \pm \left|\overline{D}^0\right\rangle\right), \tag{3.1}$$

where *CP*+ refers to the *CP*-even eigenstate, and *CP*– to the odd one. For now let's focus on the *CP*+ eigenstate only. From Equation 3.1 one expects the following two *CP* conjugate relations to hold:

$$\begin{aligned}
\sqrt{2}A(B^+ \to D^0_{CP+}K^+) &= A(B^+ \to D^0K^+) + A(B^+ \to \overline{D}^0K^+), \\
\sqrt{2}A(B^- \to D^0_{CP+}K^-) &= A(B^- \to \overline{D}^0K^-) + A(B^- \to D^0K^-).
\end{aligned} \tag{3.2}$$

We assume that there is no direct *CP* violation in the decay modes on the right hand side of Equation 3.2,

$$|A(B^+ \to D^0K^+)| = |A(B^- \to \overline{D}^0K^-)|, \tag{3.3}$$

$$|A(B^+ \to \overline{D}^0K^+)| = |A(B^- \to D^0K^-)|. \tag{3.4}$$

This is a fairly safe assumption in the Standard Model: both amplitudes in Eq. 3.3 contain the same weak phase, so do both amplitudes in Eq. 3.4. But without different weak phases, there is no direct *CP* violation because of Eq. 2.4. The decay modes on the left hand side of Eq. 3.2, however, can in principle violate *CP*, because their amplitudes are the sum of two amplitudes with different weak phases:

$$|A(B^+ \to D^0_{CP+}K^+)| \neq |A(B^- \to D^0_{CP+}K^-)|. \tag{3.5}$$

Introducing the magnitudes A and $\bar{A}$ of the amplitudes in Eqns. 3.3 and 3.4, respectively, one can rewrite Eqns 3.2 in terms of the weak phase γ and the strong phases δ and $\bar{\delta}$:

$$\begin{aligned}
\sqrt{2}A(B^+ \to D^0_{CP+}K^+) &= |A|e^{i\gamma}e^{i\delta} + |\bar{A}|e^{i\bar{\delta}}, \\
\sqrt{2}A(B^- \to D^0_{CP+}K^-) &= |A|e^{-i\gamma}e^{i\delta} + |\bar{A}|e^{i\bar{\delta}}.
\end{aligned} \tag{3.6}$$

This shows explicitly how in the amplitudes containing V^*_{ub} the weak phase acquires a sign under *CP* transformation, while the other two amplitudes remain unaffected. Equations 3.6 form two triangles in the complex plane, depicted in Figure 3.2, the sides of which are related through Equations 3.3-3.4. The amplitudes $A(B^+ \to D^0K^+)$ and $A(B^- \to \overline{D}^0K^-)$ enclose the angle 2γ:

$$A(B^+ \to D^0K^+) = e^{2i\gamma}A(B^- \to \overline{D}^0K^-). \tag{3.7}$$

Simple geometry yields an expression which relates γ to the length of the sides of the triangles, up to the aforementioned eight-fold ambiguities (the possibility to confuse $\gamma \leftrightarrow \delta - \bar{\delta}$ is not shown):

$$\begin{aligned}
\sin(\gamma) = \frac{1}{4|A\bar{A}|}\Bigg(&\pm\sqrt{\left[(|A|+|\bar{A}|)^2 - 2|A^+_{CP+}|^2\right] \times \left[2|A^-_{CP+}|^2 - (|A|+|\bar{A}|)^2\right]} \\
&\pm\sqrt{\left[(|A|+|\bar{A}|)^2 - 2|A^-_{CP+}|^2\right] \times \left[2|A^+_{CP+}|^2 - (|A|+|\bar{A}|)^2\right]}\Bigg),
\end{aligned}$$

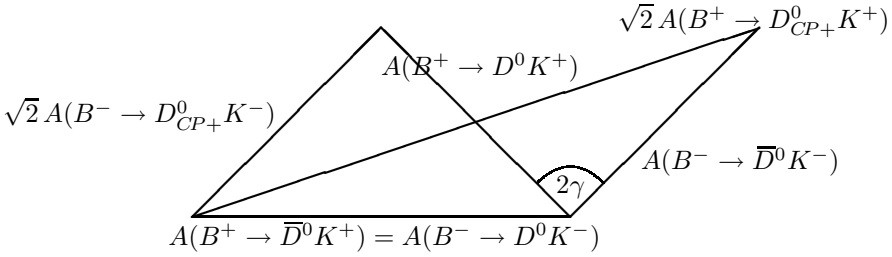

Figure 3.2: GLW triangle of related $B \to D^0K$ amplitudes, enclosing the angle 2γ.

where $A^+_{CP+} \equiv A(B^+ \to D^0_{CP+}K^+)$, $A^-_{CP+} \equiv A(B^- \to D^0_{CP+}K^-)$. The geometrical interpretation of the ambiguities is that the orientation of the triangles is unknown. Each of them could be reflected either through its basis or vertically—the enclosed angle 2γ would be different.

In principle, the measurement of γ is now straightforward. One has to measure six branching ratios, two of which are equal, and determine the length of the sides of the triangles from it. But in praxis, it is basically impossible to measure the suppressed rates of $B^- \to \overline{D}^0K^-$ and $B^+ \to D^0K^+$. If one uses semileptonic decays to tag the flavor of the $D^0/\overline{D}^0$, such as $D^0 \to l^+\nu_l X_s$, the background rates from the direct decay $B^+ \to l^+\nu_l X_{\bar{c}}$ are too high and the signal reconstruction efficiency is too low. On the other hand, if one uses hadronic decays like $\overline{D}^0 \to K^+\pi^-$, then one has interference from $B^- \to D^0K^-$, $D^0 \to K^+\pi^-$: on the one hand $B^- \to \overline{D}^0K^-$ is doubly Cabibbo suppressed but $\overline{D}^0 \to K^+\,\pi^-$ is not. On the other hand $D^0 \to K^+\,\pi^-$ is doubly Cabibbo suppressed but $B^- \to D^0K^-$ is not. As a consequence, the interference is large.

The solution is to use the modified GLW method [14]. It uses both *CP*+ and *CP*– eigenstates, whereas the original method only required one. The D^0_{CP+}-meson can be reconstructed for instance in the *CP*-even final states K^+K^- and $\pi^+\pi^-$. The D^0_{CP-}-meson can be identified in the channels $K^0_S\pi^0$, $K^0_S\omega$, or $K^0_S\phi$.[1] The modified method doesn't require to directly measure the suppressed amplitudes of the B-meson, but only experimentally accessible branching fractions. Then, the D^0 can be identified in flavor specific state final states such as $K^-(n\cdot\pi)^+$, $n \in \mathbb{N}$, neglecting the doubly Cabibbo suppressed decay of $\overline{D}^0$ to the same final state. To this extent, we define four observables, namely the branching fraction ratios $R_{CP\pm}$ and the direct *CP* asymmetries $A_{CP\pm}$

$$R_{CP\pm} = \frac{\Gamma(B^- \to D^0_{CP\pm}K^-) + \Gamma(B^+ \to D^0_{CP\pm}K^+)}{\left[\Gamma(B^- \to D^0K^-) + \Gamma(B^+ \to \overline{D}^0K^+)\right]/2}\,, \tag{3.8}$$

$$A_{CP\pm} = \frac{\Gamma(B^- \to D^0_{CP\pm}K^-) - \Gamma(B^+ \to D^0_{CP\pm}K^+)}{\Gamma(B^- \to D^0_{CP\pm}K^-) + \Gamma(B^+ \to D^0_{CP\pm}K^+)}\,. \tag{3.9}$$

The latter asymmetries $A_{CP\pm}$ are of special relevance because they would indicate, if significantly different from zero, direct *CP* violation in charged B decays. The new observables are related to the three quantities γ, the magnitude ratio r

$$r = \frac{|A(B^+ \to D^0K^+)|}{|A(B^+ \to \overline{D}^0K^+)|}, \tag{3.10}$$

[1] These final states are, in fact, only true *CP*-odd eigenstates if one neglects *CP* violation in the K^0 system: $|K^0_S\rangle = |K_1\rangle + \varepsilon\,|K_2\rangle$, $|K_1\rangle$ is the *CP*+ eigenstate $|K_1\rangle = (|K^0\rangle + |\bar{K}^0\rangle)/\sqrt{2}$, and ε is on the order of 3×10^{-3}.

and the strong phase difference $\delta_s = \delta - \bar{\delta}$ of these amplitudes through the relations

$$R_{CP\pm} = 1 + r^2 \pm 2r\cos\delta_s\cos\gamma, \tag{3.11}$$

$$A_{CP\pm} = \frac{\pm 2r\sin\delta_s\sin\gamma}{R_{CP\pm}}. \tag{3.12}$$

Only three of the four observables are independent, and the following identity holds:

$$A_{CP+}R_{CP+} = A_{CP-}R_{CP-}. \tag{3.13}$$

The r parameter plays a crucial role, as can be seen from Equations 3.11 and 3.12: the sensitivity to γ scales like $1/r$. The ratio is predicted to be on the order of 0.1 [2, 3], in agreement with recent results by *BABAR* ($r = 0.091 \pm 0.059$ [15]) and Belle ($r = 0.159 \pm 0.074$ [16]), obtained through the study of $B \to D^0K$, $D^0 \to K^+\pi^-\pi^0$ and $D^0 \to K^0_S\pi^+\pi^-$ decays. This is the main drawback of the GLW method, because it limits the *CP* violating interference effect.

Another potential difficulty arises from non-negligible D^0–$\overline{D}^0$ mixing. The GLW method assumes that it is possible to determine the *CP* eigenstate of a D^0-meson from the final state it decays to. This is not reliably possible in case of sizable D^0–$\overline{D}^0$ mixing. Because in that case, a D^0 produced in a B decay could oscillate into a $\overline{D}^0$ before it decays, diluting the D^0 sample. Particle/anti-particle oscillations are measured in terms of the quantity x: the probability for a D^0 to oscillate into a $\overline{D}^0$-meson is given by $P(D^0 \to \overline{D}^0) = 1/2\, e^{-t/\tau}[1-\cos(x_D t/\tau)]$, where τ denotes the lifetime of the D^0 state. Meca and Silva have shown [17] that a value of x_D on the order of 10^{-2} will affect the decay rates by as much as 10%. Recent work of the *BABAR* Collaboration (2008) [18, 19] sees evidence of D^0–$\overline{D}^0$ mixing. The observed value of x_D in the $K^-\pi^+\pi^0$ final state is $x_D = [2.61^{+0.57}_{-0.68}(\text{stat}) \pm 0.39(\text{syst})] \times 10^{-2}$, the statistical significance of this result is 3.2σ. The Belle Collaboration also measures x_D [20]. But considering the current level of statistics available for $B \to D^0K$ decays, the GLW method remains unaffected from D^0–$\overline{D}^0$ mixing.

There are two important alternatives to the GLW method, still relying on $B \to D^0K$ decays. The first one was proposed in 1997 by Atwood, Dunietz, and Soni (ADS, [21, 22]). The idea is to perform the analysis with non-*CP* eigenstates and to chose D^0 final states such, that $D^0 \to f_i$ is doubly Cabibbo suppressed, while $\overline{D}^0 \to f_i$ is Cabibbo allowed. Then, the two interfering amplitudes become comparable in size, maximizing the observable *CP* violation effects. One can use for example $K^+\pi^-$ as the final state—the exact opposite of the one used in the GLW method. However, the ADS method comes at the prize of hadronic uncertainties. In that sense, it is complementary to the clean GLW method. Currently the ADS analyses give the most precise measurements of the ratio r.

The second alternative was proposed by Giri, Grossman, Soffer, and Zupan (GGSZ, [23]) in 2003. Their idea is to observe the interference between $B \to D^0K$ and $B \to \overline{D}^0K$ decays when the D^0 and $\overline{D}^0$ decay to the self-conjugate Cabibbo-allowed final states $K^0_S\pi^+\pi^-$ and $K^0_SK^+K^-$. Then one determines γ from the comparison between the observed Dalitz-plot distribution of $K^0_Sh^+h^-$, $(h = \pi, K)$ produced in $B \to D^0K$, $D^0 \to K^0_Sh^+h^-$, with the Dalitz-plot distributions of $K^0_Sh^+h^-$ observed in decays of flavor-tagged D^0 and $\overline{D}^0$-mesons. A Dalitz plot shows on each of its axes the square of an invariant mass formed of two out of the three particles in the three-body final states: for instance $m^2_- = m(K^0_S\pi^-)^2$ and $m^2_+ = m(K^0_S\pi^+)^2$ in case of the $D^0 \to K^0_S\pi^+\pi^-$ decay. In this kind of plot the resonant substructure of the decay becomes clearly visible, which is different (intrinsically reflected) for a D^0 and $\overline{D}^0$. The GGSZ method yields the most precise single measurement of γ, see for instance the recent work by the *BABAR* Collaboration [24]: they measure $\gamma = (76 \pm 22 \pm 5 \pm 5)^\circ$, where the errors are statistical, systematic, and related to the Dalitz model, respectively.

The current experimental situation on the GLW parameters is summarized in Table 3.1 and Figure 3.3. The values of $R_{CP\pm}$ and $A_{CP\pm}$ have been measured by the *BABAR*, Belle, and CDF collaborations. Figure 3.3 compares the results also with the values measured in the excited final states, $B^{\pm}\to D^{(*)0}K^{(*)\pm}$. It is interesting to note that in the $B \to D^0_{CP+}K$ channel a combined direct-*CP*-violating asymmetry is found of more than 3σ significance.

Table 3.1: Values of $A_{CP\pm}$ and $R_{CP\pm}$ measured by *BABAR*, Belle, and CDF for the GLW method as of September 2008.

	BABAR [5] $N_{B\overline{B}} = 382\times10^6$	Belle [25] $N_{B\overline{B}} = 275\times10^6$	CDF [26] 1fb^{-1}	HFAG [27] Average
A_{CP+}	$+0.27\pm0.09\pm0.04$	$+0.06\pm0.14\pm0.05$	$+0.39\pm0.17\pm0.04$	$+0.24\pm0.07$
A_{CP-}	$-0.09\pm0.09\pm0.02$	$-0.12\pm0.14\pm0.05$	–	-0.10 ± 0.08
R_{CP+}	$1.06\pm0.10\pm0.05$	$1.13\pm0.16\pm0.08$	$1.30\pm0.24\pm0.12$	1.10 ± 0.09
R_{CP-}	$1.03\pm0.10\pm0.05$	$1.17\pm0.14\pm0.14$	–	1.06 ± 0.10

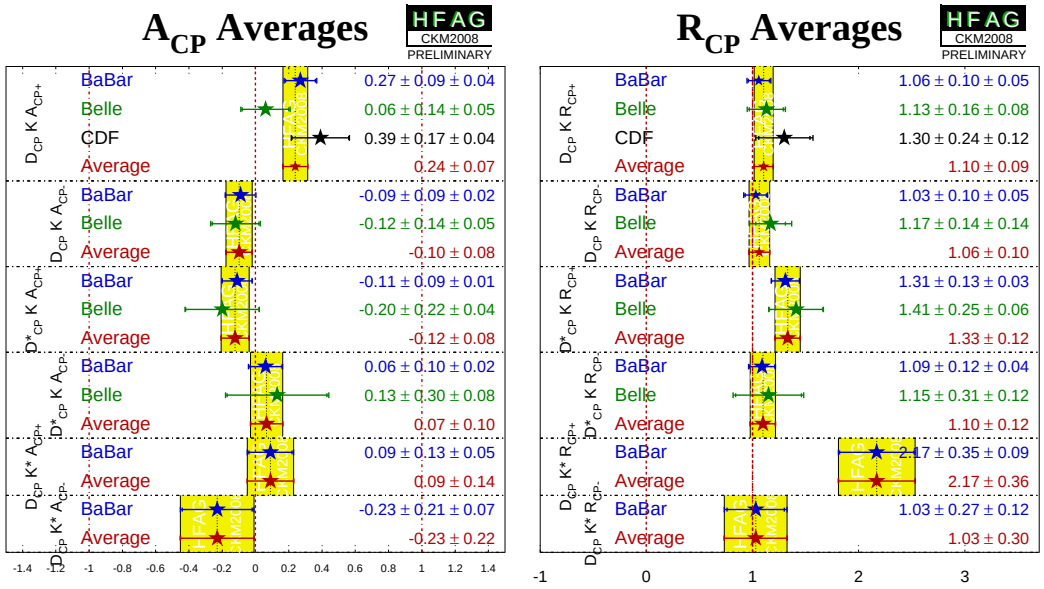

Figure 3.3: Values of $A_{CP\pm}$ and $R_{CP\pm}$ measured by *BABAR*, Belle and CDF in $B \to DK$, D^*K and DK^* decays.

3.2 Analysis Description

We now give an outline of the analysis strategy leading to a measurement of the four GLW parameters $A_{CP\pm}$ and $R_{CP\pm}$. When considering $B \to D^0K$ decays one can benefit from the flavor SU(2) related $B \to D^0\pi$ decays, which are kinematically very similar to the signal. Typically their branching ratio is about twelve times larger than the one of $B \to D^0K$ decays, therefore we have a powerful control sample at hand. We can use it to reduce the systematic

uncertainty when measuring $R_{CP\pm}$. To this extent we construct double ratios of the signal channel and the control channel. Then those systematic uncertainties cancel, which enter as a factor, and affect both samples at the same time. These are, for instance, uncertainties from the D^0 branching fractions, and reconstruction efficiencies of different D^0 channels. In detail, we define the ratios $R^{\pm}_{K/\pi}$ and $R^{\text{flv}}_{K/\pi}$:

$$R^{\pm}_{K/\pi} \equiv \frac{\mathcal{B}(B^- \to D^0_{CP\pm}K^-) + \mathcal{B}(B^+ \to D^0_{CP\pm}K^+)}{\mathcal{B}(B^- \to D^0_{CP\pm}\pi^-) + \mathcal{B}(B^+ \to D^0_{CP\pm}\pi^+)}\ , \tag{3.14}$$

$$R^{\text{flv}}_{K/\pi} \equiv \frac{\mathcal{B}(B^- \to D^0K^-) + \mathcal{B}(B^+ \to \overline{D}^0K^+)}{\mathcal{B}(B^- \to D^0\pi^-) + \mathcal{B}(B^+ \to \overline{D}^0\pi^+)}. \tag{3.15}$$

Then we approximate $R_{CP\pm}$ by the double branching fraction ratios

$$R_{CP\pm} \approx \frac{R^{\pm}_{K/\pi}}{R^{\text{flv}}_{K/\pi}}. \tag{3.16}$$

This approximation is exact in the limit $A(B^+ \to D^0_{CP\pm}\pi^+) = A(B^- \to D^0_{CP\pm}\pi^-) \approx \frac{1}{\sqrt{2}}\,A(B^- \to D^0\pi^-)$, else one discards a term $r \times |V_{us}V_{cd}/V_{ud}V_{cs}| \approx 0.01$. A systematic uncertainty will be assigned due to this approximation when quoting the final numbers for $R_{CP\pm}$.

The outline of this analysis is now straightforward. To determine the four parameters $A_{CP\pm}$ and $R_{CP\pm}$, we have to measure how many events of a certain type are contained in a sufficiently large dataset of decaying B-mesons. In particular, we will focus on eight D^0 final states: K^+K^- and $\pi^+\pi^-$ (CP+), $K^0_S\pi^0$, $K^0_S\omega$, and $K^0_S\phi$ (CP–), $K^-\pi^+$, $K^-\pi^+\pi^0$, and $K^-\pi^+\pi^+\pi^-$ (non-CP).

The analysis has been implemented in the ROOT framework [28]. It is widely used in the field, mostly because it has a strong developer base at CERN. Among many other features, ROOT re-implements in `C++` the software package MINUIT [29], which since its creation in the 1970s has become the standard package to perform function minimizations in particle physics. Such minimizations are a crucial ingredient to regression (or *fitting*) algorithms. The wide use of ROOT has inspired other authors to develop extension modules. Two of which play an important role in this work: TMVA [30] and ROOFIT [31].

The TMVA package, for *Toolkit for Multivariate Data Analysis with Root*, implements many different multivariate analysis techniques, such as neural networks, boosted decision trees, and others. We will use TMVA to optimize our event selection, and to form a *Fisher discriminant*.

The ROOFIT package is an interface to ROOT's implementation of MINUIT. It takes care of many steps necessary in performing fits, such as building the probability density functions, normalizing them, steering MINUIT, and many more. A second functionality of ROOFIT is to visualize the data and the models used to describe it. This non-trivial task is highly automated. We will just highlight one topic: statistical meaningful error bars when plotting histograms of datasets. ROOFIT implements the *classical central* intervals as described in [32]. These correspond to the 68% confidence intervals for Poisson statistics, which are generally asymmetric.[2] All fits performed in this project have been implemented using the ROOFIT package.

[2] In case of a bin entry being zero, this results in a non-zero error bar.

4 The BABAR Experiment

We have seen in the previous chapters how B-Factories play an important role in understanding the *CP* sector of the Standard Model. We shall now review the main features of the *BABAR* Experiment.

The *BABAR* detector was designed to efficiently detect the decay products of B-mesons, which are produced in e^+e^- collisions at the center-of-mass energy of 10.58 GeV. The particle beams are accelerated by the laboratory's 3 km long linear accelerator and stored in the PEP-II storage rings. The full accelerator can reach energies up to 50 GeV, which is too high for a B-Factory. This is why the beams are extracted before the end of the accelerator. Their energies are tuned to be 9.0 GeV and 3.1 GeV for the electrons and positrons, respectively. This results in a Lorentz boost $\beta\gamma = 0.56$ of the center-of-mass system. The boost is needed in order to separate the two B decay vertices ($\beta\gamma c\tau \approx 270\,\mu$m) to be able to measure the proper time difference between the two decays, which is a crucial ingredient for some of the golden measurements of *BABAR*.

The data taking phase of *BABAR* ended in April 2008, concluding a very successful running period of almost ten years. The design luminosity of $3 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ was reached in October 2000, and in October 2005 a record peak luminosity of $1 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ was delivered to the *BABAR* experiment. The period of active running was divided into seven parts, Run 1-7. During the first six Runs, $\Upsilon(4S) \to B\overline{B}$ data was collected at 10.58 GeV, Run 7 explored other energy regimes and the $\Upsilon(2S)$ and $\Upsilon(3S)$ resonances. About 10 % of the data was collected approximately 40 MeV below the $\Upsilon(4S)$ resonance peak. This *off-peak* data doesn't contain B-mesons and can be used to study continuum background processes from $e^+e^- \to q\overline{q}$, $q = u, d, s, c$. Data samples collected at the $\Upsilon(4S)$ peak energy will be referred to as *on-peak* samples. The integrated luminosity for all data taking periods is shown in Figure 4.1. Table 4.1 shows the size of the data samples collected in each Run. An overview of the *BABAR* detector is shown in Figure 4.2. We shall now briefly describe the major subsystems as they are traversed by a particle produced at the interaction point—a more detailed description of various components can be found in [33].

Table 4.1: Data sample used for the analysis. The measured $B\overline{B}$ cross-section is 1.098 ± 0.012 nb.

Sample	$\int \mathcal{L}dt$ (fb^{-1})	N($B\overline{B}$)/10^6
Run 1	20.40	22.4 ± 0.3
Run 2	61.08	67.4 ± 0.7
Run 3	32.28	35.6 ± 0.4
Run 4	100.28	110 ± 1
Run 5	133.26	147 ± 2
Run 6	78.37	84 ± 1
Total on-peak	425.68	467 ± 5
Total off-peak	44.44	

The innermost part is the Silicon Vertex Tracker (SVT). It is one of the two *BABAR* tracking devices and is designed to reconstruct decay vertices with high resolution. The resolution is

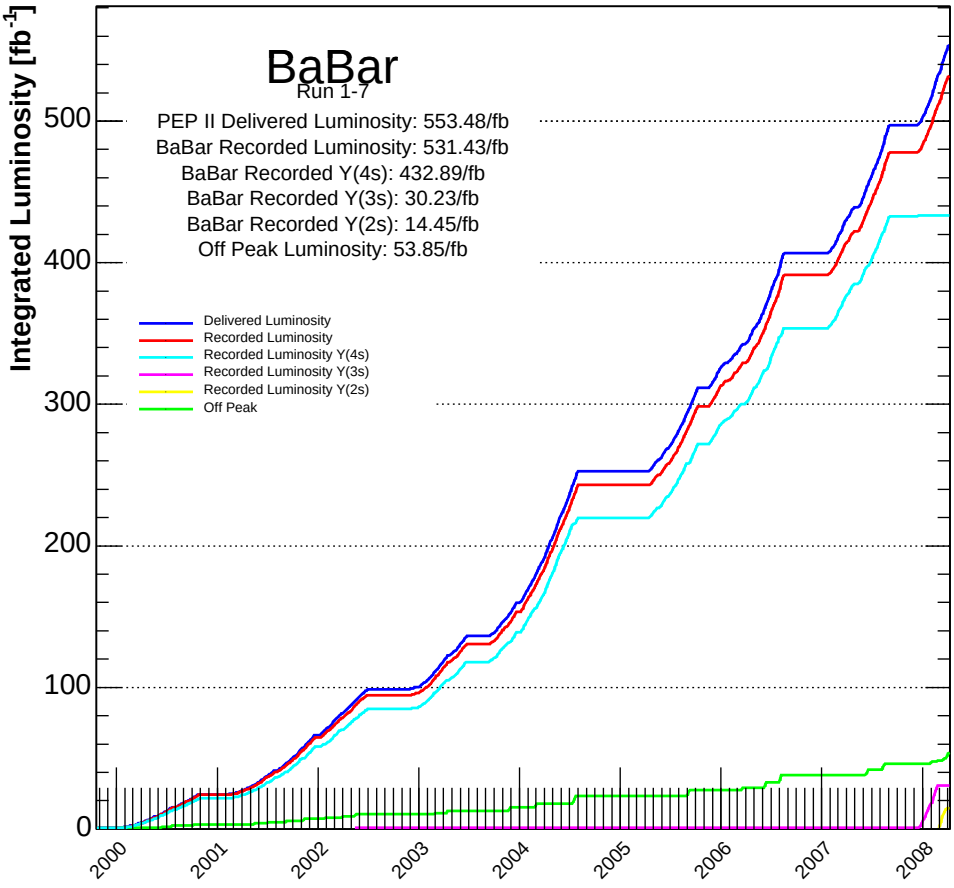

Figure 4.1: *BABAR* and PEP-II integrated luminosity in units of fb^{-1}.

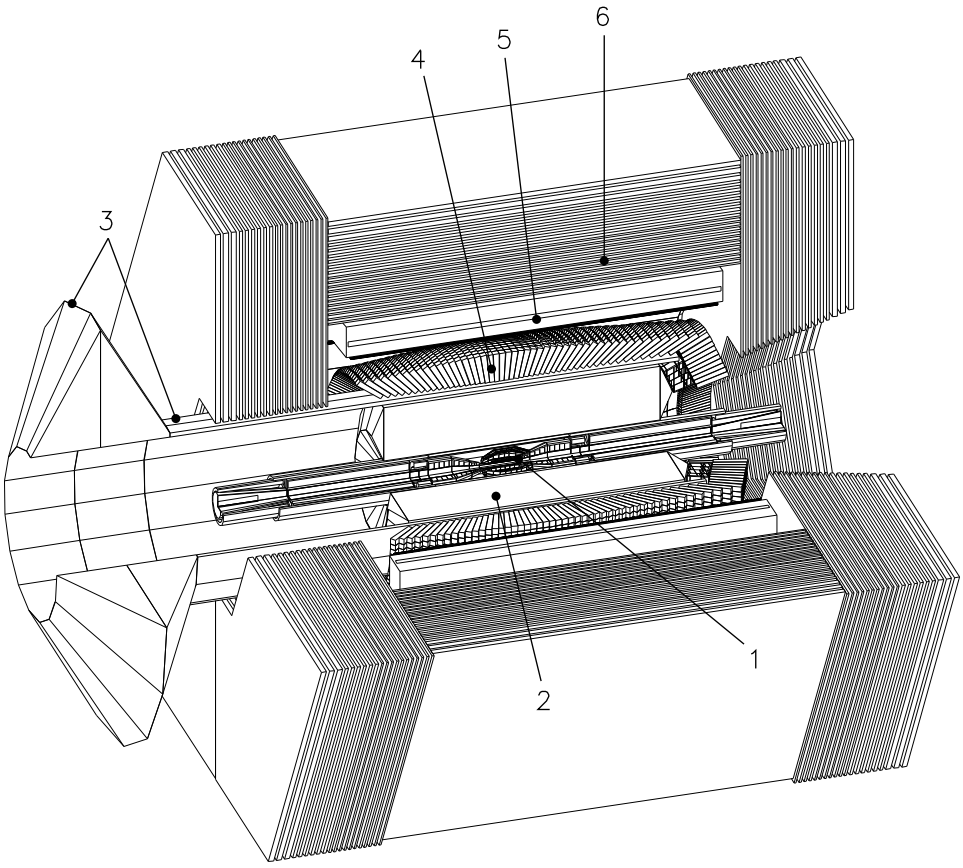

Figure 4.2: Layout of the *BABAR* detector: (1) Silicon Vertex Tracker (SVT); (2) Drift Chamber (DCH); (3) Cherenkov Detector (DIRC); (4) Electromagnetic Calorimeter (EMC) (5) Magnet Coil; (6) Instrumented Flux Return (IFR). The high energy electrons enter from the left side and the low energy positrons from the right side of the figure.

about 80 μm for fully reconstructed B-mesons, and about 180 μm for partially reconstructed ones. The SVT consists of five cylindrical layers of double-sided silicon micro strip detectors, with a total of approximately 150.000 readout channels. For charged tracks with transversal momenta lower than $p_t \approx 100\,\mathrm{MeV}/c$, the SVT provides the only tracking information. The inner radius of the SVT is 32 mm, the outer radius is 144 mm, the total strip length is 26 cm.

Charged tracks with momenta greater than $100\,\mathrm{MeV}/c$ do reach the next layer of the *BABAR* detector, the Drift Chamber (DCH). The DCH is a cylinder with a length of 280 cm, an inner radius of 24 cm, and an outer radius of 81 cm. It is filled with an Helium-isobutane gas mixture, in which a traversing charged particle produces ions and electrons. The latter drift towards one of the 7104 sense wires, where they produce an electric signal. Each of the 20 μm thin sense wires is surrounded by six field wires, providing the electrical field causing the drift. The wires are organized in 40 layers of drift cells. The geometrical acceptance of the DCH is 90 %.

After leaving the DCH, a particle reaches the long synthetic quartz bars of *BABAR*'s ring imaging Cherenkov detector (DIRC, Detector of Internally Reflected Cherenkov light). The DIRC is designed to provide excellent particle identification, in particular to separate kaons from pions. The DIRC is most efficient for track momenta basically from threshold up to $4\,\mathrm{GeV}/c$. Separating kaons from pions is of special interest for the analysis presented in this document, as we will have to find a way to distinguish $B \to D^0 K$ and $B \to D^0 \pi$ events. The DIRC makes use of the fact that charged particles moving faster than the speed of light c within a radiator of refractive index n emit Cherenkov light under a well-defined angle θ_C,

$$\cos\theta_C = \frac{1}{n\beta}, \tag{4.1}$$

where $\beta = \frac{v}{c}$ and v is the speed of the particle. So given a particle's momentum is known, one can infer its mass. The radiator consists of 144 straight fused silica bars ($n = 1.473$) with rectangular cross section. Each bar has a profile of $1.7\,\mathrm{cm} \times 3.5\,\mathrm{cm}$ and a length of 4.9 m. The light emitted by the charged particles travels through the radiator bars and after multiple internal reflections it reaches the rear end, where the Cherenkov image is allowed to expand. The expansion region is filled with 6000 liters of purified water, its back is closely packed with 10572 photomultiplier tubes. The Cherenkov light forms rings in the back plane, from which pattern recognition algorithms deduce θ_C. The geometrical acceptance of the DIRC is 80 %, which is a bit smaller compared to the DCH.

The next detector component a particle will reach is the Electromagnetic Calorimeter (EMC). It is the main device for electron-pion separation and neutral pion and photon reconstruction. It consists of 6580 Thallium-doped Cesium iodide crystals. The crystals are of trapezoidal shape, their typical transverse dimensions are $5 \times 5\,\mathrm{cm}^2$ at the front, flaring out to about $6 \times 6\,\mathrm{cm}^2$ at the back. Their typical length is 30 cm, representing about $X_0 = 17$ radiation lengths. The EMC features a dedicated calibration system, which uses a neutron generator to activate liquid fluorinert. The fluorinert is pumped through tubes in front of the crystals, the crystals then detect the 6.1 MeV photons from the ^{16}N β-γ cascade. The EMC also features a monitoring system, the light pulser, which distributes the light of a single xenon lamp to each individual crystal and is able to quickly test the readout chain. Each of the crystals is being read-out by two photodiodes glued to the rear face, which are not affected by the strong magnetic field they lie in.

The strong magnetic field (1.5 T) in which all of the above components lie, is provided by a super-conducting solenoid. The flux return yoke is instrumented (IFR, Instrumented Flux Return) to identify muons and to detect neutral hadrons. The yoke consists of 18 layers of iron plates with increasing thickness, representing a total thickness of 65 cm. The 17 gaps

between them house two kinds of detector systems, Resistive Plate Chambers (RPC) and Limited Streamer Tubes (LST), the total active area exceeding 1000 m^2. The initial design of the IFR was entirely based on RPC layers. However, they had to be replaced in 2004 and 2006 by the LST layers, because the RPCs developed large inefficiency for muon reconstruction due to unforeseen environmental conditions in the detector hall.

To summarize the physics parameters of the *BABAR* detector, Table 4.2 gives an overview of the most important design parameters.

Table 4.2: Design parameters of the *BABAR* detector. Acceptance coverage in the e^+e^- center-of-mass system.

Parameter	Value [33]
Tracking coverage ($/4\pi$)	0.92
$\sigma_{p_t}/p_t(\%)$ (1 GeV pions at 90 deg)	0.36
$\sigma_{z_0}(\mu\text{m})$ (1 GeV pions at 90 deg)	52
Calorimeter coverage ($/4\pi$)	0.90
X_0 in front of Calorimeter (at 90 deg)	0.25
$\sigma_E/E(\%)$ (1 GeV γ at all angles)	1.8
γ efficiency within acceptance (at 100 MeV)	0.92
Charged Hadron ID coverage ($/4\pi$)	0.84

At high energy physics experiments the interaction rate is usually much higher than the rate limits given by the detector readout system and by off-line computing. A fast trigger system is needed, pre-selecting interesting events. For *BABAR*, however, the situation is not so bad. The *BABAR* trigger has an efficiency greater than 99% for $B\overline{B}$ events while keeping the output rate below 400 Hz. This rate basically contains all hadronic (95 %) and tau (90 %) events. Only a simple pre-scaling is needed to suppress the rates of QED events: Bhabha scattering, dimuon events, and diphoton events.

Finally, events are reconstructed off-line from the raw information provided by all detector components. Such information includes for instance hits in the SVT, charges on the sense wires of the DCH, and so on. From this information tracks are reconstructed, which yield physical properties of the particles such as mass, energy and momentum. However, since one can't be sure about the true nature of a track, they are considered particle *candidates*. At reconstruction level, these candidates are grouped into lists corresponding to certain basic selection criteria, for instance, the `ChargedTracks` list contains all tracks obtained by fitting a trajectory with track hits from SVT or DCH. The higher-level analysis software uses these lists as input. To identify a charged track as a candidate pion, kaon, electron or muon, additional particle identification (PID) information is used. Besides the aforementioned Cherenkov angle provided by the DIRC, this is mainly the energy loss $\mathrm{d}E/\mathrm{d}x$ along the track in the tracking detectors, SVT and DCH. For tracks that don't have enough momentum to reach the DIRC, the SVT and DCH in fact provide the only PID information. All PID information is combined into particle selectors, which particle candidates can either pass or fail. The `KLHVeryTight` selector plays a special role in this analysis, it selects kaons based on a likelihood ratio approach (see Section 6.1).

The physics analysis of datasets collected by modern particle detectors such as *BABAR* requires a safe knowledge of the whole experimental setup. The way to gather this knowledge is long. One performs calibrations of the subsystems, takes cosmic data, takes data from test beams, collects data control samples, just to name a few things. This knowledge is implemented into Monte Carlo simulations. Then one generates large simulated data samples of statistics comparable to the real dataset. Such simulation is carried out stepwise: at first, an event

generator simulates the primary e^+e^- reaction and the decays of the produced unstable particles, taking into account up-to-date knowledge of the decay properties. For instance, `EvtGen` [34] is used for exclusive B decays. Then the particles are traced through the detector volume, simulating the interactions with the active and passive material [35]. Finally, the detector response is simulated. To get a realistic description of the underlying background processes, noise and beam background events, as picked up during data taking, are added. Simulated events undergo the same reconstruction process as the data. In addition, a so called *truth match* procedure is applied, which tries to match each reconstructed track to a particle on generator level. A candidate particle will be called truth-matched, if its origin is most likely a corresponding generated particle. This is a very powerful tool to understand the MC samples.

In this analysis, we will distinguish a set of various dedicated MC samples: signal MC samples, generic $B\overline{B}$ MC samples, and generic $q\overline{q}$ MC samples. In signal MC, only $\Upsilon(4S)$ decays are generated, and one of the two B-mesons produced in the $\Upsilon(4S)$ decay is forced to decay into the signal mode. The result are large samples with extremely high purity. They are used to study the properties of the signal in great detail, such as reconstruction efficiencies and distribution shapes. In $B\overline{B}$ MC, both B-mesons decay generically into final states, taking into account full present knowledge of the branching ratios. In $q\overline{q}$ MC, only non-resonant $e^+e^- \to q\overline{q}$ processes are simulated, with $q\overline{q} = u, d, s, c$. A Monte Carlo sample which closely mimics real data therefore consists of a cocktail of signal, $q\overline{q}$, and $B\overline{B}$ Monte Carlo, weighted according to their equivalent luminosities.

5 Partial Result

A partial result of this work was obtained in 2007 by an analysis group in which the author participated. The result was published in 2008 [5]. It is based on 382×10^6 $B\overline{B}$ decays, contained in the Run 1-5 datasets. The full Run 1-6 analysis implements many improvements compared to this partial result. But some parts of it still follow the same lines. In this chapter we review the partial result in some detail, serving as an outline for the following chapters. We reconstruct $B^\pm \to D^0 h^\pm$ decays, where the prompt track $h^\pm$ is either a kaon or a pion, and D^0 is either a D^0 or $\overline{D}^0$-meson. The D^0 candidates are reconstructed in the CP-even eigenstates $\pi^+\pi^-$ and K^+K^- (D^0_{CP+}), in the CP-odd eigenstates $K^0_S\pi^0$ and $K^0_S\omega$ (D^0_{CP-}), and in the (non-CP) flavor eigenstate $K^-\pi^+$. When reconstructing B-meson candidates by combining a D^0 candidate with a track h in the $D^0 \to K^-\pi^+$ mode, we require that the charge of the track h matches that of the kaon from the D^0-meson decay, selecting $b \to c$ mediated B decays.

We optimize the event selection in order to minimize the statistical error on the expected $B \to D^0 K$ signal yield, determined for each D^0 decay channel using simulated signal and background events. We apply cuts on invariant masses of intermediate particles. They are required to be within $2.5\,\sigma$ of the mean fitted mass. In case of the invariant mass of a D^0 candidate, for instance, σ is ranging from 4 to $20\,\mathrm{MeV}/c^2$ depending on the D^0 decay mode. We also apply cuts to the particle-ID information of charged kaons and pions produced in D^0 decays, and in some cases we exploit different angular distributions between signal and background.

To reduce background from $e^+e^- \to q\overline{q}$ events (with $q = u, d, s, c$), we exploit the fact that B-meson decays tend to be spherical in the e^+e^- center-of-mass system because they are almost at rest, whereas $q\overline{q}$ events tend to be jet-like. We construct a linear Fisher ($\mathcal{F}$) discriminant based on four event-shape quantities, and cut on the value of $\mathcal{F}$. We also cut on the beam-energy substituted mass, m_{ES}, as defined in Equation 2.12. For events with multiple $B \to D^0 h$ candidates, we choose the B candidate with the smallest χ^2, defined in Eq. 6.8. It is formed from the measured and true masses of the composite candidates, and the resolutions and widths of the reconstructed mass distributions. This requirement enhances the purity of the sample because the probability of having two real signal decays in the same event is negligible, while that of having one or more additional fake candidates is not. By removing the additional candidates while keeping the one with the smallest χ^2 the background is reduced. The total reconstruction efficiencies are in the range of 6 %–36 %, depending on the D^0 final state. In Chapter 6 we describe the event selection of the main analysis.

We select B-meson candidates using the energy difference ΔE, as introduced in Equation 2.13. The ΔE distribution depends on the mass assigned to the prompt track. We evaluate ΔE with the kaon mass hypothesis so that the peaks of the distributions are centered near zero for $B \to D^0 K$ events and shifted by approximately 50 MeV for $B \to D^0\pi$ events. This provides a measure to separate both decays. The ΔE resolution for $B \to D^0 K$ events is typically 16 MeV for all D^0 decay modes under study.

In addition to background from $q\overline{q}$ events, there are contributions from $B\overline{B}$ events. They come from the processes $B^- \to D^* h^-$, $B^- \to D^0\rho^-$, misreconstructed $B \to D^0 h$, and from charmless

B decays to the same final state as the signal: for instance, the process $B^-\to K^-K^+K^-$ is a background for $B^-\to D^0K^-$, $D^0\to K^-K^+$. These charmless backgrounds have similar ΔE distribution as the D^0K^- signal and are therefore referred to as peaking $B\overline{B}$ backgrounds ($B^-\to X_1X_2h^-$). These peaking backgrounds are carefully estimated from D^0 mass sidebands. In case of $D^0\to K^-\pi^+$, charmless backgrounds from $B\to K\pi h$ are negligible compared to the much more abundant signal.

We determine the signal and background yields for each D^0 decay mode independently from a two-dimensional extended unbinned maximum-likelihood fit to the selected data events. The fit is performed simultaneously on the B^+ and B^- subsamples. The input variables to the fit are ΔE and the Cherenkov angle θ_C of the prompt track as measured by the DIRC. The extended likelihood $\mathcal{L}$ for N candidates is given by the product of the probabilities for each individual candidate i and a Poisson factor:

$$\mathcal{L} = \frac{e^{-N}(N)^n}{n!}\prod_{i=1}^{n}\mathcal{P}_i(\Delta E,\theta_C). \tag{5.1}$$

The probability $\mathcal{P}_i$ is the sum of the signal and background terms,

$$\mathcal{P}_i(\Delta E,\theta_C) = \sum_J \frac{N_J}{n}\,\mathcal{P}^J_{\Delta E,i}\,\mathcal{P}^J_{\theta_C,i}\quad , \tag{5.2}$$

where J denotes the eight signal and background hypotheses D^0h, $q\bar{q}(h)$, $B\bar{B}(h)$, and X_1X_2h. The total event yield n is estimated by the fit, and N_J is the event yield in each category. The ΔE distribution for $B\to D^0K$ signal is parameterized with a double Gaussian function. The ΔE probability density function (PDF) for $B\to D^0\pi$ is the same as the $B\to D^0K$ one, but with an additional shift, $\Delta E_{\texttt{shift}}$, which arises from the wrong mass assignment to the prompt track. The shift is computed event by event as a function of the prompt track momentum p and a Lorentz factor characterizing the boost to the e^+e^- center-of-mass frame. This crucial point will be explained in more detail in Section 9.1. The particle identification PDF is a double Gaussian as a function of $\theta_C^{\rm pull}$, which is the difference between the measured Cherenkov angle θ_C and its expected value for a given mass hypothesis, divided by the estimated error. In summary, the floating parameters in each of the five fits are the D^0K and $D^0\pi$ signal yield asymmetries, the total number of signal events in $D^0\pi$, the appropriate ratios R and $R_\pm$, eight background yields (one for each charge), and two parameters of the ΔE signal shape (common for positive and negative samples). The remaining shape parameters of the PDF are fixed to values obtained from Monte Carlo simulation, the peaking background yields are fixed to their values obtained from the $m(D^0)$ sideband study. The reliability of the Monte Carlo simulation has been evaluated in control samples. For instance, the parameters of the $\theta_C^{\rm pull}$ PDF can be checked on a dedicated $D^{*0}\to D^0\pi^0$, $D^0\to K^-\pi^+$ sample, which contains kaon and pion tracks of high purity.

The results of the fits are summarized in Table 5.1. Figure 5.1 shows the distributions of ΔE for the $K^-\pi^+$, *CP*+ and *CP*– modes after enhancing the $B\to D^0K$ purity by requiring that the prompt track be consistent with the kaon hypothesis: its Cherenkov angle be within $2\,\sigma$ of the expected value for kaons, and outside $3\,\sigma$ of the expected value for pions.

In case of $D^0\to K^0_S\omega$, $\omega\to\pi^+\pi^-\pi^0$, the values of $R_{CP-}^{K^0_S\omega}$ and $A_{CP-}^{K^0_S\omega}$ need to be corrected to take into account a possible dilution from a non-resonant *CP*-even background arising from $B^-\to D^0h^-$, $D^0\to K^0_S(\pi^-\pi^+\pi^0)_{\rm non-\omega}$ decays. We estimate the corrections using a fit to the ω helicity angle in the selected data events. The ω helicity angle is the angle between the normal to the decay plane and the direction of the ω in the D^0 rest frame. We apply the corrections to the according values of A_{CP} and R_{CP}. Then, the quantities $R_\pm/R$ and $A_{CP\pm}$

Table 5.1: Uncorrected yields as obtained from a maximum likelihood fit to Run 1-5 data. The quoted uncertainties are statistical.

D^0	CP	$N(D\pi^+)$	$N(D\pi^-)$	$N(DK^+)$	$N(DK^-)$
$K^-\pi^+$		12745 ± 120	12338 ± 120	954 ± 36	918 ± 36
K^-K^+	$+$	1109 ± 36	1051 ± 35	51 ± 10	113 ± 13
$\pi^-\pi^+$	$+$	390 ± 24	378 ± 24	39 ± 9	36 ± 9
$K^0_S\pi^0$	$-$	1102 ± 37	1134 ± 38	100 ± 13	88 ± 12
$K^0_S\omega$	$-$	422 ± 24	403 ± 26	29 ± 8	18 ± 8

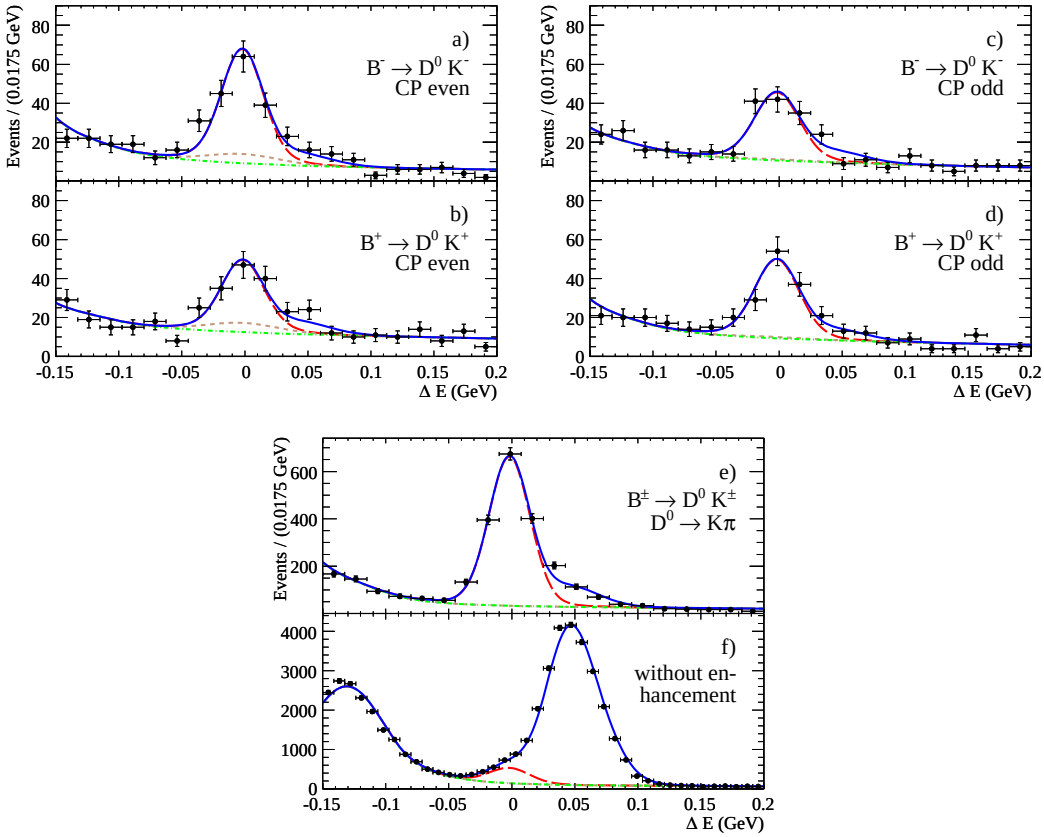

Figure 5.1: Distributions of ΔE for events enhanced in $B^\pm \to D^0K^\pm$ signal: a) $B^- \to D^0_{CP+}K^-$; b) $B^+ \to D^0_{CP+}K^+$; c) $B^- \to D^0_{CP-}K^-$; d) $B^+ \to D^0_{CP-}K^+$; $B^\pm \to D^0K^\pm$, $D^0 \to K^\pm\pi^\mp$ with (e) and without (f) signal enhancement. Blue (continuous) curve: projection of the full PDF of the maximum likelihood fit. Red (long-dashed): $B^\pm \to D^0K^\pm$ signal on all backgrounds. Brown (short-dashed): peaking component on $q\overline{q}$ and $B\overline{B}$ background. Green (dash-dotted): $q\overline{q}$ and $B\overline{B}$ background.

are computed by means of a weighted average over the $CP+$ and $CP-$ modes. The results for the $CP+$ and $CP-$ combinations are reported in Table 5.2.

For the branching fraction ratios $R_{CP\pm}$, a source of systematic uncertainty is associated with the assumption that $R_{CP\pm} = R_\pm/R$. This assumption holds only if the magnitude of the ratio $r_\pi \lesssim 0.012$ between the amplitudes of the $B^- \to \overline{D}^0\pi^-$ and $B^- \to D^0\pi^-$ processes is neglected. All considered systematic uncertainties in $A_{CP\pm}$ and $R_{CP\pm}$ are listed in Table 5.3.

Table 5.2: Measured ratios $R_{CP\pm}$ and $A_{CP\pm}$ for CP-even ($CP+$) and CP-odd ($CP-$) D decay modes. The first error is statistical; the second is systematic.

D^0 mode	R_{CP}	A_{CP}
$CP+$	$1.06 \pm 0.10 \pm 0.05$	$0.27 \pm 0.09 \pm 0.04$
$CP-$	$1.03 \pm 0.10 \pm 0.05$	$-0.09 \pm 0.09 \pm 0.02$

Table 5.3: Systematic uncertainties on the observables $R_{CP\pm}$ and $A_{CP\pm}$ in absolute terms. The total is the squared sum of the contributions.

source	ΔR_{CP+}	ΔR_{CP-}	ΔA_{CP+}	ΔA_{CP-}
fixed fit parameters	0.036	0.019	0.010	0.002
peaking background	0.029	0.037	0.031	0.003
detector charge asym.	-	-	0.022	0.022
opp. CP bkg. in $K^0_S\omega$	-	0.002	-	0.007
$R_{CP\pm}$ vs. $R_\pm$	0.026	0.025	-	-
K/π efficiency	0.002	0.007	-	-
total	0.053	0.049	0.039	0.023

In conclusion, the combined uncertainties we measure for $A_{CP\pm}$ ($R_{CP\pm}$) are smaller by a factor of 0.7 (0.9) and 0.6 (0.6) than the previous *BABAR* [4] and Belle [36] measurements, respectively. We also find A_{CP+} to deviate by 2.8 standard deviations from zero.

