SLAC-R-833
August 2006

Measurement of CP Parameters in B- --> D(pi+pi-pi0)K- and Study of the X(3872) in B --> J/psi pi+ pi- K with the BaBar Detector

By Frank Winklmeier

SLAC-R-833

Prepared for the Department of Energy
under contract number DE-AC02-76SF00515

SLAC-R-833

DISSERTATION

MEASUREMENT OF CP PARAMETERS IN $B^- \to D_{\pi^+\pi^-\pi^0} K^-$ AND STUDY OF THE $X(3872)$ IN $B \to J/\psi\, \pi^+\pi^- K$ WITH THE BABAR DETECTOR

Submitted by

Frank Winklmeier

Physics Department

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2006

ABSTRACT OF DISSERTATION

MEASUREMENT OF CP PARAMETERS IN $B^- \rightarrow D_{\pi^+\pi^-\pi^0} K^-$
AND STUDY OF THE $X(3872)$ IN $B \rightarrow J/\psi\,\pi^+\pi^- K$
WITH THE BABAR DETECTOR

This dissertation presents two analyses performed on data collected with the *BABAR* detector at the SLAC PEP-II e^+e^- asymmetric-energy B Factory. First, a Dalitz analysis is shown that performs the first measurement of CP violation parameters in the decay $B^- \rightarrow D_{\pi^+\pi^-\pi^0} K^-$ using the decay rate asymmetry and $D^0 - \overline{D}^0$ interference. The results can be used to further constrain the value of the CKM angle γ. The second analysis studies the properties of the $X(3872)$ in neutral and charged $B \rightarrow J/\psi\,\pi^+\pi^- K$ decays. Measurements of the branching ratio and mass are presented as well as the search for additional resonances at higher masses.

Frank Winklmeier
Physics Department
Colorado State University
Fort Collins, CO 80523
Fall 2006

PREFACE AND ACKNOWLEDGMENTS

This dissertation describes part of my research that I conducted during my PhD studies within the Colorado State University (CSU) High Energy Physics group working on the *BABAR* experiment at the Stanford Linear Accelerator Center (SLAC). It described two analysis in separate areas of the *BABAR* physics program: *CP* violation and rare decays. Unavoidably, the dissertation is therefore split into two more or less independent parts.

Chapter 1 gives an introduction to *CP* violation and the CKM quark mixing matrix. After summarizing the current status of the measurements, three different methods to measure the CKM angle γ are presented. The last part of this chapter is dedicated to the experimental and theoretical results for the $X(3872)$, a *new state* that continues to challenge the established quark models since its discovery almost two years ago. Chapter 2 gives an overview of analysis techniques employed in the two analysis of this dissertation. Besides introducing methods specifically used in High Energy Physics experiments it also presents general statistical methods as maximum likelihood estimators and neural networks. Chapter 3 covers the measurement of *CP* parameters in the decay $B^- \to (\pi^0\pi^+\pi^-)_D K^-$ using a Dalitz analysis. This is the first measurement in this mode and the first analysis that combines information from the Dalitz shape and the decay rate asymmetry to increase the sensitivity on γ. Chapter 4 describes the study of the $X(3872)$ in $B \to J/\psi\,\pi^+\pi^- K$ decays. It is the first analysis where the neutral and charged B decays were analyzed separately to uncover possible different properties of the $X(3872)$ in these two decay modes. Due to lack of space and the fact, that it has been written up in my Master thesis [1], this dissertation does not include the usual detector part with a description of the *BABAR* detector. Please see [2] for detailed technical information.