Compared to the strategy outlined above, the main analysis presented in this document implements a list of improvements. First of all, it benefits from a 22 % increase in number of $B\overline{B}$ pairs (from 382×10^6 to 467×10^6), representing the final *BABAR* dataset. Secondly, the event selection has been improved: there is no longer a cut on the value of $\mathcal{F}$, instead it will be included in the fit. The efficiency of the previous cut on $\mathcal{F}$ was ranging from 74 % to 78 % for $B \to D^0K$ signal events [37], so we can retain a sizable part of signal events. At third, a more sophisticated fit strategy was adopted. Apart from including $\mathcal{F}$ in the fit, we also include $m_{\rm ES}$ and we drop θ_C. Instead we distinguish between $B \to D^0K$ and $B \to D^0\pi$ events by means of a dedicated likelihood selector, `KLHVeryTight`, which is based on $\mathrm{d}E/\mathrm{d}x$ measurements in the tracking detectors in addition to θ_C. This choice increases the particle-identification efficiency for prompt tracks at low momentum and outside of the geometrical acceptance of the DIRC. Lastly, we will reconstruct the D^0 in three additional final states: the flavor channels $D^0 \to K^-\pi^+\pi^0$ and $K^-\pi^+\pi^+\pi^-$, and the $CP-$ channel $K^0_S\phi$. In the partial result presented in this section, the latter mode was not included to allow for an easy combination of the GLW result with the one obtained by a *BABAR* GGSZ analysis of $B \to D^0K$, $D^0 \to K^0_SK^+K^-$ decays (see Section 3 and Reference [24]).

In combination, all of these measures will reduce the uncertainty of the $A_{CP\pm}$ and $R_{CP\pm}$ parameters by a factor of ≈ 0.6. They will represent the most precise measurement of the GLW parameters achievable with the $B \to D^0K$ *BABAR* dataset.

6 Event Selection

After the partial result has been presented in the previous chapter, we shall now turn to the full analysis. The full analysis has been performed using the Run 1-6 datasets, corresponding to 467.4 ± 5.2 million $B\overline{B}$ pairs. This represents the final *BABAR* on-peak dataset. We will reconstruct the signal decays $B \to D^0 h$, where the bachelor track h refers to both $h = \pi, K$. The D^0-meson will be reconstructed in eight different final states, the *CP*-even K^+K^- and $\pi^+\pi^-$, the *CP*-odd $K^0_S\pi^0$, $K^0_S\omega$, and $K^0_S\phi$, and the non-*CP* states $K^-\pi^+$, $K^-\pi^+\pi^0$, and $K^-\pi^+\pi^+\pi^-$. We consequently use a set of 16 signal Monte Carlo samples, each with at least 10^5 events. They are listed in Table 6.1, together with the decay modes used to reconstruct the unstable D^0 decay products which decay outside the detector volume. Table 6.2 lists the generic Monte Carlo samples used in the analysis.

We shall quickly review the experimental situation of the branching fractions relevant for this analysis. The branching fraction for the $B \to D^0 h$ processes are $\mathcal{B}(B^- \to D^0\pi^-) = (4.84 \pm 0.15) \times 10^{-3}$ and $\mathcal{B}(B^- \to D^0 K^-) = (4.02 \pm 0.21) \times 10^{-4}$ [8]. Taking into account these numbers and the secondary branching fractions for the decays of the D^0-mesons and of their daughters, the total branching fractions for the $B \to D^0 K$ decays (D^0 decaying into a flavor mode) that are reconstructed in this analysis are in the range $1.6 - 5.5 \times 10^{-5}$. The total branching fractions for the $B \to D^0_{CP\pm} h$ decays are in the range $6 - 40 \times 10^{-7}$. The overall branching fractions are summarized in Table 6.3.

Table 6.1: Signal MC samples used in the analysis. The number of generated events is driven by computing performance. The equivalent integrated luminosity is calculated taking into account the branching fractions from Table 2.1.

D^0 mode	$B \to D^0 K$		$B \to D^0 \pi$	
	Events $\times 10^3$	$\int \mathcal{L} dt$ (fb^{-1})	Events $\times 10^3$	$\int \mathcal{L} dt$ (fb^{-1})
K^+K^-	387	202814	387	18530
$\pi^+\pi^-$	387	570550	387	52127
$K^0_S\pi^0,\ K^0_S \to \pi\pi$	387	97974	387	8626
$K^0_S\omega,\ K^0_S \to \pi\pi,\ \omega \to \pi\pi\pi^0$	387	121968	387	10738
$K^0_S\phi,\ K^0_S \to \pi\pi,\ \phi \to KK$	387	550821	387	48495
$K^-\pi^+$	387	22539	1941	9389
$K^-\pi^+\pi^0$	387	6308	1941	2628
$K^-\pi^+\pi^+\pi^-$	387	10824	387	899

The reconstruction and selection of events is performed in three major steps: skimming, pre-selection, and final selection. In each of the steps the selection requirements get tighter, until the final samples have the highest purity. The reason for this stepwise approach is the limited performance of the computing facilities: at skim level, the total size of the samples is on the order of terabytes, pre-selection reduces it to about 100 GB, the final samples are on the order of 100 MB. In the following we will not discuss the skimming process, which applies a looser (and faster in terms of CPU cycles) selection compared to the one used in

Table 6.2: Generic MC samples used for the analysis: number of generated events and corresponding integrated luminosity, according to the cross sections listed in the last column.

Sample	Events/10^6	$\int \mathcal{L}dt$ (fb^{-1})	Cross section (nb)
$e^+e^- \to q\overline{q},\ q = u, d, s$	938	449	2.09
$e^+e^- \to c\overline{c}$	1132	871	1.30
$e^+e^- \to \Upsilon(4S) \to B^0\overline{B}^0$	736	1340	0.549
$e^+e^- \to \Upsilon(4S) \to B^+B^-$	731	1332	0.549

Table 6.3: Total branching fractions for the $B^- \to D^0\pi^-$ decays that are reconstructed in this analysis.

D^0 decay mode	Overall $\mathcal{B}(B^- \to D^0\pi^-)$	Overall $\mathcal{B}(B^- \to D^0K^-)$
$K^-\pi^+$	$(18.8 \pm 0.6) \times 10^{-5}$	$(15.6 \pm 0.8) \times 10^{-6}$
$K^-\pi^+\pi^0(\gamma\gamma)$	$(66.5 \pm 3.2) \times 10^{-5}$	$(55.2 \pm 3.5) \times 10^{-6}$
$K^-\pi^+\pi^+\pi^-$	$(39.2 \pm 1.6) \times 10^{-5}$	$(32.6 \pm 1.9) \times 10^{-6}$
K^+K^-	$(19.0 \pm 0.7) \times 10^{-6}$	$(17.4 \pm 1.7) \times 10^{-7}$
$\pi^+\pi^-$	$(67.6 \pm 2.5) \times 10^{-7}$	$(61.8 \pm 6.1) \times 10^{-8}$
$K^0_S(\pi^+\pi^-)\ \pi^0(\gamma\gamma)$	$(40.4 \pm 2.3) \times 10^{-6}$	$(35.5 \pm 4.2) \times 10^{-7}$
$K^0_S(\pi^+\pi^-)\ \phi(K^+K^-)$	$(72.7 \pm 5.5) \times 10^{-7}$	$(64.0 \pm 8.2) \times 10^{-8}$
$K^0_S(\pi^+\pi^-)\ \omega(\pi^-\pi^+\pi^0(\gamma\gamma))$	$(32.4 \pm 6.0) \times 10^{-6}$	$(28.6 \pm 6.1) \times 10^{-7}$

the following steps of the analysis. We will focus on the pre-selection, and, most importantly, on the final selection.

6.1 Particle Identification

A very important concept in event selection within the *BABAR* analysis framework is the concept of particle identification (PID) selectors. They combine all the information the sub-detectors provide regarding the nature of a particle, and supply an easy-to-use mechanism to access this information. A particle candidate can be required to pass a certain selector. The higher-class the selector it passes, the higher the probability of a candidate's particle hypothesis to be correct. Particle selectors bear speaking names: if a candidate kaon passes the `VeryTight` selector, for instance, it has a higher probability to be a true kaon than if it only passes the `VeryLoose` selector. One defines the efficiency ε of a particle selector as the probability that a true particle passes the selector. The probability that another particle is mis-identified is called m. Typical values of the `KLHVeryTight` selector, which stands for kaon likelihood `VeryTight`, are $\varepsilon \approx 75\,\%$ and $m \approx 2\,\%$.

There are many different implementations of PID selectors, but the ones most important for this analysis are the selectors based on likelihood ratios. The idea is to calculate the likelihood L_i for each particle hypothesis i:

$$L_i = L_i^{\mathrm{DIRC}} \cdot L_i^{\mathrm{DCH}} \cdot L_i^{\mathrm{SVT}}. \tag{6.1}$$

The DCH and SVT likelihoods are calculated by comparing the measured energy loss per track length, $\mathrm{d}E/\mathrm{d}x$ (*meas.*), against the expected $\mathrm{d}E/\mathrm{d}x$ from the Bethe-Bloch (*B.B.*) formula. One defines the pull p, which in theory is distributed normally:

$$p = \frac{1}{\sigma}\left(\left.\frac{\mathrm{d}E}{\mathrm{d}x}\right|_{\mathrm{meas.}} - \left.\frac{\mathrm{d}E}{\mathrm{d}x}\right|_{\mathrm{B.B.}}\right), \tag{6.2}$$

where σ is the measured uncertainty. Then the DCH likelihood is calculated based on a Gaussian probability density, and the SVT likelihood is calculated based on an asymmetric Gaussian with different widths on each side. The DIRC likelihood cannot be calculated in this way, since there are significant tails on the distribution of the Cherenkov angle θ_C. To alleviate these tails, a new likelihood is constructed from θ_C, the number of photons detected in the DIRC, and the number of hits a track has in the SVT and the DCH. Finally, the likelihood selectors consist of different requirements on the relevant likelihood ratios of particles l and m:

$$\mathcal{L}^r_{l/m} = \frac{\mathcal{L}_l}{\mathcal{L}_l + \mathcal{L}_m}. \tag{6.3}$$

Table 6.4 introduces the names of the kaon likelihood selectors, together with the defining requirements on the kaon-pion likelihood ratio, the kaon-proton likelihood ratio, and the pion-proton likelihood ratio. The selectors also incorporate information from the electromagnetic calorimeter and the muon system. These are used to exclude the electron or muon hypotheses.

Table 6.4: Kaon likelihood based PID selectors

K selector	$\mathcal{L}^r_{K/\pi}$	$\mathcal{L}^r_{K/p}$	$\mathcal{L}^r_{\pi/p}$	$\vec{p}$ [GeV/c]	additional requirements
`NotAPion`	> 0.2	-	> 0.02	< 0.40 or not an $e^\pm$	
`VeryLoose`	> 0.5	> 0.018	-	< 0.40 or not an $e^\pm$	
`Loose`	> 0.8176	> 0.018	-	< 0.40 or not an $e^\pm$	
`Tight`	> 0.9	> 0.20	-	< 0.40 or not an $e^\pm$	
`VeryTight`	> 0.9	> 0.20	-	< 0.40 or not an $e^\pm$	not identified as $\mu^\pm$ by the IFR

6.2 Event Pre-Selection

In this section we describe how the unstable particles in the decay chains are reconstructed from the stable ones, and which pre-selection requirements are applied.

Photon candidates, which are used to reconstruct π^0-mesons, are clusters in the electromagnetic calorimeter that are not matched to any charged track, have a raw energy greater than 30 MeV and a lateral shower shape consistent with the expected pattern of energy deposit from an electromagnetic shower. The lateral energy distribution LAT is defined as

$$\mathrm{LAT} = \frac{\sum_{i=3}^{n} E_i r_i^2}{\sum_{i=3}^{n} E_i r_i^2 + E_1 r_0^2 + E_2 r_0^2}, \tag{6.4}$$

with ordered crystal energies $E_1 \geqslant E_2 \geqslant \ldots \geqslant E_n$. The sum extends over all crystals in a shower, $r_0 = 5\,\mathrm{cm}$ is the average crystal width and r_i is the distance between crystal i and the shower center, computed as the center-of-gravity with linear energy weighting of every crystal. We require $\mathrm{LAT} < 0.8$.

We reconstruct K^0_S candidates in the decay mode $K^0_S \to \pi^+\pi^-$. We select pairs of zero total charge formed from all tracks reconstructed in the drift chamber and/or the vertex detector (`ChargedTracks` list). The two charged tracks are assigned the pion mass hypothesis and are constrained to have a common vertex. We accept candidates whose invariant mass lies inside a $\pm 25\,\mathrm{MeV}/c^2$ window around the nominal K^0_S mass [8]. The $\pi^+\pi^-$ invariant mass is then constrained to the nominal K^0_S mass in order to improve the momentum resolution.

The π^0 candidates are reconstructed in the decay mode $\pi^0 \to \gamma\gamma$. Pairs of photon candidates (selected as described above) with invariant mass within the range 115–150 MeV/c^2 and total energy greater than 200 MeV are considered π^0 candidates. At first, the event primary vertex is assumed to be the origin of the two photons. To improve the momentum resolution, the π^0 candidates are then fit kinematically with their masses constrained to the nominal π^0 mass [8], while the origin of the photons is constrained to be the decay vertex of the π^0 mother (for instance, the vertex of the two charged pions in $\omega \to \pi^+\pi^-\pi^0$).

To reconstruct ϕ candidates in the decay mode $\phi \to K^+K^-$, we select every pair of opposite-charge tracks with invariant mass inside a $\pm 30\,\mathrm{MeV}/c^2$ mass window[1] around the nominal ϕ mass [8]. The two charged tracks are assigned the kaon mass hypothesis and are constrained to have a common vertex. They are taken from the `GoodTracksVeryLoose` list, that contains all the reconstructed tracks which satisfy additional requirements: (a) the track momentum measured in the laboratory frame must be less than 10 GeV/c; (b) the distance of closest approach to the nominal interaction point must be less than 1.5 cm in the transverse plane and less than 10 cm along the beam axis. The kaon candidates must also pass the `NotAPion` PID selector.

We reconstruct ω candidates in the decay mode $\omega \to \pi^-\pi^+\pi^0$. Two opposite-charge tracks are selected from the `GoodTracksVeryLoose` list, are assigned the pion mass hypothesis and are constrained to have a common vertex. π^0 candidates are selected in the same way described above, with the additional constraint that the point of origin of the two photons be the common vertex of the two charged tracks. We require the invariant mass of the ω candidates to be inside a $\pm 50\,\mathrm{MeV}/c^2$ window[2] around the ω nominal mass [8].

We reconstruct D^0 candidates as follows. At first, all D^0 daughters are constrained to have a common vertex, assuming the correct mass hypotheses. Then we select D^0 candidates with center-of-mass momentum $p^* > 1.3\,\mathrm{GeV}/c$ and invariant mass inside a $\pm 70\,\mathrm{MeV}/c^2$ window around the nominal D^0 mass [8] for the $K^-\pi^+$, $K^-\pi^+\pi^0$, K^+K^-, $\pi^+\pi^-$, $K^0_S\phi$, and $K^0_S\omega$ modes, $\pm 40\,\mathrm{MeV}/c^2$ for the $K^-\pi^+\pi^+\pi^-$ mode, and $\pm 110\,\mathrm{MeV}/c^2$ for the $K^0_S\pi^0$ mode. To improve the momentum resolution, selected D^0 candidates are refitted with a constraint to the nominal D^0 mass. Only candidates are retained for which the probability of the vertex fit is greater than 0.0001.

Combining a D^0 candidate with a charged track h (*bachelor* or *prompt* track) finally gives the $B \to D^0 h$ candidates. The prompt track is taken from the `GoodTracksVeryLoose` list. It is assigned the kaon mass hypothesis regardless of the true nature of the bachelor track. This has important consequences, as already mentioned in Chapter 5: the ΔE distribution depends on the mass assigned to the prompt track. Due to the chosen mass assignment ΔE peaks near zero for $B \to D^0 K$ events, and at approximately 50 MeV for $B \to D^0\pi$ events, providing a measure to separate both decays. Additionally we require the prompt track's center-of-mass momentum be greater than 0.5 GeV/c. Both the prompt track and the D^0-meson are constrained to have a common vertex. A beam-spot constraint is applied to the B production vertex, in order to improve the quality of the fit to the B decay tree in the $D^0 \to K^0_S\pi^0$ channel. For the flavor modes ($D^0 \to K^-\pi^+$, $K^-\pi^+\pi^0$, $K^-\pi^+\pi^+\pi^-$), the charge of the track h must match that of the kaon from the D^0 decay, selecting $b \to c$ mediated B decays. We pre-select B candidates requiring $5.2 < m_{\mathrm{ES}} < 5.3\,\mathrm{GeV}/c^2$ and $-0.3 < \Delta E < 0.3\,\mathrm{GeV}$. We also apply a pre-selection cut on the quantity R_2, which is a measure of the event shape and will be introduced in Section 7. To reduce a significant portion of $e^+e^- \to q\overline{q}$ continuum background, we require $R_2 < 0.6$.

[1] The ϕ width is $\Gamma_\phi = 4.26 \pm 0.05$ MeV/c^2

[2] The ω width is $\Gamma_\omega = 8.49 \pm 0.08$ MeV/c^2

The pre-selection efficiencies for all D^0 decay modes has been evaluated on signal Monte Carlo, requiring that the reconstructed B candidates are fully truth-matched to the generated ones. Table 6.5 lists the numbers.

Table 6.5: $B \to D^0 h$ reconstruction efficiency after pre-selection, evaluated using signal MC.

D^0 decay	$\varepsilon(D^0K)(\%)$	$\varepsilon(D^0\pi)$ (%)
$D^0 \to K^-\pi^+$	60.42 ± 0.08	60.78 ± 0.04
$D^0 \to K^-\pi^+\pi^0$	28.04 ± 0.07	28.39 ± 0.03
$D^0 \to K^-\pi^+\pi^+\pi^-$	37.91 ± 0.08	38.15 ± 0.08
$D^0 \to K^+K^-$	57.80 ± 0.08	58.39 ± 0.08
$D^0 \to \pi^+\pi^-$	64.36 ± 0.08	64.80 ± 0.08
$D^0 \to K^0_S\pi^0$	32.15 ± 0.08	32.54 ± 0.08
$D^0 \to K^0_S\omega$	15.35 ± 0.06	15.43 ± 0.06
$D^0 \to K^0_S\phi$	35.39 ± 0.08	35.79 ± 0.08

6.3 Final Selection

We describe in this section the final criteria adopted to select samples enriched in $B \to D^0K$ and $B \to D^0\pi$ events, whose yields will be extracted through the final fit described in Section 9. In general, it is a non-trivial task to find suitable discriminating variables and to obtain optimal cuts on them. One usually focuses on maximizing the selection significance, which is defined as

$$S = \frac{N_S}{\sqrt{N_S + N_B}}, \tag{6.5}$$

where N_S is the expected number of signal $B \to D^0K$ events, and N_B is the expected number of background events in the final sample which populate the same region as the signal. Maximizing S corresponds to minimizing the final statistical uncertainties on the $B \to D^0K$ yields, and therefore on $A_{CP\pm}$ and $R_{CP\pm}$. We are looking at an optimization problem in a multi-dimensional parameter space, which in case of sizable correlations among the parameters quickly gets hard to solve. Usually one starts with initial requirements on a set of discriminating variables and optimizes them iteratively. However, considering the number of different final states and the file sizes of the associated samples, this approach seems impracticable. We will instead use an approach based on the *toolkit for multivariate analysis* (TMVA, [30]). The advantage of TMVA is, that it is capable of optimizing rectangular cuts on correlated variables using genetic algorithms to find an optimal working point in the non-unique solution space. This spares the aforementioned complications associated with an iterative procedure.

We will distinguish three sets of discriminating variables, the requirements on each of them optimized in a different manner: the invariant masses of the intermediate particles, the particle ID of intermediate particles, and quantities other than the masses and the PID variables which will be passed to TMVA. We will call the corresponding selection requirements *mass cuts*, *PID requirements*, and TMVA requirements.

All requirements are optimized on the $B \to D^0\pi$ control sample, which is kinematically similar to the $B \to D^0K$ sample. This allows us to finalize the selection without looking at the signal, minimizing any possible experimenter bias. It also takes advantage from the higher statistics of the control sample. To simulate the selection power of the final fit, we impose requirements on the three fit variables ΔE, $m_{\rm ES}$, and the Fisher discriminant. The

latter needs to be tuned at this stage of the analysis (see Section 7). The cut values are $\Delta E \in [60, 125]\,\mathrm{MeV}$, $m_{\mathrm{ES}} \in [5.267, 5.289]\,\mathrm{GeV}/c^2$, $\mathcal{F} > -0.08$. For the whole optimization procedure we pick one candidate at random, if multiple candidates survive the pre-selection.

There is also one selection requirement that we apply in advance. It is designed to remove potentially peaking backgrounds: in the channel $B^- \to D^0 K^-$, $D^0 \to \pi^+\pi^-$, a significant contribution comes from the much more abundant processes $B^- \to D^0\pi^-$, $D^0 \to K^-\pi^+$, and $B^- \to \bar{K}^{*0}_X[K^-\pi^+]\pi^-$ (where K^{*0}_X can be $K^{*0}(892)$, $K^{*0}(1430)$ or nearby resonances), where the prompt pion is incorrectly identified as a D^0 daughter and the charged kaon from the D^0 or K^{*0}_X is incorrectly identified as a B daughter. We remove this background contribution by requiring that the invariant mass of the system given by the prompt track and the pion from D^0 with opposite charge be greater than $1.9\,\mathrm{GeV}/c^2$.

The following enumeration outlines the path of the optimization procedure. In the following sections we will describe the selection requirements on each set of variables in more detail.

1. Extract the cuts on the invariant masses from truth-matched signal MC. Since we expect the mass distributions to be slightly different in data and Monte Carlo, we will use these cuts on Monte Carlo only.
2. Apply the MC mass cuts to background MC composed of the generic MC samples, and to signal MC. Use these new samples to optimize the PID requirements and the TMVA requirements.
3. Apply the so obtained PID and TMVA requirements, but not the Monte Carlo mass cuts, to on-peak data. Use this sample to extract the mass cuts on data.

6.3.1 Mass Cuts on Signal Monte Carlo

We obtain the set of mass cuts we shall use on Monte Carlo by fitting the invariant mass distributions in truth-matched $B \to D^0\pi$ signal MC samples. The fit function is a double Gaussian for all masses but the ϕ mass. The ϕ-meson has a non-negligible width. We use a double Voigtian[3] to describe its mass distribution. All parameters of the fit functions are floating in the fit. For the masses described by a double Gaussian, the selection cut is derived from the narrow component: we select the $2.5\,\sigma$ interval around the mean. In case of the π^0 and ϕ masses, there is no (obvious) narrow component. In this case, we take the intervals which enclose 95 % of the integral of the double Gaussian, both on the left and right hand side of the maximum of the double Gaussian. Figure 6.1 shows the fits on signal MC, in representative D^0 channels, including the fit parameters and the extracted cut values. The cuts intervals are also listed in Table 6.6.

6.3.2 PID Requirements

The requirements on the particle ID of charged kaon and pion candidates reconstructed in the D^0 decay tree are optimized straightforwardly by maximizing S under the assumption of vanishing correlations across the PID variables. We consider all charged kaons and pions except the pions from K^0_S decays, because here the TMVA requirements are powerful enough. We use truth-matched $B \to D^0\pi$ signal MC over a cocktail of generic background MC. We calculate the significance (Eq. 6.5) for each possible PID requirement and use the one which maximizes the significance. The optimal PID selectors are summarized in Table 6.7.

[3] A Voigtian is a Breit-Wigner shape describing the resonance, convolved with a Gaussian describing detector resolution.

Table 6.6: Optimized cuts on $\Delta m = m - m_{\mathrm{PDG}}$. Values in MeV/$c^2$. In case of $K^0_S\omega$, the π^0 is from $\omega \to \pi^+\pi^-\pi^0$. These values are being used in the final selection on MC.

$D \to$	$\Delta m(D^0)$	$\Delta m(K_S)$	$\Delta m(\pi^0)$	$\Delta m(\phi)$	$\Delta m(\omega)$
KK	$[-15.2, 15.0]$	-	-	-	-
$\pi\pi$	$[-18.9, 18.6]$	-	-	-	-
$K^0_S\pi^0$	$[-44.2, 37.0]$	$[-5.7, 5.7]$	$[-16.4, 12.1]$	-	-
$K^0_S\omega$	$[-22.8, 19.7]$	$[-5.5, 5.6]$	$[-18.7, 12.6]$	-	$[-22.4, 19.6]$
$K^0_S\phi$	$[-7.4, 7.4]$	$[-5.4, 5.5]$	-	$[-6.6, 8.2]$	-
$K\pi$	$[-16.6, 16.5]$	-	-	-	-
$K\pi\pi^0$	$[-26.7, 23.3]$	-	$[-15.2, 11.2]$	-	-
$K\pi\pi\pi$	$[-12.0, 11.9]$	-	-	-	-

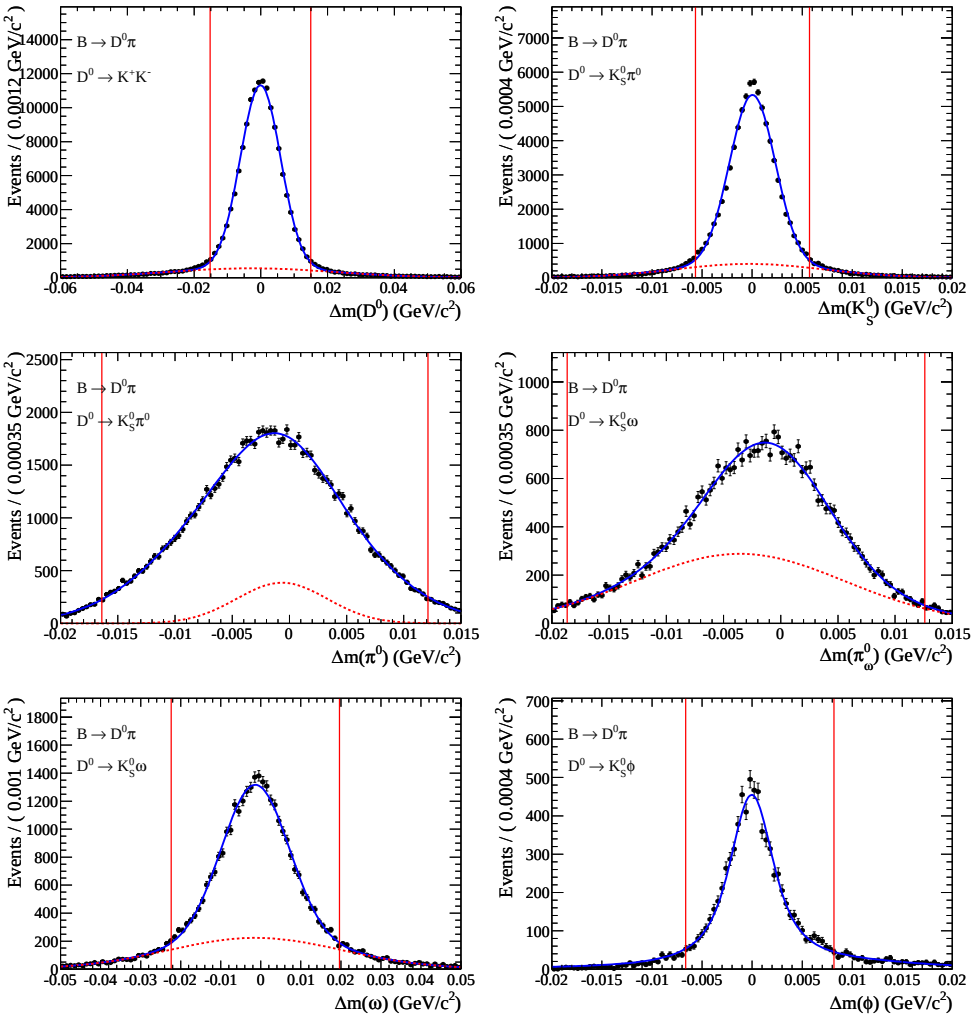

Figure 6.1: $m - m_{\mathrm{PDG}}$ fits on truth-matched signal MC to extract cuts on the invariant masses. The red lines indicate the obtained cut values.

Figure 6.2 shows the fractions of signal and background events passing each PID selector for $B \to D^0 K$ events reconstructed in the representative $D^0 \to K^- K^+$ channel.

Table 6.7: Optimal PID selectors. In case of $K^0_S\omega$ the π's are from $\omega \to \pi^+\pi^-\pi^0$, in case of $K^0_S\phi$ the K's are from $\phi \to K^+K^-$. The subscripts enumerate the kaons and pions in the D^0 final state. See Table 6.4 for the selector definitions.

$D \to$	$K_1^\pm$	$K_2^\pm$	$\pi_1^\pm$	$\pi_2^\pm$	$\pi_3^\pm$
K^+K^-	`VeryLoose`	`VeryLoose`	-	-	-
$\pi^+\pi^-$	-	-	`Loose`	`Loose`	-
$K^0_S\pi^0$	-	-	-	-	-
$K^0_S\omega$	-	-	`Loose`	`Loose`	-
$K^0_S\phi$	`NotAPion`	`NotAPion`	-	-	-
$K^-\pi^+$	`NotAPion`	-	-	-	-
$K\pi\pi^0$	`VeryLoose`	-	`Loose`	-	-
$K\pi\pi\pi$	`VeryLoose`	-	`Loose`	`VeryLoose`	`Loose`

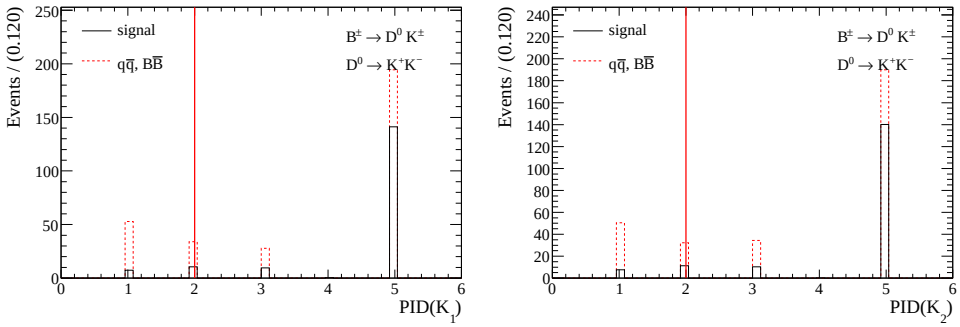

Figure 6.2: Fractions of signal and background events passing each PID selector. The optimal selector is indicated by the red vertical line. The mapping is 1–`NotAPion`, 2–`VeryLoose`, 3–`Loose`, 4–`Tight`, 5–`VeryTight`. See Table 6.4 for the selector definitions.

6.3.3 TMVA Requirements

We will consider requirements on five additional kinematic variables to enhance the purity of the final samples: the D^0 helicity angle, the K^0_S significant flight length, the ω helicity angles, the ϕ helicity angle, and the energy of the π^0. In the following, we describe each of these variables in more detail, before we focus on how to find optimal cuts using TMVA.

The decay angle $\theta_{\rm hel}(D^0)$ of the D^0-meson, defined as the angle between the direction of one of the D^0 daughters in the D^0 rest frame and the direction of the D^0-meson in the B rest frame, is expected to follow a flat $\cos\theta_{\rm hel}(D^0)$ distribution, because of angular momentum conservation: each of the B, D, and h-mesons is scalar. In background $e^+e^- \to q\overline{q}$ events, where a fake D^0 is picked from a random combination of tracks and neutral objects in the two light quark jets, $\cos\theta_{\rm hel}(D^0)$ is peaked at ± 1. This is shown in Figure 6.3 for the $D^0 \to \pi^+\pi^-$ mode.

We exploit the relatively long lifetime of the K^0_S-meson by requiring its decay vertex to be significantly displaced from the D^0 decay vertex. The flight significance S_K, defined as the ratio between the separation of the decay vertices of the K^0_S-meson and the D^0-meson in the

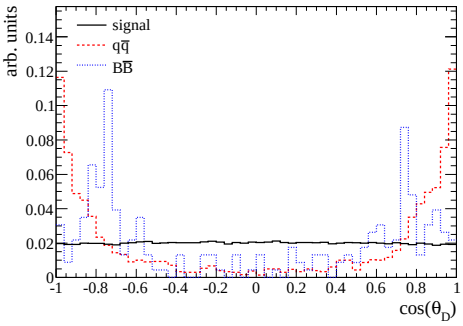

Figure 6.3: Distribution of the decay angle of the candidate D^0 reconstructed in the $B \to D^0 K$, $D^0 \to \pi^+\pi^-$ analysis in signal and background ($q\overline{q}$ and $B\overline{B}$) simulated events.

plane transverse to the beam axis, and its error, is a suitable kinematic variable to cut on. In addition, the angle α_K between the flight direction of the K^0_S, defined by a connecting line between the D^0 and K^0_S vertices, and the total momentum of the two pions in the transverse plane is expected to be zero for correctly reconstructed K^0_S-mesons. We will require this angle to be less than 90° by defining the signed flight significance, $S'_K = \mathrm{sign}(\cos\alpha_K) \cdot S_K$, and retaining only positive values of S'_K. These variables are powerful enough so we don't need to apply PID requirements to the K^0_S daughters.

In the $\omega \to \pi^-\pi^+\pi^0$ decay the three daughter pions are produced in a plane in the ω rest frame. The normal helicity angle θ_N is the angle between the normal to this plane and the flight direction of the ω-meson in the D^0 rest frame. It is distributed as $\cos^2\theta_N$: because of angular momentum conservation the decay of the scalar D^0-meson to a scalar K^0_S-meson and the ω-vector-meson ($J^{PC} = 1^{--}$), the ω-meson has helicity 0, $|S, S_z\rangle_\omega = |1, 0\rangle_\omega$, where the quantization axis is oriented along the K^0_S and ω flight line in the D^0 rest frame. Then the ω-meson decays into three pions. Let's call L the angular momentum of the $\pi^+\pi^-$ system and l the relative angular momentum between the π^0 and the $\pi^+\pi^-$ system. The $\pi^+\pi^-$ system must have $L = 1$ for the following reasons: (i) Angular momentum conservation yields $S(\omega) = 1 = L(\pi^+\pi^-) + l(\pi^0, \pi^+\pi^-)$ and thus $|L - l| \leq 1 \leq L + l$. (ii) Since parity is conserved in strong decays we have $P(\omega) = -1 = (-1)^L \cdot (-1)^l \cdot (-1)^3$, because for charged pions it is $J^P = 0^-$ and neutral ones have $J^{PC} = 0^{-+}$. Thus we infer $L + l$ is even. (iii) The $\pi^+\pi^-$ system is bosonic, so the wave function must be symmetric under exchange of the two pions: $+1 = (-1)^L \cdot C(\pi^+\pi^-)$. (iv) At last we consider charge conjugation: $C(\omega) = -1 = C(\pi^+\pi^-) \cdot C(\pi^0)$. We conclude $C(\pi^+\pi^-) = -1$. Together with (iii) we learn L is odd, and with (ii) l must be odd, too. The natural choice is therefore $L = l = 1$, since there are no known $\pi^+\pi^-$ resonances with spin-3 and invariant mass below $m(\omega)$. Coming back to the $\omega \to \pi^-\pi^+\pi^0$ decay, since $S_z = 0$, we have $|1, 0\rangle_\omega = |1, L_z\rangle_{\pi^+\pi^-} \oplus |1, l_z\rangle_{\pi^0}$. Using the Clebsch-Gordan coefficients gives $|l, l_z\rangle = |1, \pm 1\rangle$ and $|L, L_z\rangle = |1, \mp 1\rangle$. This implies that the angular distribution of the angle θ between the momentum of the $\pi^+\pi^-$ system and the direction of the ω-meson in the ω rest frame is proportional to $\sin^2\theta$, thus it follows for the normal helicity angle $\cos^2(\pi/2 - \theta) = \cos^2\theta_N$. For unpolarized ω-mesons or for fake ω candidates reconstructed from random combinations of two tracks and a π^0-meson, the $\cos\theta_N$ distribution is roughly flat (see Figure 6.4). Another angular variable which characterizes the three-pion system is the Dalitz angle $\theta_{\pi\pi}$ between the flight direction of one of the three pions in the ω rest frame and the flight direction of one of the two other pions in their center-of-mass frame. In true $\omega \to \pi^-\pi^+\pi^0$ decays, $\cos\theta_{\pi\pi}$ is distributed like $\sin^2\theta_{\pi\pi}$, as we have seen before. Fake ω candidates reconstructed from random combinations of two tracks and a π^0-meson

exhibit an oppositely convexed $\cos\theta_{\pi\pi}$ distribution (see Figure 6.4). We will consider a cut on the quantity $\cos^2\theta_N \sin^2\theta_{\pi\pi}$.

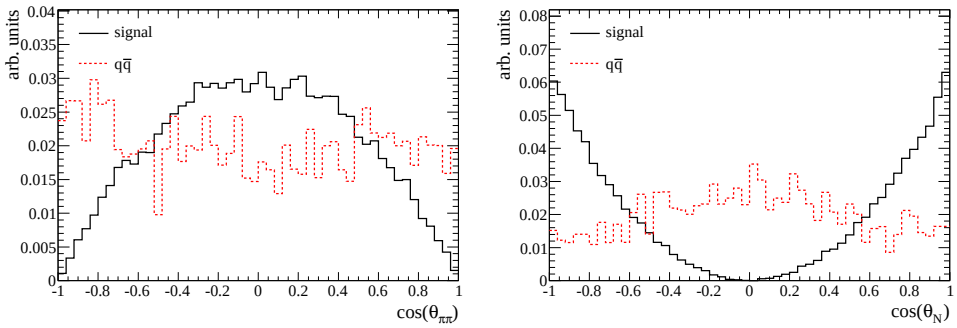

Figure 6.4: Distribution of the Dalitz angle $\cos\theta_{\pi\pi}$ and of the normal helicity angle $\cos\theta_N$ for simulated signal and simulated $q\overline{q}$ background events reconstructed in the $B^- \to D^0 K^-$, $D^0 \to K^0_S\omega$ channel. We select (fake) background ω candidates in the $\pi^+\pi^-\pi^0$ mass sidebands. The distributions are normalized to unit area.

In the case of $D^0 \to K^0_S\phi$, similar arguments hold as in the $D^0 \to K^0_S\omega$ decay. The ϕ-meson is a vector meson, so angular momentum conservation requires that the ϕ-meson be produced with helicity 0. The subsequent decay of the ϕ-meson into two scalar kaons yields a distribution of the cosine of the ϕ decay angle $\cos\theta_H$ (the angle of the kaon in the ϕ rest frame with respect to the direction of the ϕ-meson in the D^0 rest frame) which shows a characteristic $\cos^2\theta_H$ behavior. In background events the $\cos\theta_H$ distribution is flat for fake ϕ candidates reconstructed from random combinations of charged tracks, and is a sum of a constant and a $\cos^2\theta_H$ distribution—due to partial polarization—for true ϕ candidates. The $\cos\theta_H$ distribution for signal and background events is shown in Figure 6.5. We will consider a cut on $|\cos\theta_H|$.

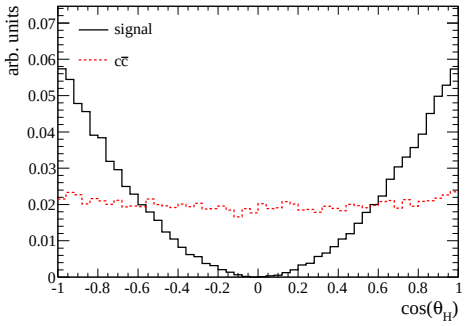

Figure 6.5: Helicity angle distribution of $\phi \to K^+K^-$ candidates reconstructed in the $B^- \to D^0 K^-$, $D^0 \to K^0_S\phi$ decay chain. The distributions are normalized to unit area.

Lastly, we will impose requirements on the total energy of the π^0 in case of the $D^0 \to K^0_S\pi^0$, $D^0 \to K^-\pi^+\pi^0$, and $D^0 \to K^0_S\omega$ ($\omega \to \pi\pi\pi^0$) decays.

We use the TMVA package to find optimal cuts on the kinematic variables defined above. TMVA works with two event samples, signal and background. We use truth-matched signal MC and a cocktail of generic background MC, as we did before to optimize the PID

requirements. TMVA maximizes the background rejection r for a given signal efficiency ε_S by scanning the parameter space of the cuts in question using a genetic algorithm. This results in the so called *Receiver Operating Characteristics* (ROC) curve, $r \equiv r(\varepsilon_S)$. The cut configuration which maximizes the significance, Eq. 6.5, corresponds to a particular working point on the ROC curve. This point can be found by rewriting Eq. 6.5 in terms of $\varepsilon_B = 1 - r$ and ε_S:

$$S = \frac{\varepsilon_S N'_S}{\sqrt{\varepsilon_S N'_S + (1 - r(\varepsilon_S)) N'_B}}, \tag{6.6}$$

where $N'_{S(B)}$ is the signal (background) yield before applying the cuts. One can now find the maximum by evaluating

$$\frac{\mathrm{d}S}{\mathrm{d}\varepsilon_S} = 0. \tag{6.7}$$

TMVA provides data points of the ROC curve rather than an analytic form. To be able to solve Eq. 6.7 we fit these points to a polynom of fifth grade. Once the optimal working point is known, the corresponding cut configuration is obtained from tables also provided by TMVA. Figure 6.6 shows the distributions of selected input variables for $B \to D^0 K$ events, together with the optimal cut values obtained from the $B \to D^0 \pi$ control sample. Figure 6.7 shows a representative ROC curve, the fitted fifth order polynom, and the optimal working point calculated from Eq. 6.7. In case of the ϕ helicity angle it turns out the MC samples are not large enough to allow for the TMVA tuning procedure. Therefore we find the optimal cut by maximizing S (Eq. 6.5) neglecting any correlations. We obtain $|\cos\theta_H > 0.4|$. Table 6.8 summarizes the optimal selection requirements.

Table 6.8: Optimized TMVA cuts. In case of $K^0_S\omega$, $E(\pi^0)$ is from inside the ω. In case of $K^-\pi^+\pi^0$, the optimization shows no cut needs to be applied.

$D^0 \to$	$\varepsilon_S^{\mathrm{opt}}$	$\lvert\cos_{\mathrm{hel}}(D^0)\rvert$	S'_K	$E(\pi^0)$	ω angles	$\lvert\cos\theta_H\rvert$
$\pi^+\pi^-$	0.88	> 0.74	-	-	-	-
$K^0_S\pi^0$	0.99	> 0.986	> 1.9	$> 241\,\mathrm{MeV}$	-	-
$K^0_S\omega$	0.94	-	> 2.2	$> 210\,\mathrm{MeV}$	> 0.046	-
$K^-\pi^+\pi^0$	1.00	-	-	don't apply	-	-
$K^0_S\phi$	0.64	-	-	-	-	> 0.4

6.3.4 Mass Cuts on Data

Optimizing the mass cuts on on-peak data is slightly different from the method we used on MC because now the sample contains background. This background has to be taken into account when fitting the invariant mass distributions. The signal is again parameterized with a double Gaussian shape (Voigtian shape for $m(\phi)$). For the background we use a straight line. Again all parameters are floating in the fit. The method to determine the cut values is the same as before, where in case of the π^0 and ϕ masses the background component is not considered when finding the 95 % intervals. We first validate the fit procedure on cocktail MC composed of the generic MC samples. We find good agreement between the cut values obtained from signal MC and those obtained from generic MC, allowing us to apply the procedure to on-peak data: Figure 6.8 shows the fits on on-peak data, and Table 6.9 summarizes the extracted mass intervals to be selected.

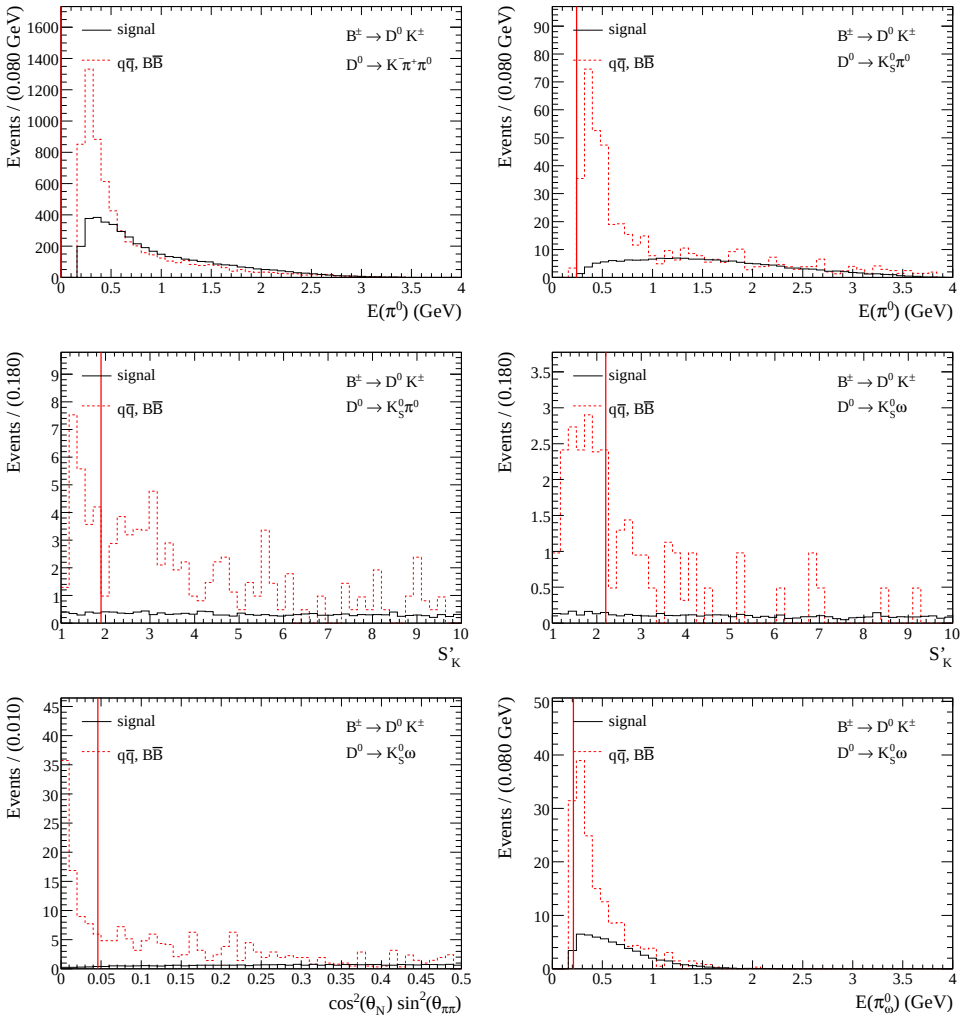

Figure 6.6: Distribution of variables on which selection requirements are optimized by TMVA: the energy of a π^0 either from a D^0 or ω decay, the signed K^0_S flight significance, and the combined ω helicity angles. Shown are $B \to D^0 K$ events. The distributions are normalized to data luminosity.

Table 6.9: Optimized cuts on $\Delta m = m - m_{\rm PDG}$. Values in MeV/$c^2$. In case of $K^0_S\omega$, the π^0 is from $\omega \to \pi^+\pi^-\pi^0$. These values are being used in the final selection on data.

$D \to$	$\Delta m(D^0)$	$\Delta m(K_S)$	$\Delta m(\pi^0)$	$\Delta m(\phi)$	$\Delta m(\omega)$
KK	$[-16.9, 15.4]$	-	-	-	-
$\pi\pi$	$[-17.9, 16.8]$	-	-	-	-
$K^0_S\pi^0$	$[-43.1, 35.9]$	$[-6.5, 5.9]$	$[-17.1, 12.8]$	-	-
$K^0_S\omega$	$[-22.6, 20.0]$	$[-6.1, 5.6]$	$[-20.0, 14.8]$	-	$[-19.0, 16.1]$
$K^0_S\phi$	$[-6.2, 6.3]$	$[-4.4, 4.2]$	-	$[-5.8, 7.5]$	-
$K\pi$	$[-17.2, 15.5]$	-	-	-	-
$K\pi\pi^0$	$[-28.5, 24.5]$	-	$[-19.6, 14.8]$	-	-
$K\pi\pi\pi$	$[-12.6, 10.9]$	-	-	-	-

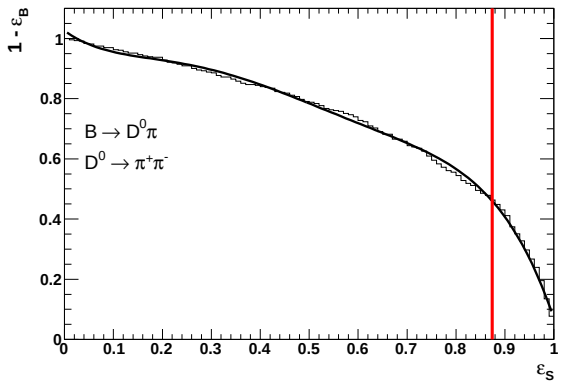

Figure 6.7: Fit of the ROC curve in $D^0 \to \pi^+\pi^-$. The vertical line depicts the optimal working point.

6.4 Arbitration of Multiple Candidates

When reconstructing B candidates, it is possible that more than one combination of tracks satisfies the selection criteria in the same event. The rate at which this occurs depends on the reconstructed decay mode and on the selection cuts. It is very small when the D^0-meson is selected in the $K^-\pi^+$, K^+K^-, $\pi^+\pi^-$ and $K^0_S\phi$ final states, but it increases when $D^0 \to K^0_S\pi^0$ or $D^0 \to K^0_S\omega$ is reconstructed, because the probability to select a fake candidate (especially due to misreconstructed ω or π^0 candidates) is higher. The fraction of events with multiple candidates ranges from 0.2 % to 8 %, depending on the channel. In order to select only one candidate per event it is necessary to define a criterion that permits to identify, as good as possible, the combination with the larger probability to be a true signal $B \to D^0h$ decay. For such a criterion the invariant masses are the first choice. We define a χ^2-function as

$$\chi^2 = \sum_i \frac{(m_i - \langle m_i \rangle)^2}{\sigma^2_{m_i} + \Gamma^2_i}, \tag{6.8}$$

where the sum ranges over all unstable particles (D^0, K^0_S, π^0, ϕ, and ω), σ is the resolution of the reconstructed invariant mass distribution, and Γ is the width of the particle (only Γ_ω is significantly different from zero). In events with multiplicity greater than one, the candidate with the minimum value of χ^2 is selected. The right choice rate of this algorithm ranges from 71 % to 94 %. In the rare case two candidates have the same minimal χ^2 value, we pick one randomly.

6.5 Summary of the Final Selection Criteria

In Tables 6.10-6.13 we summarize the selection criteria for the various $B \to D^0h$ candidates. The selection efficiencies are given in Table 6.14, together with the $B \to D^0K$ signal significance calculated from Eq. 6.5. In conclusion we must confess the TMVA approach turned out to be less practical than we expected it to be. Especially the decision to optimize on the $B \to D^0\pi$ control sample rather than the $B \to D^0K$ signal sample raised suspicions in the *BABAR* collaboration. This decision was in part motivated by the fact TMVA performs

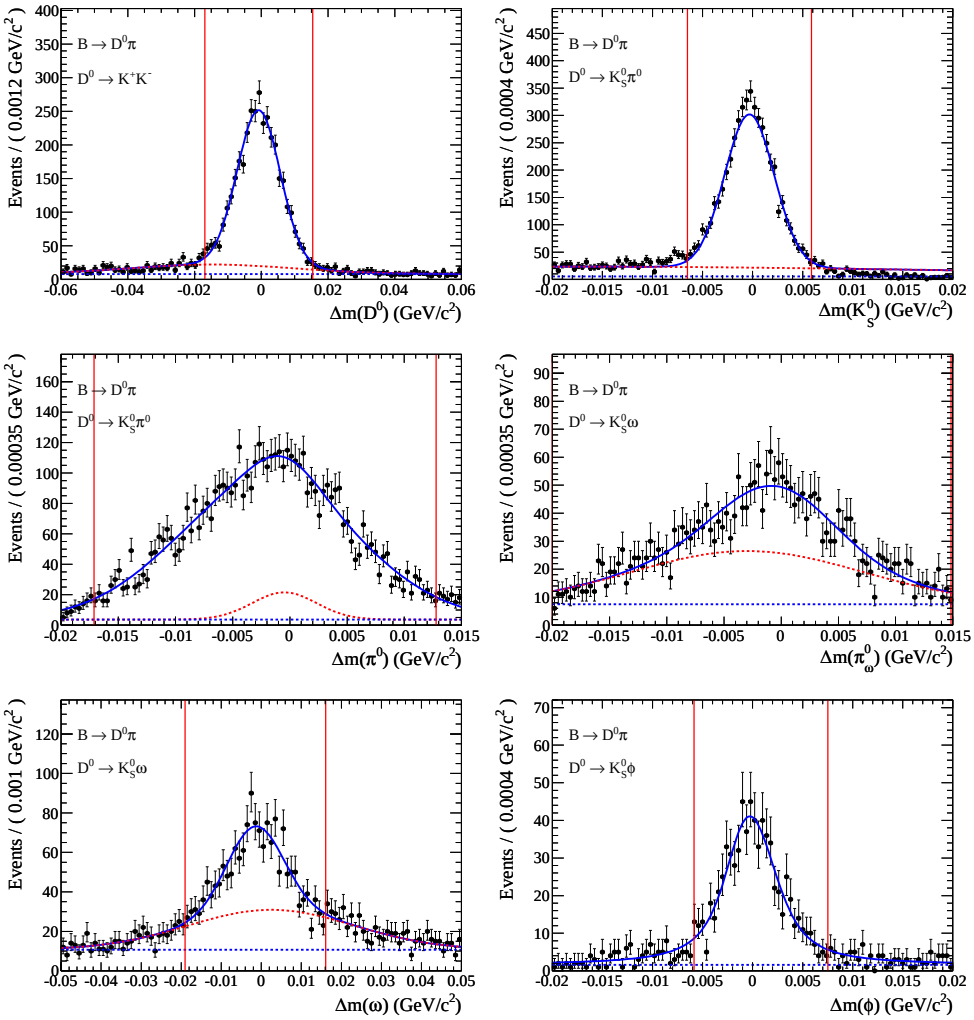

Figure 6.8: $m - m_{\rm PDG}$ fits on on-peak data to extract cuts on these masses. The red lines indicate the optimal mass intervals to be selected.

poorly on low statistics samples. We decided to perform the iterative procedure afterwards as a crosscheck. Although the iterative procedure indeed gives slightly better values of the $B \to D^0 K$ signal significance, the improvement is not enough to justify reprocessing the full analysis chain.