This work was only possible due to the enormous effort of the entire *BABAR* collaboration in running the machine, the detector, the computing infrastructure, the people in the analysis working groups and tools groups, and many more. I want to especially thank my advisor Walter Toki for giving me the opportunity to work on three exciting analyses, for the excellent support, guidance and supervision I received from him throughout the program, for all the cheerful moments and for the culinary experiences on our travels and at his own place. During my stay at SLAC and the work on the *CP* analysis I was working very closely with Abi Soffer, who I want to thank for his tremendous help, for sharing his never ending optimism when we faced yet another problem in the analysis, for countless discussions about object-oriented data modeling and for teaching me the subtleties of C++. It was a true pleasure to work with him. Of course, there were countless other people who made this work enjoyable and fun and who I would like to thank at this place: my committee members John Harton, Robert Wilson and Alexander Hulpke for reading this document on a two week notice, the faculty at CSU for welcoming me in their department from the very first day, Prof. Wolfgang Hüttner in Ulm for his help in initiating the contact with CSU, the wonderful people from the Fort Collins International Center for creating such a welcoming atmosphere for international students, my office mates at CSU and SLAC for distracting chats about anything but Physics, my friends in Colorado and California for all the unforgettable moments and adventures, my friends in Germany for staying in touch over the last three years and anybody else that I forgot to mention.

Besonderer Dank gebührt meinen Eltern und meiner Schwester Kristina, für ihre Unterstützung während der vergangenen drei Jahre, für ihre Besuche hier in den USA, für ihr Verständnis, wenn ich mich mehrere Wochen nicht gemeldet hatte, für die herzlichen Wiedersehen im Sommer und an Weihnachten und für die unzähligen Brief und Pakete, die sie über den Ozean geschickt haben.

Contents

Chapter 1

Introduction

1.1 Weak interactions and CP violation

The invariance of Quantum Field Theory (QFT) under the continuous symmetry transformations of the Poincaré group (the relativistic equivalent to the Galilei group in classical mechanics) gives rise to the usual ten conservation laws. In addition, we have the following potential discrete symmetry operations:

- C : particle $\mapsto$antiparticle (charge conjugation)
- $P : (t, \mathbf{x}) \mapsto (t, -\mathbf{x})$ (parity or space inversion)
- $T : (t, \mathbf{x}) \mapsto (-t, \mathbf{x})$ (time reversal).

The CPT-theorem states that the combined operation of C, P and T is a symmetry of every local, Lorentz invariant QFT with a hermitian Hamiltonian.[1] Direct consequences (and possibilities for experimental tests) are the equality of particle and antiparticle masses as well as their lifetimes. If CPT is indeed conserved in nature it follows that CP and T separately, are either both conserved or both violated for a given interaction.

Since the weak interaction only couples left-handed particles (or right-handed antiparticles) both C and P are violated. The combined operation CP that turns left-handed

[1] For a pedagogical derivation of the CPT theorem and historic references see [3].

particle into right-handed anti-particles on the other side, was thought to be a good symmetry for weak interactions until 1964 when Cronin and Fitch discovered *CP* violation in the neutral kaon system [4]. Even larger *CP* violating effects in the B meson system observed by the *BABAR* and Belle experiment establish *CP* (and T) violation for the weak interaction. In the following we will see under which conditions the Standard Model (or any field theory) allows for *CP* violating effects.

First, we need to establish the *CP* transformation of an arbitrary complex field ϕ under the *CP* operator U_{CP}. A heuristic argument uses the fact that $(CP) = T^{-1}$ in case the *CPT* theorem holds. Using $\phi(t, \mathbf{x}) = \phi_0\, e^{i(\mathbf{px}-Et)}$ and $T\mathbf{p} = -\mathbf{p}$ we get for the time reversal

$$T\phi(t, \mathbf{x}) = \phi_0\, e^{i(-\mathbf{px}+Et)} = \phi(t, \mathbf{x})^*, \tag{1.1}$$

i.e. the complex conjugate of the original field. Hence, the *CP* conjugate of ϕ is given by the hermitian conjugate ϕ^* and we allow for an arbitray phase factor

$$U_{CP}\, \phi\, U_{CP}^\dagger = e^{i\alpha}\phi^*. \tag{1.2}$$

Next, we consider a simple "toy" field theory with the hermitian Hamilton density