Table 6.10: $B \to D^0 h$ selection criteria that are common to all the eight D^0 decay modes under investigation.

Parameter	Criterion
ΔE (MeV)	$[-80, 120]$
$m_{\rm ES}$ (GeV/c^2)	$[5.2095, 5.29]$
$\mathcal{F}$	$[-1.5, 1.5]$
$\lvert m(D^0) - m(D^0)_{\rm PDG}\rvert$	$< 2.5\,\sigma$
D^0 vertex probability	$> 0.01\%$

Table 6.11: Selection criteria used to select the candidate samples on which the unbinned maximum likelihood fit is performed.

$D^0 \to$	$K^-\pi^+$	$K^-\pi^+\pi^+\pi^-$	$K^-\pi^+\pi^0$
PID K from D^0	`NotAPion`	`VeryLoose`	`VeryLoose`
PID π('s) from D^0	-	`Loose`	`Loose`
$\lvert m(\pi^0) - m(\pi^0)_{\rm PDG}\rvert$ (MeV)	-	-	$[-19.3, 12.1]$

Table 6.12: Selection criteria used to select the candidate samples on which the unbinned maximum likelihood fit is performed.

$D^0 \to$	K^+K^-	$\pi^+\pi^-$
D^0 $\lvert\cos\theta_{\rm hel}\rvert$	-	< 0.74
PID 1st D^0 daughter	`VeryLoose`	`Loose`
PID 2nd D^0 daughter	`VeryLoose`	`Loose`
other requirements	-	$m(h\pi) > 1.9$ GeV/c^2 (m_h=m_K)

Table 6.13: Selection criteria used to select the candidate samples on which the unbinned maximum likelihood fit is performed. Masses in MeV/c^2.

$D^0 \to$	$K^0_S\pi^0$	$K^0_S\phi$	$K^0_S\omega$
D^0 $\lvert\cos\theta_{\rm hel}\rvert$	< 0.98	-	-
$\lvert m(K^0_S) - m(K^0_S)_{\rm PDG}\rvert$	$[-5.6, 5.2]$	$[-5.6, 5.2]$	$[-5.6, 5.2]$
K^0_S flight-length	$> 3\sigma$	$> 3\sigma$	$> 3\sigma$
$E(\pi^0)$ (MeV)	> 220	-	> 207
$\lvert m(\pi^0) - m(\pi^0)_{\rm PDG}\rvert$	$[-18.4, 14.0]$	-	$[-17.3, 10.0]$
$\lvert m(\phi) - m(\phi)_{\rm PDG}\rvert$	-	$[-6.5, 6.4]$	-
$\lvert\cos\theta_H(\phi)\rvert$	-	> 0.4	-
PID ϕ daughters	-	NotAPion	-
$\lvert m(\omega) - m(\omega)_{\rm PDG}\rvert$	-	-	$[-17.9, 15.3]$
PID ω daughters	-	-	Loose
$\cos^2\theta_N \sin^2\theta_{\pi\pi}$	-	-	> 0.046

Table 6.14: $B \to D^0 h$ reconstruction efficiency after final selection, evaluated using signal MC. Errors are based on the Poissonian errors due to the sample size. We also give the values of the significance S on $B \to D^0 K$ events, both for the TMVA and the iterative approach.

D^0 decay	$\varepsilon(D^0K)(\%)$	$\varepsilon(D^0\pi)(\%)$	$\varepsilon(D^0K)/\varepsilon(D^0\pi)$	S (TMVA)	S (it.)
$D^0 \to K^-\pi^+$	52.53 ± 0.08	52.84 ± 0.04	0.994 ± 0.002	7.9	8.0
$D^0 \to K^-\pi^+\pi^0$	25.12 ± 0.07	25.45 ± 0.03	0.987 ± 0.003	2.9	3.1
$D^0 \to K^-\pi^+\pi^+\pi^-$	30.36 ± 0.07	30.70 ± 0.07	0.989 ± 0.003	7.1	8.6
$D^0 \to K^+K^-$	44.75 ± 0.08	45.19 ± 0.08	0.990 ± 0.002	3.8	4.5
$D^0 \to \pi^+\pi^-$	38.58 ± 0.08	39.08 ± 0.08	0.987 ± 0.003	4.2	4.7
$D^0 \to K^0_S\pi^0$	24.29 ± 0.07	24.70 ± 0.07	0.983 ± 0.004	35.8	36.4
$D^0 \to K^0_S\omega$	9.28 ± 0.05	9.32 ± 0.05	0.996 ± 0.008	42.4	44.7
$D^0 \to K^0_S\phi$	17.69 ± 0.06	17.97 ± 0.06	0.984 ± 0.005	45.3	45.4

7 Fisher Discriminant

A large fraction of background candidates arise from $e^+e^- \to q\overline{q}$ ($q = u, d, s, c$) events, in which random combinations of tracks and photons in the event are picked up in the reconstruction. In order to separate these $q\overline{q}$ events from signal events we make use of several event shape variables. Event shape variables exploit the different topologies of $q\overline{q}$ and $B\overline{B}$ events in the e^+e^- center-of-mass frame. The two B mesons produced in $\Upsilon(4S) \to B\overline{B}$ decays are in fact almost at rest, so there is no direction preferred by their decay products, thus the $B\overline{B}$ events are spherical. On the other hand, the light quarks from $e^+e^- \to q\overline{q}$ are produced with significant momentum, so their decay products are contained in two more or less collimated back-to-back jets. We combine several event shape variables into a *Fisher discriminant* [38], which gives a single value for each event, $\mathcal{F}$, that we include in the final fit.

A Fisher discriminant is a linear combination of the considered input parameters, where the coefficients are chosen in such a way, that the separation between the $\mathcal{F}$ distributions of $q\overline{q}$ and $B\overline{B}$ events is maximal. To obtain the coefficients, one first linearly decorrelates the input parameters. Then one determines an axis in the uncorrelated parameter space. The axis is chosen such that the signal and background distributions, when projected upon this axis, are separated as far from each other as possible. The metric used to define the distance in the uncorrelated space is the covariance matrix of the input parameters. The distribution of $\mathcal{F}$ is shifted towards -1 for $q\overline{q}$ events, and towards $+1$ for $B\overline{B}$ events, the latter include the signal.

Before we list the exact shape variables in use, we define two concepts: the *rest-of-event* (ROE) is the set of all detected tracks and photons in the event that have not been used to reconstruct the $B \to D^0 h$ candidate. The *thrust* T, and its axis $\hat{T}$ are defined via

$$T = \max \left(\frac{\sum_i |\hat{T} \cdot \mathbf{p}_i^*|}{\sum_i |\mathbf{p}_i^*|} \right), \tag{7.1}$$

where $\hat{T}$ is the direction in the center-of-mass frame that maximizes T, and $\mathbf{p}_i^*$ is the CM momentum of the i-th (charged or neutral) particle. The thrust is computed by summing over all tracks and photons belonging to the B decay tree. We consider the following four event-shape variables, example distributions are given in Figure 7.2.

First, we use the ratio of the second and zeroth Legendre monomial, L_2/L_0. Legendre monomials are a set of momentum-weighted sums of the tracks and neutral calorimetric energy deposits in the rest-of-event:

$$L_j = \sum_i^{\text{ROE}} p_i^* \times |\cos(\theta_i^*)|^j, \tag{7.2}$$

where θ_i^* is the CM angle between $\mathbf{p}_i^*$ and the thrust axis. We have considered only the (L_0, L_2) pair, since many analyses in *BABAR* have shown that adding other L_j ($j \neq 0, 2$) to the set of discriminating variables does not improve the signal/background separating power.

Second, we use the ratio of the Fox-Wolfram moments $R_2 = H_2/H_0$, computed using tracks and photons in the rest-of-event. The l^{th} Fox-Wolfram moment is defined as [39]:

$$H_l \equiv \sum_{i,j} \frac{|\mathbf{p}_i^*||\mathbf{p}^*{}_j|}{E_{\rm vis}^{*2}} P_l(\cos\theta_{ij}^*), \tag{7.3}$$

where the P_l are the Legendre polynomials, θ_{ij} is the opening angle between $\mathbf{p}_i^*$ and $\mathbf{p}_j^*$, and $E_{\rm vis}^*$ is the total visible energy of the event. For jet-like continuum events $\cos\theta_{ij}^*$ is peaked at ± 1, while for spherical $B\overline{B}$ events it is more uniformly distributed. Since the second Legendre polynomial is $P_2(\cos\theta_{ij}^*) = \frac{1}{2}(3\cos\theta_{ij}^{*2} - 1)$, the ratio R_2 is shifted towards one in $q\overline{q}$ events and towards zero in $B\overline{B}$ events. Previous studies in *BABAR* [40] have shown it is advantageous to calculate R_2 in the rest-of-event rather than the full event.

Third, we use the cosine of the angle of the B candidate momentum with respect to the beam axis, $|\cos(\theta_B)|$. In $\Upsilon(4S) \to B\overline{B}$ decays it follows a $\sin^2(\theta_B)$ distribution because the $\Upsilon(4S)$ are produced almost completely with helicity ± 1. It is almost flat in $q\overline{q}$ events.

Fourth, we use $|\cos(\theta_T)|$, the cosine of the angle of the thrust axis with respect to the beam axis. Signal events have an almost uniform distribution. It is not perfectly flat because of the small—but non-zero—B momentum in the CM frame. Background events follow a $1 + \cos^2(\theta_T)$ shape, which is reminiscent of the $1 + \cos^2(\theta_{q\overline{q}})$ distribution of the $q\overline{q}$ pair produced in the QED $e^+e^- \to q\overline{q}$ reaction. At values of $|\cos(\theta_T)| > 0.7$ both signal and background distributions suffer from a detection efficiency drop-off.

To construct the Fisher discriminant, to find its optimal coefficients, and to calculate the values of $\mathcal{F}$ for each event, we use again the TMVA software package. The coefficients have been optimized using MC samples of true signal events, and off-peak data. Events in these samples are only required to pass the pre-selection (see chapter 6.2); if multiple candidates survive in an event, we select one at random to exclude any possible bias on the coefficients due to the algorithm selecting the best candidate. Each D^0 final state has been optimized independently. An example ROC curve, plotting the signal efficiency ε_S versus $q\overline{q}$ background rejection $1 - \varepsilon_B$, is shown in Figure 7.1. Figure 7.3 shows the distributions of $\mathcal{F}$ in the representative $D^0 \to K^-K^+$ channel.

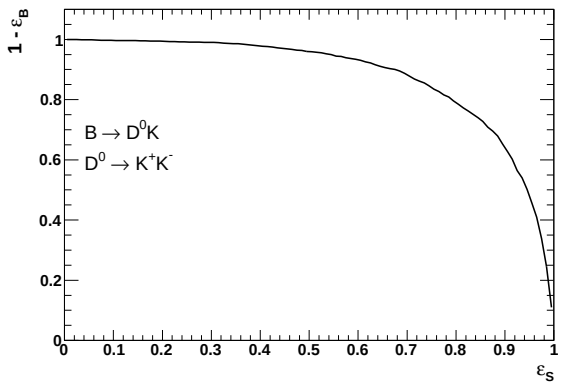

Figure 7.1: Signal efficiency ε_S versus continuum background rejection $1 - \varepsilon_B$ of the optimized Fisher discriminant, evaluated on simulated $B \to D^0K$, $D^0 \to K^-K^+$ events and off-peak data.

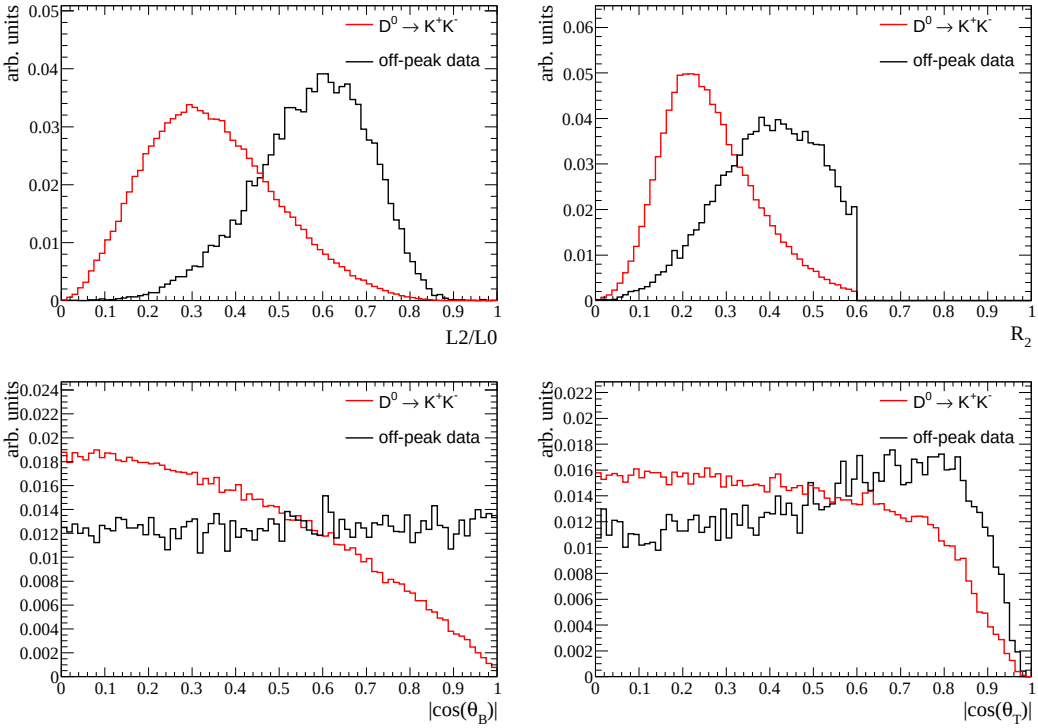

Figure 7.2: Distributions of the four event-shape variables considered in the Fisher discriminant, in simulated $B^- \to D^0 K^-$ and off-peak events, $D^0 \to K^- K^+$ decay mode. Only pre-selection criteria are required.

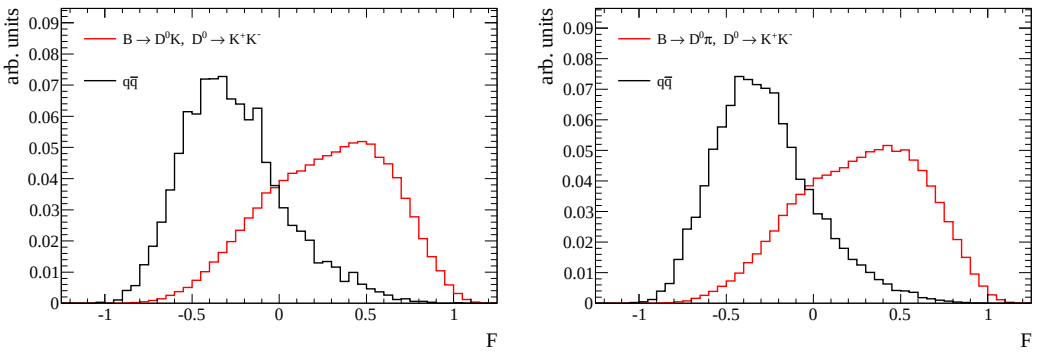

Figure 7.3: $\mathcal{F}$ distribution of truth-matched $B \to D^0 h$ candidates and $q\overline{q}$ events after the final selection, normalized to unit area. Left: $h = K$; right $h = \pi$.

8 Data-MC Comparison

The main component of this analysis is the final fit to data, extracting the $B \to D^0 h$ signal yields. As we will see in the next chapter, many parameters of the fit will be fixed to values obtained from the MC simulation. In this chapter we shall therefore investigate the reliability of the simulation. We will perform two cross-checks before we discuss measures to account for discrepancies.

As a first cross-check, we plot the ΔE, $m_{\rm ES}$, and $\mathcal{F}$ distributions in both data and MC samples after the final selection was applied. We present the plots of the $D^0 \to K^-K^+$ and $D^0 \to K^-\pi^+\pi^0$ channels in Figure 8.1: $D^0 \to K^-K^+$ is representative for all channels other than $D^0 \to K^-\pi^+\pi^0$, which is the one channel with the most abundant background contribution. In Figure 8.2 we show the same plots, but with the additional requirement that the bachelor track passes the `KLHVeryTight` selector, greatly enhancing the $B \to D^0 K$ signal component. The only two visible discrepancies are the overall normalization, and small shifts in the ΔE and $m_{\rm ES}$ signal distributions.

As a second cross-check, we fit the invariant mass distributions of composite particles in both data and Monte Carlo. We use an asymmetric Gaussian for the fit function, rather than the double Gaussian we used before to extract the mass cuts (Section 6.3.4). This way we are more sensitive to small asymmetries in the distributions. In order to minimize any possible experimenter bias, we perform this cross-check on the $B \to D^0\pi$ control sample. To this extent we require $|\Delta E(\pi)| < 40\,\text{MeV}$, $|m_{\rm ES} - m_B| < 10\,\text{MeV}/c^2$, and $\mathcal{F} > -0.8$, in addition to the final selection criteria but the one on the quantity that is being fit. $\Delta E(\pi)$ is computed using the pion mass hypothesis for the bachelor track. The particles are the B ($m_{\rm ES}$) and D^0-mesons; K^0_S, ω, and ϕ-mesons from D^0 decays; and π^0-mesons from both D^0 and ω decays. The corresponding plots for representative D^0 final states are shown in Figures 8.3-8.5. In addition to the invariant masses, we show a comparison of the $\Delta E(\pi)$ variable. The agreement is good, and only small shifts of fractions of MeV($/c^2$) are seen in some cases.

We will account for the observed discrepancies by floating the associated parameters in the final fit. This way we make sure the model we use to describe the data is accurate. We also emphasize we measure yield ratios, in which small efficiency differences between data and Monte Carlo simulation cancel.

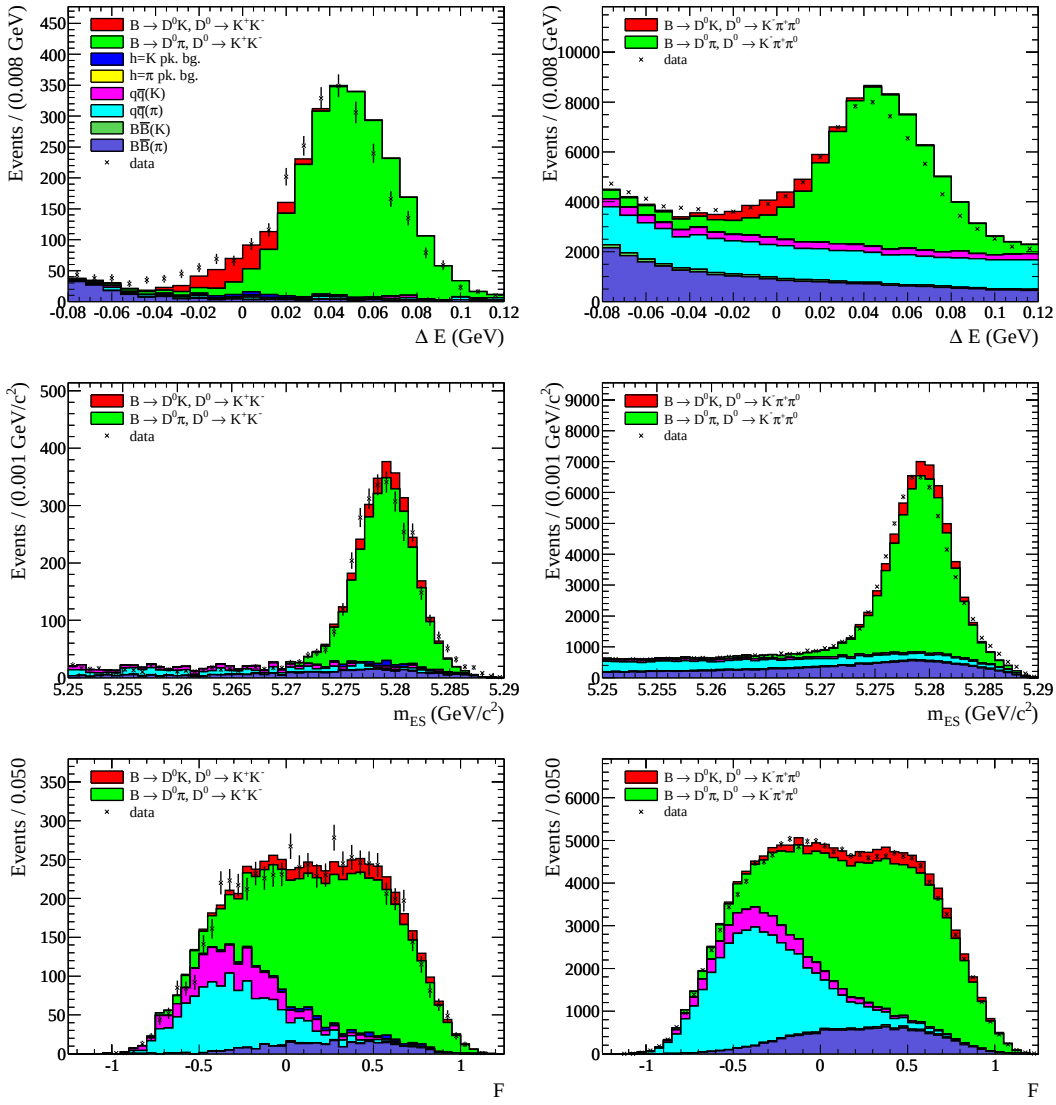

Figure 8.1: ΔE, $m_{\rm ES}$, and $\mathcal{F}$ distributions in both data and MC in the representative $D^0 \to K^-K^+$ channel and the one channel with most abundant background, $D^0 \to K^-\pi^+\pi^0$. Additional signal range cuts have been applied on the fit variables not plotted: $\mathcal{F} > 0.2$, $m_{\rm ES} > 5.275\,{\rm GeV}/c^2$.

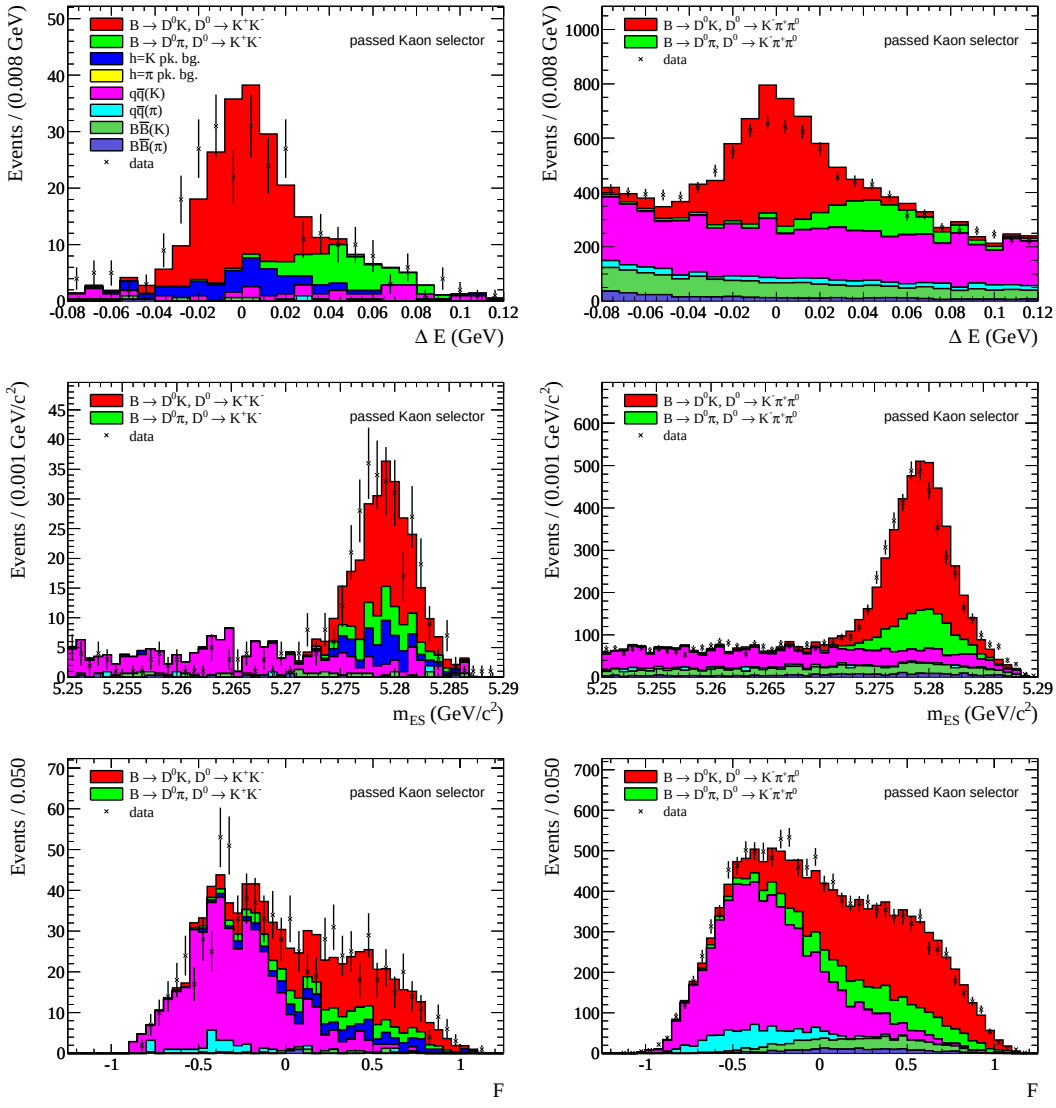

Figure 8.2: ΔE, $m_{\rm ES}$, and $\mathcal{F}$ distributions in both data and MC in the representative $D^0 \to K^-K^+$ and $D^0 \to K^-\pi^+\pi^0$ channels. The prompt track is required to pass the `KLHVeryTight` selector. Additional signal range cuts have been applied on the fit variables not plotted: $\mathcal{F} > 0.2$, $m_{\rm ES} > 5.275\,{\rm GeV}/c^2$.

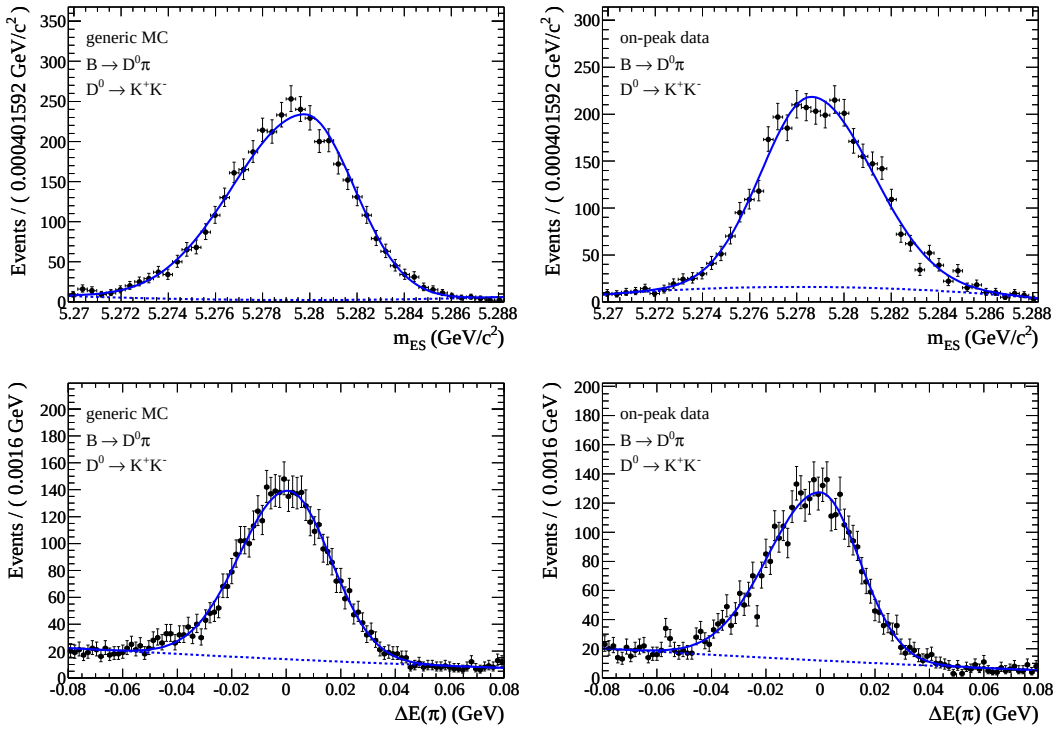

Figure 8.3: Distribution of ΔE and $m_{\rm ES}$ in the $B \to D^0\pi$ control sample, $D^0 \to K^-K^+$ channel. Left column: generic MC, scaled to data luminosity. Right column: on-peak data. The distributions are fitted with the sum of an asymmetric Gaussian and a $2^{\rm nd}$ order polynomial.

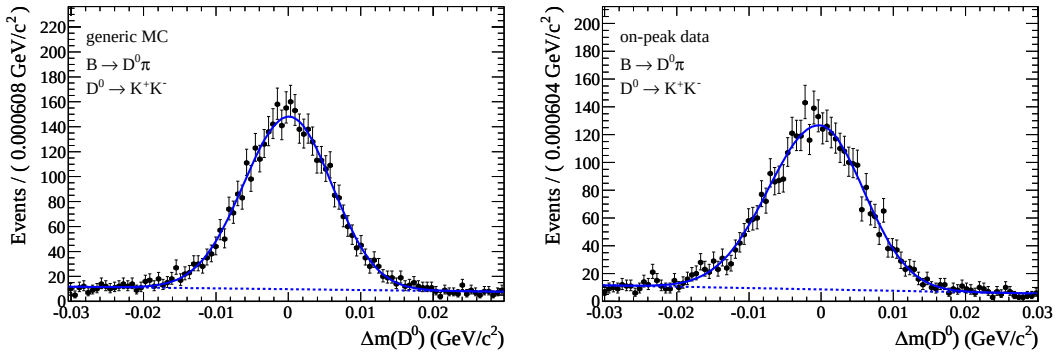

Figure 8.4: Distribution of the D^0 invariant mass ($\Delta m = m_{\rm rec} - m_{\rm PDG}$) in the $B \to D^0\pi$ control sample, $D^0 \to K^-K^+$ channel. Left column: generic MC, scaled to data luminosity. Right column: on-peak data. The distributions are fitted with the sum of an asymmetric Gaussian and a linear background.

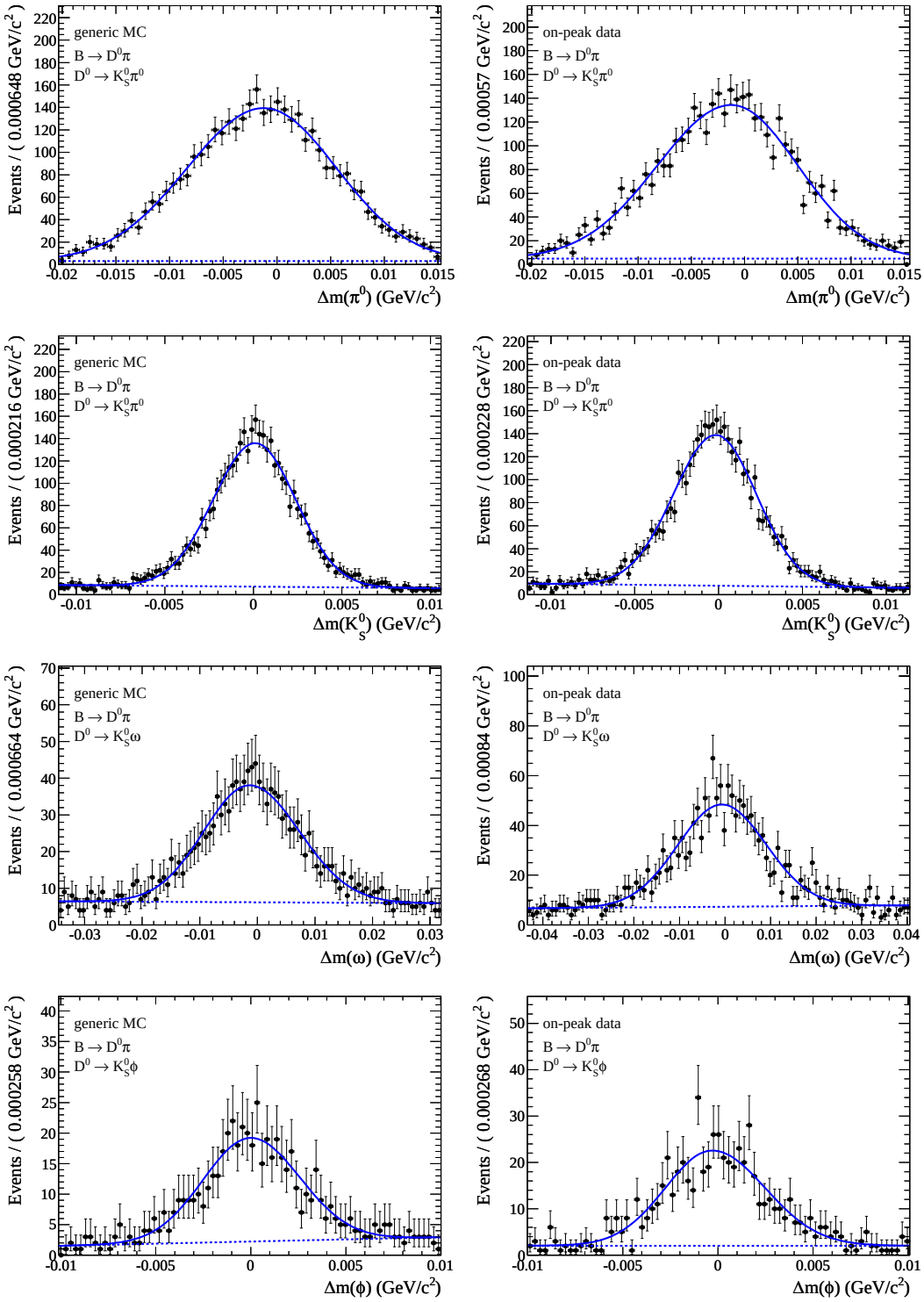

Figure 8.5: Distribution of invariant masses ($\Delta m = m_{\rm rec} - m_{\rm PDG}$) of composite D^0 daughters in the $B \to D^0\pi$ control sample. Left column: generic MC, scaled to data luminosity. Right column: on-peak data. From top to bottom: π^0 ($D^0 \to K^0_S\pi^0$), K^0_S ($D^0 \to K^0_S\pi^0$), ω, ϕ. The distributions are fitted with the sum of an asymmetric Gaussian and a linear or a 2nd order background.

9 Fit Procedure

The goal of this work is to measure the four GLW parameters $R_{CP\pm}$ and $A_{CP\pm}$. The charge-averaged ratios R_{CP} are being expressed through the ratios $R_{K/\pi} = \mathcal{B}(B \to D^0 K)/\mathcal{B}(B \to D^0\pi)$, as described in Section 3, so we are left with the determination of A_{CP+}, A_{CP-}, and the five values of $R_{K/\pi}$: one for the $B \to D^0_{CP+}h$ channels, one for the $B \to D^0_{CP-}h$ channels, and three for the flavor channels. We shall extract these parameters by means of a *three-dimensional simultaneous extended maximum likelihood fit*.

The basic concept of a maximum likelihood fit is the following [41]: suppose we have n measurements of a variable $\vec{x}$, the underlying probability density function (PDF), $f(\vec{x};\alpha)$, shall be known up to the exact values of the parameters α. The maximum likelihood fit will find an optimal estimate $\hat{\alpha}$ for the parameters α. One defines the *likelihood function*

$$\mathcal{L}(\alpha) = \prod_{i=1}^{n} f(\vec{x_i};\alpha), \tag{9.1}$$

which represents, for a given sample $\vec{x_i}$, the likelihood to obtain this exact sample for a certain choice of α. The optimal estimate $\hat{\alpha}$ corresponds to a maximum of $\mathcal{L}(\alpha)$, hence the name. For practical reasons it has become common to define the negative log-likelihood function,

$$l(\alpha) = -\ln \mathcal{L}(\alpha), \tag{9.2}$$

and to search the minimum, which, due to the monotone nature of the logarithm, corresponds to $\hat{\alpha}$, too. In case of one single parameter, the 68% confidence interval $[\hat{\alpha} - \sigma_l,\, \hat{\alpha} + \sigma_r]$ is defined by

$$l(\hat{\alpha} + \sigma_l) = l(\hat{\alpha}) + \frac{1}{2}\ , \qquad l(\hat{\alpha} + \sigma_r) = l(\hat{\alpha}) + \frac{1}{2}. \tag{9.3}$$

In case of many parameters, however, one has to evaluate Eqns. 9.3 for a certain parameter α_i while staying at the minimum with respect to all other parameters. This quickly becomes numerically demanding. An algorithm performing this calculation is implemented in MINUIT under the name MINOS [29].

One also defines the significance S of a measurement, as the square root of the difference of the minimum values of two log-likelihood functions, one including the signal PDF, the second excluding it (null hypothesis):

$$S = \sqrt{2|l(\hat{\alpha}) - l_{\text{null}}(\hat{\alpha}_{\text{null}})|}. \tag{9.4}$$

In case of decaying particles, where one measures a certain number of events, the number of expected measurements N is itself a parameter. To account for this, one can extend the definition of the likelihood function by a Poisson factor reflecting the probability of observing n events while N have been expected *(extended likelihood)*:

$$\mathcal{L}(\alpha) = \frac{e^{-N}(N)^n}{n!} \prod_{i=1}^{n} f(\vec{x_i};\alpha). \tag{9.5}$$

The final dataset consists of disjoint subsamples *(slices)*. Each slice is defined by a unique set of values of appropriate discrete category variables. A certain slice contains information on PDF parameters unique to this slice. However, there are also *shared* parameters, to which more slices contribute information to. An obvious example are the slices defined by the D^0 final state. Both the K^+K^- and the $\pi^+\pi^-$ slice contribute information to the A_{CP+} parameter, whereas the background yields can be different in both samples. In principle, one could fit each slice independently, and combine the shared parameters by means of a weighted average. By performing a simultaneous fit, one saves the averaging procedure and the associated treatment of correlated uncertainties. One also gets an easy access to the combined significance of the fit result through Equation 9.4.

We shall now look into the details of the final fit used in this analysis. The fit is three-dimensional, the three fit variables are $m_{\rm ES}$, ΔE, and $\mathcal{F}$. The final dataset is split into 32 slices by the D^0 final state ($\times 8$), the charge of the B-meson ($\times 2$), and whether or not the bachelor track passes the `KLHVeryTight` selector ($\times 2$). The 32 slices are fit simultaneously by five independent fits: $32 = 8(CP+) + 12(CP-) + 3 \cdot 4(\text{flavor})$. The generic likelihood function for one of these simultaneous fits is given by

$$\mathcal{L}(\alpha) = \frac{e^{-N}(N)^n}{n!} \prod_s \prod_{i=1}^{N_s} \mathcal{P}(m_{{\rm ES}i}, \Delta E_i, \mathcal{F}_i; \alpha)_s, \tag{9.6}$$

where s ranges over the slices under consideration, N_s is the number of events in slice s, n is the total number of events in the dataset $n = \sum_s N_s$, and N is the expected number of events. The probability $\mathcal{P}_{i,s} \equiv \mathcal{P}(m_{{\rm ES}i}, \Delta E_i, \mathcal{F}_i)_s$ for an event i is the sum of six signal and background components: $B \to D^0 K$ signal, $B \to D^0 \pi$ signal, $e^+e^- \to q\overline{q}$ continuum background, combinatoric background of B decays, and peaking background arising from charmless $B \to fK$ and $B \to f\pi$ decays, which will be discussed in more detail in Chapter 11:

$$\begin{aligned} N\mathcal{P}_{i,s} \;=\; & N_s^{{\rm sig}(\pi)} \mathcal{P}_{s,i}^{{\rm sig}(\pi)} + N_s^{{\rm sig}(K)} \mathcal{P}_{s,i}^{{\rm sig}(K)} + \\ & N_s^{q\overline{q}} \mathcal{P}_{s,i}^{q\overline{q}} + N_s^{B\overline{B}} \mathcal{P}_{s,i}^{B\overline{B}} + \\ & N_s^{{\rm pk}(\pi)} \mathcal{P}_{s,i}^{{\rm pk}(\pi)} + N_s^{{\rm pk}(K)} \mathcal{P}_{s,i}^{{\rm pk}(K)}, \end{aligned} \tag{9.7}$$

with the normalization condition

$$N = \sum_{s,j} N_s^j, \tag{9.8}$$

where j ranges over the six components and the N_s^j are the expected yields in each component. In case of negligible correlations among the fit variables, each $\mathcal{P}$ (we omit component and slice indices for the moment) factors into

$$\mathcal{P}(m_{\rm ES}, \Delta E, \mathcal{F}) = \mathcal{P}(m_{\rm ES})\, \mathcal{P}(\Delta E)\, \mathcal{P}(\mathcal{F}). \tag{9.9}$$

In most cases this is a good description of the data. Only in the $(m_{\rm ES}, \Delta E)$-plane of the $B\overline{B}$ background there are significant correlations. We account for them by using a non-factoring PDF, as we will describe in more detail in Section 9.2:

$$\begin{aligned} \mathcal{P}(m_{\rm ES}, \Delta E, \mathcal{F}) &= \mathcal{P}(m_{\rm ES}, \Delta E) \cdot \mathcal{P}(\mathcal{F}) \\ \mathcal{P}(m_{\rm ES}, \Delta E) &= f\, \mathcal{P}_1(m_{\rm ES}) \mathcal{P}_1(\Delta E) + (1-f)\, \mathcal{P}_2(m_{\rm ES}) \mathcal{P}_2(\Delta E). \end{aligned}$$

To perform the simultaneous fit of the charge slices, we rewrite the yields N_s^j in equation 9.7 by means of the CP asymmetries A_{CP}^j, and the total number of events N^j:

$$N_\pm^j = \frac{1}{2}(1 \mp A_{CP}^j) N^j \tag{9.10}$$

To be able to fit simultaneously the slices defined by the `KLHVeryTight` selector, we introduce the efficiency ε and the mis-id rate m of the `KLHVeryTight` selector for $B \to D^0 h$ signal events and $B \to fh$ peaking background events. Thus the signal event yields for events where the bachelor track passes (p) or fails (f) the selector become

$$\begin{aligned} N_p^{\mathrm{sig}(K)} &= \varepsilon\, N^{\mathrm{sig}(K)}\,, \qquad N_f^{\mathrm{sig}(K)} = (1-\varepsilon)\, N^{\mathrm{sig}(K)} \\ N_p^{\mathrm{sig}(\pi)} &= m\, N^{\mathrm{sig}(\pi)}\,, \qquad N_f^{\mathrm{sig}(\pi)} = (1-m)\, N^{\mathrm{sig}(\pi)}. \end{aligned} \tag{9.11}$$

At last, we introduce the ratios $R_{K/\pi}$ through

$$N^{\mathrm{sig}(K)} = N^{\mathrm{sig}(\pi)} \cdot R_{K/\pi}, \tag{9.12}$$

assuming the ratio of selection efficiencies of $B \to D^0 K$ and $B \to D^0 \pi$ events is compatible with unity (compare Table 6.14). Combining equations 9.10, 9.11, and 9.12 we get the expressions for all the N_s^j of Equation 9.7. The yield of the $B^- \to D^0 K^-$, $D^0 \to K^- K^+$ signal events which fail the `KLHVeryTight` selector is, for example,

$$N_{KK,-,f}^{\mathrm{sig}(K)} = \frac{1}{2}\left(1 + A_{CP}^{\mathrm{sig}(K)}\right)\, N_{\mathrm{tot}}^{\mathrm{sig}(\pi)}\, R_{K/\pi}\,(1-\varepsilon), \tag{9.13}$$

where $N_{\mathrm{tot}}^{\mathrm{sig}(\pi)}$ refers to the total number of $B \to D^0\pi$, $D^0 \to K^- K^+$ signal events, summed over all charge and `KLHVeryTight` slices.

As mentioned before, we perform five independent fits: a fit on the *CP*+ final states of the D^0 (K^+K^-, $\pi^+\pi^-$), a fit on the *CP*– states ($K^0_S\pi^0$, $K^0_S\omega$, $K^0_S\phi$), and three fits on the flavor final states ($K^-\pi^+$, $K^-\pi^+\pi^0$, $K^-\pi^+\pi^+\pi^-$). The latter can't be fit simultaneously because due to the high statistics in these modes we hit performance limits. Instead we will average the $R_{K/\pi}^{\mathrm{flv}}$ parameters, that in principle would be shared, by means of a weighted average. Of the flavor modes, only this quantity enters the final result. In the *CP*± fits, the shared parameters are the asymmetries $A_{CP}^{\mathrm{sig}(K)} \equiv A_{CP}$ and $A_{CP}^{\mathrm{sig}(\pi)}$, and the ratios $R_{K/\pi}$. The ratios will lead to $R_{CP\pm}$ when combined with the averaged value of $R_{K/\pi}^{\mathrm{flv}}$ as obtained from the flavor modes. The ε and m parameters are expected to be the same in all D^0 modes. They are therefore shared in the *CP*± fits.

In principle it is desirable to float as many parameters as possible in the final fit. The more parameters can be determined directly from data, the less a-priori knowledge must be put in, and the lower the associated systematic uncertainty will be. However, it is difficult to control a fit with too many floating parameters. We decide for each parameter whether to float it or to fix it to a value obtained from MC simulation.

There are two classes of parameters in the final PDF: yield related parameters, introduced through Eq. 9.13 (and analogous ones), and shape parameters. The shape parameters will be introduced in the following sections.

As for the yield parameters, we float all parameters related to the signal yields, and therefore to the GLW parameters, except ε. A dedicated working group at *BABAR* provides the ε and m values for the various PID selectors. Since the final fit doesn't reach their precision on ε we'll fix this parameter to the provided value: $\varepsilon = 0.77 \pm 0.01$. On the other hand, the final fit is able to determine m to a better precision than the working group, so m will be left free. In addition, we float as many background yields and *CP*-asymmetries as possible, because from the previous Chapter 8 we know the MC samples don't describe the overall normalization very well. As a matter of fact, we can't float all background yields, because the fit becomes unstable if too many yield parameters are allowed to fluctuate to negative

values. We therefore fix $N^{\rm pk}$ to values obtained in Chapter 11; $A_{CP}^{\rm pk} = 0$; $A_{CP,p}^{B\overline{B}} = 0$; $A_{CP,f}^{B\overline{B}} = 0$ for *CP*– modes; $N_{K_S^0\phi,p}^{B\overline{B}}$ to the MC value.

As for the shape parameters, we observed shifts between data and MC in the signal shapes of ΔE and $m_{\rm ES}$, so we float the according shape parameters. We also float most parameters describing the distribution of the $q\overline{q}$ background component in the ΔE and $\mathcal{F}$ variables.

Table 9.1 gives an overview of the multiplicity of the floating parameters, both yield related and shape parameters. The following sections will introduce the exact shapes an their floating parameters. We will assign a systematic uncertainty due to the fixed parameters.

Table 9.1: Floating parameters of the final fit PDF.

Parameter	multiplicity				
	CP+	*CP*–	$K\pi$	$K\pi\pi^0$	$K\pi\pi\pi$
$A_{CP\pm}^{{\rm sig}(K)}$, $A_{CP\pm}^{{\rm sig}(\pi)}$	2	2	2	2	2
$N_{\rm tot}^{{\rm sig}(\pi)}$	2	3	1	1	1
$R_{K/\pi}$	1	1	1	1	1
$N_p^{q\overline{q}}$, $N_f^{q\overline{q}}$	4	6	2	2	2
$N_p^{B\overline{B}}$, $N_f^{B\overline{B}}$	4	5	2	2	2
$A_{CP,p}^{q\overline{q}}$, $A_{CP,f}^{q\overline{q}}$	4	6	2	2	2
$A_{CP,f}^{B\overline{B}}$	2	-	1	1	1
m	1	1	1	1	1
signal ΔE shape	4	6	2	2	2
signal $m_{\rm ES}$ shape	6	9	3	3	3
$q\overline{q}$ ΔE shape	2	3	1	2	1
$q\overline{q}$ $\mathcal{F}$ shape	8	11	4	4	5
$B\overline{B}$ ΔE shape	-	-	1	-	-
total	40	53	23	23	23

Finding reasonable parameterizations can be challenging. We will rely on the MC samples to determine the functional form of the PDFs used to parameterize the six components in the final fit. Out of the many possibilities to describe the observed shapes, we chose the one with the lowest number of parameters still well describing the data. The agreement is quantified by the probability $P(\chi^2, n_{\rm dof})$, see for instance chapter *Probability* in Reference [8]. The probability P is evaluated on histograms of the distributions, scaled to data luminosity. The number of degrees of freedom of the PDF, $n_{\rm dof}$, is given by the number of non-zero bins minus the number of free parameters of the PDF. The χ^2 value is given by

$$\chi^2 = \sum_i \left(\frac{y_i - {\rm PDF}(x_i)}{\sigma_i} \right)^2 , \tag{9.14}$$

the sum ranging over all non-zero bins x_i with bin entries y_i. The probability $P(\chi^2, n_{\rm dof})$ is given by

$$P(\chi^2, n) = \frac{1}{2^{n/2}\,\Gamma(n/2)} \int_{\chi^2}^{\infty} z^{n/2-1} e^{-z/2} {\rm d}z, \tag{9.15}$$

$\Gamma(k)$ being the gamma function. One usually considers P values of greater than 0.01 good agreement. However, in the high statistics case one has to trade off the complexity of the PDF and the P value.

We simplify the complexity of the final PDF by the assumption the shapes differ neither across the charged subsamples, nor across the subsamples defined by the `KLHVeryTight` selector. Again, we validate this assumption by means of $P(\chi^2, n_{\rm dof})$.

For each component, an extraction fit is run on all available MC statistics to get the most precise estimate of the parameters which eventually will be fixed in the final fit. However, the MC is only an approximation to the real data, and when we obtain shapes from the MC that are going to be fixed in the final fit, we potentially bias the fit result. We will estimate these biases and correct for them in Section 10.2.

The following sections describe which functional forms we use for each component. The shapes are fairly similar for all D^0 final states under study. For brevity's sake, we only show the corresponding plots for the representative $D^0 \to K^-K^+$ channel, and we also omit tables with the exact parameter values and plots which overlay the shapes to the charge and `KLHVeryTight` slices. The full set of plots and tables is contained in [42].

9.1 Signal Parameterization

The $m_{\rm ES}$ signal components are being parameterized using a bifurcated Gaussian shape, which is an asymmetric Gaussian function with different widths on each side. Its mean $\mu_{m_{\rm ES}}$ and left and right widths, $\sigma_{m_{\rm ES},l}$, $\sigma_{m_{\rm ES},r}$, will be floating in the final fit. We use the same shape for $B \to D^0K$ and $B \to D^0\pi$, so the $m_{\rm ES}$ signal shape will mostly be determined by the more abundant control sample. In the $D^0 \to K^-\pi^+\pi^0$ and the $D^0 \to K^-\pi^+\pi^+\pi^-$ channels we add a single Gaussian to account for a small self cross-feed component: the ratios of the integral of the single Gaussian to the integral of the bifurcated Gaussian are 8.6 % and 2.7 %, in both channels respectively. In the final fit the parameters of the single Gaussian will be fixed to values obtained from signal MC. Figure 9.1 overlays the $m_{\rm ES}$ signal shape to $D^0 \to K^-K^+$ signal MC.

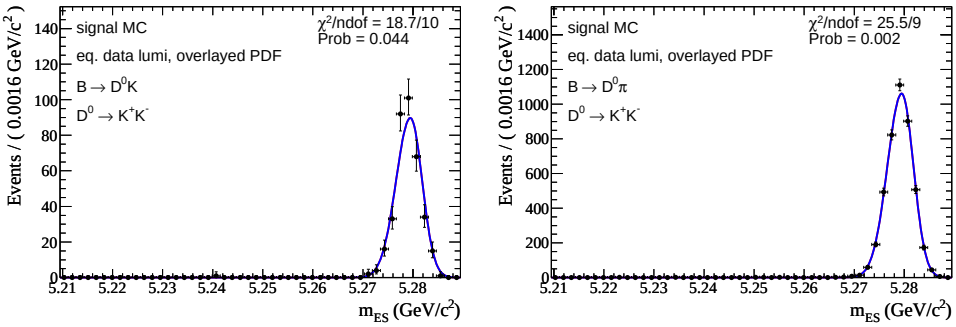

Figure 9.1: $m_{\rm ES}$ signal PDF overlaid to a subsample of $B \to D^0K$, $D^0 \to K^-K^+$ signal MC. The shape is a bifurcated Gaussian.

The $B \to D^0K$ ΔE signal component is parameterized with a double Gaussian shape. To be able to determine the crucial parameters from the fit to data, we fix the parameters of the wide component to values obtained from the signal MC samples. In particular, we fix the difference between the means of the wide and the narrow component, $\mu_s = \mu_1 - \mu_2$, the ratio of the widths of both components, $r_\sigma = \sigma_1/\sigma_2$, and the ratio of the integrals, f. The remaining two parameters, mean and width of the narrow Gaussian, $\mu_{\Delta E}$ and $\sigma_{\Delta E}$, will be left free in the fit.

The $B \to D^0\pi$ ΔE shape is not double Gaussian. This is due to the fact that we assigned the kaon mass to the prompt track; if we would have used the pion mass to reconstruct $B \to D^0\pi$

events, the shapes would be similar (see Section 6.2). The wrong mass assignment introduces a shift to ΔE, which depends on the momentum $\vec{p}$ of the prompt track in the laboratory frame,

$$\Delta E_{\texttt{shift}}(p) = \gamma_{\mathrm{PEP-II}} \left(\sqrt{m_K^2 + \vec{p}^2} - \sqrt{m_\pi^2 + \vec{p}^2} \right), \tag{9.16}$$

where $\gamma_{\mathrm{PEP-II}} = E_{ee}/E_{ee}^*$ characterizes the boost to the e^+e^- CM frame. So the shapes of $B \to D^0 K$ and $B \to D^0 \pi$ are in fact similar if the first one is evaluated as a function of ΔE, while the second one is evaluated as a function of $\Delta E - \Delta E_{\texttt{shift}}$. Since we want the more abundant $B \to D^0 \pi$ control sample to dominate the $B \to D^0 K$ signal shape, we parameterize the $B \to D^0 \pi$ ΔE signal component with a conditional double Gaussian, whose mean is shifted by $\Delta E_{\texttt{shift}}$, computed event-per-event as

$$\mu(D^0 \pi) = \mu(D^0 K) + \Delta E_{\texttt{shift}}(p), \tag{9.17}$$

$\mu(D^0 K)$ and $\mu(D^0 \pi)$ denoting the means of the narrow Gaussians. We take the floating width and mean of the narrow component, $\mu_{\Delta E}$ and $\sigma_{\Delta E}$, to be the same in both the $B \to D^0 K$ and $B \to D^0 \pi$ PDFs.

In case of the high statistics flavor modes, we add a straight line to the double Gaussian shape to account for a self cross-feed component, which peaks in m_{ES} but not in ΔE. This is most important in $D^0 \to K^- \pi^+ \pi^0$ events.

Figure 9.2 overlays the PDF to scaled-down signal MC samples matching the data luminosity. Since in the final fit $\mu_{\Delta E}$ and $\sigma_{\Delta E}$ will mostly be determined by the $B \to D^0 \pi$ component, we use the shape of the narrow Gaussian that was extracted from $B \to D^0 \pi$ also for the $B \to D^0 K$ plot.

We emphasize the $B \to D^0 \pi$ ΔE PDF depends on an external parameter, $\Delta E_{\texttt{shift}}$, which is different for each event. Such PDFs are called *conditional* PDFs. The difference to regular, non-conditional PDFs is the normalization condition. A regular (two-dimensional) PDF G is normalized as

$$\int G(x, y)\, \mathrm{d}x\mathrm{d}y \equiv 1, \tag{9.18}$$

whereas a conditional PDF is normalized in x *for each value of* y as

$$\int G(x|y)\, \mathrm{d}x\mathrm{d}y \equiv 1, \qquad \forall y. \tag{9.19}$$

Fitting with conditional PDFs is well behaved, since it is straightforward to construct the likelihood function. However, plotting is not straightforward. In order to visualize the distribution of x in data and overlay the fitted PDF, one has to project (i.e. integrate) over the variable y, requiring knowledge of the full PDF

$$H(x, y) = G(x|y) \cdot F(y) \tag{9.20}$$

in order to compute the integral

$$H(x) = \int H(x, y)\, \mathrm{d}y = \int G(x|y) F(y)\, \mathrm{d}y. \tag{9.21}$$

To this purpose one can either determine an analytic expression for $F(y)$, or in alternative use the value of $F(y_i)$ for n events ($i \in [1, n]$) of an external (*prototype*) dataset to approximate $F(y)$ by an histogram of infinitesimal binning

$$F(y) \approx \frac{1}{n} \sum_i \delta(y - y_i), \tag{9.22}$$

so that

$$H(x) = \frac{1}{n}\sum_i \int G(x|y)\delta(y-y_i)\,\mathrm{d}y = \frac{1}{n}\sum_i G(x|y_i). \tag{9.23}$$

An important consequence is, that the resulting curve depends on the external dataset. We will use the first approach based on Eq. 9.20 for plotting and toy MC purposes as described in Section 10.1.

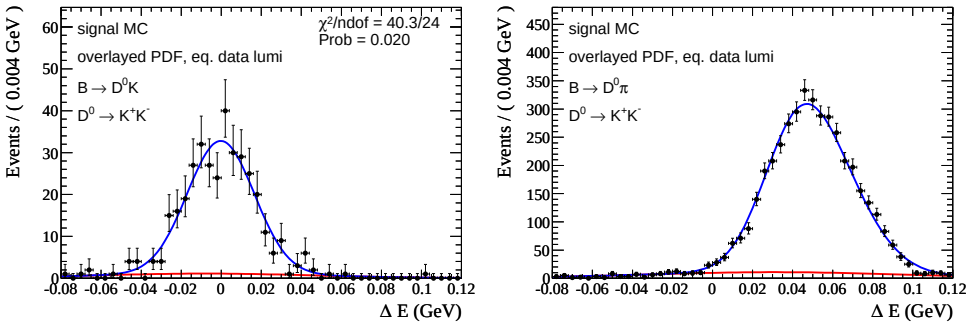

Figure 9.2: ΔE signal shape extracted applied to both $B \to D^0 K$ and $B \to D^0 \pi$ samples of same equivalent luminosity as data. The shape is a double (conditional) Gaussian.

As for the third variable, $\mathcal{F}$, the $B \to D^0\pi$ signal component is parameterized by a double bifurcated Gaussian. In case of the $D^0 \to K^0_S\phi$ channel, it can be simplified to a double Gaussian shape. The $B \to D^0 K$ signal component is parameterized using the shapes extracted from $B \to D^0\pi$. Figure 9.3 shows the $\mathcal{F}$ signal shape.

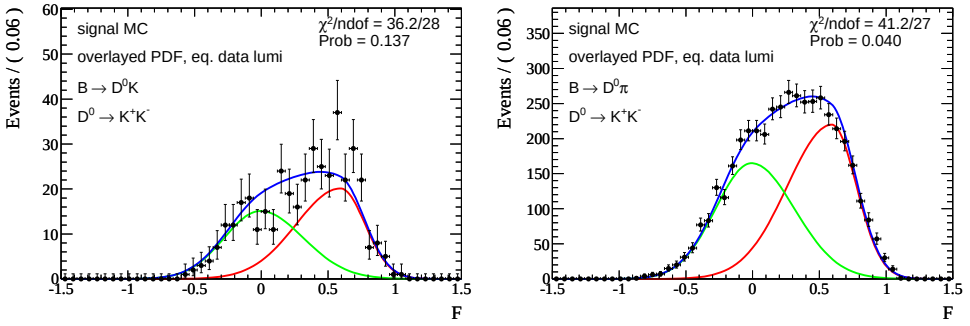

Figure 9.3: $\mathcal{F}$ signal shape extracted from a fit to $B \to D^0\pi$ signal MC, applied to both $B \to D^0 K$ and $B \to D^0\pi$ samples of same equivalent luminosity as data. The shape is a double bifurcated Gaussian.

9.2 $B\overline{B}$ Background Parameterization

There are significant correlations among the (ΔE, m_{ES}) variables arising from the fact that there is a peak sitting on a continuous background from random combinations of particles in the event. The peak is mainly due to $B \to D^* h$ and $D^0\rho^0$ events, where a soft neutral pion

is missed when reconstructing the B decay tree: for this reason, $m_{\rm ES}$ (which depends on the momentum of the B but not on its energy) is peaked at $m(B)$ as for the signal, while ΔE (which depends on the reconstructed energy of the B candidate) is shifted towards negative values by about 130 MeV due to the missing π^0-meson. As a consequence, the PDF doesn't factorize. As mentioned before, we parameterize it by means of two factoring components:

$$\begin{aligned}\mathcal{P}_{B\overline{B}}(\Delta E, m_{\rm ES}) &= f\ g_{\rm peak}(m_{\rm ES})\, h_{\rm peak}(\Delta E) \\ &\quad +(1-f)\ g_{\rm cont}(m_{\rm ES})\, h_{\rm cont}(\Delta E). \end{aligned} \tag{9.24}$$

The exact functional forms for the g and h components are listed in Table 9.2. The fit result of the extraction fit in the $D^0 \to K^-K^+$ channel is shown in Figure 9.4 (projections), and in Figures 9.5 (two-dimensional plots).

Virtually all parameters of the ΔE-$m_{\rm ES}$ PDF will be fixed to values obtained from MC, there are only two exceptions: the width $\sigma^{B\overline{B}}_{\Delta E, m_{\rm ES}}$ of the Landau distribution [43] in $D^0 \to K^-\pi^+$ is not modelled well by the MC sample [42], we will float this parameter. The second exception is the parameter describing the kinematic cut-off in $m_{\rm ES}$, $m_{{\rm ES}c}$. We fix this parameter to a value obtained from a dedicated fit to on-peak data in the $B \to D^0\pi$, $D^0 \to K^-\pi^+$ channel: $m_{{\rm ES}c} = 5289.28 \pm 0.49\,{\rm MeV}/c^2$.

The $\mathcal{F}$ $B\overline{B}$ background component is parameterized following the same lines as the signal component, the extraction fit result is depicted in Figure 9.6.

Table 9.2: Details of the ΔE-$m_{\rm ES}$ $B\overline{B}$ background parameterization. Crystal Ball (CB) [44], Gaussian (G), Novosibirsk distribution (N) [45], Landau distribution (L) [43], Exponential (Exp), Argus (A) [46], polynom of first order (P).

Component	K^+K^-	$\pi^+\pi^-$	$K^0_S\pi^0$	$K^0_S\omega$	$K^0_S\phi$	$K^-\pi^+$	$K^-\pi^+\pi^0$	$K^-\pi^+\pi^+\pi^-$
$g_{\rm peak}(m_{\rm ES})$	CB	G	CB	G	G	N	CB	CB
$h_{\rm peak}(\Delta E)$	Exp	Exp	Exp	Exp	Exp	L	Exp	Exp
$g_{\rm cont}(m_{\rm ES})$	A	A	A	A	A	A	A	A
$h_{\rm cont}(\Delta E)$	P	P	P	P	P	P	P	P

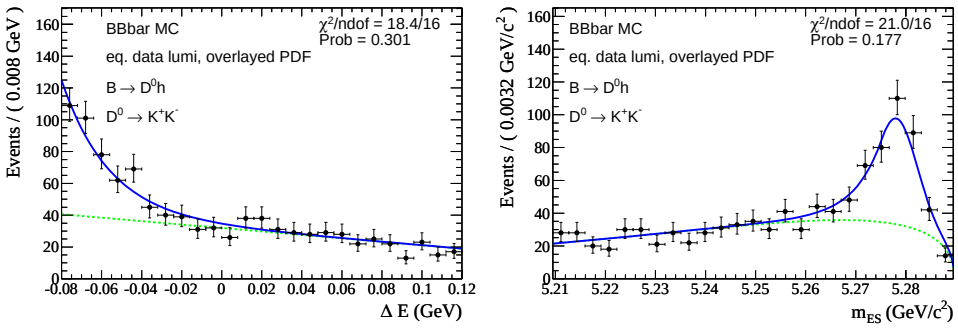

Figure 9.4: ΔE-$m_{\rm ES}$ $B\overline{B}$ PDF, $D^0 \to K^-K^+$ channel, applied to a subsample of same equivalent luminosity as data. Left: ΔE projection. Right: $m_{\rm ES}$ projection. The dashed line shows the continuous component.

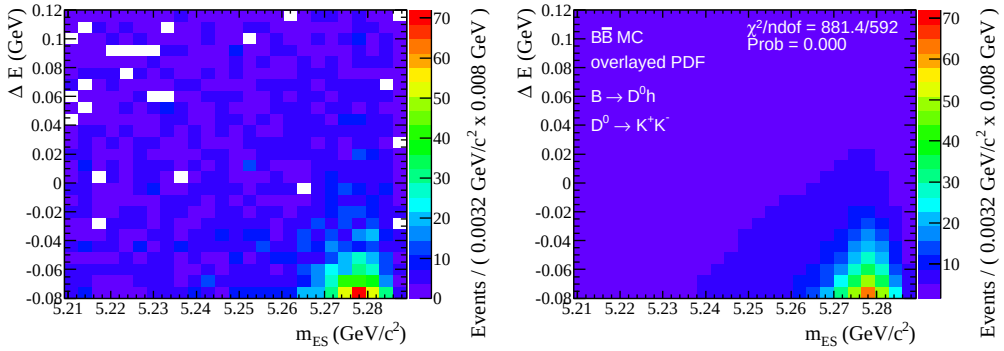

Figure 9.5: ΔE-$m_{\rm ES}$ $B\overline{B}$ PDF, $D^0 \to K^-K^+$ channel, two-dimensional plots. Left: generic $B\overline{B}$ MC, right: fitted PDF.

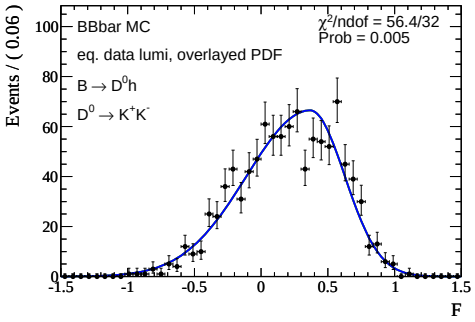

Figure 9.6: $\mathcal{F}$ $B\overline{B}$ PDF, $D^0 \to K^-K^+$ channel

9.3 $q\overline{q}$ Background Parameterization

The $m_{\rm ES}$ $q\overline{q}$ background component is being parameterized by an Argus function [46]. Only in case of $B \to D^0\pi$, $D^0 \to K^-\pi^+\pi^0$, we need to use a different PDF to describe the shape. Here we use an arc tangent multiplied by a power law, $f(m_{\rm ES}) = -\text{atan}((m_{\rm ES} - m_{{\rm ES}c})\,C) \cdot {m_{\rm ES}}^P$. Other than for the components described previously, for the $q\overline{q}$ background component we have a data control sample at hand to validate the extracted shapes. For this purpose we can use the off-peak data sample, which doesn't contain B-mesons. We apply the $q\overline{q}$ PDF to this sample, and calculate the χ^2 probability values. All comparisons show good agreement between the shapes obtained from generic $q\overline{q}$ MC and the off-peak data sample. The $m_{\rm ES}$ plots are shown in Figure 9.7. All parameters of the $q\overline{q}$ $m_{\rm ES}$ shape are fixed in the final fit. The $m_{\rm ES}$ cut-off parameter, $m_{{\rm ES}c}$, is taken to be the same as in the previous section.

The ΔE $q\overline{q}$ background component is being parameterized by a linear shape. Only in the channel $B \to D^0\pi$, $D^0 \to K^-\pi^+\pi^0$, a linear shape is not sufficient, and we use a second order polynom. The associated parameters, $a^{q\overline{q}}_{\Delta E}$ and $b^{q\overline{q}}_{\Delta E}$, are left free in the fit. The ΔE plots are shown in Figure 9.8.

The $\mathcal{F}$ $q\overline{q}$ background component is being parameterized by a double bifurcated Gaussian. In $D^0 \to K^0_S\phi$ we simplify to a single bifurcated Gaussian. A double bifurcated Gaussian has

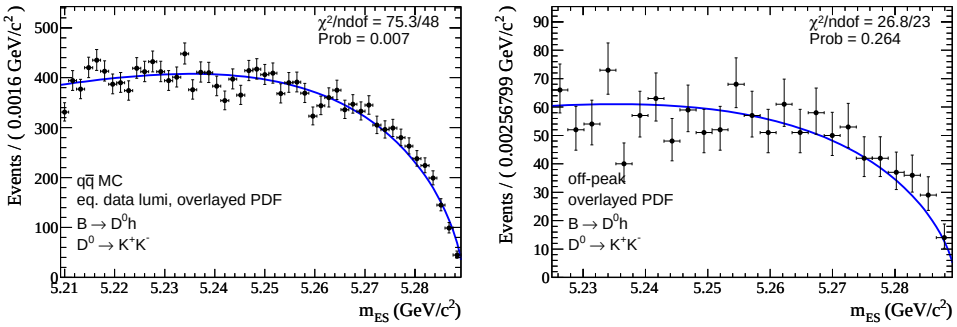

Figure 9.7: $m_{\rm ES}$ shape extracted from $q\overline{q}$ MC, applied to both a subsample of same equivalent luminosity as data, and to off-peak data. The $m_{\rm ES}$ range is reduced in the off-peak sample.

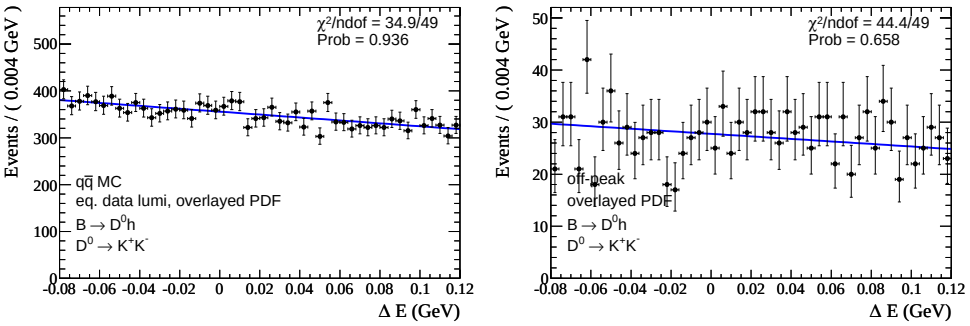

Figure 9.8: ΔE shape extracted from $q\overline{q}$ MC, applied to both a subsample of same equivalent luminosity as data, and to off-peak data.

seven parameters, two means, four widths $\sigma^{q\overline{q}}_{\mathcal{F},i,j}$, $i = \{l, r\}$, $j = \{1, 2\}$, and the ratio of the integrals $f^{q\overline{q}}_{\mathcal{F}}$. We fix the means in all channels to values obtained from MC, and we also fix $f^{q\overline{q}}_{\mathcal{F}}$ in case of $D^0 \to K^0_S\pi^0$ and $D^0 \to K^0_S\omega$. The remaining parameters are left free in the fit. The $\mathcal{F}$ plots are shown in Figure 9.9.

9.4 Peaking Background Parameterization

The $B \to fh$ charmless peaking background is discussed in detail in Chapter 11. Its $m_{\rm ES}$ distribution is parameterized by re-using the $m_{\rm ES}$ signal shape.

The ΔE $B \to fh$ background components are being parameterized by a single Gaussian. Its parameters are obtained from a fit to $B \to fh$ events in generic $B\overline{B}$ MC, in the $B \to D^0K$, $D^0 \to K^-K^+$ channel. The width of this distribution is 38 ± 2 MeV, which is about twice as wide compared to the one of $B \to D^0h$ signal (16 MeV). This is due to the fact that in the reconstruction of the signal decay tree a mass constraint is applied to the f system forming the D^0. The constraint improves the energy resolution of the f system if it originates from a true D^0, otherwise it worsens the resolution.

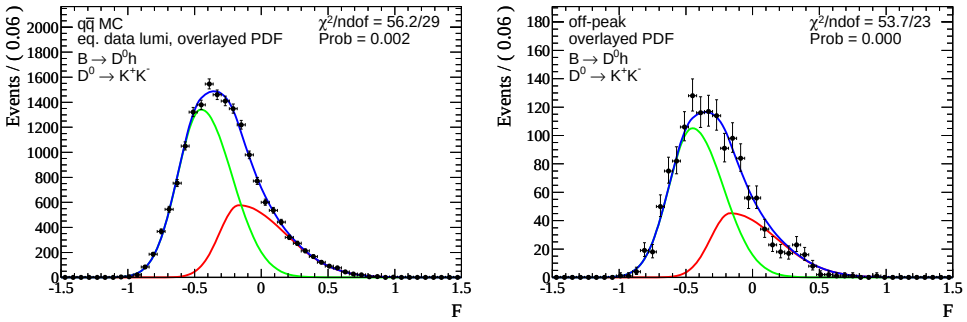

Figure 9.9: $\mathcal{F}$ shape extracted from $q\overline{q}$ MC, applied to both a subsample of same equivalent luminosity as data, and to off-peak data.

The $\mathcal{F}$ $B \to fh$ background component is being parameterized by a double bifurcated Gaussian. Its parameters are obtained from a fit to $B \to fh$ events in generic $B\overline{B}$ MC, in the $B \to D^0\pi$, $D^0 \to K^-\pi^+$ channel. Figure 9.10 shows the plots of all three distributions.

9.5 Start Parameters

To summarize this chapter, we give the start values of all floating parameters of the final fits in Tables 9.3 and 9.4. The expected yields are extracted from the Monte Carlo samples by applying the final selection and scaling to data luminosity, the same is true for the mis-id parameter of the `KLHVeryTight` selector. The start values of the shape parameters are extracted by fitting the corresponding MC samples, as described in the previous sections.

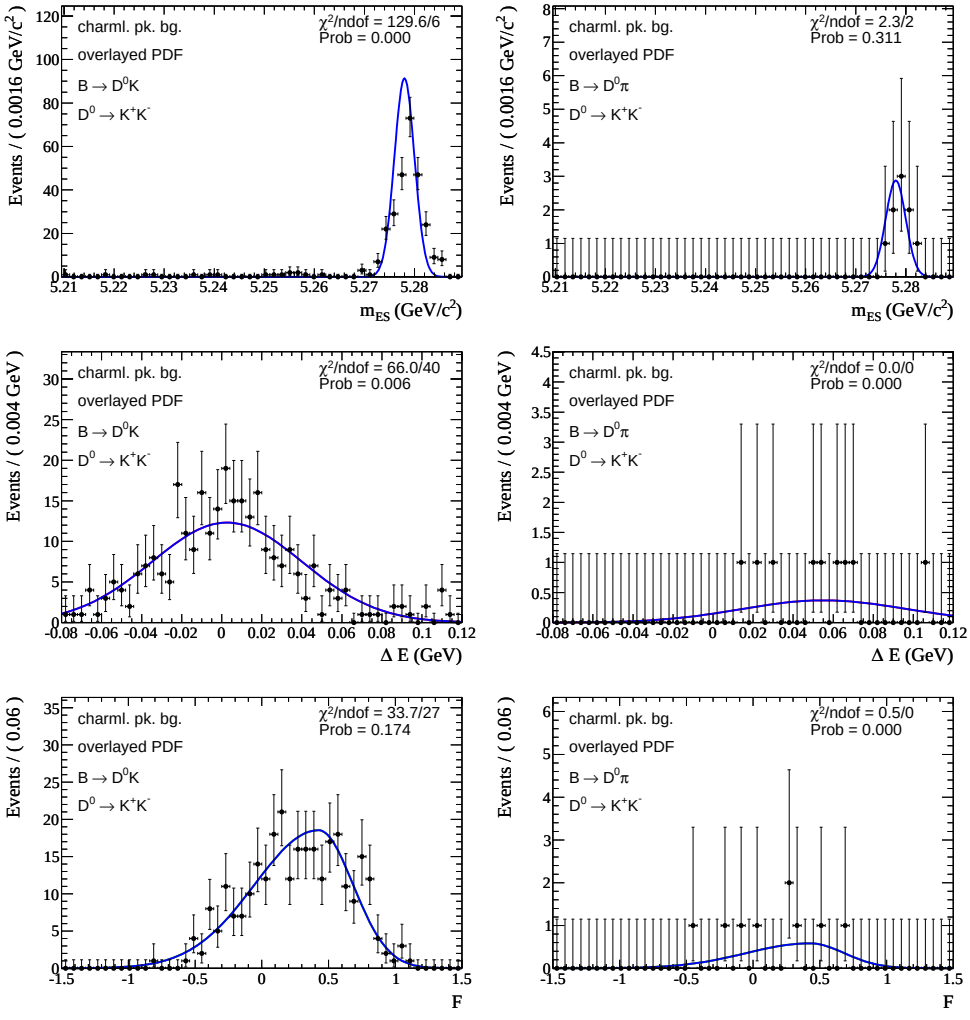

Figure 9.10: The $m_{\rm ES}$, ΔE, and $\mathcal{F}$ shapes of $B \to fh$ events extracted from $B\overline{B}$ MC.

Table 9.3: Start parameters of the final PDF of the *CP*+ and *CP*− fits. Means and widths of ΔE and $m_{\rm ES}$ parameters in MeV (MeV/c^2). The slope $a^{q\overline{q}}_{\Delta E}$ is in units of GeV^{-1}.

Parameter	KK	$\pi\pi$	$K^0_S\pi^0$	$K^0_S\omega$	$K^0_S\phi$
$A_{CP}^{\mathrm{sig}(K)}$	0.2			0.0	
$A_{CP}^{\mathrm{sig}(\pi)}$	0.0			0.0	
$R_{K/\pi}$	0.08			0.08	
m	0.0247			0.0247	
$A_{CP,f}^{B\overline{B}}$	0.0	0.0	0.0	0.0	0.0
$A_{CP,p}^{q\overline{q}}$	0.0	0.0	0.0	0.0	0.0
$A_{CP,f}^{q\overline{q}}$	0.0	0.0	0.0	0.0	0.0
$f_{\mathcal{F}}^{q\overline{q}}$	0.343	0.494104	0.440	0.265287	n/a
$\sigma_{\mathcal{F},l,1}^{q\overline{q}}$	0.147	0.190	0.179	0.149	0.298
$\sigma_{\mathcal{F},l,2}^{q\overline{q}}$	0.1756	0.1930	0.1542	0.1957	n/a
$\sigma_{\mathcal{F},r,1}^{q\overline{q}}$	0.348	0.3610	0.3866	0.343	0.455
$\sigma_{\mathcal{F},r,2}^{q\overline{q}}$	0.232	0.278	0.294	0.240	n/a
$a_{\Delta E}^{q\overline{q}}$	−0.88	−0.89	−1.386	−1.19	−0.477452
$\mu_{\Delta E}$	−0.222	−0.300	−0.948	−1.16	−0.376
$\sigma_{\Delta E}$	16.689	16.439	18.037	18.15	16.827
$\mu_{m_{\rm ES}}$	5279.496	5279.459	5279.499	5279.579	5279.516
$\sigma_{m_{\rm ES},l}$	2.871	2.866	2.956	3.080	2.889
$\sigma_{m_{\rm ES},r}$	2.349	2.380	2.411	2.400	2.355
$N_p^{B\overline{B}}$	77	141	86	89	3
$N_f^{B\overline{B}}$	925	651	816	1080	59
$N_p^{q\overline{q}}$	4282	3585	8114	3042	364
$N_f^{q\overline{q}}$	13173	17739	23148	8118	807
$N_{\rm tot}^{\mathrm{sig}(K)}$	374	115	408	140	60.0
$N_{\rm tot}^{\mathrm{sig}(\pi)}$	4284	1324	4886	1656	724

Table 9.4: Start parameters of the final PDF. Means and widths of ΔE and $m_{\rm ES}$ parameters in MeV (MeV/c^2). The parameters $a^{q\overline{q}}_{\Delta E}$ and $b^{q\overline{q}}_{\Delta E}$ are in units of GeV^{-1} and GeV^{-2}, respectively.

Parameter	$K\pi$	$K\pi\pi^0$	$K\pi\pi\pi$
$A_{CP}^{\mathrm{sig}(K)}$	0.0	0.0	0.0
$A_{CP}^{\mathrm{sig}(\pi)}$	0.0	0.0	0.0
$R_{K/\pi}$	0.08	0.08	0.08
m	0.0247	0.0247	0.0247
$A_{CP,f}^{B\overline{B}}$	0.0	0.0	0.0
$A_{CP,p}^{q\overline{q}}$	0.0	0.0	0.0
$A_{CP,f}^{q\overline{q}}$	0.0	0.0	0.0
$f_{\mathcal{F}}^{q\overline{q}}$	0.369	0.3698	0.373
$\sigma_{\mathcal{F},l,1}^{q\overline{q}}$	0.189	0.1676	0.1752
$\sigma_{\mathcal{F},l,2}^{q\overline{q}}$	0.1919	0.17716	0.1886
$\sigma_{\mathcal{F},r,1}^{q\overline{q}}$	0.3532	0.3676	0.3693
$\sigma_{\mathcal{F},r,2}^{q\overline{q}}$	0.257	0.2487	0.2657
$\sigma_{\Delta E,m_{\rm ES}}^{B\overline{B}}$	0.00325	n/a	n/a
$a_{\Delta E}^{q\overline{q}}$	-1.093	-1.629	-1.173
$b_{\Delta E}^{q\overline{q}}$	n/a	4.96	n/a
$\mu_{\Delta E}$	-0.191	-0.843	-0.067
$\sigma_{\Delta E}$	15.996	17.402	16.13
$\mu_{m_{\rm ES}}$	5279.6359	5279.605	5279.512
$\sigma_{m_{\rm ES},l}$	2.8814	3.0140	2.856
$\sigma_{m_{\rm ES},r}$	2.3369	2.4682	2.327
$N_p^{B\overline{B}}$	177	1977	827
$N_f^{B\overline{B}}$	3184	28829	13325
$N_p^{q\overline{q}}$	4789	24468	8432
$N_f^{q\overline{q}}$	32033	170968	60205
$N_{\rm tot}^{\mathrm{sig}(K)}$	3952	6664	4732
$N_{\rm tot}^{\mathrm{sig}(\pi)}$	49667	84499	59552

10 Fit Validation

Before trusting the outcome of a complex fitting scenario one should do a series of crosschecks and validations. In this analysis three checks are performed: a Pure Toy Monte Carlo study (Section 10.1) to validate the implementation and the abilities of the fit, an Embedded Toy Monte Carlo study (Section 10.2) to estimate biases, and a fit of Cocktail Monte Carlo (Section 10.3) as the most realistic simulation of real data.

10.1 Toy Monte Carlo

The first crosscheck is to fit a large number of datasets which have been generated according to the final PDF, so called *(pure) toy datasets* or *(pure) toy experiments*. By doing so one can study pull distributions and conclude whether or not the fitting algorithm is intrinsically biased. In theory, such a study serves as a crosscheck of the implementation, since there is only little reason why a maximum likelihood fit should be intrinsically biased (other than in the low-statistics case). In praxis, there are many possibilities to acquire biases in complex implementations: machine accuracy, parameter range issues, issues in the event generation process, just to name a few.

The pull of a fit parameter α is defined as the ratio of the difference e between the true value $<\alpha>$ and the parameter, and the parameter's standard deviation σ_α, as obtained from the fit:

$$p_\alpha = \frac{\alpha - <\alpha>}{\sigma_\alpha} = \frac{e}{\sigma_\alpha}. \tag{10.1}$$

In case of asymmetric errors $\sigma_\alpha^\pm$ (obtained by MINOS), one generalizes $\sigma_\alpha = \sigma_\alpha^+$, if $e > 0$, and $\sigma_\alpha = \sigma_\alpha^-$ if $e < 0$. One expects p be distributed normally, that is to follow a Gaussian distribution with mean $\mu = 0$ and width $\sigma = 1$. It is possible to use the mean and the width of the observed pull distributions to obtain unbiased estimates of the fit parameters. Consider a parameter A, for which the final fit gives an estimate $a \pm \sigma_a$. The fit of the pull distribution to a Gaussian shape shall give values of μ^p and σ^p. Then one calculates an unbiased corrected value $a^c \pm \sigma_a^c$ according to

$$\begin{aligned} a^c &= a - \sigma_a \mu^p, \\ \sigma_a^c &= \sigma_a \sigma^p. \end{aligned} \tag{10.2}$$

Before we can draw events from the final PDF we need to parameterize the $\Delta E_{\texttt{shift}}$ distribution (compare Section 9.1). We chose a Fermi function,

$$f(\Delta E_{\texttt{shift}}) = \frac{1}{e^{b\,(\mu_f - \Delta E_{\texttt{shift}})} + 1}, \tag{10.3}$$

multiplied by a Gaussian shape $G(\mu_g, \sigma_g)$. We obtain the parameters from a fit to generic $B\overline{B}$ Monte Carlo in the $D^0 \to K^-K^+$ channel. The corresponding extraction fit is shown in the left hand side of Figure 10.1. The resulting parameters are $b = 1641 \pm 69\,\text{GeV}^{-1}$,

$\mu_g = 0.029 \pm 0.001\,\text{GeV}$, $\mu_f = 0.03393 \pm 0.00005\,\text{GeV}$, $\sigma_g = 0.0240 \pm 0.0004\,\text{GeV}$. This description is not perfect, because it doesn't take into account the different shapes observed in $B\overline{B}$ and $q\overline{q}$ events. The latter case is shown on the right hand side of Figure 10.1. But since the likelihood function of the final fit does not depend on the shape of the conditional variable at all, this still is a reasonable approximation. We also tried to use a more accurate description, consisting of three bifurcated Gaussians. However, due to a technical problem in RooFit it is not possible to obtain reasonable fit results when using the more complex shape. This problem is still under investigation [47].

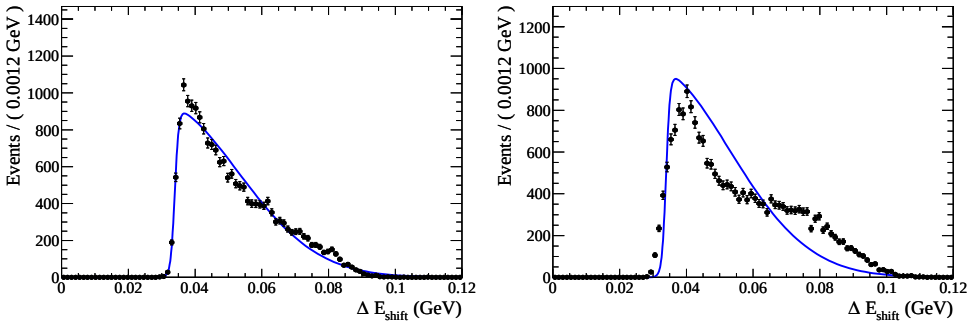

Figure 10.1: Parameterization of $\Delta E_{\texttt{shift}}$ in $B\overline{B}$ MC (left), and overlaid to $q\overline{q}$ MC (right), both in the $D^0 \to K^-K^+$ channel.

Coming back to the pure toy experiments, we performed a few hundred toy MC experiments for each of the five final fits: $CP+$ (188), $CP-$ (190), $K^-\pi^+$ (186), $K^-\pi^+\pi^0$ (135), $K^-\pi^+\pi^+\pi^-$ (90). The resulting pull distributions have been fit to a Gaussian shape, the results for selected parameters are given in Table 10.1. Figure 10.2 shows plots of the pull distribution of four representative parameters. The full set of plots and tables is contained in Reference [42]. In Table 10.1 we also show the value of the relative corrections, $\hat{c}$, which one would have to apply to correct for the observed biases. They are defined as

$$\hat{c} = \frac{a - a^c}{\sigma_a^c} = \frac{\mu^p}{\sigma^p}, \tag{10.4}$$

following the notation of Equation 10.2. We will, however, not correct for the biases obtained from the pure toy MC study. Instead we will apply the corrections found in the *embedded* toy MC study described in the next section. In summary the pulls look reasonable. None of the crucial parameters, that is parameters that will enter the final result, is affected by a significant intrinsic bias.

10.2 Embedded Toy Monte Carlo

In the on-peak dataset, the fit variables are correlated. We consider a sizable correlation explicitly in the final PDF, but we neglect the small correlations. This is expected to bias the fit result. In addition, the PDF may be an imperfect model even if one doesn't take into account correlations. Especially in the high statistics modes one easily finds small, but statistically significant deviations. These can introduce biases, too. The sizes of such biases are reflected by the pulls of an embedded toy MC study.

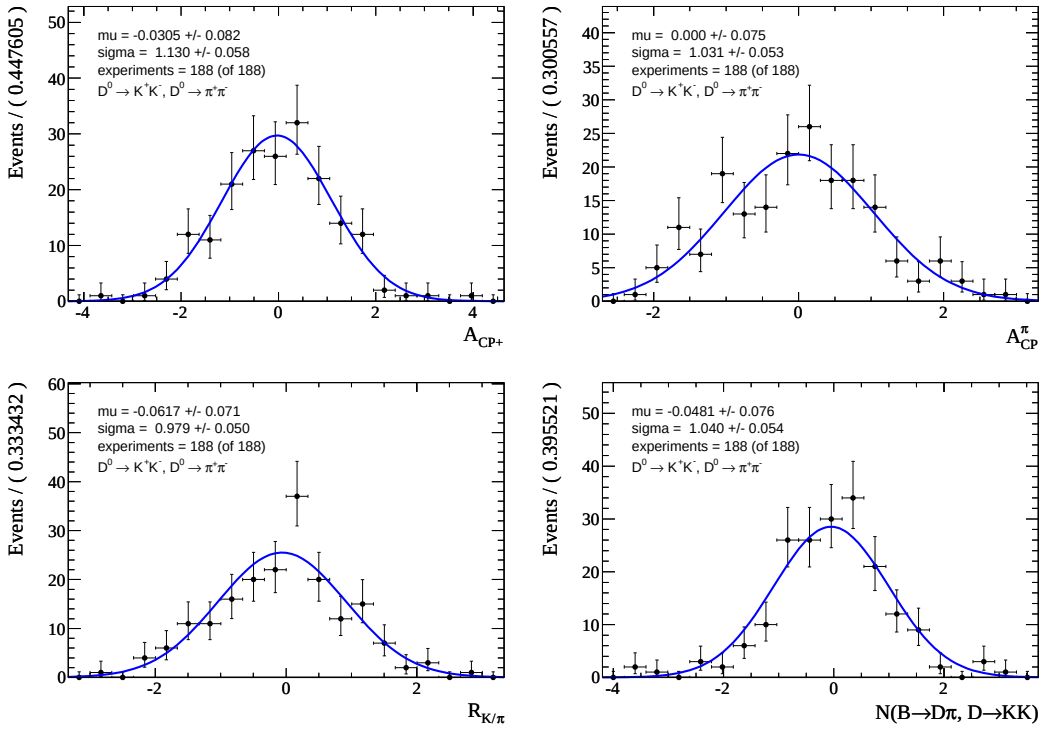

Figure 10.2: Pull distributions of four representative parameters, as obtained from a pure toy MC study.

The second crosscheck is therefore to fit a large number of *embedded toy MC datasets*. Such datasets ideally are drawn from the full MC samples, meaning events which were produced by an event generator and then put through full detector simulation and reconstruction. In praxis, it is only possible to draw a large number of independent datasets for the $B \to D^0 K$ signal component, because here the equivalent luminosity of the signal MC samples exceeds that of the on-peak dataset typically by factors on the order of 100. The simulated background samples, however, are only 2-3 times larger than the on-peak dataset. By drawing from samples of insufficient statistics, one inevitably introduces artificial biases due to the fact one ends up using a large fraction of events multiple times. For this reason one usually chooses to draw the signal components from the large enough signal MC samples, and to draw the background components from their analytic PDFs. Hence the name *embedded* toy Monte Carlo.

However, we don't expect the backgrounds to be entirely uncorrelated. As a matter of fact we take the correlations into account explicitly in case of the ΔE-m_{ES} $B\overline{B}$ PDF. So we will use a different approach to overcome the limited statistics problem: we will construct a binned PDF from the available MC samples, and we will draw embedded toy datasets from it using an accept-reject method. There are several shortcomings to this method as well, but they seem well controllable:

Technical difficulties. RooFit can't generate datasets from a binned PDF of four dimensions without printing warnings. That's why we chose to only draw the fit variables (ΔE, m_{ES}, $\mathcal{F}$) from the binned PDF, and to draw the conditional variable, $\Delta E_{\texttt{shift}}$, from its analytic

Table 10.1: Means μ^p and widths σ^p of the pull distributions of important parameters as obtained in the Pure Toy MC study. We also give the relative correction $\hat{c}$ defined in Equation 10.4.

Fit	Parameter	μ^p	σ^p	rel. corr.
CP+	$A_{CP}^{\mathrm{sig}(K)}$	-0.03 ± 0.08	1.13 ± 0.06	-0.03
	$A_{CP}^{\mathrm{sig}(\pi)}$	0.00 ± 0.08	1.03 ± 0.05	0.00
	$R_{K/\pi}$	-0.06 ± 0.07	0.98 ± 0.05	-0.06
	m	-0.17 ± 0.07	0.98 ± 0.05	-0.17
CP–	$A_{CP}^{\mathrm{sig}(K)}$	0.02 ± 0.07	0.98 ± 0.05	0.02
	$A_{CP}^{\mathrm{sig}(\pi)}$	0.07 ± 0.08	1.09 ± 0.06	0.06
	$R_{K/\pi}$	0.02 ± 0.07	1.00 ± 0.05	0.02
	m	-0.25 ± 0.07	1.03 ± 0.05	-0.24
$K^-\pi^+$	$A_{CP}^{\mathrm{sig}(K)}$	-0.07 ± 0.07	0.98 ± 0.05	-0.07
	$A_{CP}^{\mathrm{sig}(\pi)}$	0.05 ± 0.08	1.03 ± 0.05	0.05
	$R_{K/\pi}$	0.02 ± 0.07	0.96 ± 0.05	0.02
	m	0.04 ± 0.07	1.00 ± 0.05	0.04
$K^-\pi^+\pi^0$	$A_{CP}^{\mathrm{sig}(K)}$	-0.08 ± 0.09	1.02 ± 0.06	-0.08
	$A_{CP}^{\mathrm{sig}(\pi)}$	0.04 ± 0.08	0.90 ± 0.05	0.05
	$R_{K/\pi}$	-0.08 ± 0.09	1.06 ± 0.06	-0.07
	m	0.02 ± 0.09	1.03 ± 0.06	0.02
$K^-\pi^+\pi^+\pi^-$	$A_{CP}^{\mathrm{sig}(K)}$	-0.08 ± 0.11	1.09 ± 0.08	-0.08
	$A_{CP}^{\mathrm{sig}(\pi)}$	-0.07 ± 0.10	0.99 ± 0.07	-0.07
	$R_{K/\pi}$	0.02 ± 0.09	0.89 ± 0.07	0.03
	m	-0.01 ± 0.11	1.08 ± 0.08	-0.01

PDF. To get the correct relation between ΔE and $\Delta E_{\texttt{shift}}$ in case of the $B \to D^0\pi$ signal component, we draw $\Delta E(\pi)$ rather than $\Delta E(K)$, and compute $\Delta E = \Delta E(\pi) + \Delta E_{\texttt{shift}}$ in a second step.

Computing resources. It is a non-trivial task to perform accept-reject on a three-dimensional binned PDF. Since by construction nothing factors out, one cannot limit the parameter space a priori. That's why the dataset generation demands strong computing resources.

Finite binning. In principle the finite binning of the binned PDF can introduce biases. We chose 50 bins in ΔE, 50 bins in $\mathcal{F}$, and 75 bins in m_{ES}. We emphasize we still perform an unbinned fit on the embedded toy datasets.

The results of the pull fits are given in Table 10.2. The number of experiments is 215 (*CP*+), 186 (*CP*–), 244 ($K^-\pi^+$), 243 ($K^-\pi^+\pi^0$), 198 ($K^-\pi^+\pi^+\pi^-$). We will correct the final values of the parameters A_{CP} and $R_{K/\pi}$ by the values obtained in this study using Equations 10.2. A systematic uncertainty will be assigned based on the errors of the corrections (see Section 12.3).

10.3 Cocktail Monte Carlo

The third crosscheck is to run the final fit on a luminosity weighted mixture of generic $q\overline{q}$ MC, generic $B\overline{B}$ MC, and signal MC, which is supposed to mimic on-peak data as closely as possible. This mixture is referred to as *Cocktail Monte Carlo*. From the fit results on Cocktail MC, given in Tables A.1-A.3 in the Appendix A, we calculate the GLW parameters $A_{CP\pm}$ and $R_{CP\pm}$. First, we correct each fit result by the observed bias using Equations 10.2 and the appropriate values from Table 10.2. Then we average the values for $R_{K/\pi}^{\mathrm{flv}}$ obtained from

Table 10.2: Means μ^p and widths σ^p of the pull distributions of important parameters as obtained in the Embedded Toy MC study. We will correct the final result for these biases. We also give the relative correction $\hat{c}$ defined in Equation 10.4.

Fit	Parameter	μ^p	σ^p	rel. corr.
CP+	$A_{CP}^{\mathrm{sig(K)}}$	0.70 ± 0.06	0.93 ± 0.04	0.75
	$A_{CP}^{\mathrm{sig}(\pi)}$	-0.18 ± 0.07	0.96 ± 0.05	-0.19
	$R_{K/\pi}$	0.66 ± 0.06	0.95 ± 0.05	0.69
	m	-0.63 ± 0.07	1.08 ± 0.05	-0.59
CP–	$A_{CP}^{\mathrm{sig(K)}}$	0.11 ± 0.07	0.97 ± 0.05	0.12
	$A_{CP}^{\mathrm{sig}(\pi)}$	-0.11 ± 0.07	1.00 ± 0.05	-0.11
	$R_{K/\pi}$	0.38 ± 0.07	0.91 ± 0.05	0.42
	m	-0.42 ± 0.08	1.08 ± 0.06	-0.39
$K^-\pi^+$	$A_{CP}^{\mathrm{sig(K)}}$	-0.22 ± 0.07	1.03 ± 0.05	-0.22
	$A_{CP}^{\mathrm{sig}(\pi)}$	-0.11 ± 0.06	0.99 ± 0.04	-0.11
	$R_{K/\pi}$	0.88 ± 0.06	1.00 ± 0.05	0.88
	m	-1.79 ± 0.06	1.01 ± 0.05	-1.77
$K^-\pi^+\pi^0$	$A_{CP}^{\mathrm{sig(K)}}$	-0.42 ± 0.06	0.97 ± 0.04	-0.43
	$A_{CP}^{\mathrm{sig}(\pi)}$	-0.31 ± 0.07	1.11 ± 0.05	-0.28
	$R_{K/\pi}$	0.76 ± 0.06	1.01 ± 0.05	0.75
	m	-3.34 ± 0.06	0.98 ± 0.04	-3.40
$K^-\pi^+\pi^+\pi^-$	$A_{CP}^{\mathrm{sig(K)}}$	-0.63 ± 0.07	0.94 ± 0.05	-0.67
	$A_{CP}^{\mathrm{sig}(\pi)}$	0.25 ± 0.07	0.95 ± 0.05	0.26
	$R_{K/\pi}$	0.85 ± 0.08	1.07 ± 0.05	0.79
	m	-1.72 ± 0.07	1.05 ± 0.05	-1.64

the $D^0 \to K^-\pi^+$, $K^-\pi^+\pi^0$, and $K^-\pi^+\pi^+\pi^-$ channels. Finally we calculate $R_{CP\pm}$ using Equation 3.16. The final result on Cocktail MC is given in Table 10.3. The numbers agree well with the expected values which have been used to generate the cocktail MC sample: $A_{CP+} = 0.2$, $A_{CP-} = 0.0$ and $R_{CP\pm} = 1$.

Table 10.3: Measured ratios $R_{CP\pm}$ and A_{CP-} in cocktail MC. Generated values: $A_{CP+} = 0.2$, $A_{CP+} = 0.0$ and $R_{CP\pm} = 1$.

D^0 mode	R_{CP}	A_{CP}
CP+	0.998 ± 0.072	$+0.207 \pm 0.061$
CP–	0.956 ± 0.060	-0.001 ± 0.058

There are various possibilities to visualize the fit result of simultaneous multi-dimensional fits. We present twelve projections per final state of the D^0, corresponding to the twelve slices defined by fit variable, charge, and `KLHVeryTight` selector. They are shown in Figures 10.3 and 10.4 for the representative $D^0 \to K^-K^+$ channel. In order to better reflect the true signal significance, signal range cuts have been applied to the variables not plotted: $\Delta E \in [-0.04, 0.1]\,\mathrm{GeV}$, $m_{\mathrm{ES}} \in [5.275, 5.285]\,\mathrm{GeV}/c^2$, $\mathcal{F} \in [0.2, 1.0]$. Table 10.4 introduces the color code used in all plots showing results of the final fit.

For projecting out the conditional $\Delta E_{\texttt{shift}}$ (compare Section 9.1) we use the more accurate parameterization already mentioned in Section 10.1: a triple bifurcated Gaussian, different for each of the three components $q\overline{q}$, $B\overline{B}$, and $B \to D^0 h$ signal. We obtain the parameters from a fit to the respective MC samples, in the representative $B \to D^0\pi$, $D^0 \to K^-\pi^+\pi^0$ channel. In the $D^0 \to K^-\pi^+\pi^0$ channel, there is a relatively large fraction of $q\overline{q}$ events

compared to the signal events. Therefore in this channel the shape of $\Delta E_{\texttt{shift}}$ has the most influence when producing plots in the ΔE projection.

Table 10.4: Color code for plots which show results of the final fit.

component	color
overall PDF	blue
$B \to D^0 K$ signal	light red, stacked on all backgrounds
$B \to D^0 \pi$ signal	dark red, stacked on all backgrounds
$B \to D^0 K$ $q\overline{q}$	light green, stacked
$B \to D^0 \pi$ $q\overline{q}$	dark green, stacked
$B \to D^0 K$ $B\overline{B}$	light orange, stacked
$B \to D^0 \pi$ $B\overline{B}$	dark orange, stacked

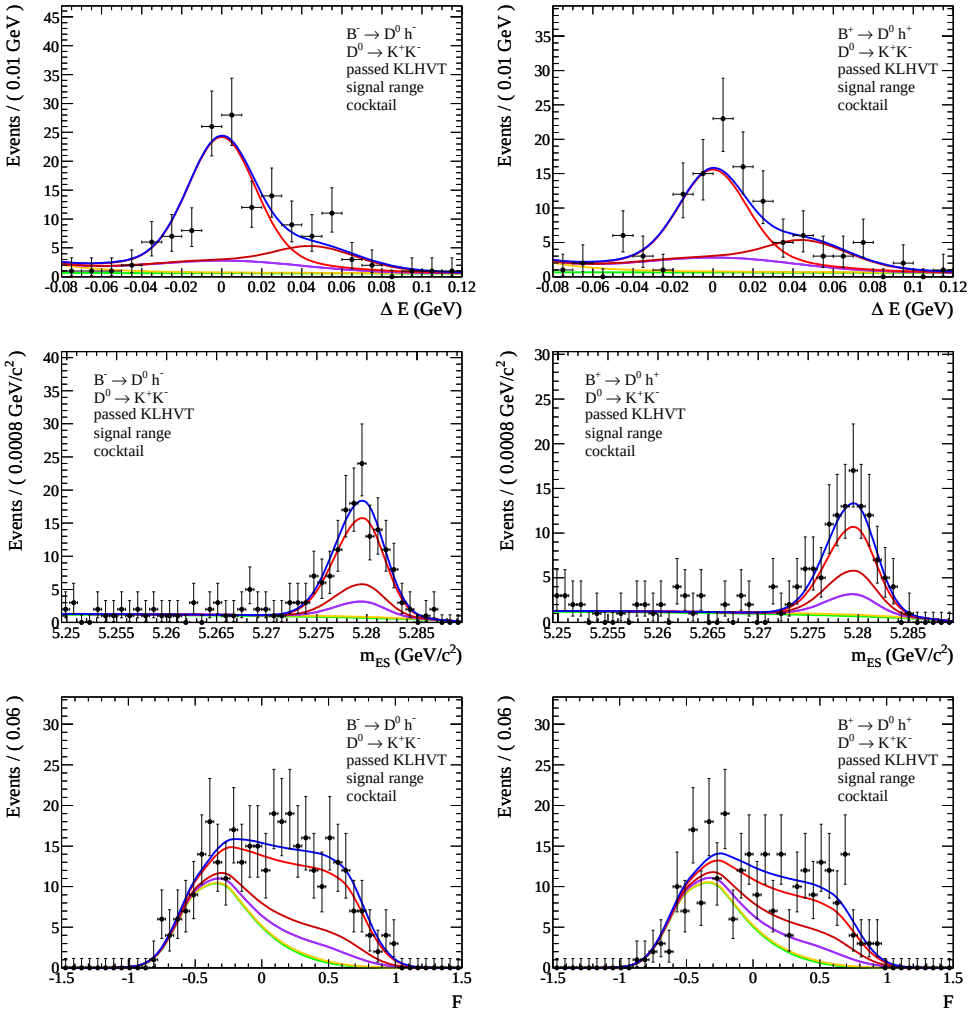

Figure 10.3: *CP*+ fit of cocktail MC, events where the bachelor track passes the `KLHVeryTight` selector, $D^0 \to K^- K^+$ projection. Color code in Table 10.4.