$$\mathcal{H}_1 = g\, \phi\, O + g^*\, \phi^*\, O^*, \tag{1.3}$$

where g is a coupling constant and O an arbitrary operator [5]. The *CP* conjugate of this Hamiltonian is readily obtained as

$$U_{CP}\, \mathcal{H}_1\, U_{CP}^\dagger = g\ e^{i\alpha}\phi^*\, O^* + g^*\ e^{-i\alpha}\, \phi\, O \tag{1.4}$$

and if we choose $\alpha = -2\arg(g)$ the system is invariant under U_{CP}. In other words, the phase factor of U_{CP} is fixed by the coupling constant. If we consider a system with two coupling constants, like

$$
\begin{aligned}
\mathcal{H}_2 &= g_1\,\phi\,O_1 + g_2\,\phi\,O_2 + h.c. \\
U_{CP}\,\mathcal{H}_2\,U_{CP}^{\dagger} &= g_1\;e^{i\alpha}\phi^*\,O_1^* + g_2\;e^{i\alpha}\phi^*\,O_2^* + h.c.
\end{aligned}
\tag{1.5}
$$

we can no longer absorb the phase into the coupling constant if the phases of g_1 and g_2 are different. Whatever α we choose, U_{CP} will not be a symmetry of the system giving rise to *CP* violating effects.

In the Standard Model, the charged current weak interaction for N generations is

$$
\mathcal{H}_{cc} = \frac{\sqrt{2}g}{2} W_\mu^+\,\overline{\psi}_i\,\gamma^\mu(1-\gamma^5)V_{ij}\,\psi_j + h.c.\,, \tag{1.6}
$$

where ψ_j are the down-type particle mass eigenstates, $\overline{\psi}_i$ are the up-type antiparticle mass eigenstates and $V_{ij}\,\psi_j$ are the weak eigenstates connected to the mass eigenstates by the unitary $N \times N$ weak mixing matrix V. The coupling between generation i and j is therefore proportional to $g\,V_{ij}$. If V contains irreducible complex phases, the theory will allow for *CP* violating processes. The following section, will investigate this for the currently known three quark generations.

1.2 CKM quark mixing matrix

For three quark generations, the quark mixing matrix V is commonly expressed as a 3x3 unitary matrix operating on the down-type quark mass eigenstates

$$
\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.7}
$$

and is called the Cabibbo-Kobayashi-Maskawa (CKM) quark-mixing matrix [6, 7]. It is instructive to check whether a 3×3 unitary matrix allows for a (CP violating) complex phase. An $N \times N$ complex matrix contains $2N^2$ real numbers. The unitarity requirement of V

$$\sum_{k=1}^{N} V_{ki} V_{kj}^{*} = \delta_{ij} \tag{1.8}$$

consists of N^2 constraints on the matrix elements, reducing the number of independent parameters to N^2. For $2N$ quark fields, there are $(2N-1)$ independent phases that can be absorbed into the quark fields, leaving $N^2 - (2N-1) = (N-1)^2$ free parameters. Out of these $(N-1)^2$ parameters $\binom{N}{2} = N(N-1)/2$ can be chosen as mixing angles between the different quark pairs. The remaining $(N-1)(N-2)/2$ parameters are complex phases. For $N = 3$ quark generations, the CKM matrix V can therefore be parameterized using three real mixing angles and one complex phase giving rise to possible CP violation in the quark sector. For an explicit confirmation for the cases $N = 2, 3$ see [5].

A "standard" parameterization [8] of V using three (Euler) angles and a phase is given by

$$V = \begin{pmatrix} c_{12}s_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix} \tag{1.9}$$

with $c_{ij} = cos\theta_{ij}$ and $s_{ij} = sin\theta_{ij}$ for the three generations $i, j = 1, 2, 3$. This parameterization is exact to all orders and contains the three real mixing angles $\theta_{12}, \theta_{23}, \theta_{13}$ and the phase δ. Another parameterization that makes use of the relative size of the mixing angles $s_{s12} \gg s_{23} \gg s_{13}$ was proposed by Wolfenstein and is expressed in powers of $\lambda \equiv s_{12}$:

$$V = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \tag{1.10}$$

with A, ρ and η being real numbers of order unity.

Out of the unitarity constraints (1.8) of the CKM matrix the following describes best the CP violation in B meson decays:

$$\sum_{k=1}^{3} V_{kd}V_{kb}^* = V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \tag{1.11}$$

In the complex plane this equation represents a triangle, the Unitarity Triangle (UT). Rotating one of the sides ($V_{cd}V_{cb}^*$) onto the real axis and rescaling it to unity yields

$$\frac{V_{ud}V_{ub}^*}{|V_{cd}V_{cb}^*|} + 1 + \frac{V_{td}V_{tb}^*}{|V_{cd}V_{cb}^*|} = 0\,. \tag{1.12}$$

Figure 1.1 shows a picture of the rescaled UT. The three angles of the UT are denoted by α, β and γ:

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \qquad \beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \qquad \gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \tag{1.13}$$

Using the Wolfenstein approximation (1.10) the angle γ can be expressed in terms of η and ρ:

$$\begin{aligned} \gamma &= \arg\left(-\frac{(1-\lambda^2/2)A\lambda^3(\rho+i\eta)}{\lambda A\lambda^2}\right) \\ &= \arg\big((1-\lambda^2)(1-i\eta/\rho)\big) \\ &= \tan^{-1}\frac{\eta}{\rho}\,, \end{aligned} \tag{1.14}$$

which is also indicated in Fig. 1.1 by the apex of the triangle. Moreover, since $s_{13}e^{-i\delta} = A\lambda^3(\rho-i\eta)+\mathcal{O}(\lambda^4)$ we can identify $\delta \approx -\arg(A\lambda^3(\rho-i\eta)) = \tan^{-1}\frac{\eta}{\rho} = \gamma$ up to corrections of λ^4. This also justifies γ being called the *weak phase*.

Fig. 1.2 summarizes the current experimental results [9] for the apex of the UT. The different color bands indicate confidence regions from different measurements. The nu-

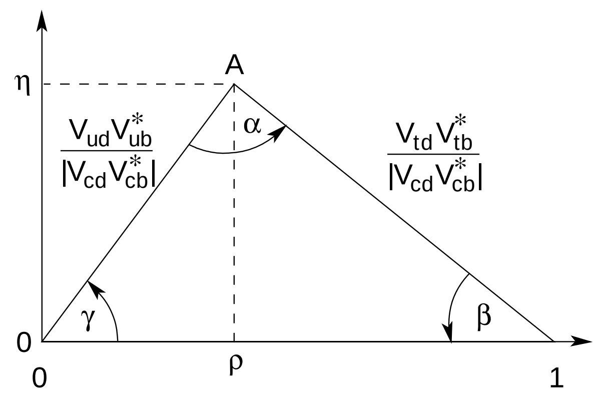


Figure 1.1: The rescaled Unitarity Triangle described by Eq. (1.12).

merical values of some observables obtained by the global fit shown in the figure are:

$$\begin{aligned}
\lambda &= 0.2272 \pm 0.0010 \\
A &= 0.809 \pm 0.0014 \\
\bar{\rho} \equiv \rho(1-\lambda^2/2) &= 0.197 \pm 0.030 \\
\bar{\eta} \equiv \eta(1-\lambda^2/2) &= 0.339 \pm 0.019 \\
\alpha &= 97.3 \pm 5.0 \\
\beta &= 22.9 \pm 1.0 \\
\gamma &= 59.8 \pm 4.9
\end{aligned} \tag{1.15}$$

Note that these numbers are obtained by a global fit of all measurements. Direct measurements might have larger errors. For example, the world average on β is $\sin(2\beta) = 0.687 \pm 0.032$ [10] and the best direct measurement of γ from *BABAR* is currently $\gamma = 67 \pm 33$ [11]. Despite the fact that we already have fairly good constraints on the apex of the UT, it is still indispensable to measure each observable directly, and thus to over-constrain the triangle. New Physics beyond the Standard Model (beyond CKM) could show up in deviations from the UT and therefore it is important to measure all sides and angles of the triangle to very high accuracy.