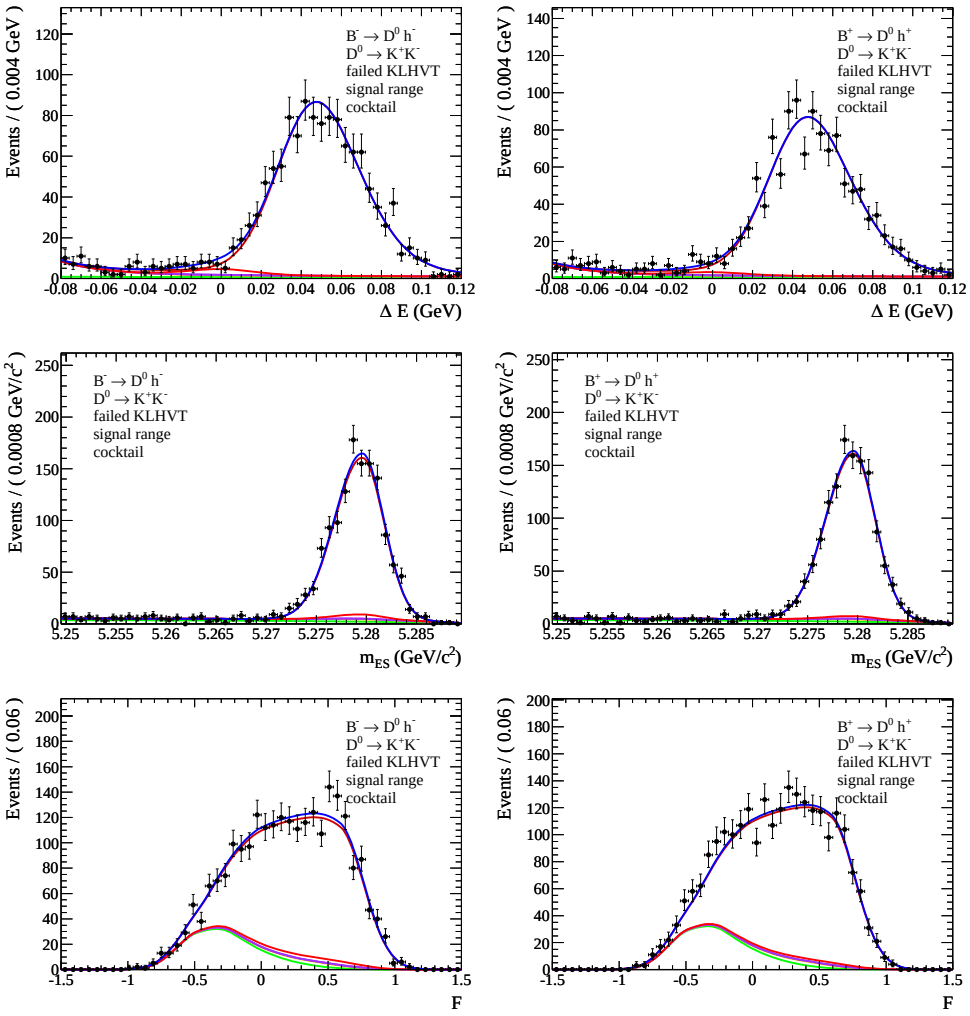

Figure 10.4: *CP*+ fit of cocktail MC, events where the bachelor track fails the `KLHVeryTight` selector, $D^0 \to K^-K^+$ projection. Color code in Table 10.4.

11 Peaking Backgrounds

An irreducible type of peaking background arises from charmless $B^- \to fh^-$ decays, which have the same final states as the $B^- \to D^0(\to f)h^-$ signal. When exploiting the ΔE, $m_{\rm ES}$, and $\mathcal{F}$ variables, this background is therefore indistinguishable from the signal. The according yields are fixed in the final fit. For this reason, the peaking background events are a source of sizable systematic errors, both for the R_{CP} and A_{CP} parameters. The better the knowledge of the peaking backgrounds, the more precise the final result will be. In this chapter we develop a method to estimate the peaking background yields and present the resulting yields.

In the D^0 flavor modes, the presence of significant charmless peaking background can be excluded a priori. In case of the $D^0 \to K^-\pi^+$ mode, for instance, the branching ratio of the according charmless decays are $\mathcal{B}(B^- \to K^-\pi^+K^-) < 1.3 \times 10^{-6}$ at 90 % confidence level, and $\mathcal{B}(B^- \to K^-\pi^+\pi^-) = (5.5 \pm 0.7) \times 10^{-5}$. Taking into account a conservative estimate of the selection efficiency of 5% which is 10% of the signal selection efficiency[1], the expected peaking events are about 20 for $h = K$ and 600 for $h = \pi$. This is negligible compared to more than 3900 $B \to D^0K$ and more than 49000 $B \to D^0\pi$ expected signal events. Table 11.1 summarizes the known branching ratios of charmless B decays potentially peaking in the D^0 flavor channels.

In the *CP* modes, where the signal yields are lower than in the $K^-\pi^+$ case by a factor of greater than ten, the upper limits for $B \to fh$ are at the 10^{-5} level, and with selection efficiencies around 5% we cannot exclude a relevant peaking background contribution a priori. In fact, in the previous Run 1-5 analysis, we found significant peaking components in the $D^0 \to K^-K^+$ and $D^0 \to K^0_S\pi^0$ channels. If we take the previous numbers and scale them up by the ratio of integrated luminosities, $\mathcal{L}_{1-6}/\mathcal{L}_{1-5} = 1.23$, and by the ratio of selection efficiencies $\varepsilon_{1-6}/\varepsilon_{1-5}$, the latter being in the range of 1.5, we get an estimate of what to expect in the present analysis. The scaled yields are shown in Table 11.2.

An independent estimate can be obtained from generic $B\overline{B}$ MC, by counting how many simulated $B \to fh$ decays survive the final selection. These numbers are given in Table 11.3. However, a few of the charmless decays that have been flagged as potentially peaking at reconstruction level do not actually peak in all variables. Table 11.3 shows the potentially peaking yields, not the actually peaking ones. Therefore one shouldn't take the yields obtained from generic $B\overline{B}$ MC too seriously. But nevertheless they mostly agree with the expectation from the Run 1-5 analysis.

To estimate the peaking background yields the basic assumption is, that these peaking backgrounds show the same behavior in both the signal and sideband regions of the D^0 invariant mass. Therefore, one can investigate the sidebands, where the signal is largely suppressed, to get a handle on the peaking background yields. Since this is possible on data, we won't rely on MC simulations too much for this important component of the analysis.

[1] The ratio between the selection efficiency of $B^- \to K^-\pi^+\pi^-$ compared to $B^- \to D^0K^-$ should be given at first order by the fraction of $B^- \to K^-\pi^+\pi^-$ events in the Dalitz plot where $m(K\pi)$ is within the D^0 mass window we are selecting. Fig. 2 in [48] shows this plot. The slice around $m(K\pi)^2 = M(D^0)^2 = 3.5\,\mathrm{GeV}/c^2$ contains much less than 10% of events in the Dalitz plot (those are accumulating in the $m(K\pi)^2 < 3\,\mathrm{GeV}/c^2$ and $m(\pi\pi)^2 < 2\,\mathrm{GeV}/c^2$ regions).

Table 11.1: Branching ratios of potentially peaking charmless final states.

$B^- \to D^0 h$	Branching Ratio [8]	
$K^-\pi^+$, $h = K^-$	$\mathcal{B}(B^- \to K^-\pi^+K^-)$	$< 1.3 \times 10^{-6}$, CL=90%
$K^-\pi^+$, $h = \pi^-$	$\mathcal{B}(B^- \to K^-\pi^+\pi^-)$	$= (5.5 \pm 0.7) \times 10^{-5}$
$K^-\pi^+\pi^0$, $h = K^-$	$\mathcal{B}(B^- \to K^{*-}\pi^+K^-) \cdot 2^{nd}\mathcal{B}$	$< 1.8 \times 10^{-6}$, CL=90%
$K^-\pi^+\pi^0$, $h = K^-$	$\mathcal{B}(B^- \to K^{*-}K^{*0}) \cdot 2^{nd}\mathcal{B}$	$< 6.4 \times 10^{-6}$, CL=90%
$K^-\pi^+\pi^0$, $h = \pi^-$	$\mathcal{B}(B^- \to K^{*-}\pi^+\pi^-) \cdot 2^{nd}\mathcal{B}$	$= (2.3 \pm 0.3) \times 10^{-5}$
$K^-\pi^+\pi^0$, $h = \pi^-$	$\mathcal{B}(B^- \to \omega K^-) \cdot 2^{nd}\mathcal{B}$	$= (6.0 \pm 0.7) \times 10^{-6}$
$K^-\pi^+\pi^+\pi^-$, $h = K^-$	$\mathcal{B}(B^- \to K^{*-}K^{*0}) \cdot 2^{nd}\mathcal{B}$	$< 5 \times 10^{-8}$, CL=90%
$K^-\pi^+\pi^+\pi^-$, $h = \pi^-$	*not seen*	
secondary	$\mathcal{B}(K^{*-} \to K^-\pi^0)$	≈ 0.33
branching ratios	$\mathcal{B}(K^{*0} \to K^+\pi^-)$	≈ 0.67
	$\mathcal{B}(K^* \to K\pi\pi)$	$< 7 \times 10^{-4}$, CL=95%
	$\mathcal{B}(\omega \to \pi^+\pi^-\pi^0)$	≈ 0.89

Table 11.2: Scaled peaking background yields in the $m(D^0)$ signal region, obtained by scaling the numbers we found in the previous analysis of Run 1-5 data [5], which used a different method than the one described here. The negative numbers are reasonably consistent with zero.

D^0 mode	prompt track $K^\pm$	prompt track $\pi^\pm$
K^+K^-	85.8 ± 11.6	-30.5 ± 12.6
$\pi^+\pi^-$	-14.2 ± 10.0	-45.8 ± 13.4
$K^0_S\pi^0$	-37.6 ± 18.8	21.8 ± 35.8
$K^0_S\omega$	8.4 ± 5.9	-3.9 ± 10.5
$K^0_S\phi$	-12.0 ± 5.9	-0.7 ± 1.6

Table 11.3: Expected peaking background yields in the $m(D^0)$ signal region, obtained from generic $B\overline{B}$ MC. These numbers are only a rough estimate for the true peaking background yields (see text).

D^0 mode	prompt track $K^\pm$	prompt track $\pi^\pm$
K^+K^-	89	3
$\pi^+\pi^-$	22	19
$K^0_S\pi^0$	6	47
$K^0_S\omega$	4	7
$K^0_S\phi$	0	1

We will perform a fit of the $m(D^0)$ sidebands in data, and scale the yields by the ratios of the widths of the $m(D^0)$ sideband and signal windows. The sideband definitions, together with the resulting scale factors, are given in Table 11.4. Figure 11.1 shows the $m(D^0)$ distributions in generic $B\overline{B}$ MC, together with the sideband definitions.

The fit strategy we are using for the $m(D^0)$ sidebands closely resembles the final fit described in Section 9. We briefly summarize the key features: we fit three variables, (ΔE, $m_{\rm ES}$, $\mathcal{F}$), simultaneously in four slices (positively/negatively charged, passed/failed the `KLHVeryTight` selector). The PDF consists of six components: signal(K) leaking into the $m(D^0)$ sideband, leaking signal(π), $q\overline{q}$, $B\overline{B}$, peaking(K), and peaking(π), where K and π refer to $B \to D^0K$ and $B \to D^0\pi$ candidates respectively.

It is important to note that, when exploring the D^0 sidebands, one needs to calculate ΔE without constraining the D^0 mass to its PDG value: $\Delta E_{\tt unc}$. The advantage of using $\Delta E_{\tt unc}$

Table 11.4: D^0 mass sideband definitions, the scale factor defined as the ratio of the widths of the D^0 mass signal and sideband regions.

D^0 mode	$m(D^0)$ sideband region (MeV/c^2)	scale factor
K^+K^-	$[1794.5-1834.5]$, $[1884.5-1914.5]$	0.43
$\pi^+\pi^-$	$[1814.5-1839.5]$, $[1889.5-1934.5]$	0.48
$K^0_S\pi^0$	$[1774.5-1804.5]$, $[1924.5-1954.5]$	1.67
$K^0_S\omega$	$[1794.5-1829.5]$, $[1899.5-1934.5]$	0.69
$K^0_S\phi$	$[1794.5-1834.5]$, $[1894.5-1934.5]$	0.28

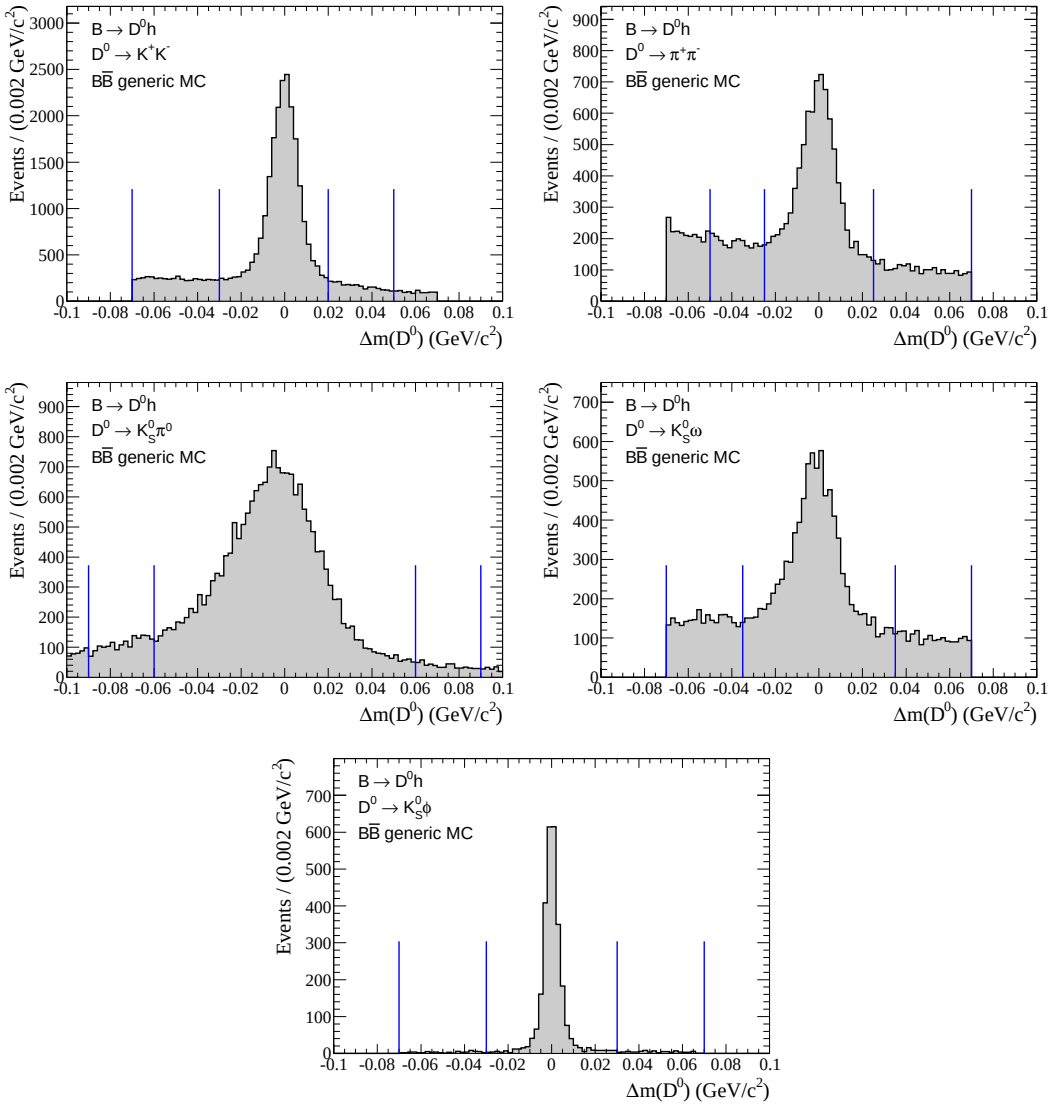

Figure 11.1: D^0 mass sideband definitions in generic $B\overline{B}$ MC.

over ΔE is that the distribution of charmless B candidates is not distorted. On the other hand, however, the distribution of signal events leaking into the sidebands will be shifted to higher or lower $\Delta E_{\mathtt{unc}}$ values, depending on which sideband, upper or lower, the candidate originates from. Figures 11.2-11.4 show the distributions of the three variables $\Delta E_{\mathtt{unc}}$, m_{ES}, $\mathcal{F}$, in cocktail MC and on-peak data. As already shown in Chapter 8, the overall normalization of the MC doesn't match the data very well, which is visible in the sidebands, too. We account for this by floating all yields other than the $B\overline{B}$ yields.

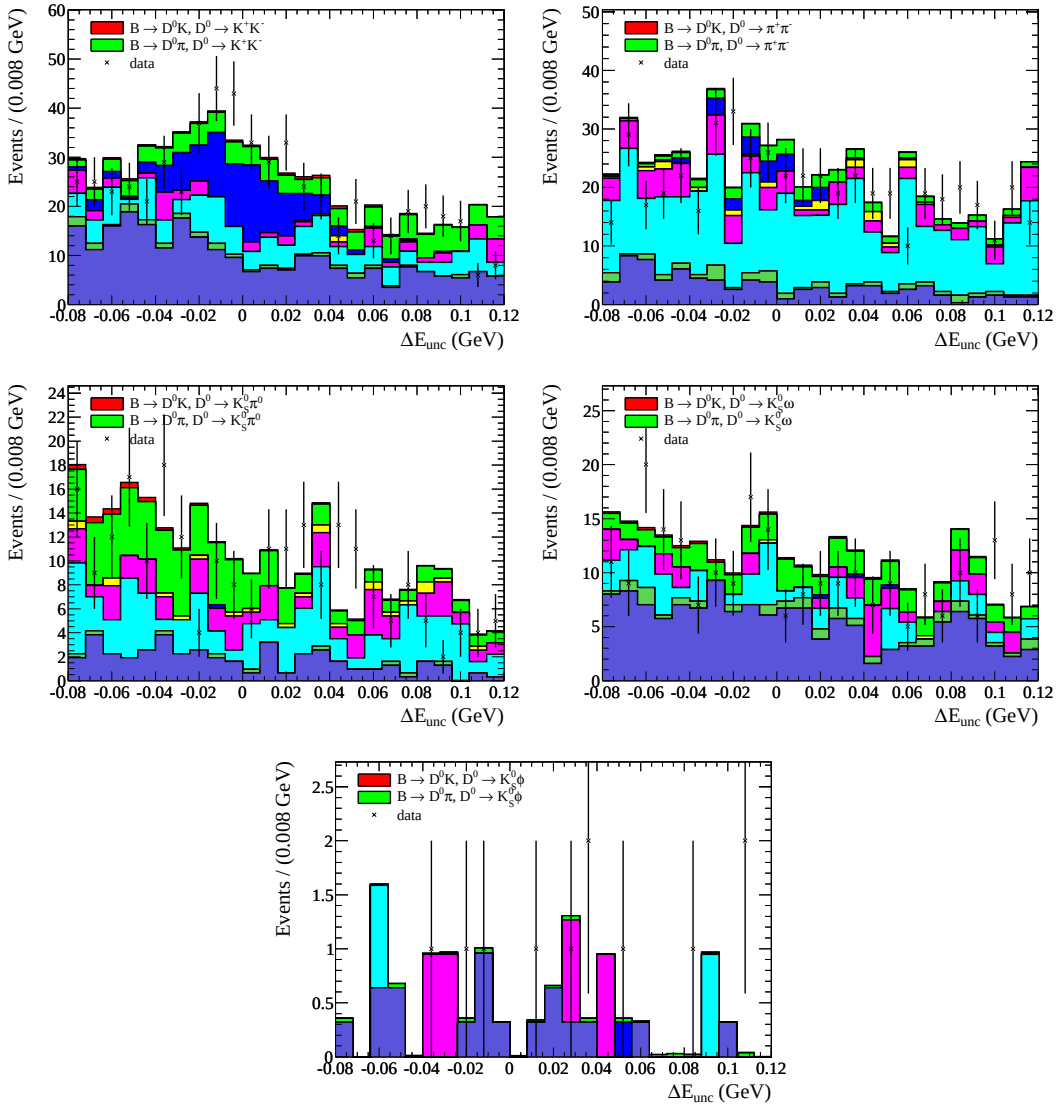

Figure 11.2: $m(D^0)$ sideband distribution of $\Delta E_{\mathtt{unc}}$ in both MC and data. The color code of the backgrounds is the same as in Figure 8.1 on page 46.

The following section describes the shapes of the PDFs of the six components introduced above. The $q\overline{q}$ component is by far the most abundant one in the sidebands, and we can afford to float all according parameters. We'll fix the remaining shape parameters, however, to values obtained from MC simulation. As for the yields, we will float the $q\overline{q}$ yields and

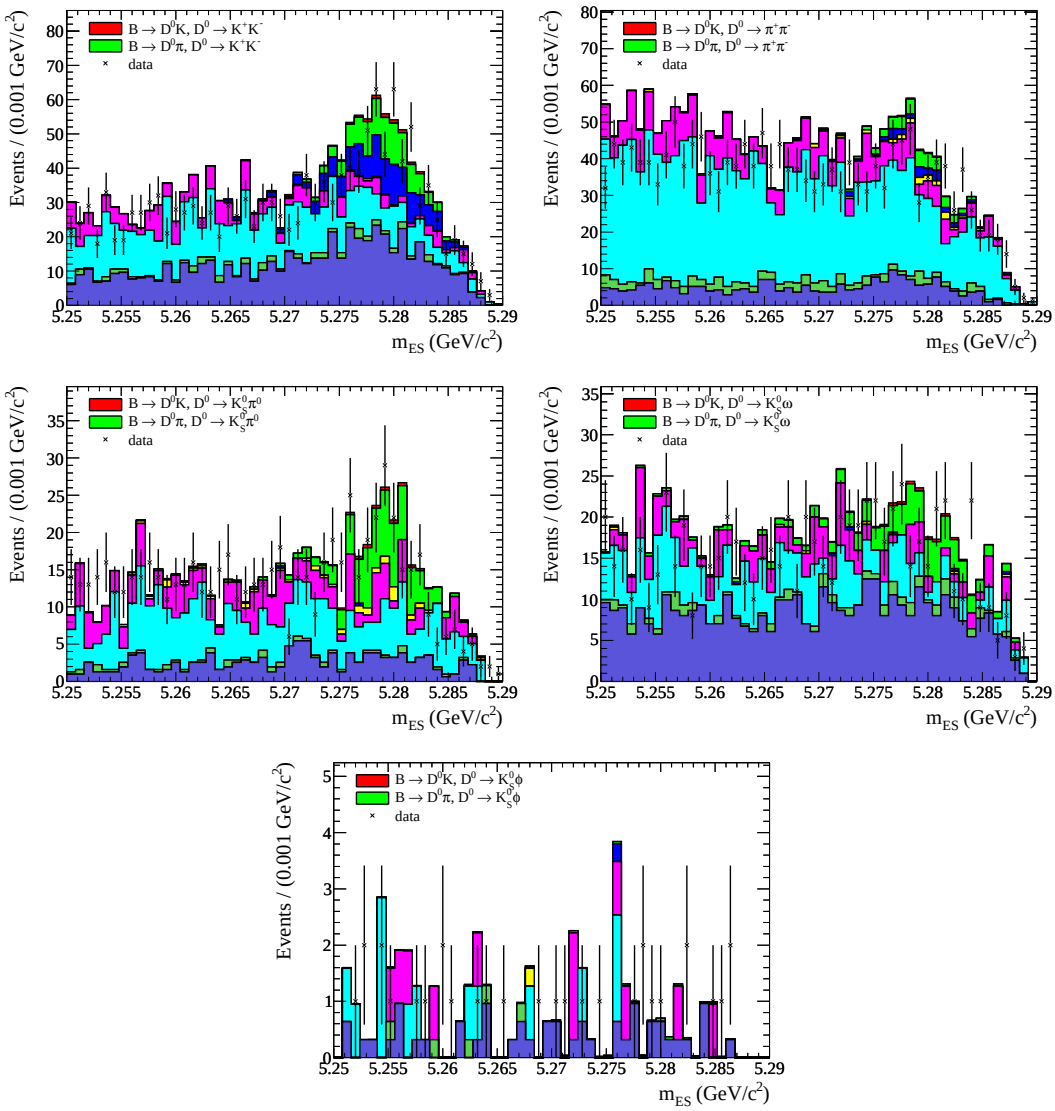

Figure 11.3: $m(D^0)$ sideband distribution of m_{ES} in both MC and data. The color code of the backgrounds is the same as in Figure 8.1 on page 46.

the leaking $B \to D^0\pi$ signal yields; while the ratio of $B \to D^0\pi$ and $B \to D^0K$ signal will be fixed to $R_{K/\pi} = 0.083$ [8]. The $B\overline{B}$ yields will be fixed to their expectation from generic $B\overline{B}$ MC. All yields have been expressed through their charge asymmetries A_{CP}. Except for the $q\overline{q}$ component, we fix A_{CP} to zero. We'll also fix the efficiency and mis-id parameters of the `KLHVeryTight` selector, which enter through the yield definitions of the slices (see Equation 9.11 on page 52), to values obtained from MC: $\varepsilon = 75\%$, $m = 1.8\%$. Table 11.5 lists the expected $m(D^0)$ sideband yields for all of the six components. As already mentioned, not all events considered peaking do actually peak. We'll therefore add the according yields to the regular $B\overline{B}$ background. That's why Table 11.5 seemingly disagrees with Table 11.3.

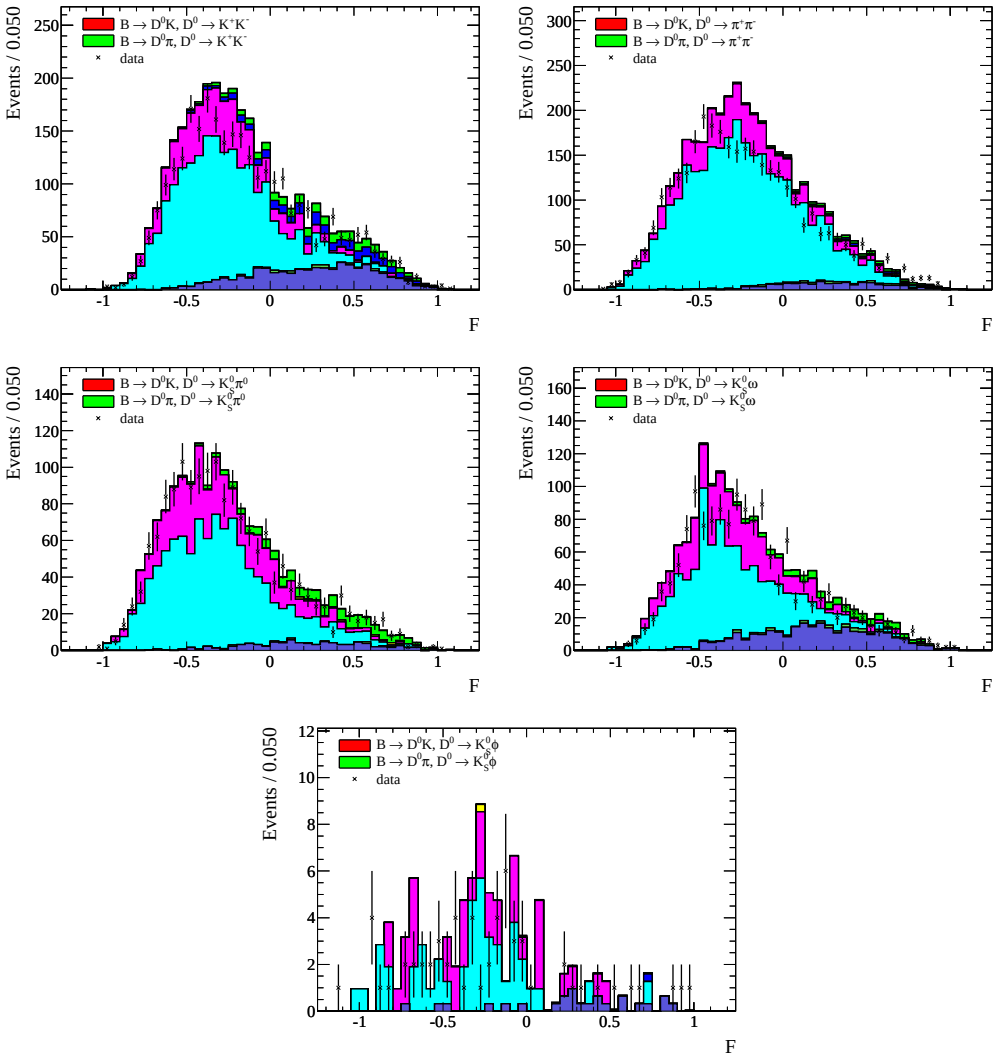

Figure 11.4: $m(D^0)$ sideband distribution of $\mathcal{F}$ in both MC and data. The color code of the backgrounds is the same as in Figure 8.1 on page 46.

Table 11.5: Expected $m(D^0)$ sideband yields. Peaking background yields and $B\overline{B}$ yields are obtained from generic $B\overline{B}$ MC, leaking signal yields from signal MC. The $B\overline{B}$ yields are fixed to these values in the sideband fit.

Parameter	KK	$\pi\pi$	$K^0_S\pi^0$	$K^0_S\omega$	$K^0_S\phi$
$N^{\mathrm{pk}(K)}_{\mathrm{tot}}$	233	41	1	6	1
$N^{\mathrm{pk}(\pi)}_{\mathrm{tot}}$	5	0	0	4	1
$N^{q\overline{q}}_{p}$	4972	5282	3624	3468	216
$N^{q\overline{q}}_{f}$	19186	28630	11040	9910	476
$N^{B\overline{B}}_{p}$	157	280	46	125	5
$N^{B\overline{B}}_{f}$	1791	988	430	1468	43
$N^{\mathrm{sig}(K)}_{\mathrm{tot}}$	16	5	9	10	0
$N^{\mathrm{sig}(\pi)}_{\mathrm{tot}}$	187	60	153	135	2

11.1 D^0 Mass Sideband PDF

The shape of the sideband PDF is very similar to the shape of the PDF used to describe the signal region. We'll therefore use similar parameterizations for the six components: charmless peaking background $B \to fh$ ($h = K, \pi$), leaking $B \to D^0 h$ signal ($h = K, \pi$), $q\overline{q}$, and $B\overline{B}$.

To parameterize the $B \to fh$ shape in the sideband, we will actually re-use the final fit signal parameterizations for $m_{\rm ES}$ and $\mathcal{F}$. This is preferable even though the $m_{\rm ES}$ shape doesn't describe the $B \to fh$ events contained in generic $B\overline{B}$ MC too well, because eventually we need to estimate how many background events will get picked up by the signal PDF. For $\Delta E_{\tt unc}$, however, we cannot re-use ΔE signal shape to describe $B \to fh$ events, because without the mass constraint on the D^0 mass the shapes are distorted. Instead we'll rely on the $\Delta E_{\tt unc}$ distribution of the $B \to fh$ events contained in generic $B\overline{B}$ MC. Because of the lack of MC statistics of the peaking background in channels other than $D^0 \to K^-K^+$ and $D^0 \to \pi^+\pi^-$ we'll re-use the shapes obtained in these modes for the lower statistics channels.

The leaking signal component in $\Delta E_{\tt unc}$ develops a double peak structure, as mentioned in the previous section. We will parameterize this by two Gaussians and a linear shape.

All remaining shapes follow closely the parameterization of the signal band. The functional form of each component of the $D^0 \to K^-K^+$ channel is summarized in Table 11.6, as well as how the shape parameters are obtained. The according extraction fits are shown in Figure 11.5. The situation of the remaining modes is fairly similar [42].

Table 11.6: Functional form of the PDF used to fit the D^0 mass sidebands, $D^0 \to K^-K^+$ channel. Abbreviations: G–Gaussian, P–first order polynom, D–double, A–Argus [46], CB–Crystal Ball [44], BF–bifurcated. As for the source of the fixed parameters: SB–fit to sideband MC, FF–same shape as in the final fit.

component	$\Delta E_{\tt unc}$	$m_{\rm ES}$	$\mathcal{F}$
leaking signal(K)	DG+P (SB)	same as lk. sig(π)	FF sig(K)
leaking signal(π)	DG+P (SB)	DG (SB)	FF sig(π)
$q\overline{q}$	P (floating)	A (floating)	DBFG (floating)
$B\overline{B}$	P (SB)	A+CB (SB)	DG (SB)
peaking(K)	DG (SB)	FF sig(K)	FF sig(K)
peaking(π)	G (SB)	FF sig(π)	FF sig(π)

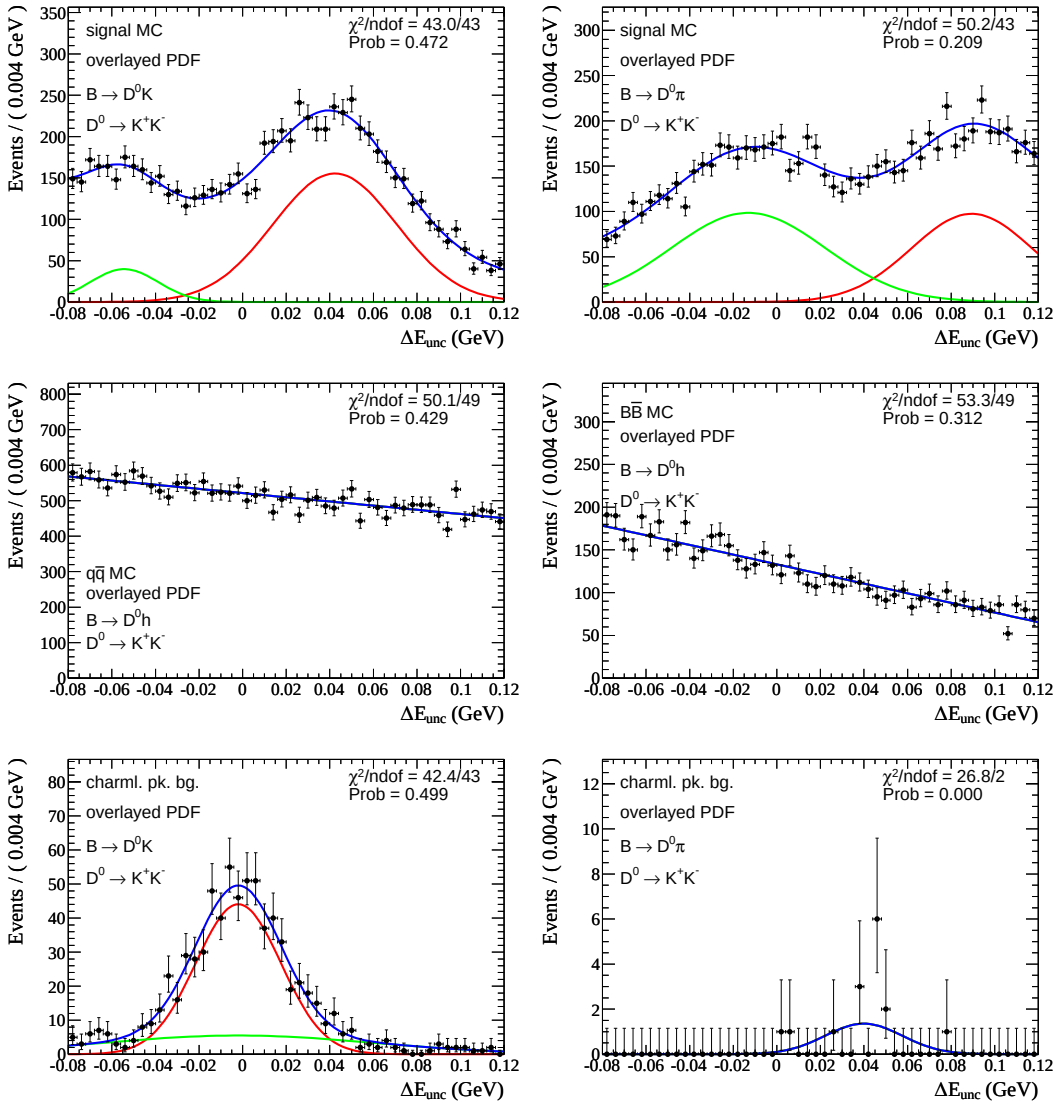

Figure 11.5: $\Delta E_{\rm unc}$ D^0 sideband PDFs, $D^0 \to K^-K^+$.

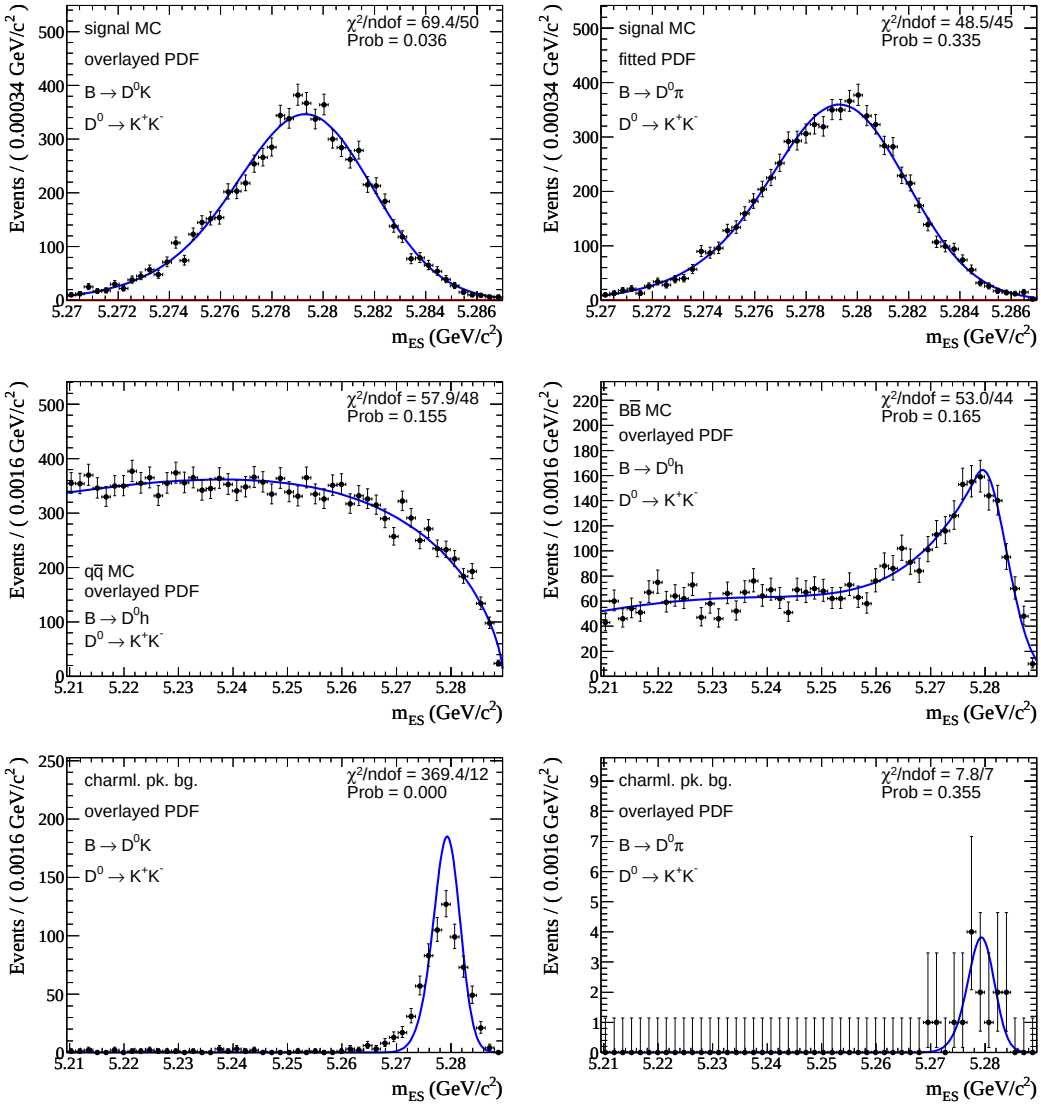

Figure 11.6: $m_{\rm ES}$ D^0 sideband PDFs, $D^0 \to K^-K^+$. The bottom plots overlay the $m_{\rm ES}$ signal shape used to fit the signal band (see text).

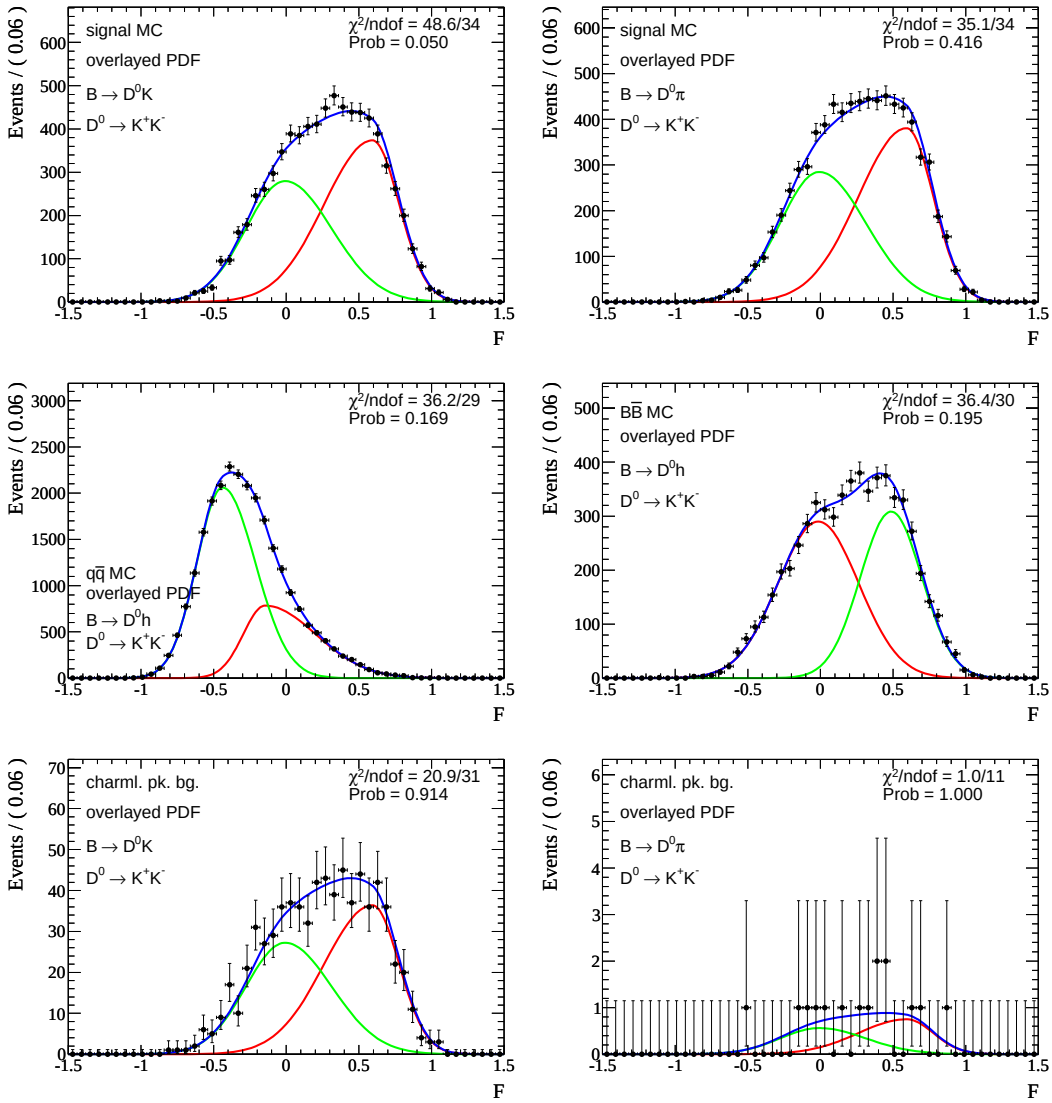

Figure 11.7: $\mathcal{F}$ D^0 sideband PDFs, $D^0 \to K^-K^+$.

11.2 Sideband Fit Validation

We validate the strategy to extract peaking background estimates from the D^0 mass sidebands on cocktail MC. Table 11.7 shows the fit results on the $D^0 \to K^-K^+$ and $D^0 \to K^0_S\pi^0$ channels, where we saw the largest contributions in the Run 1-5 analysis. Figure 11.8 shows the corresponding slice-combined plots. All reported yields agree with their expectation (see Table 11.5).

Table 11.7: Fit results on D^0 mass sidebands in cocktail MC, $D^0 \to K^-K^+$ and $D^0 \to K^0_S\pi^0$ channels. The slope $a^{q\overline{q}}_{\Delta E}$ is in units of GeV^{-1}. The parameters of the Argus function [46], $c_{m_{\mathrm{ES}}}$ and $p_{m_{\mathrm{ES}}}$, are in units of GeV/c^2 and 1, respectively.

Parameter	KK	$K^0_S\pi^0$
$A^{q\overline{q}}_{CP,p}$	0.003 ± 0.014	0.002 ± 0.017
$A^{q\overline{q}}_{CP,f}$	0.0020 ± 0.0075	0.0007 ± 0.0097
$f^{q\overline{q}}_{\mathcal{F}}$	0.311 ± 0.035	0.73 ± 0.12
$\mu^{q\overline{q}}_{\mathcal{F},l,1}$	-0.136 ± 0.032	-0.423 ± 0.036
$\mu^{q\overline{q}}_{\mathcal{F},l,2}$	-0.437 ± 0.013	-0.695 ± 0.038
$\sigma^{q\overline{q}}_{\mathcal{F},l,1}$	0.157 ± 0.026	0.167 ± 0.011
$\sigma^{q\overline{q}}_{\mathcal{F},l,2}$	0.1776 ± 0.0048	0.108 ± 0.014
$\sigma^{q\overline{q}}_{\mathcal{F},r,1}$	0.349 ± 0.011	0.411 ± 0.012
$\sigma^{q\overline{q}}_{\mathcal{F},r,2}$	0.229 ± 0.013	0.311 ± 0.052
$a^{q\overline{q}}_{\Delta E}$	-1.14 ± 0.11	-1.50 ± 0.13
$c_{m_{\mathrm{ES}}}$	-30.4 ± 2.1	-26.6 ± 2.6
$p_{m_{\mathrm{ES}}}$	0.617 ± 0.025	0.573 ± 0.030
$N^{\mathrm{pk}(K)}_{\mathrm{tot}}$	177 ± 22	0.7 ± 7.8
$N^{\mathrm{pk}(\pi)}_{\mathrm{tot}}$	-8 ± 20	3 ± 12
$N^{q\overline{q}}_{p}$	5010 ± 72	3620 ± 61
$N^{q\overline{q}}_{f}$	19199 ± 146	10980 ± 108
$N^{\mathrm{sig}(\pi)}_{\mathrm{tot}}$	215 ± 40	203 ± 25

11.3 Sideband Fit to Data

The fit results of the fit to on-peak data in the $D^0 \to K^-K^+$ and $D^0 \to K^0_S\pi^0$ channels are shown in Table 11.8 and in Figure 11.9. The peaking background yields reported from the fits have to be scaled to account for the different widths of the signal region and the sidebands. We take the ratio of the widths from Table 11.4 and show the scaled yields in Table 11.9. These yields are used in the final fit. The yields we find are in good agreement with the expectations based on the previous estimate of the Run 1-5 analysis (Table 11.2) and on generic $B\overline{B}$ MC (Table 11.3).

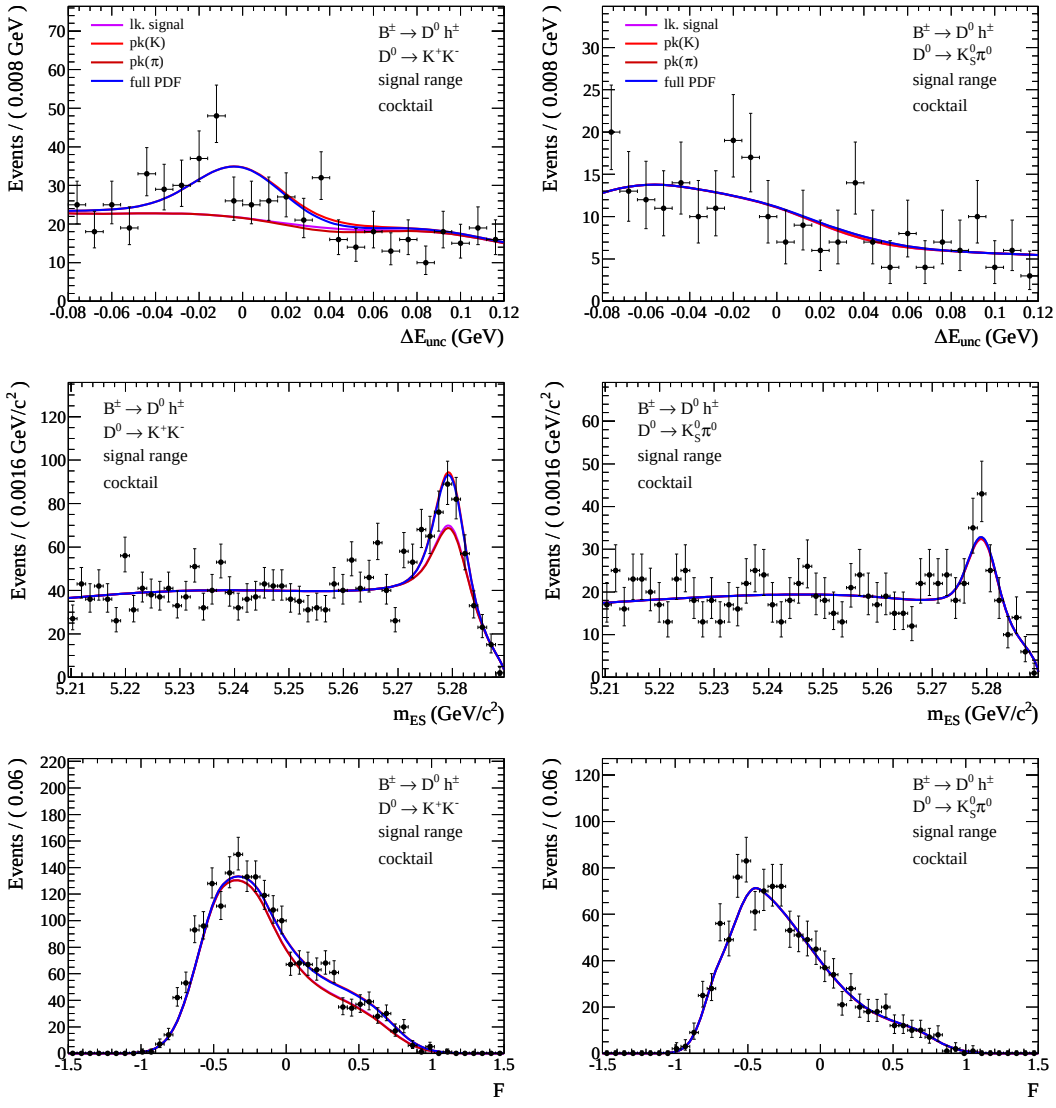

Figure 11.8: Fit of the D^0 sideband on cocktail MC, $D^0 \to K^-K^+$ and $D^0 \to K^0_S\pi^0$ channels. Depicted is the sum of all slices in the signal region of the variables not plotted. Color code: light red–peaking(K), dark red–peaking(π), violet–leaking signal, blue–full PDF. The peaking components include leaking signal. All components include $q\overline{q}$ and $B\overline{B}$ background.