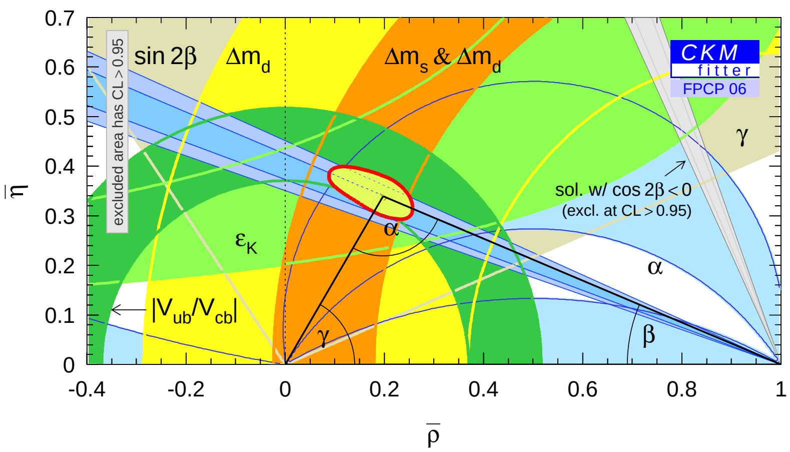


Figure 1.2: Experimental constraints [9] on the apex of the Unitarity Triangle. $\bar{\rho} = \rho(1-\lambda^2/2)$ and $\bar{\eta} = \eta(1-\lambda^2/2)$ are the rescaled Wolfenstein parameters. ϵ_K is the CP violating parameter in the kaon system and $\Delta m_{s,d}$ are measurements obtained from $B^0_{s,d}$ mixing.

1.3 Measuring the CKM angle γ

In the following sections we present different methods to directly measure the CKM angle γ. The basic idea of all methods is similar to the double slit experiment. Given two amplitudes A_1 and $A_2e^{i\phi}$ with an unknown relative phase ϕ, one can measure the phase difference through the interference pattern of the two amplitudes $|A_1 + A_2e^{i\phi}|^2 = |A_1|^2 + |A_2|^2 + 2\Re(A_1A_2e^{i\phi})$. If $A_{1,2}$ are real, the interference term becomes $2A_1A_2\cos\phi$ and by measuring $\cos\phi$ one can determine the phase difference ϕ up to a two-fold ambiguity due to $cos(\phi) = cos(\phi+\pi)$. In our case, we are not interfering light, but quantum mechanical decay amplitudes of particles.

Since γ is the phase of the V_{ub} matrix element, we necessarily need to interfere amplitudes of which at least one involves a $b \to u$ transition. The Feynman diagrams of two decays that we will encounter numerous times in what follows are shown in Fig. 1.3. The left diagram shows the color-allowed decay $B^- \to D^0K^-$ and the right diagram shows the color-suppressed[2] decay $B^- \to \overline{D}^0K^-$. With $N_c = 3$ colors and the measured values for

[2]The internal W emission in the right diagram puts constraints on the allowed colors of the $\overline{c}$ and s quark to form colorless mesons together with the pre-determined colors of the b and $\overline{u}$ quark from the B^-. This reduces the number of possible amplitudes.

the CKM matrix elements one can estimate the ratio of the amplitudes to be

$$\begin{aligned} r_B \equiv \frac{|A(B^- \to \overline{D}^0 K^-)|}{|A(B^- \to D^0 K^-)|} &\approx \frac{1}{N_c}\frac{|V_{ub}^* V_{cs}|}{|V_{cb}^* V_{us}|} \\ &\approx \frac{1}{N_c}|(\rho - i\eta)(1-\lambda^2/2)| = \frac{1}{N_c}|\bar{\rho} - i\bar{\eta}| \\ &\approx 0.39/3 \approx 0.1. \end{aligned} \tag{1.16}$$