Table 11.8: Fit results on the $m(D^0)$ sidebands in on-peak data, $D^0 \to K^-K^+$ and $D^0 \to K^0_S\pi^0$ channels. The slope $a^{q\bar{q}}_{\Delta E}$ is in units of GeV^{-1}. The parameters of the Argus function [46], $c_{m_{\mathrm{ES}}}$ and $p_{m_{\mathrm{ES}}}$, are in units of GeV/c^2 and 1, respectively.

Parameter	KK	$K^0_S\pi^0$
$A^{q\bar{q}}_{CP,p}$	0.002 ± 0.015	-0.030 ± 0.016
$A^{q\bar{q}}_{CP,f}$	-0.0042 ± 0.0084	-0.014 ± 0.010
$f^{q\bar{q}}_{\mathcal{F}}$	0.293 ± 0.028	0.848 ± 0.027
$\mu^{q\bar{q}}_{\mathcal{F},l,1}$	-0.136 ± 0.026	-0.4908 ± 0.0089
$\mu^{q\bar{q}}_{\mathcal{F},l,2}$	-0.428 ± 0.012	-0.830 ± 0.014
$\sigma^{q\bar{q}}_{\mathcal{F},l,1}$	0.137 ± 0.027	0.1611 ± 0.0058
$\sigma^{q\bar{q}}_{\mathcal{F},l,2}$	0.1829 ± 0.0047	0.0426 ± 0.0097
$\sigma^{q\bar{q}}_{\mathcal{F},r,1}$	0.3517 ± 0.0098	0.4249 ± 0.0074
$\sigma^{q\bar{q}}_{\mathcal{F},r,2}$	0.216 ± 0.012	0.513 ± 0.060
$a^{q\bar{q}}_{\Delta E}$	-1.24 ± 0.12	-1.53 ± 0.14
$c_{m_{\mathrm{ES}}}$	-21.66 ± 0.94	-24.8 ± 2.7
$p_{m_{\mathrm{ES}}}$	0.522610 ± 0.000038	0.553 ± 0.030
$N^{\mathrm{pk}(K)}_{\mathrm{tot}}$	216 ± 23	-2.4 ± 5.3
$N^{\mathrm{pk}(\pi)}_{\mathrm{tot}}$	-11 ± 19	39 ± 14
$N^{q\bar{q}}_{p}$	4651 ± 70	3847 ± 62
$N^{q\bar{q}}_{f}$	15424 ± 131	10275 ± 104
$N^{\mathrm{sig}(\pi)}_{\mathrm{tot}}$	126 ± 36	137 ± 22

Table 11.9: Result of the peaking background estimate. These numbers will be used in the final fit.

D^0 mode	prompt track $K^\pm$	prompt track $\pi^\pm$
K^+K^-	93 ± 10	-5 ± 8
$\pi^+\pi^-$	4 ± 6	0 ± 9
$K^0_S\pi^0$	-4 ± 9	65 ± 23
$K^0_S\omega$	3 ± 6	0 ± 8
$K^0_S\phi$	0.5 ± 0.7	1.4 ± 1.0

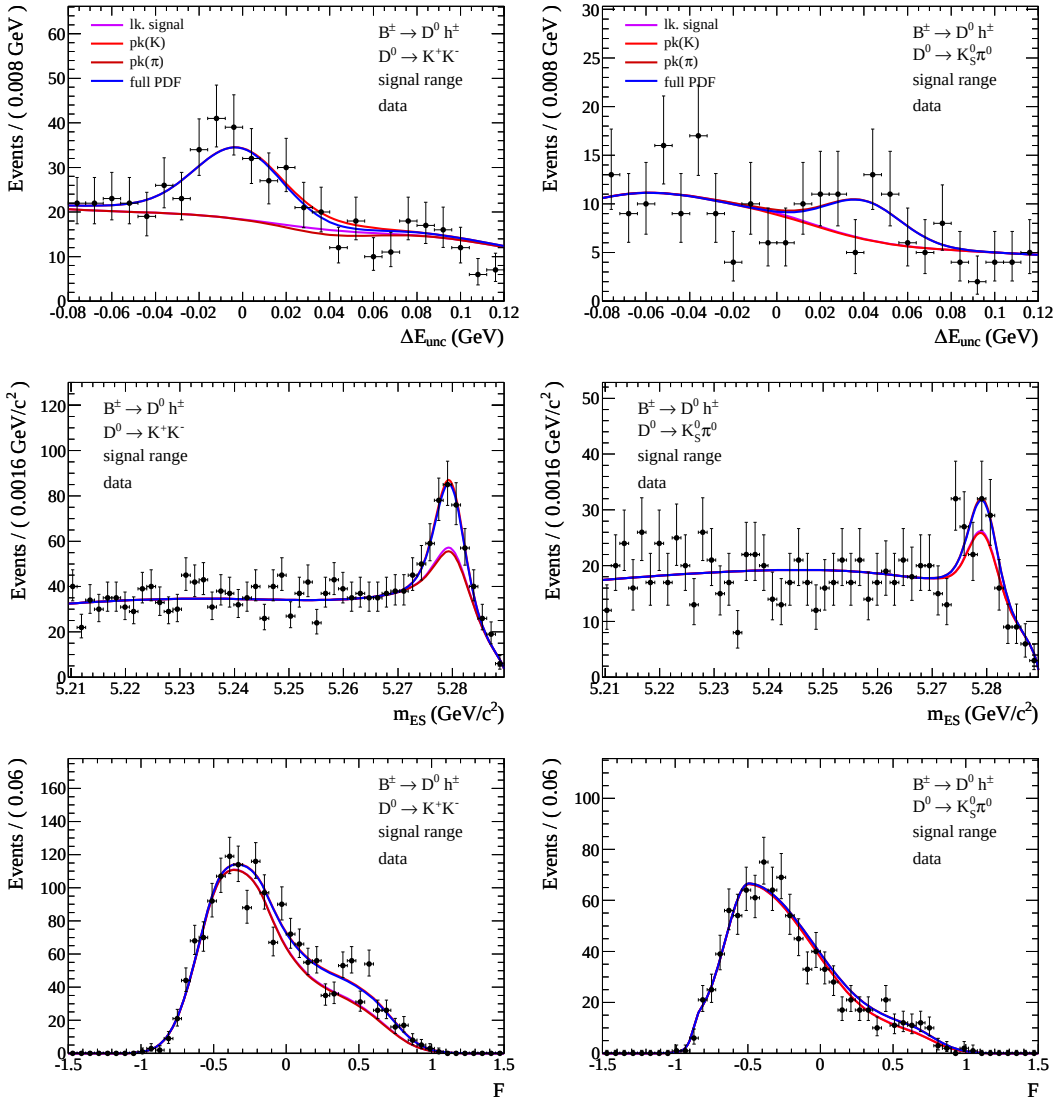

Figure 11.9: Fit of the D^0 sideband in on-peak data, $D^0 \to K^-K^+$ and $D^0 \to K^0_S\pi^0$ channels. Depicted is the sum of all slices in the signal region of the variables not plotted. Color code: light red–peaking(K), dark red–peaking(π), violet–leaking signal, blue–full PDF. The peaking components include leaking signal. All components include $q\overline{q}$ and $B\overline{B}$ background.

12 Systematic Uncertainties

In this section we consider eight sources of systematic uncertainty possibly affecting the GLW parameters $A_{CP\pm}$ and $R_{CP\pm}$. Their contributions are summarized in Table 12.1. The total systematic error has been calculated by taking into account the correlations between the $A_{CP\pm}$ and $R_{CP\pm}$ parameters for each source. We calculate the systematic covariance matrix $V_{(\mathrm{syst})}[\vec{y}]$, $\vec{y} = (A_{CP+}, A_{CP-}, R_{CP+}, R_{CP-})^T$, as the sum of the covariance matrices due to each of the eight sources. The resulting matrix is

$$V_{(\mathrm{syst})}[\vec{y}] = \begin{pmatrix} 9.4\times10^{-4} & 6.8\times10^{-6} & -5.6\times10^{-5} & 0.0 \\ 6.8\times10^{-6} & 8.5\times10^{-5} & 0.0 & 5.1\times10^{-6} \\ -5.6\times10^{-5} & 0.0 & 1.4\times10^{-3} & 3.9\times10^{-5} \\ 0.0 & 5.1\times10^{-6} & 3.9\times10^{-5} & 9.3\times10^{-4} \end{pmatrix}. \tag{12.1}$$

Table 12.1: Summary of systematic errors. The total systematic error is the square root of the diagonal elements of $V_{(\mathrm{syst})}$ of Eq.12.1.

source	A_{CP+}	A_{CP-}	R_{CP+}	R_{CP-}
fixed fit parameters	0.0039	0.0062	0.028	0.024
peaking background	0.030	0.0043	0.018	0.012
bias correction	0.0038	0.0043	0.0055	0.0048
signal self cross-feed	0.0002	0.0010	-	-
detector charge asym.	0.0026	0.0026	-	-
S-wave contribution to $CP-$ states	-	0.0004	-	0.0029
$\varepsilon(\pi)/\varepsilon(K)$	-	-	0.018	0.017
$R_{CP\pm}$ vs. $R_{\pm}$	-	-	0.015	0.014
Total	0.031	0.0092	0.037	0.030

12.1 Parameterization of the PDF

The final PDF we use to describe the data is only an approximation to the true PDF, chosen by nature, and laying at heart of the on-peak dataset. Therefore it is crucial to estimate the quality of the approximation. A good way to do this is to vary all parameters which are fixed in the final fit, by a reasonable amount. In this work, the parameters of the component PDFs are obtained from extraction fits to appropriate MC samples (Section 9). Thus their standard deviations are known, and also the correlations among the parameters belonging to a certain component PDF. To estimate the systematic error on the final result, we perform a large number of test fits on on-peak data, similar to the final fit. In each of these test fits the fixed parameters of the final PDF are varied according to their covariance matrices, correctly taking into account (first order) correlations. By diagonalizing the covariance matrix V of the parameters a_i one obtains the rotation matrix which transforms $\vec{a}$ to the uncorrelated parameter space, $\vec{a}'$. In the uncorrelated parameter space, values for the a'_i are drawn from

Gaussian functions whose means and widths are given by a'_i and $\sigma'_i = \sqrt{V'_{ii}}$. Using again the rotation matrix one transforms the generated $\vec{a}'$ back to the correlated space.[1]

The widths of the resulting distributions for the floating fit parameters are taken to be the systematic uncertainties due to the parameterization of the final PDF. Figure 12.1 shows the distributions of the crucial fit parameters $A_{CP\pm}$ and $R_{K/\pi}$ of the *CP*+ and *CP*– fits, after the nominal values have been subtracted. The distributions of the flavor fits are shown in Figure B.1 in Appendix B. Some of the distributions are poorly described by a single Gaussian. There is usually a prominent narrow peak, but also a wider continuum component. We therefore fit the distributions to a double Gaussian. There can be many reasons for this behavior. One possibility is a nearby second minimum in the likelihood function. In that case, only little variations of the fixed parameters may be sufficient to make the fit converge to the second minimum. Hence the double Gaussian structure. We define a naive total width $\sigma_{\rm tot}$ of the double Gaussian shape $f\,G(\mu_1,\sigma_1) + (1-f)\,G(\mu_2,\sigma_2)$,

$$\sigma_{\rm tot} = f\,\sigma_1 + (1-f)\,\sigma_2 + g\,|\mu_1 - \mu_2|, \tag{12.2}$$

where $g = 2\,f$ for $f \leq 0.5$, and $g = 2\,(1-f)$ otherwise[2]. The resulting uncertainties of each parameter which enters the final result are given in Table 12.2.

The covariance matrix $V_{\rm PDF}[\vec{y}]$ for the parameters $\vec{y} = (A_{CP+}, A_{CP-}, R_{CP+}, R_{CP-})^T$ is obtained in the following way. We first build the covariance matrix $V_{\rm PDF}[\vec{x}]$ for the parameters $\vec{x} = (A_{CP+}, A_{CP-}, R^+_{K/\pi}, R^-_{K/\pi}, R^{\rm flv}_{K/\pi})^T$, where $R^{\rm flv}_{K/\pi}$ is the weighted average of the $R_{K/\pi}$ values from the flavor fits (see next Chapter 13). We take the non-trivial correlations between A_{CP+} and $R^+_{K/\pi}$ (A_{CP-} and $R^-_{K/\pi}$) to be the same as in the nominal fit to data. Then we transform (see for instance [41])

$$V_{\rm PDF}[\vec{y}] = B\,V_{\rm PDF}[\vec{x}]\,B^T, \tag{12.3}$$

with B denoting the matrix of derivatives $B_{ij} = \partial y_i/\partial x_j$, calculated from Eq. 3.16. The square root of the diagonal elements of $V_{\rm PDF}[\vec{y}]$ is reported in Table 12.1.

Table 12.2: Systematic errors on the crucial parameters due to the fixed shape of the final PDF, as obtained from a series of test fits varying all fixed parameters according to their covariance matrices.

fit	$\sigma_{\rm tot}(A_{CP})$	$\sigma_{\rm tot}(R_{K/\pi})$
CP+	0.0039	0.0017
CP–	0.0062	0.0014
$K\pi$	0.0013	0.00092
$K\pi\pi^0$	0.0020	0.0011
$K\pi\pi\pi$	0.0021	0.0013

12.2 Peaking Backgrounds

The uncertainties in the numbers of peaking background events estimated from the D^0 mass sidebands introduce a systematic uncertainty in the $B \to D^0_{CP\pm}h$ yields and therefore in

[1] This procedure neglects correlations among parameters of separate components.

[2] It was checked the square root of the 2$^{\rm nd}$ moment of the double Gaussian shape yields comparable values for the width in all cases.

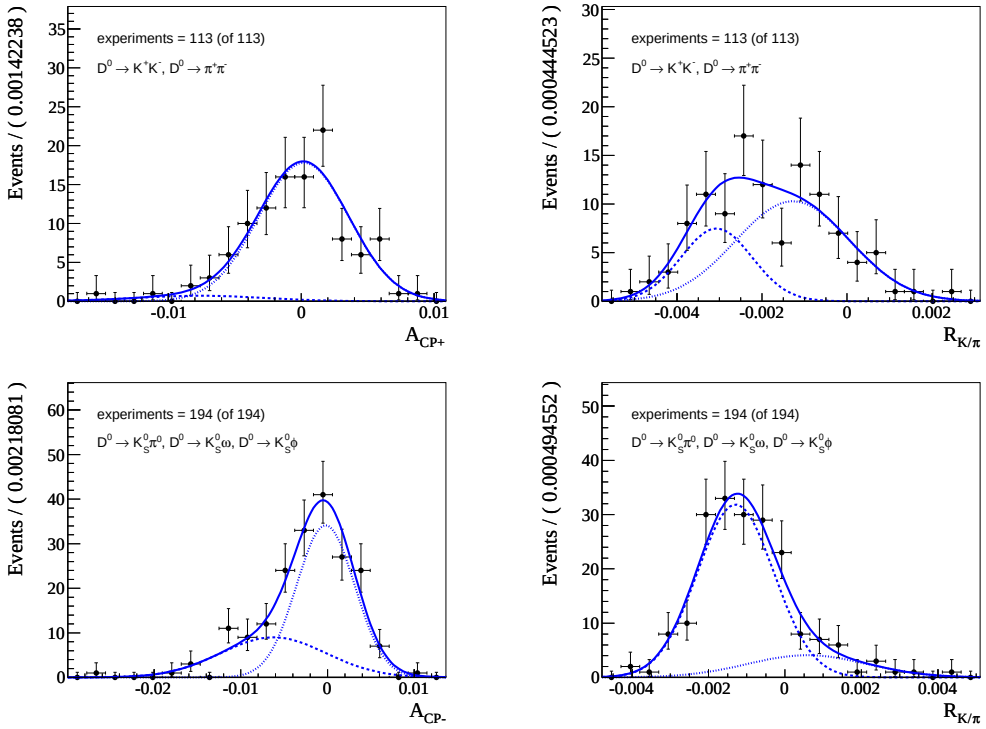

Figure 12.1: Distributions of the crucial parameters as obtained from a series of systematic fits, varying all fixed parameters according to their covariance matrices. The nominal value has been subtracted. Top row *CP*+, bottom row *CP*−. The shape is the sum of two Gaussians (dashed and dotted lines).

$R_{CP\pm}$. As for the asymmetries, any charge asymmetry in the peaking background would directly affect the measured values of $A_{CP\pm}$.

We perform a series of test fits on on-peak data floating the $B \to fh$ yields and the $B \to fh$ asymmetries by their uncertainties. While the errors on the yields are obtained from the sideband study (Table 11.9), the errors on the asymmetries could not be obtained that way because the sideband samples don't contain sufficient statistics. Here we assign an uncertainty of $\pm 20\,\%$. For comparison, the measured value of A_{CP} in $B^+ \to K^+K^-\pi^+$ is $A_{CP} = 0.00 \pm 0.10$, in $B^+ \to K^+K^-K^+$ it is $A_{CP} = -0.017 \pm 0.030$ (both values taken from [8]). Despite the fact A_{CP} depends on which part of the Dalitz plot is selected, $\pm 20\%$ is a conservative estimate that is consistent with not having observed *CP* violation in those decays.

We take the total width (Eq. 12.2) of the resulting distributions to be the systematic uncertainty. The uncertainties of each parameter which enters the final result are given in Table 12.3. Figure 12.2 shows the distributions of the A_{CP} and $R_{K/\pi}$ parameters as obtained from the *CP* test fits. We construct the covariance matrices $V_{\mathrm{PKBG}}[\vec{x}]$ and $V_{\mathrm{PKBG}}[\vec{y}]$ in full analogy to the previous section. The square root of the diagonal elements of $V_{\mathrm{PKBG}}[\vec{y}]$ is reported in Table 12.1.

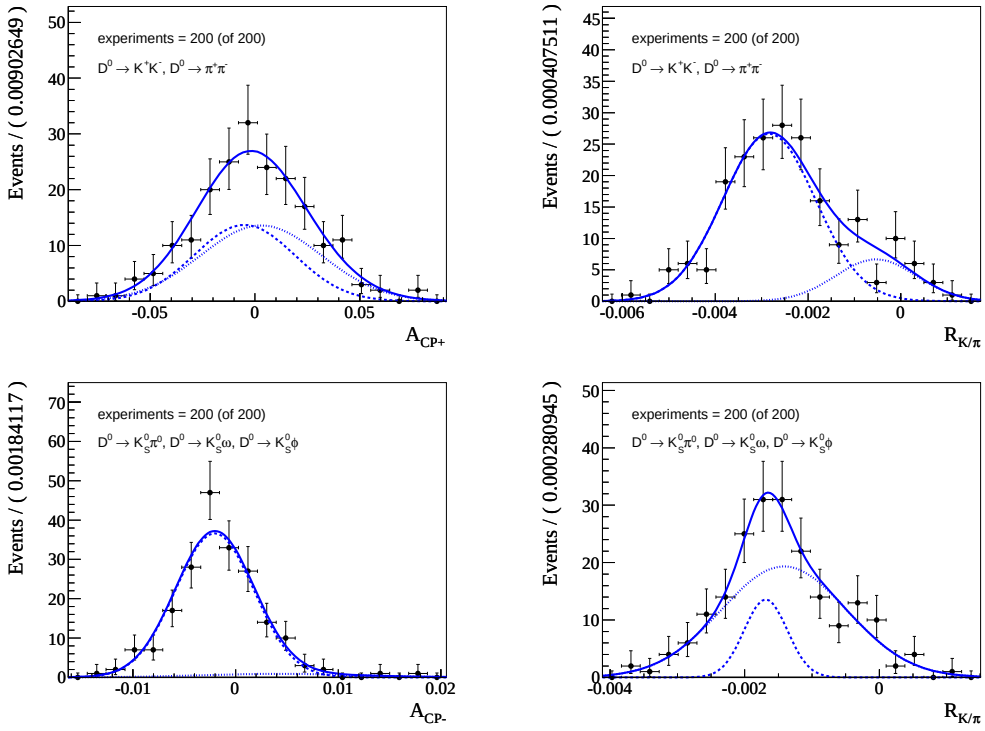

Figure 12.2: Distributions of the crucial parameters as obtained from a series of systematic fits, where the fixed peaking background parameters have been varied according to their uncertainties. The nominal value has been subtracted. Top row *CP*+, bottom row *CP*−. The shape is the sum of two Gaussians (dashed and dotted lines).

Table 12.3: Systematic errors on the crucial parameters due to the peaking background, as obtained from a series of test fits varying the $B \to fh$ yields and asymmetries based on their error estimates (Table 11.9 for the yields, $\pm 20\,\%$ for A_{CP}).

fit	$\sigma_{\rm tot}(A_{CP})$	$\sigma_{\rm tot}(R_{K/\pi})$
CP+	0.030	0.0014
CP−	0.0043	0.00087

12.3 Bias Correction

As described in Section 10.2, we correct the results of the final fits for biases. We will consider the uncertainties of the corrections to assign a systematic uncertainty due to the bias correction. The corrections are given in Table 10.2. At first we propagate their uncertainties through the correction procedure (Eq. 10.2). Then we construct the covariance matrices $V_{\rm BIAS}[\vec{x}]$ and $V_{\rm BIAS}[\vec{y}]$ in full analogy to Section 12.1. The square root of the diagonal elements of $V_{\rm BIAS}[\vec{y}]$ is reported in Table 12.1.

12.4 Detector Charge Asymmetry

A source that must be investigated arises from a potential charge asymmetry of the *BABAR* detector, due to a possible charge bias in tracking efficiency (e.g., K^+ vs K^-) and/or particle identification. In order to understand if this effect can bring a significant bias to the asymmetry measurement, a number of control samples are studied. These samples are the $B \to D^0\pi$ samples, and in addition the $B \to D^0K$ flavor modes, in both categories the *CP* asymmetry is expected to be negligible. The same check has been performed on Cocktail Monte Carlo. The results of the measured charge asymmetries are reported in Table 12.4. All results show that there is no evidence of a charge asymmetry of the *BABAR* detector. We take the uncertainty on the weighted average of the A_{CP} values as measured in the data control samples to be the systematic error: $\sigma_{A_{CP}} = 0.0025$. When building the covariance matrix for this source, we consider the errors on A_{CP+} and A_{CP-} to be 100% correlated.

Table 12.4: Charge asymmetries measured on signal Monte Carlo and data for different B decays.

decay mode	A_{CP} (data)	A_{CP} (MC)
$B \to D^0_{CP+}\pi$	0.006 ± 0.014	0.001 ± 0.014
$B \to D^0_{CP-}\pi$	-0.007 ± 0.014	-0.007 ± 0.013
$B \to D^0\pi$, $D^0 \to K^-\pi^+$	-0.0111 ± 0.0050	-0.0018 ± 0.0047
$B \to D^0\pi$, $D^0 \to K^-\pi^+\pi^0$	-0.0105 ± 0.0048	0.0027 ± 0.0051
$B \to D^0\pi$, $D^0 \to K^-\pi^+\pi^+\pi^-$	-0.0007 ± 0.0046	-0.0007 ± 0.0043
$B \to D^0K$, $D^0 \to K^-\pi^+$	-0.004 ± 0.022	-0.001 ± 0.021
$B \to D^0K$, $D^0 \to K^-\pi^+\pi^0$	-0.021 ± 0.019	-0.003 ± 0.019
$B \to D^0K$, $D^0 \to K^-\pi^+\pi^+\pi^-$	0.001 ± 0.019	-0.002 ± 0.018
weighted avg.	-0.0068 ± 0.0026	-0.0005 ± 0.0025

12.5 S-Wave Pollution in *CP*-odd Final States

The measured *CP* asymmetry in $B \to D^0K$, $D^0 \to K^0_S\phi$, can be diluted by the presence of $B \to D^0K$ decays with D^0 decaying to the same final state $K^0_SK^+K^-$ as $K^0_S\phi$, $\phi \to K^+K^-$, but with opposite *CP* content. The same can happen in the $B \to D^0K$, $D^0 \to K^0_S\omega$ analysis with backgrounds from $B \to D^0K$, $D^0 \to K^0_S\pi^-\pi^+\pi^0$. Moreover, also the measured ratio R_{CP-} can be affected by the presence of this peculiar background. It is possible to obtain correction factors to both A_{CP-} and R_{CP-} from a fit to the distributions of the relevant helicity angles, $\cos\theta_N$ and $\cos\theta_H$ for $K^0_S\omega$ and $K^0_S\phi$, respectively. These angles have been defined in Section 6.3.3. The fit is performed on dedicated $B \to D^0\pi$ samples, in which the selection requirements on the helicity angles have not been applied. It can be shown [49], that for these two final states the observed charge asymmetry should be corrected by a factor

$$A_{CP}^{\text{true}} = A_{CP}^{\text{obs}} \cdot \frac{1 + f|z|^2R'}{1 - f|z|^2}, \tag{12.4}$$

while the ratio R_{CP} should be corrected by

$$R_{CP}^{\text{true}} = R_{CP}^{\text{obs}} \cdot \frac{1 + f|z|^2}{1 + f|z|^2R'}. \tag{12.5}$$

Here, R' is the ratio of the $R_{CP\pm}$ values where R_{CP-} is calculated using the $D^0 \to K^0_S\pi^0$ final state only (as opposed to using all three $CP-$ final states under study),

$$R' = \frac{R_{CP+}}{R_{CP-}^{K^0_S\pi^0}}, \tag{12.6}$$

f is the ratio of the efficiencies of the selection criterion on the helicity angles: $f_\omega = 0.71$ obtained from the cut on the quantity $\cos^2\theta_N \sin^2\theta_{\pi\pi} > 0.046$, $f_\phi = 0.64$ obtained from the cut on the quantity $\cos\theta_H > 0.4$. The parameter $|z|^2$ can be extracted from a $|z|^2 + 3\cos^2\theta$ fit to the samples introduced above. In these samples, we apply four additional selection criteria to simulate the selection power of the final fit, and to select $B \to D^0\pi$ events: a 2.5σ cut on $\Delta E(\pi)$ around zero, the prompt track is required to pass the `KLHVeryTight` selector, $m_{\rm ES} > 5.275\,{\rm GeV}/c^2$, and $\mathcal{F} > 0.2$. We subtract the background expected from the Monte Carlo simulation. As shown in Figure 12.3, we find $|z|^2 = 0.061 \pm 0.032$ in case of $K^0_S\omega$, and $|z|^2 = 0.220 \pm 0.063$ in case of $K^0_S\phi$. The corresponding asymmetry corrections are:

$$\begin{aligned} A^{\rm true}_{CP(K^0_S\omega)} &= A^{\rm obs}_{CP(K^0_S\omega)} \times (1.093 \pm 0.055), \\ A^{\rm true}_{CP(K^0_S\phi)} &= A^{\rm obs}_{CP(K^0_S\phi)} \times (1.344 \pm 0.125). \end{aligned}$$

For the ratios it is

$$\begin{aligned} R^{\rm true}_{CP(K^0_S\omega)} &= R^{\rm obs}_{CP(K^0_S\omega)} \times (1.004 \pm 0.005), \\ R^{\rm true}_{CP(K^0_S\phi)} &= R^{\rm obs}_{CP(K^0_S\phi)} \times (1.012 \pm 0.013). \end{aligned}$$

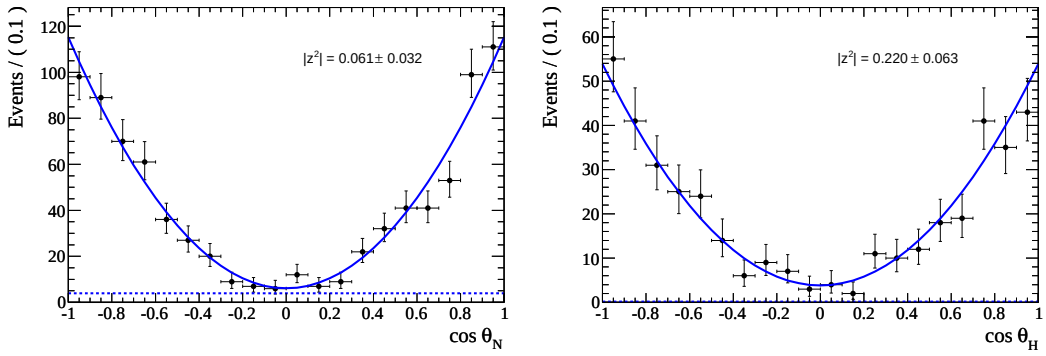

Figure 12.3: Left: $\cos\theta_N$ distribution of $B \to D^0\pi$, $D^0 \to K^0_S\omega$ candidates selected in data. Right: $\cos\theta_H$ distribution of $B \to D^0\pi$, $D^0 \to K^0_S\phi$ candidates selected in data. Background expected from $q\overline{q}$ and $B\overline{B}$ (non peaking) MC is shown with a dashed line.

We note both correction factors for the ratios R_{CP} are consistent with unity. Thus we will not apply these factors, we will, however, assign a systematic error based on the uncertainties of the correction factors. The corrections for the asymmetries are less consistent with unity. Since the final result on A_{CP-} is obtained from a maximum likelihood fit, simultaneously running on the $D^0 \to K^0_S\pi^0$, $D^0 \to K^0_S\omega$, and $D^0 \to K^0_S\phi$ subsamples, we will apply the correction factors inside the final fit procedure. This allows the likelihood fitter to correctly estimate their influence.

We assign a systematic error in the following way. At first, we fit the three $CP-$ final states independently. Then, we combine their fit results by means of a weighted average, using

the squared inverse of the statistical error as the weight. This is done for both relevant parameters, $R^-_{K/\pi}$ and A_{CP-}. Finally, we use standard error propagation to propagate the systematic uncertainties through the averaging procedure. The resulting uncertainties are reported in Table 12.1. When building the covariance matrix for this source, we consider the errors on $R^-_{K/\pi}$ and A_{CP-} to be 50% anti-correlated, since the error on $|z|^2$ dominates over the error on R' in Eqns 12.4 and 12.5.

12.6 Derivation of $R_{CP\pm}$ From $R_\pm$

The ratios $R_\pm$ are defined by:

$$\begin{aligned} R_\pm &= \frac{\mathcal{B}(B \to D^0_{CP} K)}{\mathcal{B}(B \to D^0_{CP}\pi)} \frac{\mathcal{B}(B \to D^0\pi)}{\mathcal{B}(B \to D^0 K)} \\ &= \frac{|A(B^- \to D^0 K^-)|^2 (1 + r_B^2 \pm 2 r_B \cos\delta_B \cos\gamma)}{|A(B^- \to D^0\pi^-)|^2 (1 + r_{B\pi}^2 \pm 2 r_{B\pi} \cos\delta_{B\pi}\cos\gamma)} \frac{|A(B^- \to D^0\pi-)|^2}{|A(B^- \to D^0 K^-)|^2} \\ &= \frac{R_{CP\pm}}{1 + r_{B\pi}^2 \pm 2 r_{B\pi}\cos\delta_{B\pi}\cos\gamma}, \end{aligned} \tag{12.7}$$

where $r_{B\pi} = |A(B^- \to \bar{D^0}\pi^-)/A(B^- \to D^0\pi^-)|$ and $\delta_{B\pi}$ is the strong phase difference between the two amplitudes. The denominator is usually considered equal to 1, hence $R_{CP\pm} = R_\pm$. However, the denominator is not exactly one and a systematic uncertainty related to this assumption must be evaluated when quoting the final results in terms of $R_{CP\pm}$ instead of the measured quantities $R_\pm$. The diagrams of $B^- \to D^0/\overline{D}^0\pi^-$ are the same as those of $B^- \to D^0/\overline{D}^0 K^-$, after the replacement of the s quark with a d quark. Therefore the ratio $r_{B\pi}$ is:

$$\begin{aligned} r_{B\pi} &\equiv \left|\frac{A(B^- \to \overline{D}^0\pi^-)}{A(B^- \to D^0\pi^-)}\right| \\ &= \frac{|A(B^- \to \overline{D}^0 K^-)||V_{cd}/V_{cs}|}{|A(B^- \to D^0 K^-)||V_{ud}/V_{us}|} = \left|\frac{A(B^- \to \overline{D}^0 K^-)}{A(B^- \to D^0 K^-)}\right| \cdot \tan^2\theta_C \\ &= r_B \tan^2\theta_C. \end{aligned} \tag{12.8}$$

We take $\sin\theta_C = 0.2257 \pm 0.0010$ from [8], and $r_B = 0.103^{+0.017}_{-0.023}$ from [13]. We obtain $r_{B\pi} = (5.5 \pm 1.2) \times 10^{-2}$ and a relative systematic uncertainty on $R_{CP\pm}$ of $\pm 2 \cdot 6.7 \times 10^{-3} = \pm 1.34 \times 10^{-2}$ (the term $r^2_{B\pi}$ is neglected). Note that this systematic uncertainty is 100% anti-correlated between R_{CP+} and R_{CP-}. The resulting values are summarized in Table 12.1.

12.7 Signal Self Cross-Feed

We define the signal self cross-feed (SCF) as reconstructed signal MC events which are not truth-matched. The ratio $N_{\mathrm{SCF}}/N_{\mathrm{total}}$ ranges from 0.3 % to 12 % in $B \to D^0_{CP}K$ events, depending on the channel. We treat this component as signal. By doing so, we implicitly assume, that the charge asymmetries of both components are equal. If the self cross-feed component had a significantly different charge asymmetry compared to the signal, it would dilute the observed values. We use the signal MC samples to estimate the influence of the self cross-feed on the measured values of A_{CP}. The detailed calculation is contained in Section B.2 in Appendix B. We obtain a systematic error on A_{CP} as given in Table 12.5.

We consider the errors uncorrelated and propagate them analogously to Section 12.5. The resulting uncertainties are summarized in Table 12.1.

Table 12.5: Systematic uncertainties on A_{CP} due to signal self cross-feed. Values calculated from Table B.1 and Eq. B.8.

$D^0 \to$	KK	$\pi\pi$	$K^0_S\pi^0$	$K^0_S\omega$	$K^0_S\phi$
σ_{SCF}	0.00021	0.00011	0.00082	0.0019	0.00030

12.8 Differences in Selection Efficiency

The selection efficiencies for $B \to D^0 K$ and $B \to D^0 \pi$ signal events are not exactly equal. From Table 6.14 we calculate the weighted average of the ratios $\varepsilon(D^0K)/\varepsilon(D^0\pi)$ to be 0.9896 ± 0.0010, where the error results from the limited size of the MC samples. We assign a relative systematic error to the measured values of $R_{K/\pi}$ of $1 - 0.9896 + 0.0010 = 0.0114$. This systematic uncertainty is 100% correlated between R_{CP+} and R_{CP-}. The resulting uncertainties are summarized in Table 12.1.

13 Final Result

We present the final result of the GLW parameters $A_{CP\pm}$ and $R_{CP\pm}$, as obtained from a fit to the Run 1-6 dataset of the *BABAR* Experiment. Table 13.1 shows the final values. We combine the values of $R_{K/\pi}$, as measured in the flavor modes, by means of a weighted average:

$$R^{\rm flv}_{K/\pi} = \frac{1}{w}\sum_i w_i R_{K/\pi}, i, \tag{13.1}$$

where $w_i = 1/\sigma^2_{R_{K/\pi},i}$, $w = \sum_i w_i$, and i ranges over $K^-\pi^+$, $K^-\pi^+\pi^0$, and $K^-\pi^+\pi^+\pi^-$. The combined statistical uncertainty is given by the standard deviation $\sigma_{R^{\rm flv}_{K/\pi}} = 1/\sqrt{w}$. Then we calculate the values $R_{CP\pm}$ as

$$R_{CP\pm} = \frac{R^{\pm}_{K/\pi}}{R^{\rm flv}_{K/\pi}}. \tag{13.2}$$

Table 13.1: Final result on $R_{CP\pm}$ and $A_{CP\pm}$ as measured in the Run 1-6 on-peak dataset of *BABAR*. The first error is statistical, the second systematic.

D^0 mode	R_{CP}	A_{CP}
CP+	$1.128 \pm 0.081 \pm 0.037$	$+0.197 \pm 0.060 \pm 0.031$
CP–	$1.041 \pm 0.069 \pm 0.030$	$-0.096 \pm 0.064 \pm 0.0092$

The statistical covariance matrix $V_{\rm (stat)}[\vec{y}]$, $\vec{y} = (A_{CP+}, A_{CP-}, R_{CP+}, R_{CP-})^T$, of the final result is obtained using the covariances as reported by the final fits and error propagation according to Eq. 12.3. The matrix is

$$V_{\rm (stat)}[\vec{y}] = \begin{pmatrix} 0.0036 & 0.0 & -0.00041 & 0.0 \\ 0.0 & 0.0041 & 0.0 & 0.00013 \\ -0.00041 & 0.0 & 0.0065 & 0.00022 \\ 0.0 & 0.00013 & 0.00022 & 0.0047 \end{pmatrix}. \tag{13.3}$$

The statistical significance of A_{CP+} being greater than zero is $S = 3.4$ standard deviations. It is calculated from the minimal values of the likelihood function of the nominal fit and a dedicated null-hypothesis fit, where A_{CP} was fixed to zero, using Eq. 9.4. This constitutes evidence for direct *CP* violation in charged B decays.

To constrain the CKM angle γ from the present measurement we follow Reference [24]. We define a χ^2-function as

$$\chi^2(\gamma, \delta, r) = -2\ln\mathcal{L}(\gamma, \delta, r), \tag{13.4}$$

$$\mathcal{L}(\gamma, \delta, r) = \frac{1}{(2\pi)^2\sqrt{|V|}}\exp\left(-\frac{1}{2}(\vec{y}-\vec{y_{\rm t}})V^{-1}(\vec{y}-\vec{y_{\rm t}})^T\right), \tag{13.5}$$

where V is the combined covariance matrix $V = V_{(\mathrm{stat})}[\vec{y}] + V_{(\mathrm{syst})}[\vec{y}]$. The vector $\vec{y_t} = \vec{y_t}(\gamma, \delta, r)$ is the vector of truth variables defined as a function of γ, the amplitude ratio r, and the strong phase difference δ, through Eqns. 3.11 and 3.12. The optimal values for γ, δ, and r correspond to a minimum of the χ^2-function. Due to the inherent eight-fold ambiguity of the GLW method there are eight equivalent minima. We calculate the one most consistent with the current experimental situation as summarized by the Particle Data Group (PDG, [8]) to be

$$\gamma = 1.64^{+0.12}_{-0.11}, \tag{13.6}$$
$$\delta = 2.87^{+0.10}_{-0.24}, \tag{13.7}$$
$$r = 0.29^{+0.09}_{-0.13}, \tag{13.8}$$

where the errors represent the intervals corresponding to a change in χ^2 of $\Delta\chi^2 = 1$ as obtained by MINOS. Figure 13.1 shows the χ^2 function as function of γ and δ, with r set to its value at the minimum.

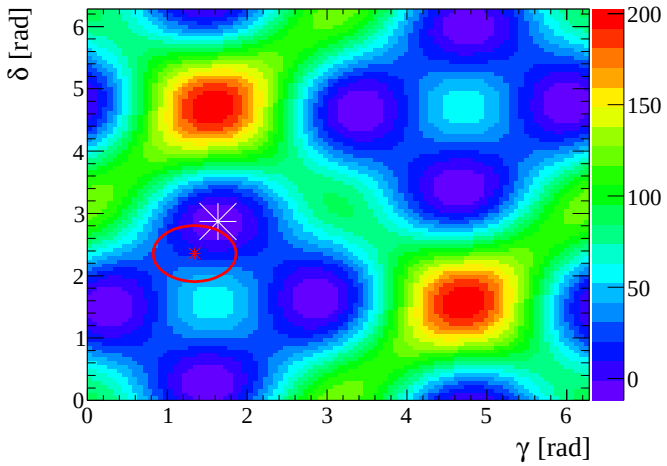

Figure 13.1: The χ^2-function (Eq. 13.4) with r fixed to its optimal value (Eq. 13.8). The white marker depicts the result on γ and δ (Eqns. 13.6, 13.7). The red marker shows the PDG values taken from Reference [8], together with a naive one-standard-deviation error ellipse.

However, one has to exercise care interpreting the errors in Eqns. 13.6-13.8. Due to the fact that two minima lie closely together at $\gamma \approx 1.5$, the errors do not reflect the true 68.3 % confidence levels (CL). Instead we perform a CL scan for γ: for each given value of γ, γ_0, we take the value of the χ^2 function at the new minimum, $\chi^2_{\mathrm{min}}(\gamma_0, \delta_0, r_0)$. Then we calculate the difference of this value and the global minimum of χ^2,

$$\Delta\chi^2 = \chi^2_{\mathrm{min}}(\gamma_0, \delta_0, r_0) - \chi^2_{\mathrm{min}}. \tag{13.9}$$

We compute the probability that $\Delta\chi^2$ is exceeded for a χ^2-distribution with one degree of freedom,

$$1 - \mathrm{CL} = \frac{1}{\sqrt{2}\,\Gamma(1/2)} \int_{\Delta\chi^2}^{\infty} e^{-t/2} t^{-1/2}\,\mathrm{d}t. \tag{13.10}$$

We obtain the 68.3 % confidence intervals of γ to be

$$\gamma \in [0.16, 0.51],\ [1.39, 1.75],\ [2.63, 2.97] \quad @68.3\%\ \mathrm{CL}, \tag{13.11}$$

the 95 % confidence interval is

$$\gamma \in [0.08, 3.06] \quad @95\% \text{ CL}, \tag{13.12}$$

where each interval shifted by π is a solution, too.[1] Figure 13.2 plots the quantity $1 - \text{CL}$ for each value of γ. From Eq. 13.12 it is clear, that this measurement does not put hard constraints on the angle γ.

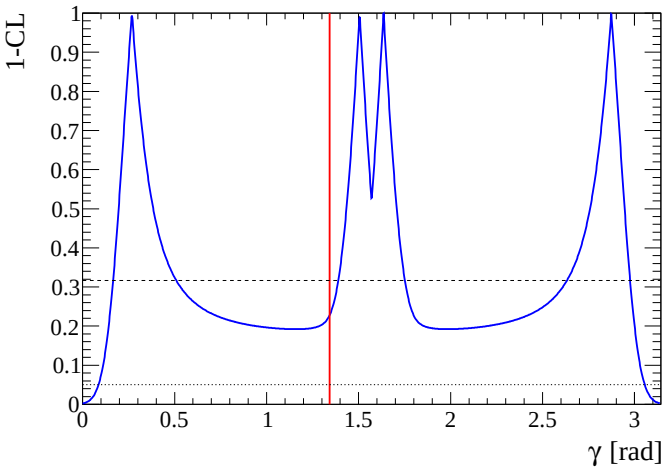

Figure 13.2: Confidence level of a given value of γ. Both statistical and systematic uncertainties are taken into account. The plot is restricted to $[0, \pi]$, it is identical in the range $[\pi, 2\pi]$. The vertical line depicts the PDG value [8]. The horizontal lines show the 68.3% CL (dashed) and the 95% CL (dotted).

It is possible, however, to combine this result with the most precise measurement of γ in *BABAR*, which follows the GGSZ method[2]. To this extent we express the *CP*-sensitive observables in terms of three independent quantities x_+, x_-, and r^2, which are part of the so called Cartesian coordinates. They are related to the *CP* parameters that are measured in the GGSZ method. The connection of the Cartesian coordinates to γ, r, and δ, is given by $x_\pm = r\cos(\delta \pm \gamma)$ and $y_\pm = r\sin(\delta \pm \gamma)$. The GLW parameters give access to x_+, x_-, and r^2 through

$$x_\pm = \frac{R_{CP+}(1 \mp A_{CP+}) - R_{CP-}(1 \mp A_{CP-})}{4}, \tag{13.13}$$

$$r^2 = x_\pm^2 + y_\pm^2 = \frac{R_{CP+} + R_{CP-} - 2}{2}. \tag{13.14}$$

We measure

$$x_+ = -0.059 \pm 0.034(\text{stat}) \pm 0.014(\text{syst}), \tag{13.15}$$

$$x_- = 0.102 \pm 0.036(\text{stat}) \pm 0.015(\text{syst}), \tag{13.16}$$

$$r^2 = 0.084 \pm 0.054(\text{stat}) \pm 0.024(\text{syst}), \tag{13.17}$$

[1] The integral in Eq. 13.10 is evaluated using the `TMath::Prob(`χ^2`,` n_{dof}`)` function in ROOT.

[2] Compare Section 3.1. Note that the result presented here contains the $D^0 \to K^0_S\phi$ channel, which is also contained in the GGSZ analysis through $D^0 \to K^+K^-K^0_S$.

with covariances $V[\vec{c}]$, $\vec{c} = (x_+, x_-, r^2)^T$,

$$V_{(\mathrm{stat})}[\vec{c}] = \begin{pmatrix} 119 & 11 & 7.5 \\ 11 & 131 & 38 \\ 7.5 & 38 & 293 \end{pmatrix} \times 10^{-5}, \quad V_{(\mathrm{syst})}[\vec{c}] = \begin{pmatrix} 21 & 5.7 & 1.7 \\ 5.7 & 23 & 9.2 \\ 1.7 & 9.2 & 59 \end{pmatrix} \times 10^{-5}. \quad (13.18)$$

At the time this work was carried out the GGSZ result on the full *BABAR* dataset was not available. Therefore the combination will be left for future publications.

The fit results of the five final fits are visualized in Figures 13.3-13.8, which show different projections of the multidimensional simultaneous maximum likelihood fits: one set of four plots for each fit variable, ΔE, m_{ES}, and $\mathcal{F}$. In each set, the left hand side corresponds to the the negatively charged subsample ($B^- \to D^0 h^-$), the right hand side to the positively charged subsample. By comparing both sides the direct *CP* asymmetry, A_{CP}, becomes visible. The upper row shows events in which the bachelor track, h, passes the `KLHVeryTight` selector and therefore has a great likelihood of originating from a true kaon. The lower row shows events in which h fails the selector. The figures show the combination of the *CP*-even channels ($D^0 \to K^-K^+$, $\pi^+\pi^-$), and the of the *CP*-odd channels ($D^0 \to K^0_S\pi^0$, $K^0_S\omega$, $K^0_S\phi$). The plots of the remaining flavor channels ($K^-\pi^+$, $K^-\pi^+\pi^0$, $K^-\pi^+\pi^+\pi^-$) are shown in Figures C.11-C.16 in Appendix C, as well as projections of each of the *CP* channels separately. Also in Appendix C are tables with the final values of each floating fit parameter, they are Tables C.1-C.3.

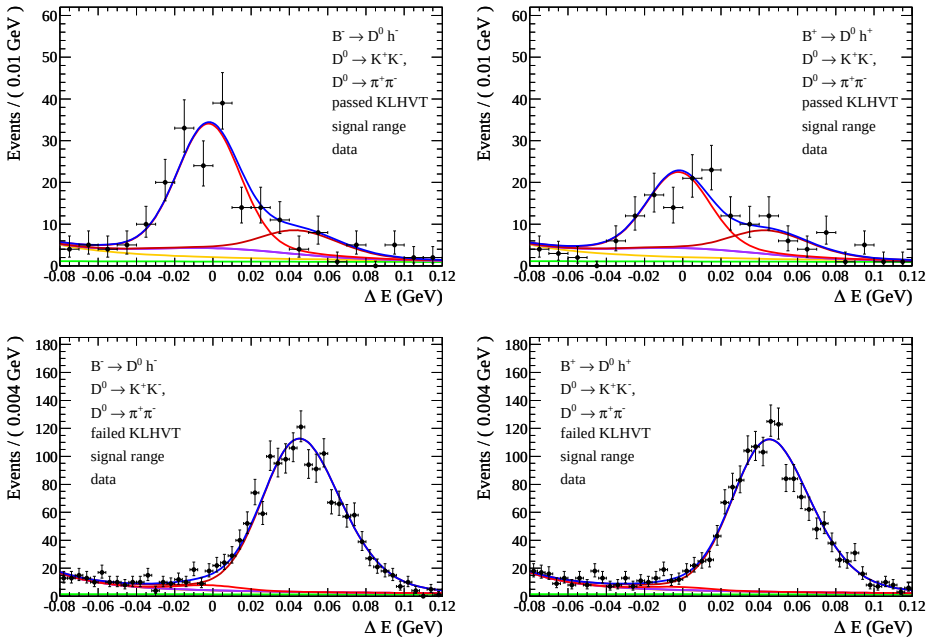

Figure 13.3: ΔE projections of the *CP*+ final fit to $B \to D^0 h$ Run 1-6 on-peak data. Top (bottom): the bachelor track h is required to pass (fail) the `KLHVeryTight` selector. Left: B^- subsample. Right: B^+ subsample. Color code in Table 10.4 on page 69.

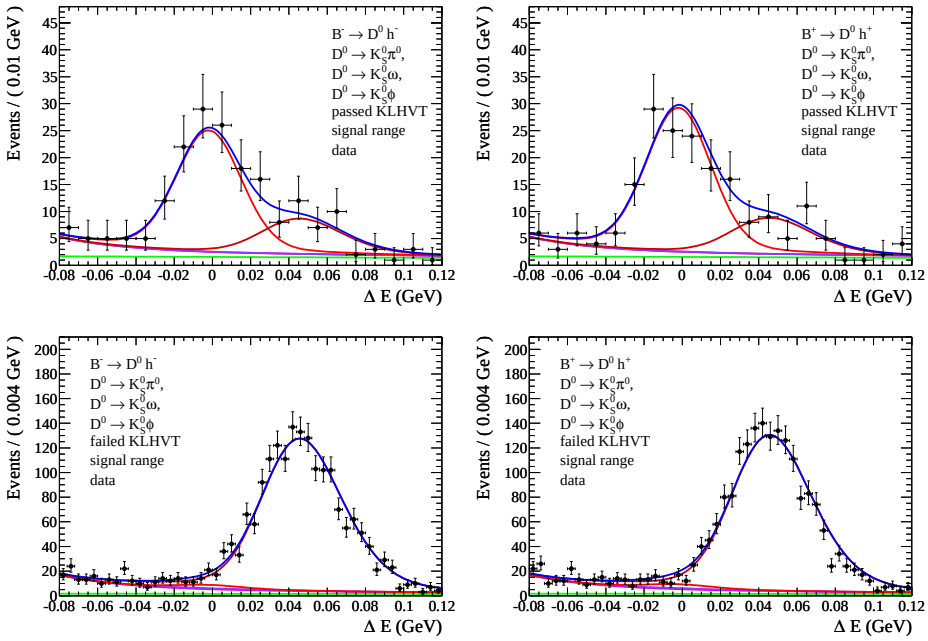

Figure 13.4: ΔE projections of the *CP*– final fit to $B \to D^0 h$ Run 1-6 on-peak data. Top (bottom): the bachelor track h is required to pass (fail) the `KLHVeryTight` selector. Left: B^- subsample. Right: B^+ subsample. Color code in Table 10.4 on page 69.

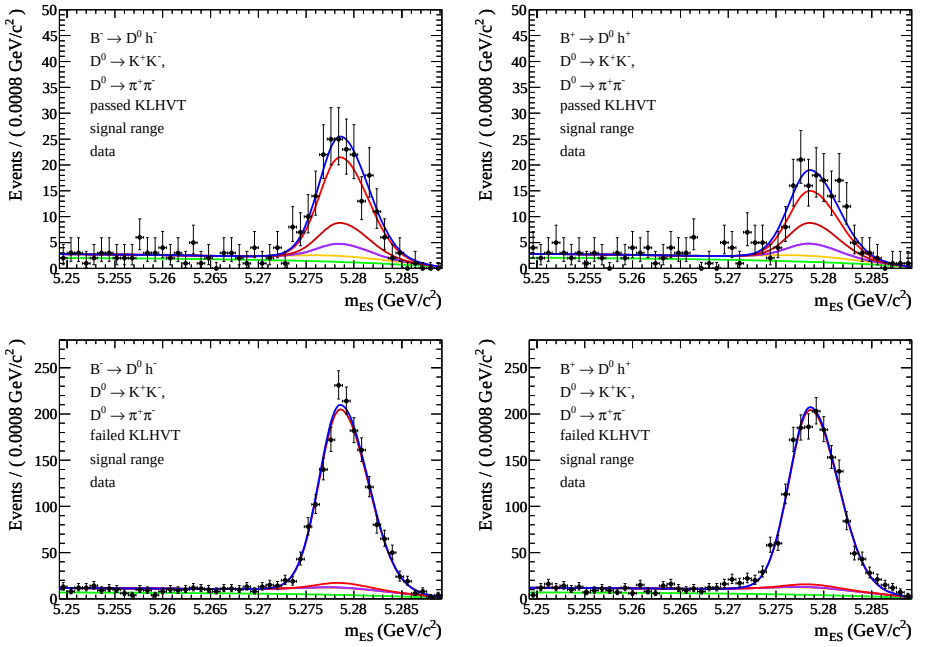

Figure 13.5: $m_{\rm ES}$ projections of the *CP*+ final fit to $B \to D^0 h$ Run 1-6 on-peak data. Top (bottom): the bachelor track h is required to pass (fail) the `KLHVeryTight` selector. Left: B^- subsample. Right: B^+ subsample. Color code in Table 10.4 on page 69.