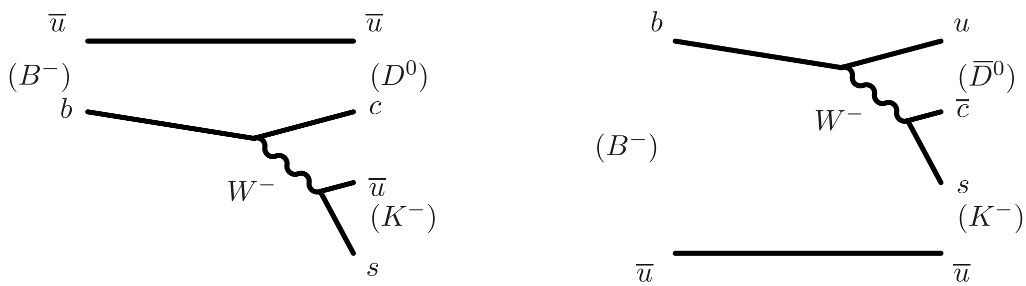


Figure 1.3: Color allowed $B^- \to D^0 K^-$ (left) and color suppressed $B^- \to \overline{D}^0 K^-$ decay (right). The $b \to u$ transition in the right decay gives rise to the weak phase $e^{-i\gamma}$.

We can then write the amplitudes of the two decays as

$$\begin{aligned} A(B^- \to D^0 K^-) &\equiv A_B \\ A(B^- \to \overline{D}^0 K^-) &\equiv A_B r_B \, e^{i\delta_B} \, e^{-i\gamma} \end{aligned} \tag{1.17}$$

where we have defined the magnitude and phases of the color suppressed decay relative to the color allowed decay and A_B is a real number. $e^{-i\gamma}$ arises due to V_{ub} in the weak decay and δ_B is an unknown (CP-conserving) strong phase. To observe interference between the two decays, the D^0 and $\overline{D}^0$ need to decay to the same final state, which we call f. We define the D-decay amplitudes as

$$\begin{aligned} A(D^0 \to f) &\equiv A_f \\ A(\overline{D}^0 \to f) &\equiv A_f r_f \, e^{i\delta_f} \end{aligned} \tag{1.18}$$

with r_f as the amplitude ratio and δ_f an unknown strong phase. The amplitude of the cascade decay including $D^0/\overline{D}^0$ interference is

$$A(B^- \to D_f K^-) = A_B A_f \left(1 + r_B r_f \, e^{i(\delta-\gamma)}\right), \tag{1.19}$$

where $\delta = \delta_B + \delta_f$ is the overall strong phase and D_f indicates that the D meson is detected via the final state f. We only observe the branching ratio $\mathcal{B}(B^- \to D_f K^-)$, which is related to the amplitude (1.19) by

$$\mathcal{B}(B^- \to D_f K^-) = |A(B^- \to D_f K^-)|^2 = A_B^2 A_f^2 \left(1 + {r_B}^2 r_f^2 + 2 r_B r_f \cos(\delta - \gamma)\right). \tag{1.20}$$

Since the only *CP*-violating quantity in (1.20) is γ, the *CP*-conjugate mode is obtained by replacing $i\gamma \to -i\gamma$ and the branching ratio for the B^+ decay becomes

$$\mathcal{B}(B^+ \to D_{\bar{f}} K^+) = |A(B^+ \to D_{\bar{f}} K^+)|^2 = A_B^2 A_f^2 \left(1 + {r_B}^2 r_f^2 + 2 r_B r_f \cos(\delta + \gamma)\right). \tag{1.21}$$