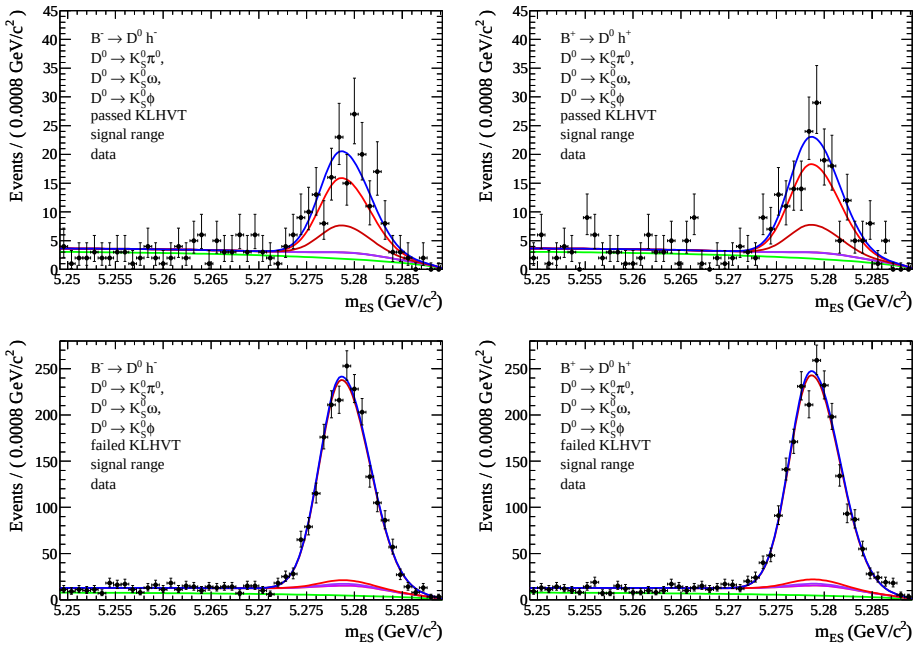

Figure 13.6: $m_{\rm ES}$ projections of the *CP*– final fit to $B \to D^0 h$ Run 1-6 on-peak data. Top (bottom): the bachelor track h is required to pass (fail) the `KLHVeryTight` selector. Left: B^- subsample. Right: B^+ subsample. Color code in Table 10.4 on page 69.

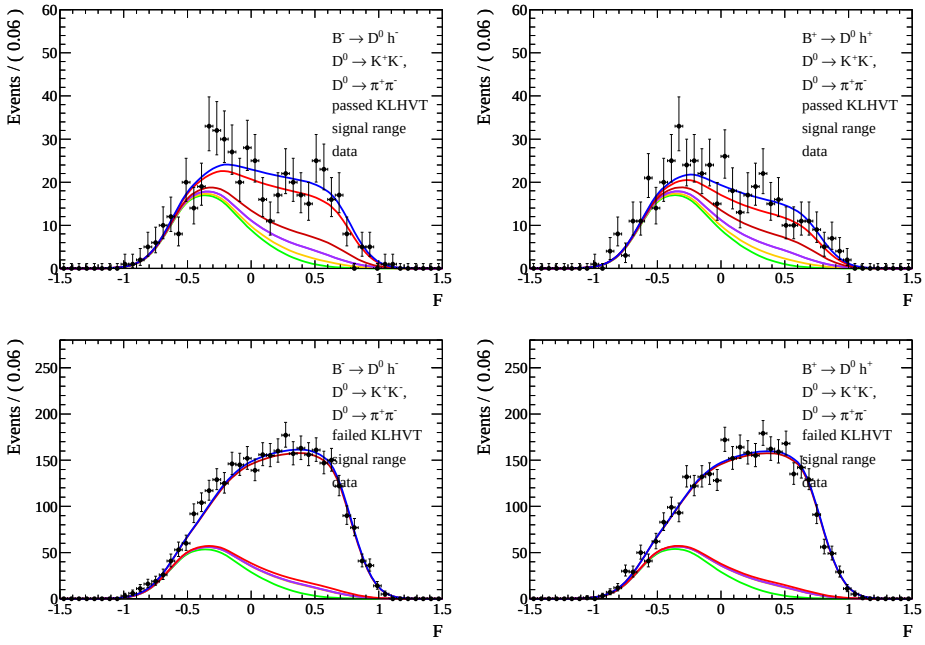

Figure 13.7: $\mathcal{F}$ projections of the *CP*+ final fit to $B \to D^0 h$ Run 1-6 on-peak data. Top (bottom): the bachelor track h is required to pass (fail) the `KLHVeryTight` selector. Left: B^- subsample. Right: B^+ subsample. Color code in Table 10.4 on page 69.

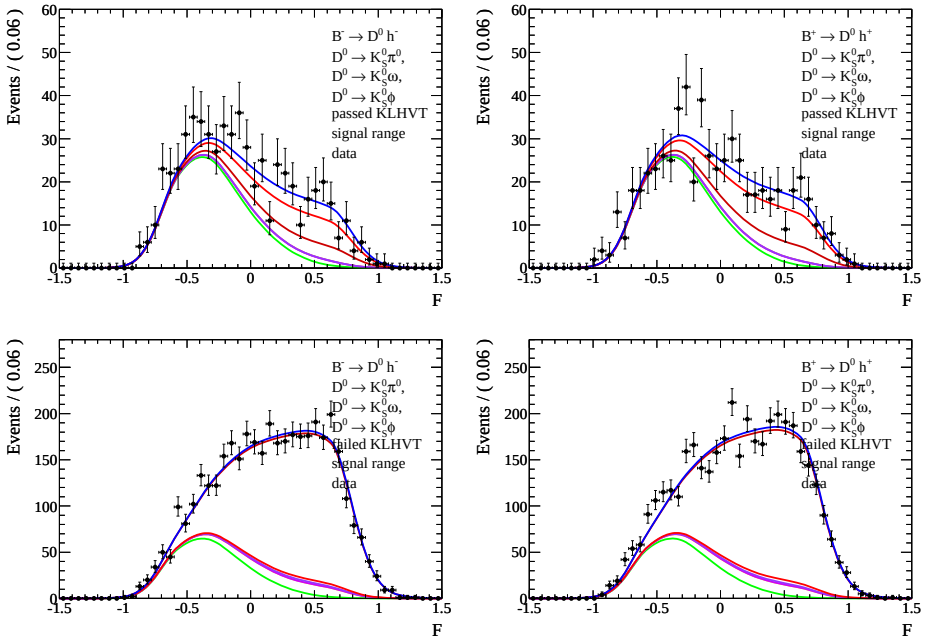

Figure 13.8: $\mathcal{F}$ projections of the *CP*– final fit to $B \to D^0 h$ Run 1-6 on-peak data. Top (bottom): the bachelor track h is required to pass (fail) the `KLHVeryTight` selector. Left: B^- subsample. Right: B^+ subsample. Color code in Table 10.4 on page 69.

14 Summary

In summary we reconstructed $B^- \to D^0\pi^-$ and $B^- \to D^0K^-$ decays and their charge conjugates. The D^0-meson was reconstructed in eight final states: the *CP*-even K^+K^- and $\pi^+\pi^-$, the *CP*-odd $K^0_S\pi^0$, $K^0_S\phi$, and $K^0_S\omega$, and the flavor eigenstates $K^-\pi^+$, $K^-\pi^+\pi^0$, and $K^-\pi^+\pi^+\pi^-$. We used a data sample of 467×10^6 $\Upsilon(4S)$ decays, representing the final $\Upsilon(4S)$ dataset of the *BABAR* Experiment.

We measured the *CP* sensitive parameters $A_{CP\pm}$ and $R_{CP\pm}$, as they are defined in the GLW method. These parameters have been measured through five maximum likelihood fits to the data: two simultaneous fits to the *CP*± states and three fits to the flavor states. The resulting values are

$$\begin{aligned} R_{CP+} &= 1.128 \pm 0.081 \pm 0.037, \\ R_{CP-} &= 1.041 \pm 0.069 \pm 0.030, \\ A_{CP+} &= +0.197 \pm 0.060 \pm 0.031, \\ A_{CP-} &= -0.096 \pm 0.064 \pm 0.009. \end{aligned}$$

We considered eight possible sources of systematic errors. The two most important ones are the uncertainty in a non-reducible charmless peaking background contribution, and the uncertainty in the shapes of the probability density functions of the five fits.

The statistical significance of A_{CP+} being non-zero is 3.4 standard deviations, constituting evidence for direct *CP* violation in charged B decays.

We also express the results in terms of the so called Cartesian coordinates x_+, x_-, and r^2: $x_+ = -0.059 \pm 0.034(\text{stat}) \pm 0.014(\text{syst})$, $x_- = 0.102 \pm 0.036(\text{stat}) \pm 0.015(\text{syst})$, $r^2 = 0.084 \pm 0.054(\text{stat}) \pm 0.024(\text{syst})$. They are related to the *CP* parameters that are measured using a Dalitz analysis of $B \to D^0K$, $D^0 \to K^0_S\pi^-\pi^+$ decays [24]. This choice allows the results of the two measurements to be expressed in a consistent manner in future publications.

This work improved the accuracy of a previous *BABAR* measurement of $A_{CP\pm}$ and $R_{CP\pm}$ [4] using additional 231×10^6 $\Upsilon(4S)$ decays as well as a new analysis strategy. Considering the statistical errors on both the ratios and asymmetries scale like the square root of the ratio of luminosities, $\sqrt{\mathcal{L}_1/\mathcal{L}_2}$, the improvement due to the new analysis strategy can be quantified to be equivalent to $\approx 600\times10^6$ additional $\Upsilon(4S)$ decays. But in spite of this improvement, the statistical uncertainties are still dominating.

The results of this work help to better constrain the phase parameter $\gamma = \arg(-V_{ud}V^*_{ub}/V_{cd}V^*_{cb})$ of the Cabibbo-Kobayashi-Maskawa quark mixing matrix. We determine γ up to a six-fold ambiguity at the 68.3 % confidence level. The value most consistent with the current experimental situation [8] is $\gamma = 1.64^{+0.11}_{-0.25} = [94^{+7}_{-14}]^\circ$, at the 68.3 % confidence level, where the errors are a combination of the systematic and statistical uncertainties.

15 Acknowledgements

It has become a nice tradition not only to acknowledge the help of the many people who made this work possible, but to also reminisce about the past three years time, during which this work was carried out. I had the chance to stay a significant part of this time in Menlo Park/Palo Alto, California, where the SLAC National Laboratory is located. For this opportunity I am grateful to Bernhard Spaan.

During the first three months of my time at SLAC I learned a lot from Vitaly Eyges, regarding both physics analysis techniques and teamwork. Without him, the partial result could not have been published so soon. As for the final result, I am especially thankful to Giovanni Marchiori, who often pointed me in the right direction when weird problems arose, and who reassured my points of view in a different aspect of teamwork. Grazie mille, Giovanni, per essermi stato veramente di grande aiuto. Senza di te non sarei mai riuscito a finire in tempo.

I was put in charge of the light pulser system of the electromagnetic calorimeter. As the commissioner, I received continuous support from Wolfgang Gradl, who still knew all the sophisticated wrinkles of the *BABAR* online production system. I'm also grateful to *BABAR*'s Technical Coordinator Bill Wisniewski, who, once you get to know him, is a wonderful person. I had the honor to take *BABAR*'s last data-taking shift with him.

Many thanks go to Miriam Fritsch, who was my office mate in the German office at SLAC. Being in the USA as a German just as me, she was a great help when it came to God's own bureaucracy: do I really need a Driver's License to get a Social Security Card, which I need to get a Driver's License? Thank you also to Florian Bernlochner, my Swiss office mate during the last three months, for great discussion and the many lessons in Swiss German.

The *BABAR* Collaboration took (and still takes) great benefit from the strong Spanish and Italian contribution. I want to thank the "Spanish and Spanish speaking" people, Diego, Neus, and Enrico for many coffees on the sunny ROB terrace, for the rock music, and for the nights in San Francisco.

I spent a big part of my spare time climbing the most beautiful crags I've ever seen. We have been to Lover's Leap, in Bishop, and many times in Yosemite and Tuolumne. Thank you to my brave climbing partners, Sarah, Rasmus, and Stefan. It was great times!

I am also very grateful to my friends in Dortmund, who, although I had been away for one and a half years, welcomed me back in the most amazing way. Thank you so much, Martin, Neki, and Yeti, it really meant a lot to me. And I'd like to thank you, Jesko, for your support during the traditionally tough times of finishing up this work. Thank you, Heiko, for your comments and valuable discussion.

At last, I wish to thank my parents for their continuous and unconditional support: Danke, Jutta und Walter.

Till Moritz Karbach, *Dortmund, 5. April 2009*

A Additional Material to Chapter 10

Table A.1: Fit result of the *CP*+ final fit to cocktail MC.

Parameter	KK	$\pi\pi$
$A_{CP}^{\mathrm{sig}(K)}$	0.252 ± 0.065	
$A_{CP}^{\mathrm{sig}(\pi)}$	-0.002 ± 0.015	
$R_{K/\pi}$	0.0847 ± 0.0060	
m	0.0166 ± 0.0027	
$A_{CP,f}^{B\overline{B}}$	0.033 ± 0.065	-0.089 ± 0.079
$A_{CP,p}^{q\overline{q}}$	-0.004 ± 0.016	-0.004 ± 0.017
$A_{CP,f}^{q\overline{q}}$	-0.0020 ± 0.0096	0.0052 ± 0.0086
$f_{\mathcal{F}}^{q\overline{q}}$	0.389 ± 0.083	0.49
$\sigma_{\mathcal{F},l,1}^{q\overline{q}}$	0.158 ± 0.019	0.201 ± 0.019
$\sigma_{\mathcal{F},l,2}^{q\overline{q}}$	0.168 ± 0.011	0.1870 ± 0.0064
$\sigma_{\mathcal{F},r,1}^{q\overline{q}}$	0.357 ± 0.022	0.362 ± 0.012
$\sigma_{\mathcal{F},r,2}^{q\overline{q}}$	0.249 ± 0.025	0.304 ± 0.036
$a_{\Delta E}^{q\overline{q}}$	-0.91 ± 0.13	-0.81 ± 0.12
$\mu_{\Delta E}$	-0.05 ± 0.31	0.31 ± 0.59
$\sigma_{\Delta E}$	16.44 ± 0.26	15.66 ± 0.50
$\mu_{m_{\mathrm{ES}}}$	5279.58 ± 0.10	5279.71 ± 0.22
$\sigma_{m_{\mathrm{ES}},l}$	2.765 ± 0.078	2.94 ± 0.17
$\sigma_{m_{\mathrm{ES}},r}$	2.164 ± 0.069	2.06 ± 0.15
$N_p^{B\overline{B}}$	72 ± 33	147 ± 44
$N_f^{B\overline{B}}$	953 ± 81	986 ± 145
$N_p^{q\overline{q}}$	4277 ± 72	3582 ± 73
$N_f^{q\overline{q}}$	13133 ± 136	17417 ± 192
$N_{\mathrm{tot}}^{\mathrm{sig}(\pi)}$	4295 ± 71	1308 ± 43

Table A.2: Fit result of the *CP*– final fit to cocktail MC.

Parameter	$K^0_S\pi^0$	$K^0_S\omega$	$K^0_S\phi$
$A_{CP}^{\mathrm{sig}(K)}$		0.006 ± 0.059	
$A_{CP}^{\mathrm{sig}(\pi)}$		-0.008 ± 0.013	
$R_{K/\pi}$		0.0794 ± 0.0052	
m		0.0193 ± 0.0026	
$A_{CP,p}^{q\bar{q}}$	0.000 ± 0.011	0.002 ± 0.019	0.007 ± 0.053
$A_{CP,f}^{q\bar{q}}$	0.0028 ± 0.0068	0.000 ± 0.012	0.011 ± 0.037
$f_{\mathcal{F}}^{q\bar{q}}$	0.766 ± 0.028	0.27	n/a
$\sigma_{\mathcal{F},l,1}^{q\bar{q}}$	0.1603 ± 0.0063	0.145 ± 0.039	0.302 ± 0.017
$\sigma_{\mathcal{F},l,2}^{q\bar{q}}$	0.0881 ± 0.0062	0.1954 ± 0.0053	n/a
$\sigma_{\mathcal{F},r,1}^{q\bar{q}}$	0.4204 ± 0.0060	0.344 ± 0.014	0.448 ± 0.021
$\sigma_{\mathcal{F},r,2}^{q\bar{q}}$	0.471 ± 0.044	0.238 ± 0.019	n/a
$a_{\Delta E}^{q\bar{q}}$	-1.379 ± 0.095	-1.24 ± 0.16	-0.48
$\mu_{\Delta E}$	-0.96 ± 0.33	-2.24 ± 0.60	-1.10 ± 0.74
$\sigma_{\Delta E}$	17.27 ± 0.28	17.58 ± 0.49	16.49 ± 0.58
$\mu_{m_{\mathrm{ES}}}$	5279.65 ± 0.10	5279.84 ± 0.17	5279.60 ± 0.23
$\sigma_{m_{\mathrm{ES}},l}$	2.961 ± 0.078	3.10 ± 0.14	2.98 ± 0.17
$\sigma_{m_{\mathrm{ES}},r}$	2.173 ± 0.070	2.09 ± 0.12	2.15 ± 0.15
$N_p^{B\overline{B}}$	94 ± 37	113 ± 46	3
$N_f^{B\overline{B}}$	1067 ± 92	948 ± 110	79 ± 18
$N_p^{q\overline{q}}$	8109 ± 97	3011 ± 69	373 ± 20
$N_f^{q\overline{q}}$	23011 ± 173	8275 ± 131	801 ± 31
$N_{\mathrm{tot}}^{\mathrm{sig}(\pi)}$	4791 ± 77	1642 ± 47	707 ± 28

Table A.3: Fit results of the three final fits to cocktail MC in the flavor final states of the D^0.

Parameter	$K\pi$	$K\pi\pi^0$	$K\pi\pi\pi$
$A_{CP}^{\mathrm{sig}(K)}$	-0.005 ± 0.020	-0.012 ± 0.020	-0.014 ± 0.019
$A_{CP}^{\mathrm{sig}(\pi)}$	-0.0024 ± 0.0048	0.0013 ± 0.0045	0.0004 ± 0.0045
$A_{CP,f}^{b\bar{b}}$	-0.012 ± 0.028	-0.036 ± 0.011	-0.001 ± 0.025
$A_{CP,p}^{q\bar{q}}$	0.000 ± 0.015	0.0024 ± 0.0068	0.007 ± 0.017
$A_{CP,f}^{q\bar{q}}$	0.0057 ± 0.0062	0.0058 ± 0.0028	0.0000 ± 0.0073
$f_{\mathcal{F}}^{q\bar{q}}$	0.624 ± 0.061	0.568 ± 0.011	0.5878 ± 0.0031
$\sigma_{\mathcal{F},l,1}^{q\bar{q}}$	0.184 ± 0.013	0.1570 ± 0.0061	0.157 ± 0.011
$\sigma_{\mathcal{F},l,2}^{q\bar{q}}$	0.1595 ± 0.0079	0.1503 ± 0.0027	0.1603 ± 0.0043
$\sigma_{\mathcal{F},r,1}^{q\bar{q}}$	0.4026 ± 0.0099	0.3931 ± 0.0047	0.3889 ± 0.0095
$\sigma_{\mathcal{F},r,2}^{q\bar{q}}$	0.294 ± 0.024	0.2811 ± 0.0091	0.278 ± 0.023
$R_{K/\pi}$	0.0811 ± 0.0017	0.0830 ± 0.0019	0.0831 ± 0.0017
$a_{\Delta E}^{q\bar{q}}$	-1.052 ± 0.093	-1.675 ± 0.058	-1.21 ± 0.10
$b_{\Delta E}^{q\bar{q}}$	n/a	5.03 ± 0.81	n/a
$\mu_{\Delta E}$	-0.215 ± 0.089	-0.662 ± 0.092	-0.12 ± 0.12
$\sigma_{\Delta E}$	15.650 ± 0.071	16.369 ± 0.075	15.565 ± 0.099
$\mu_{m_{\mathrm{ES}}}$	5279.621 ± 0.031	5279.523 ± 0.028	5279.455 ± 0.037
$\sigma_{m_{\mathrm{ES}},l}$	2.778 ± 0.022	2.785 ± 0.022	2.700 ± 0.031
$\sigma_{m_{\mathrm{ES}},r}$	2.298 ± 0.020	2.370 ± 0.020	2.288 ± 0.027
$N_p^{B\overline{B}}$	79 ± 24	2206 ± 124	661 ± 118
$N_f^{B\overline{B}}$	3479 ± 109	29694 ± 494	13194 ± 485
$N_p^{q\bar{q}}$	4905 ± 76	24497 ± 185	8550 ± 171
$N_f^{q\bar{q}}$	32026 ± 206	171442 ± 592	60659 ± 588
$N_{\mathrm{tot}}^{\mathrm{sig}(\pi)}$	49333 ± 238	82794 ± 360	59128 ± 396
m	0.02433 ± 0.00088	0.02131 ± 0.00085	0.0240 ± 0.0012

B Additional Material to Chapter 12

B.1 Parameterization of the PDF

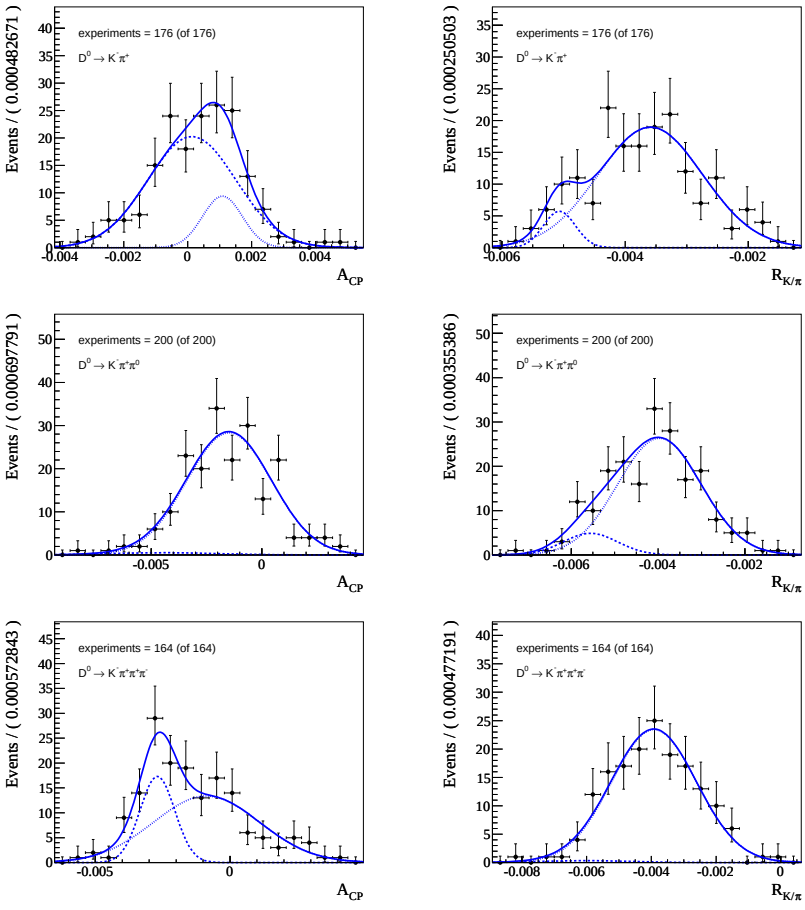

Figure B.1: Distributions of the crucial parameters as obtained from a series of systematic fits. The rows from top to bottom: $K^-\pi^+$, $K^-\pi^+\pi^0$, $K^-\pi^+\pi^+\pi^-$.

B.2 Signal Self Cross-Feed

We define the fractions f^+ (f^-) between positively (negatively) charged self cross-feed events and true signal events in Monte Carlo, will assume these ratios are the same in the data

sample:

$$f^{\pm} = \frac{N^{\pm}_{\rm scf,MC}}{N^{\pm}_{\rm sig,MC}}. \tag{B.1}$$

With the abbreviations $\hat{f} = (f^+ + f^-)/2$ and $\Delta f = (f^+ - f^-)/2$ the observed event yields in the positive and negative slices are

$$N^{\pm}_{\rm obs} = N^{\pm}_{\rm sig} + N^{\pm}_{\rm scf} = N^{\pm}_{\rm sig}(1 + f^{\pm}) = N^{\pm}_{\rm sig}(1 + \hat{f} \mp \Delta f). \tag{B.2}$$

Thus the measured charge asymmetry is a function of the true charge asymmetry,

$$A^{\rm obs}_{CP} = \frac{N^-_{\rm obs} - N^+_{\rm obs}}{N^-_{\rm obs} + N^+_{\rm obs}} = \frac{A^{\rm sig}_{CP}(1+\hat{f}) + \Delta f}{(1+\hat{f}) + A^{\rm sig}_{CP}\Delta f}. \tag{B.3}$$

From Equation B.3 we get an expression for the true charge asymmetry,

$$\Leftrightarrow A^{\rm sig}_{CP} = \frac{A^{\rm obs}_{CP}(1+\hat{f})}{(1+\hat{f})\left(1 - \frac{\Delta f A^{\rm obs}_{CP}}{(1+\hat{f})}\right)} - \frac{\Delta f}{(1+\hat{f})\left(1 - \frac{\Delta f A^{\rm obs}_{CP}}{(1+\hat{f})}\right)}. \tag{B.4}$$

To express Equation B.4 in terms of additive and multiplicative corrections, we approximate

$$\frac{1}{1-x} \approx 1 + x, \qquad \text{for} \quad x \ll 1, \tag{B.5}$$

and obtain

$$A^{\rm sig}_{CP} = A^{\rm obs}_{CP}\left(1 + \frac{\Delta f A^{\rm obs}_{CP}}{1+\hat{f}}\right) - \frac{\Delta f}{1+\hat{f}}\left(1 + \frac{\Delta f A^{\rm obs}_{CP}}{1+\hat{f}}\right) \tag{B.6}$$

$$= A^{\rm obs}_{CP} \cdot c_{\rm mult} + c_{\rm add}. \tag{B.7}$$

Since the involved numbers are small, $A^{\rm obs}_{CP} < 1$, $\Delta f/(1+\hat{f}) \approx 10^{-4}$, yielding $c_{\rm mult} - 1 \approx c_{\rm add} \approx 10^{-4}$, we conclude no correction has to be applied. Conservatively assuming $A^{\rm obs}_{CP} = 0.5$ we obtain the correction factors given in Table B.1. The dilution in the $B \to D^0\pi$ channels is negligible. From Equation B.7 we'll assign a systematic error of

$$\sigma_{\rm syst} = \sqrt{(0.5\,\sigma_{c_{\rm mult}})^2 + \sigma^2_{c_{\rm add}}}. \tag{B.8}$$

Table B.1: Correction factors to A_{CP} due to the signal self cross-feed component, $B \to D^0K$ channel. We will not apply these factors. We will, however, assign a systematic error based on their uncertainties.

$D^0 \to$	f^+ $(B \to D^0K)$	f^- $(B \to D^0K)$	$c_{\rm mult}$	$c_{\rm add}$
KK	0.0054 ± 0.0003	0.0065 ± 0.0003	1.0003 ± 0.0001	-0.0006 ± 0.0002
$\pi\pi$	0.0025 ± 0.0002	0.0030 ± 0.0002	1.0001 ± 0.0001	-0.0003 ± 0.0001
$K^0_S\pi^0$	0.0558 ± 0.0012	0.0568 ± 0.0012	1.0002 ± 0.0004	-0.0005 ± 0.0008
$K^0_S\omega$	0.1326 ± 0.0030	0.1312 ± 0.0029	0.9997 ± 0.0009	0.0006 ± 0.0018
$K^0_S\phi$	0.0053 ± 0.0004	0.0057 ± 0.0004	1.0001 ± 0.0001	-0.0002 ± 0.0003
$K\pi$	0.0074 ± 0.0003	0.0069 ± 0.0003	0.9999 ± 0.0001	0.0002 ± 0.0002
$K\pi\pi^0$	0.2423 ± 0.0028	0.2505 ± 0.0029	1.0016 ± 0.0008	-0.0033 ± 0.0016
$K\pi\pi\pi$	0.0460 ± 0.0009	0.0446 ± 0.0009	0.9997 ± 0.0003	0.0007 ± 0.0006

C Additional Material to Chapter 13

Table C.1: Fit result of the *CP*+ final fit to the Run 1-6 on-peak dataset.

Parameter	KK	$\pi\pi$
$A_{CP}^{\mathrm{sig}(K)}$	0.242 ± 0.064	
$A_{CP}^{\mathrm{sig}(\pi)}$	0.003 ± 0.015	
$R_{K/\pi}$	0.0899 ± 0.0063	
m	0.0205 ± 0.0030	
$A_{CP,f}^{B\overline{B}}$	-0.004 ± 0.045	-0.045 ± 0.047
$A_{CP,p}^{q\overline{q}}$	0.012 ± 0.016	-0.016 ± 0.018
$A_{CP,f}^{q\overline{q}}$	-0.004 ± 0.011	-0.0047 ± 0.0098
$f_{\mathcal{F}}^{q\overline{q}}$	0.326 ± 0.026	0.49
$\sigma_{\mathcal{F},l,1}^{q\overline{q}}$	0.160 ± 0.016	0.258 ± 0.024
$\sigma_{\mathcal{F},l,2}^{q\overline{q}}$	0.1742 ± 0.0020	0.2047 ± 0.0024
$\sigma_{\mathcal{F},r,1}^{q\overline{q}}$	0.312 ± 0.011	0.329 ± 0.011
$\sigma_{\mathcal{F},r,2}^{q\overline{q}}$	0.230 ± 0.014	0.269 ± 0.018
$a_{\Delta E}^{q\overline{q}}$	-0.96 ± 0.14	-0.72 ± 0.14
$\mu_{\Delta E}$	-2.62 ± 0.32	-1.36 ± 0.57
$\sigma_{\Delta E}$	16.63 ± 0.27	14.82 ± 0.49
$\mu_{m_{\mathrm{ES}}}$	5278.56 ± 0.12	5278.61 ± 0.20
$\sigma_{m_{\mathrm{ES}},l}$	2.207 ± 0.081	2.12 ± 0.15
$\sigma_{m_{\mathrm{ES}},r}$	2.896 ± 0.081	2.83 ± 0.15
$N_p^{B\overline{B}}$	79 ± 28	345 ± 52
$N_f^{B\overline{B}}$	1428 ± 81	1516 ± 143
$N_p^{q\overline{q}}$	4005 ± 69	3457 ± 76
$N_f^{q\overline{q}}$	10893 ± 125	13030 ± 176
$N_{\mathrm{tot}}^{\mathrm{sig}(\pi)}$	4091 ± 70	1230 ± 41

Table C.2: Fit result of the *CP*– final fit to the Run 1-6 on-peak dataset.

Parameter	$K^0_S\pi^0$	$K^0_S\omega$	$K^0_S\phi$
$A_{CP}^{\mathrm{sig}(K)}$		-0.088 ± 0.066	
$A_{CP}^{\mathrm{sig}(\pi)}$		-0.009 ± 0.014	
$R_{K/\pi}$		0.0812 ± 0.0056	
m		0.0207 ± 0.0029	
$A_{CP,p}^{q\overline{q}}$	-0.002 ± 0.011	0.012 ± 0.020	-0.069 ± 0.060
$A_{CP,f}^{q\overline{q}}$	0.0021 ± 0.0071	-0.004 ± 0.013	0.001 ± 0.039
$f_{\mathcal{F}}^{q\overline{q}}$	0.520 ± 0.030	0.27	n/a
$\sigma_{\mathcal{F},l,1}^{q\overline{q}}$	0.206 ± 0.014	0.175 ± 0.034	0.2758 ± 0.0092
$\sigma_{\mathcal{F},l,2}^{q\overline{q}}$	0.1546 ± 0.0015	0.1963 ± 0.0028	n/a
$\sigma_{\mathcal{F},r,1}^{q\overline{q}}$	0.3541 ± 0.0068	0.317 ± 0.019	0.447 ± 0.014
$\sigma_{\mathcal{F},r,2}^{q\overline{q}}$	0.275 ± 0.020	0.238 ± 0.013	n/a
$a_{\Delta E}^{q\overline{q}}$	-0.926 ± 0.099	-1.04 ± 0.18	-0.48
$\mu_{\Delta E}$	-1.80 ± 0.35	-2.88 ± 0.59	-0.95 ± 0.75
$\sigma_{\Delta E}$	17.01 ± 0.29	16.10 ± 0.52	15.82 ± 0.60
$\mu_{m_{\mathrm{ES}}}$	5278.62 ± 0.12	5278.50 ± 0.22	5278.99 ± 0.25
$\sigma_{m_{\mathrm{ES}},l}$	2.299 ± 0.084	2.11 ± 0.16	2.33 ± 0.17
$\sigma_{m_{\mathrm{ES}},r}$	2.922 ± 0.084	3.08 ± 0.16	2.72 ± 0.17
$N_p^{B\overline{B}}$	176 ± 43	180 ± 48	3
$N_f^{B\overline{B}}$	1929 ± 102	1195 ± 109	119 ± 20
$N_p^{q\overline{q}}$	8588 ± 101	2675 ± 68	284 ± 17
$N_f^{q\overline{q}}$	21659 ± 172	6673 ± 124	716 ± 29
$N_{\mathrm{tot}}^{\mathrm{sig}(\pi)}$	4181 ± 73	1440 ± 45	648 ± 27

Table C.3: Fit results of the three flavor final fits ($D^0 \to K^-\pi^+$, $K^-\pi^+\pi^0$, $K^-\pi^+\pi^+\pi^-$) to the Run 1-6 on-peak dataset.

Parameter	$K\pi$	$K\pi\pi^0$	$K\pi\pi\pi$
$A_{CP}^{\mathrm{sig}(K)}$	-0.008 ± 0.022	-0.030 ± 0.019	-0.012 ± 0.020
$A_{CP}^{\mathrm{sig}(\pi)}$	-0.0116 ± 0.0050	-0.0118 ± 0.0044	0.0005 ± 0.0048
$A_{CP,f}^{b\overline{b}}$	-0.043 ± 0.017	-0.0228 ± 0.0081	-0.017 ± 0.011
$A_{CP,p}^{q\overline{q}}$	-0.027 ± 0.016	-0.0218 ± 0.0069	-0.020 ± 0.011
$A_{CP,f}^{q\overline{q}}$	-0.0017 ± 0.0068	-0.0056 ± 0.0030	-0.0007 ± 0.0048
$f_{\mathcal{F}}^{q\overline{q}}$	0.397 ± 0.018	0.357 ± 0.011	0.394 ± 0.012
$\sigma_{\mathcal{F},l,1}^{q\overline{q}}$	0.198 ± 0.014	0.1646 ± 0.0058	0.1390 ± 0.0052
$\sigma_{\mathcal{F},l,2}^{q\overline{q}}$	0.1965 ± 0.0017	0.17673 ± 0.00067	0.1868 ± 0.0010
$\sigma_{\mathcal{F},r,1}^{q\overline{q}}$	0.3068 ± 0.0061	0.3197 ± 0.0044	0.2808 ± 0.0057
$\sigma_{\mathcal{F},r,2}^{q\overline{q}}$	0.237 ± 0.010	0.2419 ± 0.0066	0.1944 ± 0.0075
$R_{K/\pi}$	0.0753 ± 0.0018	0.0768 ± 0.0017	0.0806 ± 0.0018
$\sigma_{\Delta E,m_{\mathrm{ES}}}^{B\overline{B}}$	0.01048 ± 0.00057	n/a	n/a
$a_{\Delta E}^{q\overline{q}}$	-0.88 ± 0.10	-1.547 ± 0.063	-1.075 ± 0.068
$b_{\Delta E}^{q\overline{q}}$	n/a	6.33 ± 0.90	n/a
$\mu_{\Delta E}$	-1.527 ± 0.092	-2.075 ± 0.098	-2.280 ± 0.091
$\sigma_{\Delta E}$	15.424 ± 0.076	16.493 ± 0.081	15.518 ± 0.075
$\mu_{m_{\mathrm{ES}}}$	5278.586 ± 0.033	5278.591 ± 0.032	5278.549 ± 0.033
$\sigma_{m_{\mathrm{ES}},l}$	2.210 ± 0.022	2.249 ± 0.023	2.137 ± 0.023
$\sigma_{m_{\mathrm{ES}},r}$	2.852 ± 0.023	2.948 ± 0.024	2.901 ± 0.024
$N_p^{B\overline{B}}$	327 ± 40	3133 ± 127	1523 ± 85
$N_f^{B\overline{B}}$	7719 ± 170	43753 ± 487	20437 ± 337
$N_p^{q\overline{q}}$	4722 ± 77	24285 ± 183	10382 ± 123
$N_f^{q\overline{q}}$	28009 ± 205	151106 ± 551	63269 ± 380
$N_{\mathrm{tot}}^{\mathrm{sig}(\pi)}$	44630 ± 232	76848 ± 352	52723 ± 261
m	0.02143 ± 0.00089	0.02061 ± 0.00089	0.02123 ± 0.00090

C.1 $D^0 \to K^- K^+$

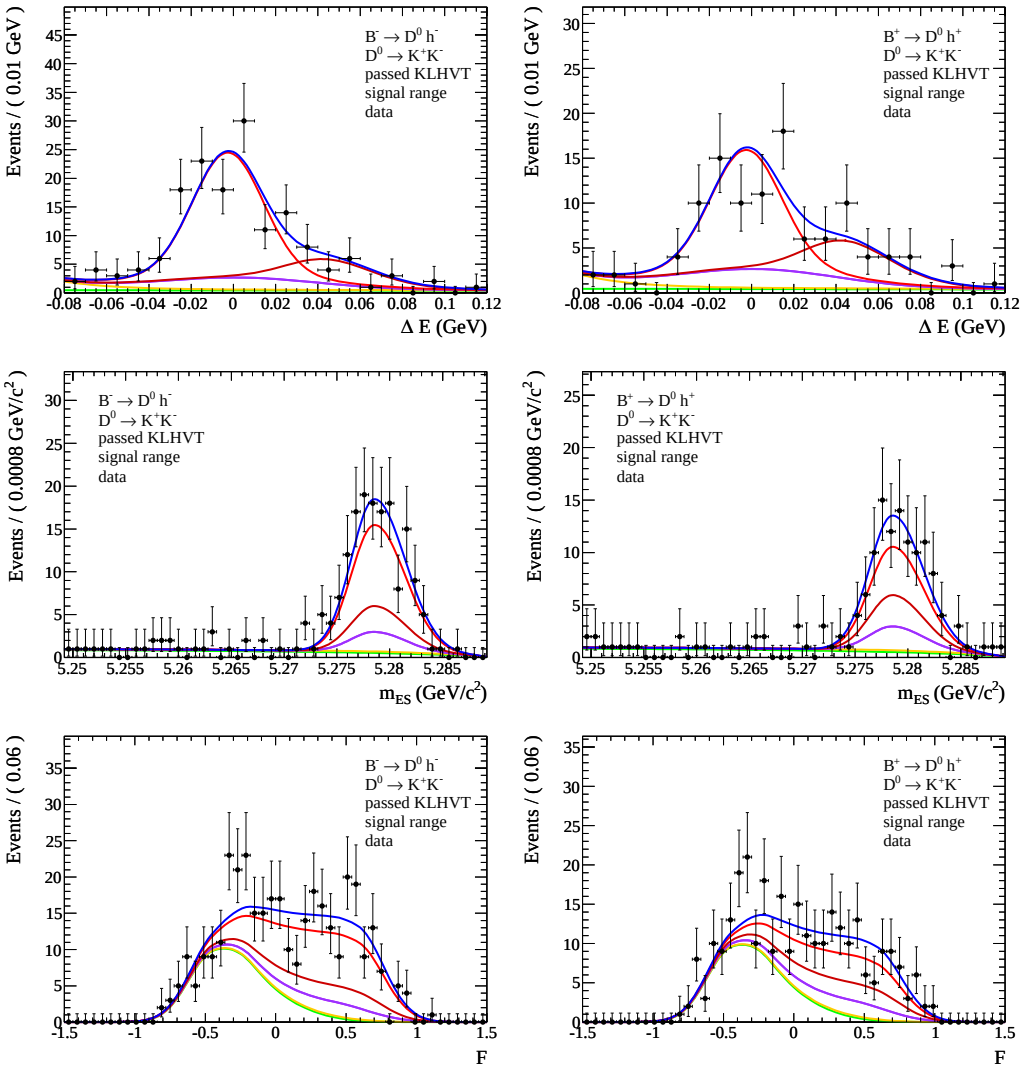

Figure C.1: Projections of the *CP*+ final fit to $B \to D^0 h$ Run 1-6 on-peak data, $D^0 \to K^- K^+$ slice. The bachelor track h is required to pass the `KLHVeryTight` selector. Left: B^- subsample. Right: B^+ subsample. Color code in Table 10.4 on page 69.

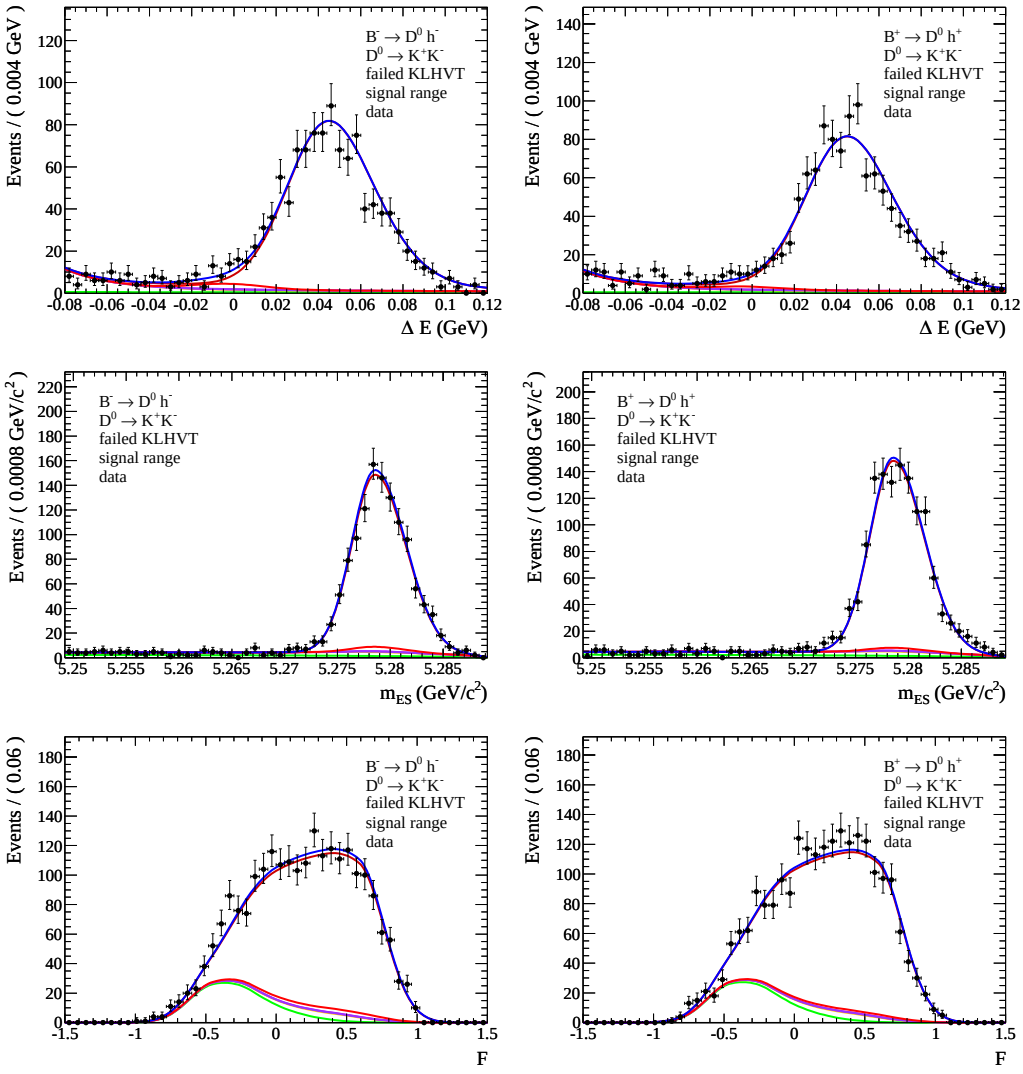

Figure C.2: Projections of the *CP*+ final fit to $B \to D^0 h$ Run 1-6 on-peak data, $D^0 \to K^- K^+$ slice. The bachelor track h is required to fail the `KLHVeryTight` selector. Left: B^- subsample. Right: B^+ subsample. Color code in Table 10.4 on page 69.

C.2 $D^0 \to \pi^+\pi^-$

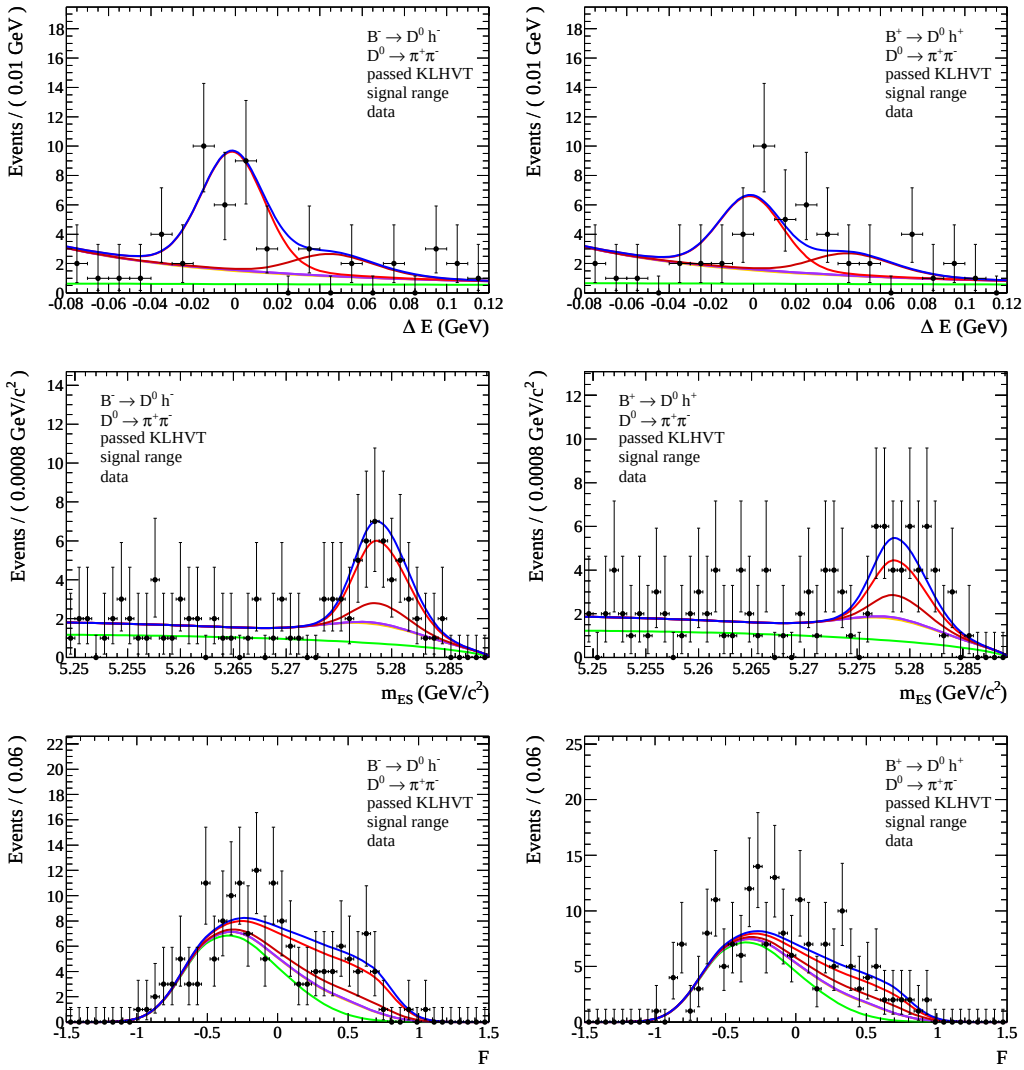

Figure C.3: Projections of the *CP*+ final fit to $B \to D^0 h$ Run 1-6 on-peak data, $D^0 \to \pi^+\pi^-$ slice. The bachelor track h is required to pass the `KLHVeryTight` selector. Left: B^- subsample. Right: B^+ subsample. Color code in Table 10.4 on page 69.

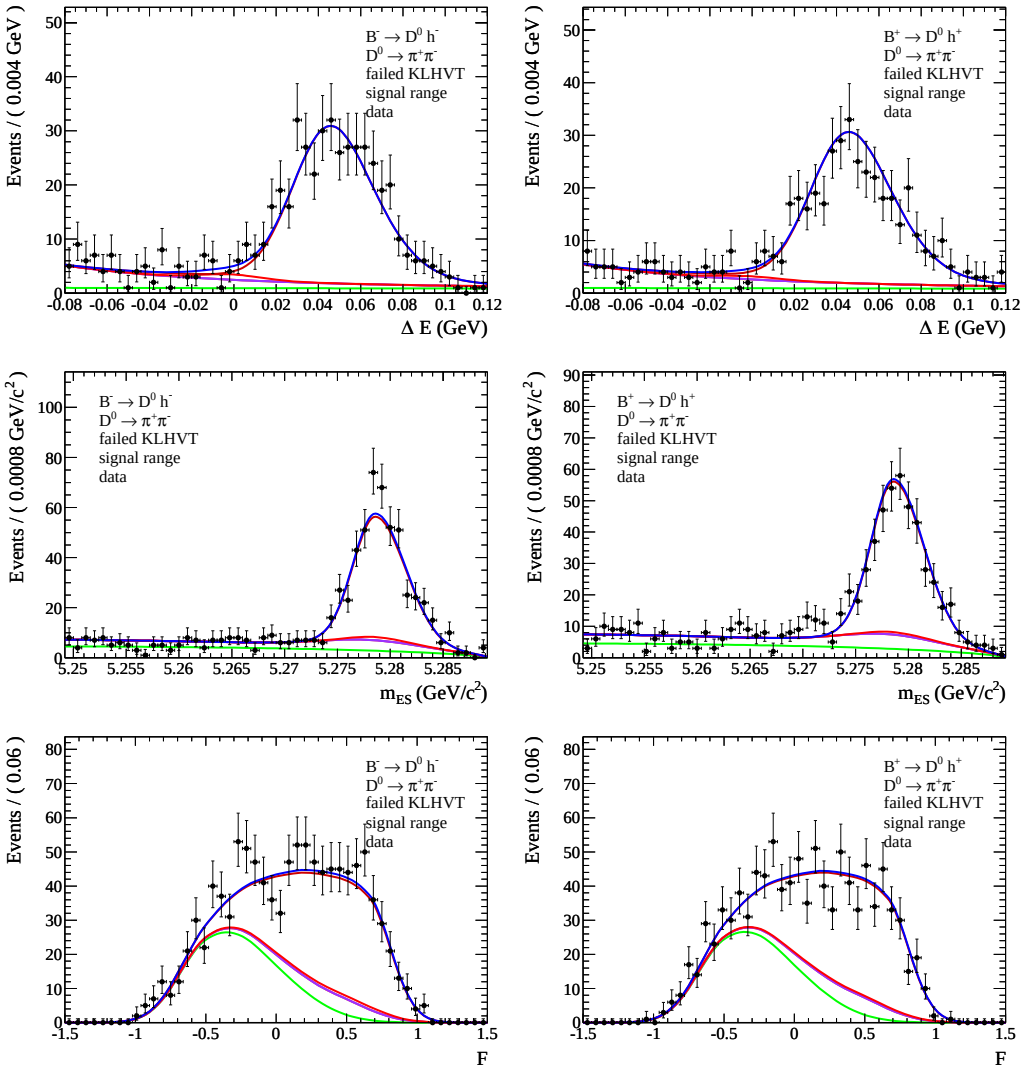

Figure C.4: Projections of the *CP*+ final fit to $B \to D^0 h$ Run 1-6 on-peak data, $D^0 \to \pi^+\pi^-$ slice. The bachelor track h is required to fail the `KLHVeryTight` selector. Left: B^- subsample. Right: B^+ subsample. Color code in Table 10.4 on page 69.

C.3 $D^0 \to K^0_S \pi^0$

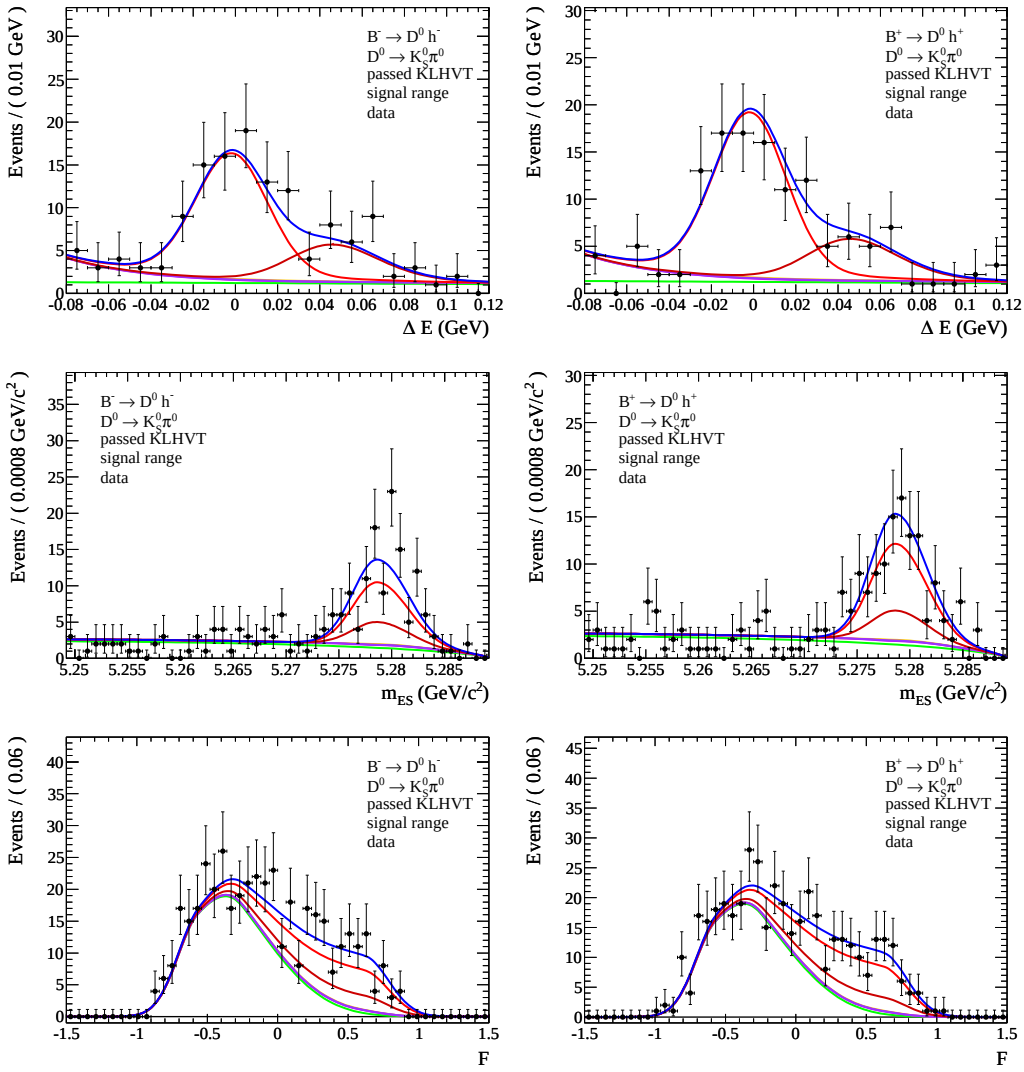

Figure C.5: Projections of the *CP*+ final fit to $B \to D^0 h$ Run 1-6 on-peak data, $D^0 \to K^0_S \pi^0$ slice. The bachelor track h is required to pass the `KLHVeryTight` selector. Left: B^- subsample. Right: B^+ subsample. Color code in Table 10.4 on page 69.

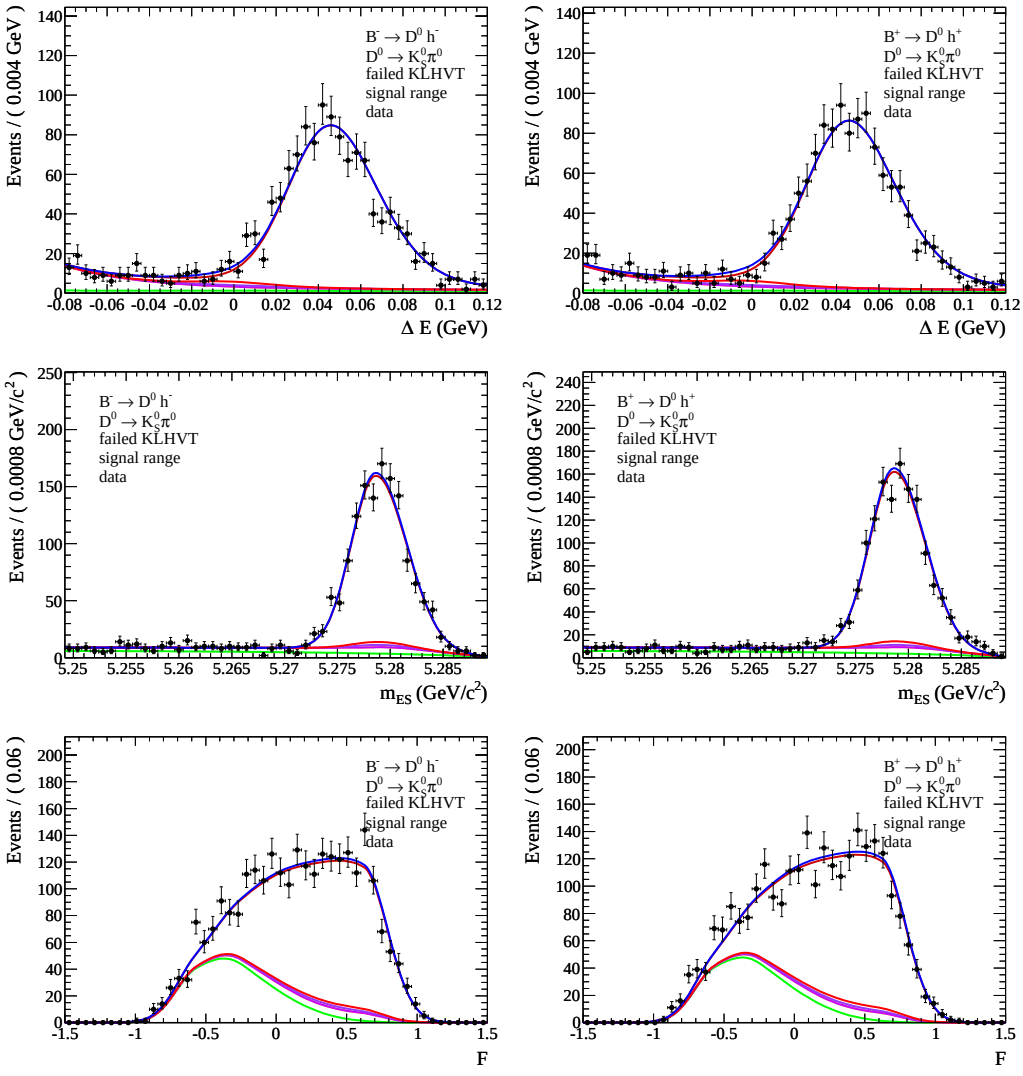

Figure C.6: Projections of the *CP*+ final fit to $B \to D^0 h$ Run 1-6 on-peak data, $D^0 \to K^0_S \pi^0$ slice. The bachelor track h is required to fail the `KLHVeryTight` selector. Left: B^- subsample. Right: B^+ subsample. Color code in Table 10.4 on page 69.

C.4 $D^0 \to K^0_S\omega$

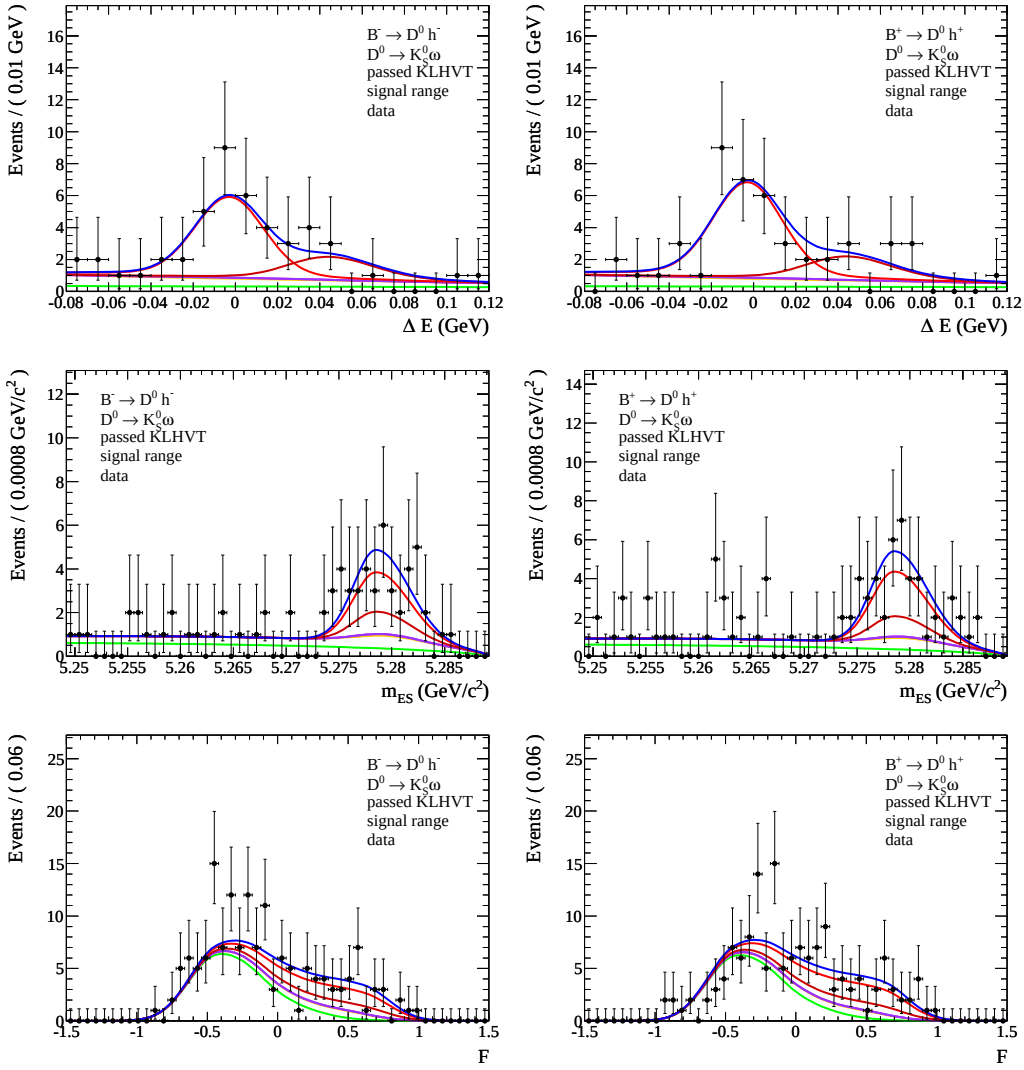

Figure C.7: Projections of the *CP*+ final fit to $B \to D^0 h$ Run 1-6 on-peak data, $D^0 \to K^0_S\omega$ slice. The bachelor track h is required to pass the `KLHVeryTight` selector. Left: B^- subsample. Right: B^+ subsample. Color code in Table 10.4 on page 69.

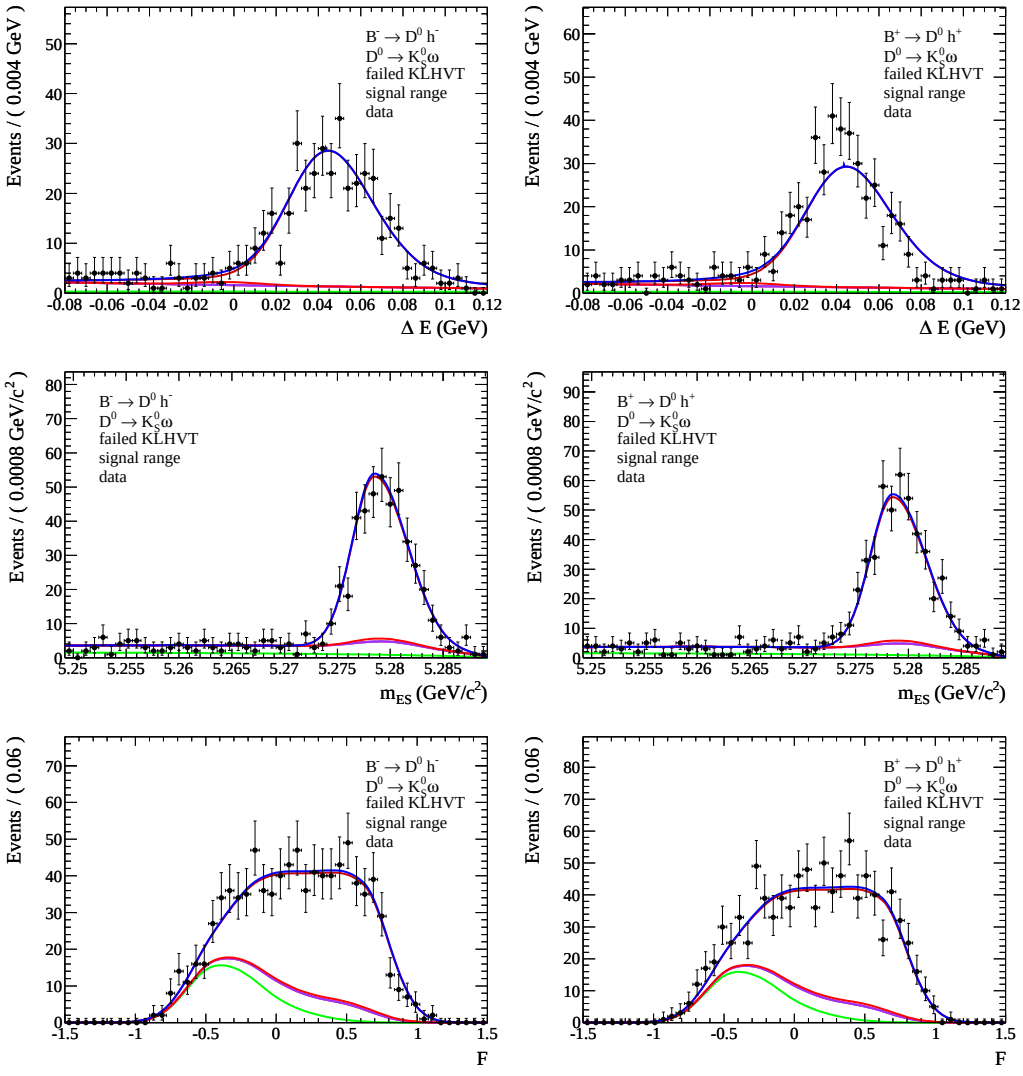

Figure C.8: Projections of the *CP*+ final fit to $B \to D^0 h$ Run 1-6 on-peak data, $D^0 \to K^0_S \omega$ slice. The bachelor track h is required to fail the `KLHVeryTight` selector. Left: B^- subsample. Right: B^+ subsample. Color code in Table 10.4 on page 69.

C.5 $D^0 \to K^0_S \phi$

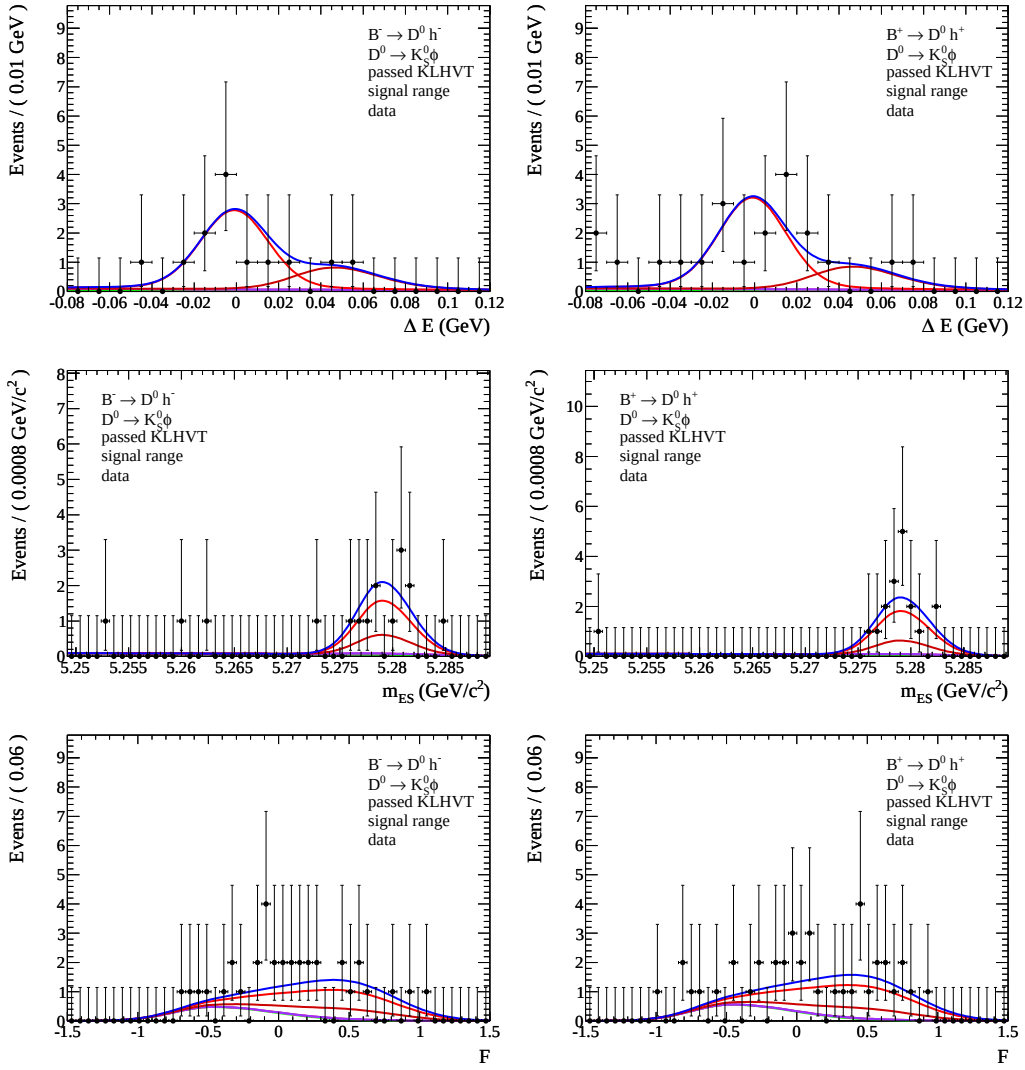

Figure C.9: Projections of the *CP*+ final fit to $B \to D^0 h$ Run 1-6 on-peak data, $D^0 \to K^0_S \phi$ slice. The bachelor track h is required to pass the `KLHVeryTight` selector. Left: B^- subsample. Right: B^+ subsample. Color code in Table 10.4 on page 69.

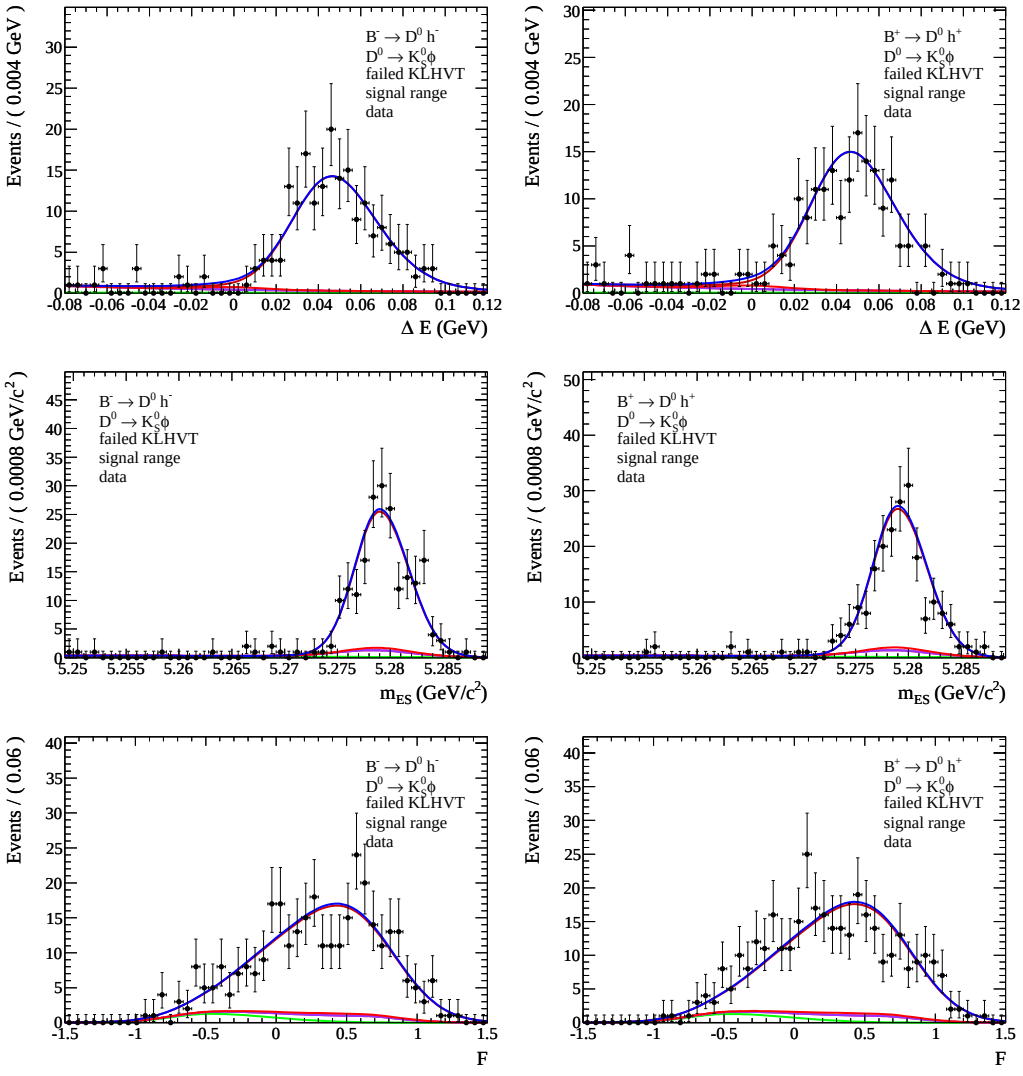

Figure C.10: Projections of the *CP*+ final fit to $B \to D^0 h$ Run 1-6 on-peak data, $D^0 \to K^0_S \phi$ slice. The bachelor track h is required to fail the `KLHVeryTight` selector. Left: B^- subsample. Right: B^+ subsample. Color code in Table 10.4 on page 69.

C.6 $D^0 \to K^-\pi^+$

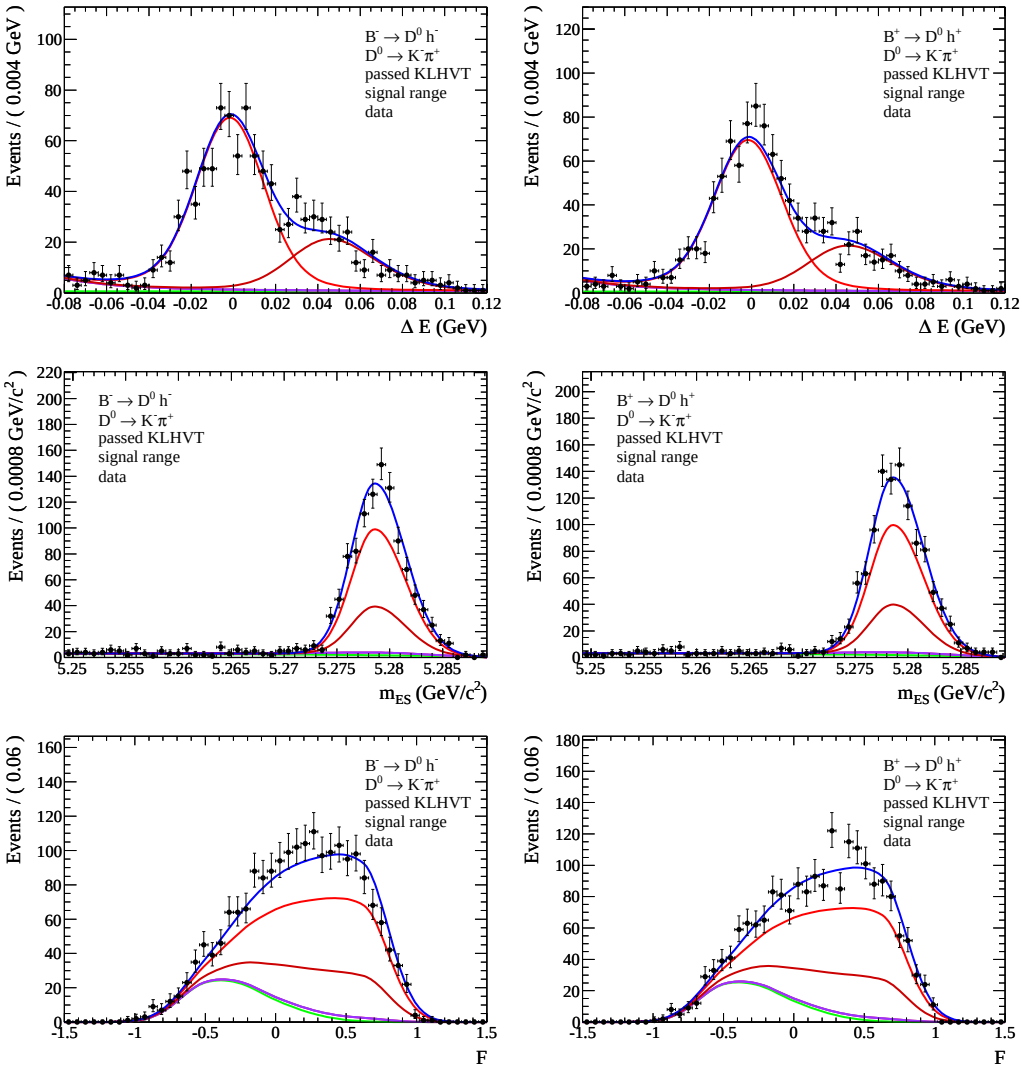

Figure C.11: Projections of the *CP*+ final fit to $B \to D^0 h$ Run 1-6 on-peak data, $D^0 \to K^-\pi^+$ slice. The bachelor track h is required to pass the `KLHVeryTight` selector. Left: B^- subsample. Right: B^+ subsample. Color code in Table 10.4 on page 69.

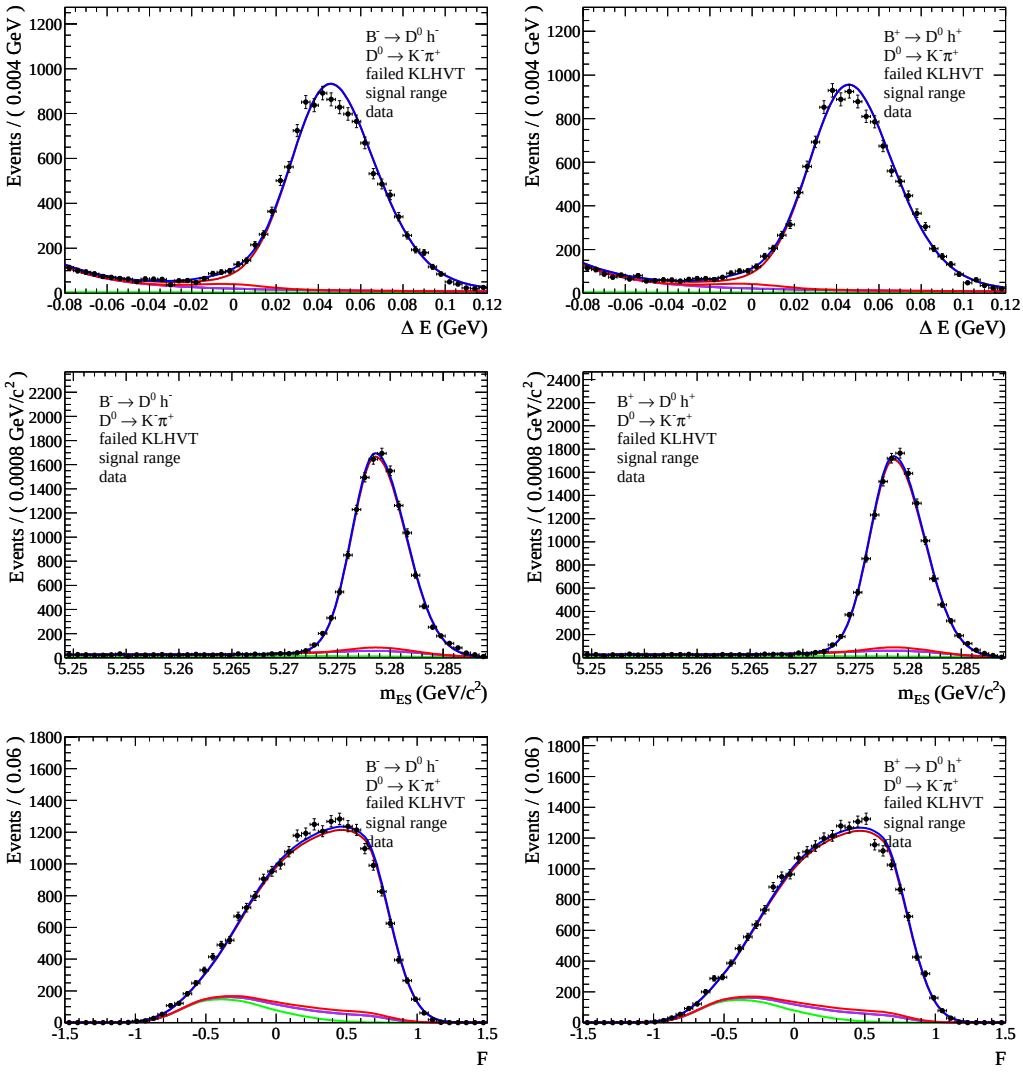

Figure C.12: Projections of the *CP*+ final fit to $B \to D^0 h$ Run 1-6 on-peak data, $D^0 \to K^-\pi^+$ slice. The bachelor track h is required to fail the `KLHVeryTight` selector. Left: B^- subsample. Right: B^+ subsample. Color code in Table 10.4 on page 69.

C.7 $D^0 \to K^-\pi^+\pi^0$

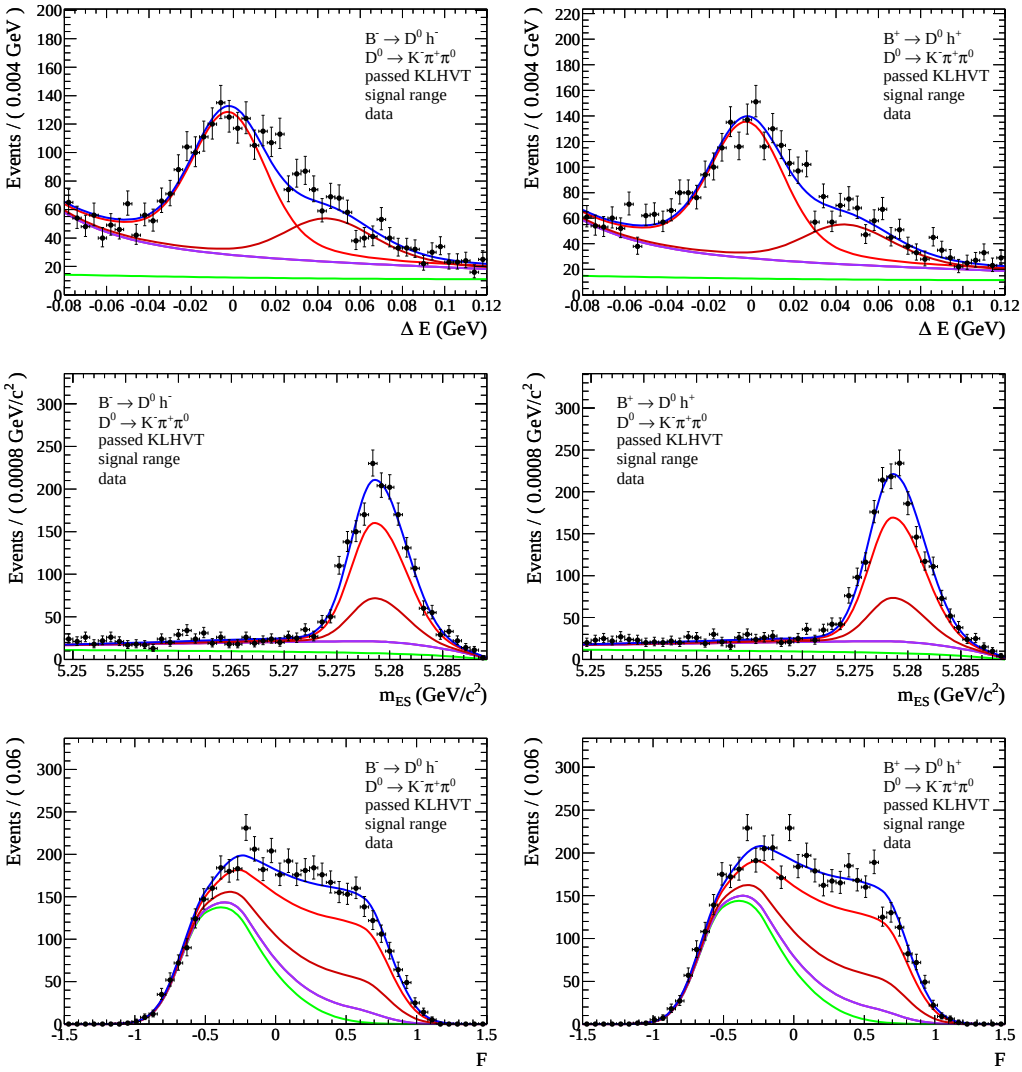

Figure C.13: Projections of the *CP*+ final fit to $B \to D^0 h$ Run 1-6 on-peak data, $D^0 \to K^-\pi^+\pi^0$ slice. The bachelor track h is required to pass the `KLHVeryTight` selector. Left: B^- subsample. Right: B^+ subsample. Color code in Table 10.4 on page 69.

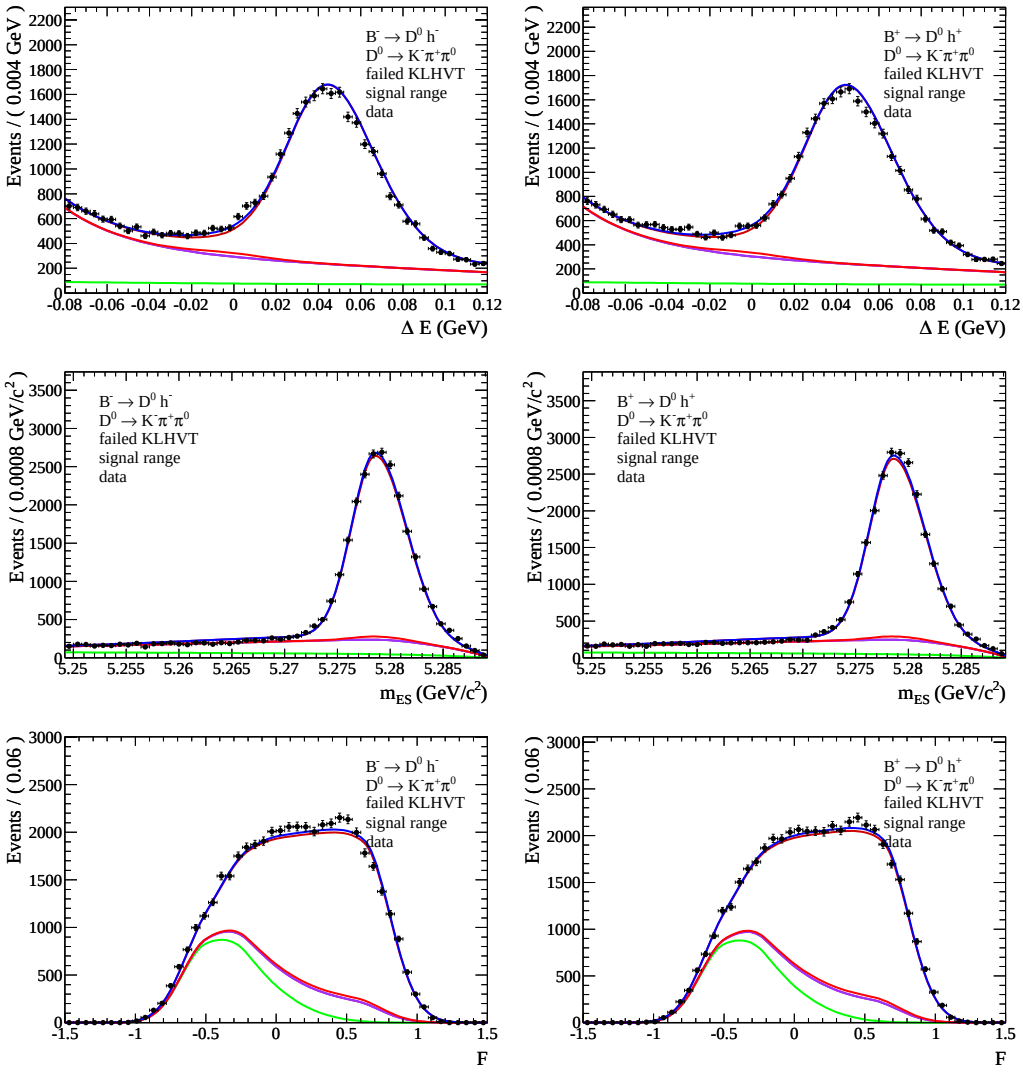

Figure C.14: Projections of the *CP*+ final fit to $B \to D^0 h$ Run 1-6 on-peak data, $D^0 \to K^-\pi^+\pi^0$ slice. The bachelor track h is required to fail the `KLHVeryTight` selector. Left: B^- subsample. Right: B^+ subsample. Color code in Table 10.4 on page 69.

C.8 $D^0 \to K^-\pi^+\pi^+\pi^-$

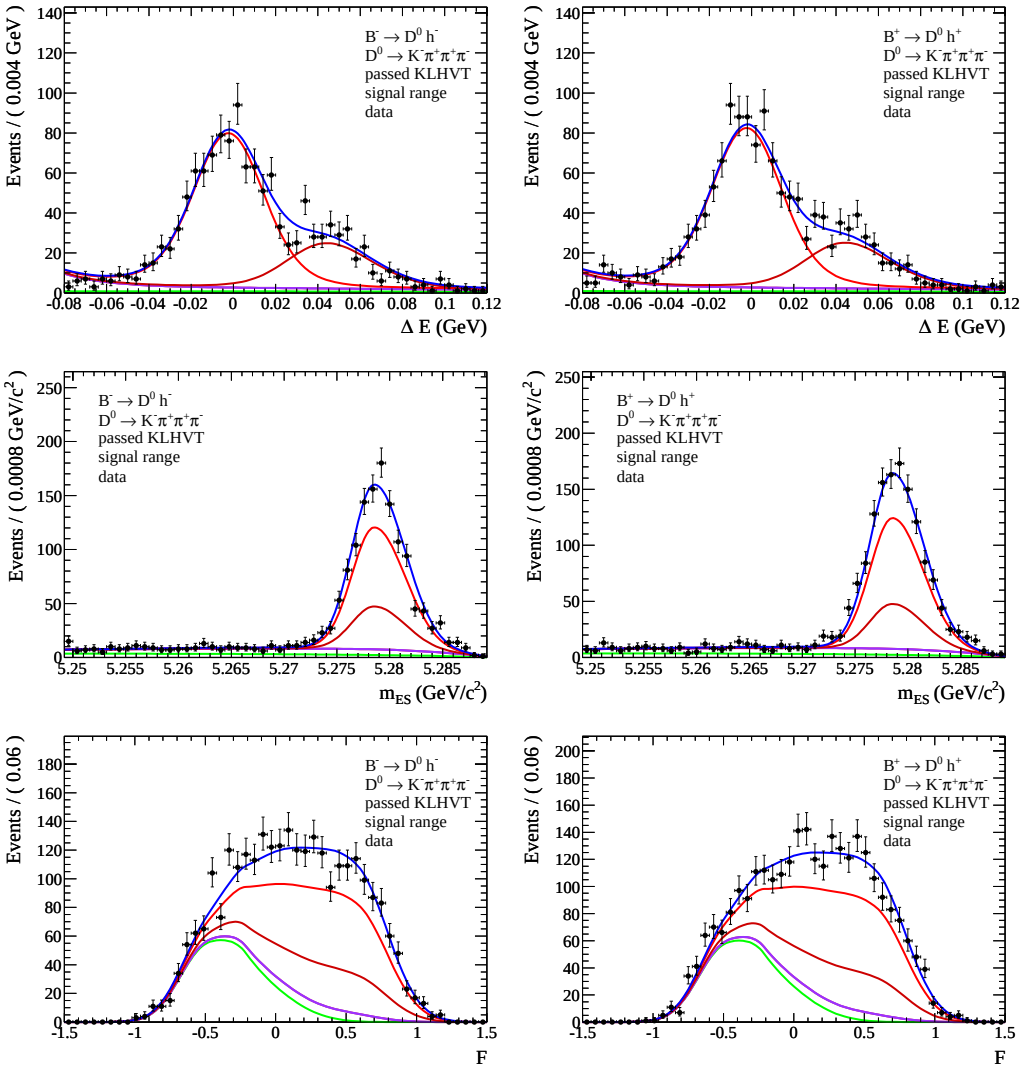

Figure C.15: Projections of the *CP*+ final fit to $B \to D^0 h$ Run 1-6 on-peak data, $D^0 \to K^-\pi^+\pi^+\pi^-$ slice. The bachelor track h is required to pass the `KLHVeryTight` selector. Left: B^- subsample. Right: B^+ subsample. Color code in Table 10.4 on page 69.

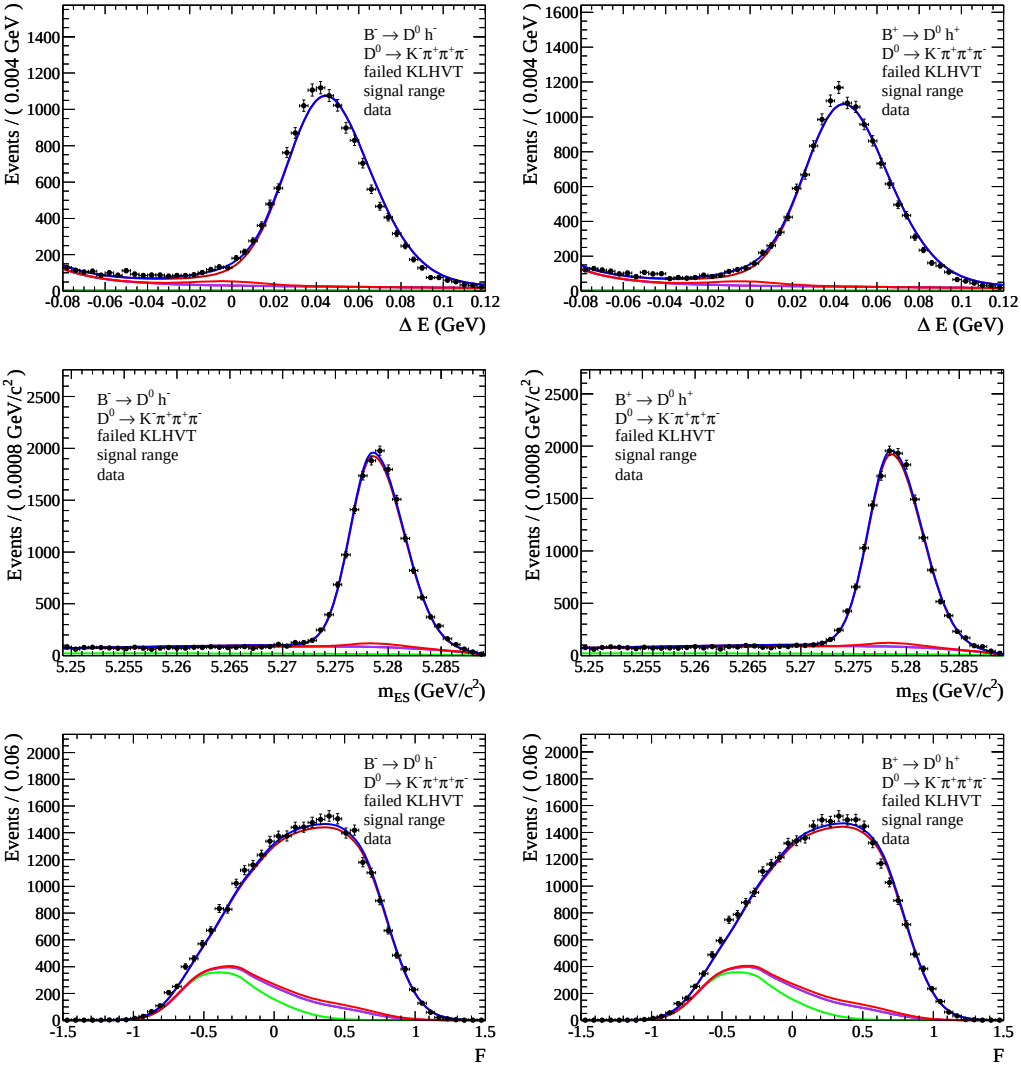

Figure C.16: Projections of the *CP*+ final fit to $B \to D^0 h$ Run 1-6 on-peak data, $D^0 \to K^-\pi^+\pi^+\pi^-$ slice. The bachelor track h is required to fail the `KLHVeryTight` selector. Left: B^- subsample. Right: B^+ subsample. Color code in Table 10.4 on page 69. The mismatch in ΔE is due to a plotting issue. The parameterization of the conditional $\Delta E_{\texttt{shift}}$ used to make the projection plots was optimized to describe the $D^0 \to K^-\pi^+\pi^0$ channel (see Section 10.3).

Bibliography

[1] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).

[2] M. Gronau and D. Wyler, Phys. Lett. **B265**, 172 (1991).

[3] M. Gronau and D. London, Phys. Lett. **B253**, 483 (1991).

[4] *BABAR*, B. A. et al., Phys. Rev. D **73**, 051105(R) (2006).

[5] *BABAR*, B. A. et al., Phys. Rev. D **77**, 111102 (2008).

[6] G. Branco, L. Lavoura, and J. Silva, *CP Violation* (Oxford Press, 1999).

[7] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983).

[8] Particle Data Group, C. A. et al., Phys. Lett. **B667**, 1 (2008).

[9] *BABAR*, B. A. et al., Phys. Rev. Lett. **99**, 171803 (2007).

[10] *BABAR*, B. A. et al., Phys. Rev. D **76**, 052007 (2007).

[11] *BABAR*, B. A. et al., (2007), arXiv:0708.1630.

[12] J. Nardulli, (2008), arXiv:0807.1670, on behalf of the LHCb Collaboration.

[13] CKMfitter Group, J. C. et al., http://ckmfitter.in2p3.fr/plots_Summer2008/#etiquette4.

[14] M. Gronau, Phys. Rev. D **58**, 037301 (1998).

[15] *BABAR*, B. A. et al., (2007), arXiv:0708.0182, submitted to Physical Review Letters.

[16] Belle, A. P. et al., Phys. Rev. D **73**, 112009 (2006).

[17] C. C. Meca and J. P. Silva, Phys. Rev. Lett. **81**, 1377 (1998).

[18] *BABAR*, B. A. et al., Phys. Rev. D **78**, 011105 (2008).

[19] *BABAR*, B. A. et al., (2008), arXiv:0807.4544.

[20] Belle, K. A. et al., Phys. Rev. Lett. **99**, 131803 (2007).

[21] D. Atwood, I. Dunietz, and A. Soni, Phys. Rev. Lett. **78**, 3257 (1997).

[22] D. Atwood, I. Dunietz, and A. Soni, Phys. Rev. D **68**, 036005 (2001).

[23] A. Giri, Y. Grossman, A. Soffer, and J. Zupan, Phys. Rev. D **68**, 054018 (2003).

[24] *BABAR*, B. A. et al., Phys. Rev. D **78**, 034023 (2008).

[25] Belle, K. A. et al., Phys. Rev. D **73**, 051106 (2006).

[26] CDF, K. G. et al., arXiv:0809.4809, ICHEP08 preliminary result.

[27] HFAG Group, http://www.slac.stanford.edu/xorg/hfag/triangle/summer2008/index.shtml.

[28] ROOT, http://root.cern.ch.

[29] Comp. Phys. Comm. **10** (1972).

[30] TMVA, http://tmva.sourceforge.net.

[31] ROOFIT, http://root.cern.ch.

[32] R. D. Cousins, Am. J. Phys. **63**, 398 (1995).

[33] *BABAR*, B. A. et al., Nucl. Instrum. Methods Phys. Res., Sect. A **479**, 1 (2002).

[34] D. J. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A **462**, 152 (2001).

[35] GEANT4 Collaboration, S. A. et al., Nucl. Instrum. Methods Phys. Res., Sect. A **506**, 250 (2003).

[36] Belle, K. A. et al., Phys. Rev. D **73**, 051106(R) (2006).

[37] V. Eyges, T. M. Karbach, V. Santoro, and G. Marchiori, *BABAR* Analysis Document 1739, 2008.

[38] R. A. Fisher, Annals Eugen. **7**, 179 (1936).

[39] G. C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978).

[40] J. O. et al., *BABAR* Analysis Document 346, 2002.

[41] V. Blobel and E. Lohrmann, *Statistische und numerische Methoden der Datenanalyse* (Teubner, 1998).

[42] T. M. Karbach and G. Marchiori, *BABAR* Analysis Document 2105, 2008.

[43] The Landau function parameterizes energy loss in material and has no analytical form: $p(x) = \frac{1}{\pi}\int_0^\infty e^{-t\ln t - xt}\sin(\pi t)\,\mathrm{d}t$.

[44] The Crystal Ball function is defined as $\frac{n^n}{|\alpha|^n}\exp\left(-|\alpha|^2/2\right) / \left(\frac{n}{|\alpha|} - |\alpha| - \bar{x}\right)^n$ for $\bar{x} < -|\alpha|$, $\exp\left(-\frac{1}{2}\bar{x}^2\right)$ for $\bar{x} \geq -|\alpha|$; $\bar{x} = \frac{x-\mu}{\sigma}$, $\bar{x} \to -\bar{x}$ for $\alpha < 0$.

[45] The Novosibirsk function is a modified Gaussian with an extra tail parameter that skews the Gaussian into an asymmetric shape with a long tail on one side and a short tail on the other: $f(m_{ES}) = A_S\exp(-0.5\ln^2[1+\Lambda\tau\cdot(m_{ES}-m_0)]/\tau^2+\tau^2)$, where $\Lambda = \sinh(\tau\sqrt{\ln 4})/(\sigma\tau\sqrt{\ln 4})$, the peak position is m_0, the width is σ, and τ is the tail parameter.

[46] ARGUS, H. Albrecht, Phys. Lett. B **241**, 278 (1990), the Argus function is an empirical formula to model the phase space of multi-body decays near threshold: $f(x) = x\cdot\left[1-(x/c)^2\right]^p\exp\left\{-\chi\cdot\left(1-(x/c)^2\right)\right\}$.

[47] G. Marchiori, S. Sitt, and W. Verkerke, private communication, 2009.

[48] *BABAR*, B. A. et al., Phys. Rev. D **78**, 012004 (2008), arXiv:0803.4451.

[49] G. Batignani, G. Marchiori, and M. Rama, *BABAR* Analysis Document 932, 2005.