Equations (1.20) and (1.21) have a total of six unknowns $(A_B, A_f, r_B, r_f, \delta, \gamma)$. How many observable are there? A_B can be determined by measuring $\mathcal{B}(B^- \to D^0 K^-) \times \mathcal{B}(D^0 \to K^- l^+ \nu_l)$ where the flavor of the D meson is tagged by the charge of the kaon. Measurements of $\mathcal{B}(D^0 \to f)$ and $\mathcal{B}(\overline{D}^0 \to f)$ give A_f and r_f. The remaining two observables, $\mathcal{B}(B^- \to D_f K^-)$ and $\mathcal{B}(B^+ \to D_{\bar{f}} K^+)$, depend on δ, γ and r_B. We realize that we have six unknowns but only five observables and therefore the problem is under-determined. However, this can easily be overcome by adding a second decay mode $D \to f'$. This will result in four new observables, $\mathcal{B}(B^- \to D_{f'} K^-)$, $\mathcal{B}(B^+ \to D_{\bar{f}'} K^+)$, $\mathcal{B}(D^0 \to f')$ and $\mathcal{B}(\overline{D}^0 \to f')$, but only three new unknowns ($r_{f'}$, $A_{f'}$, $\delta_{f'}$). Therefore, it allows the extraction of the strong phase δ, the amplitude ratio r_B and the weak phase γ, simultaneously.

A. Soffer first pointed out that *CP*-violation measurements performed by this method suffer from ambiguity which is at least 8-fold [12]. Since the measured decay widths contain the term $\cos(\delta \pm \gamma)$, from which we hope to extract γ, the value of $\cos(\delta \pm \gamma)$ will be invariant under the following symmetry operations

$$\begin{aligned} S_{\pm} &: \quad \gamma \mapsto -\gamma, \quad \delta \mapsto -\delta \\ S_{\leftrightarrow} &: \quad \gamma \leftrightarrow \delta \\ S_{\pi} &: \quad \gamma \mapsto \gamma + \pi, \quad \delta \mapsto \delta - \pi \end{aligned}$$

leading to an 8-fold ambiguity. This holds for all *CP*-violation measurements in which the measurable widths depend only on trigonometric functions of the sum of a weak phase and a *CP*-conserving phase.

So far, we have not specified which final states f should be used. In the following, we present in chronological order three different proposals that have been put forward to realize the above measurement.

1.3.1 Gronau-London-Wyler method (GLW)

Gronau, London and Wyler (GLW [13, 14]) proposed in 1991 to use D decays into *CP* eigenstates for the direct measurement of γ. The *CP* even or odd eigenstates of the D meson are $D^0_{\pm} = (D^0 \pm \overline{D}^0)/\sqrt{2}$. These can be identified by their *CP* even ($K^0_S\,\pi^0$, $K^0_S\,\rho^0$, $K^0_S\omega$) or *CP* odd ($\pi^+\pi^-$, K^+K^-) decay products. Since *CP* violation in the D meson system in negligible small in the Standard Model it can be safely disregarded. The idea is to extract γ from the decay rate asymmetry of $B^- \to D^0_{\pm}K^-$ and $B^+ \to D^0_{\pm}K^+$. With the definitions from (1.17) we can write the amplitude of the B decay as

$$\begin{aligned} \sqrt{2}A(B^- \to D^0_{\pm}K^-) &= A(B^- \to D^0K^-) \pm A(B^- \to \overline{D}^0K^-) = A_B\left(1 \pm r_B\, e^{i(\delta_B-\gamma)}\right) \\ \sqrt{2}A(B^+ \to D^0_{\pm}K^+) &= A(B^+ \to D^0K^+) \pm A(B^+ \to \overline{D}^0K^+) = A_B\left(1 \pm r_B\, e^{i(\delta_B+\gamma)}\right) \end{aligned}$$

and the decay rate asymmetry is

$$|A(B^+ \to D^0_{\pm}K^+)|^2 - |A(B^- \to D^0_{\pm}K^-)|^2 = \frac{A_B^2}{2}\left(|1 \pm r_B\, e^{i(\delta_B+\gamma)}|^2 - |1 \pm r_B\, e^{i(\delta_B-\gamma)}|^2\right)$$

$$
\begin{aligned}
&= A_B^2 r_B \left(\pm\cos(\delta_B+\gamma) \mp \cos(\delta_B-\gamma)\right) \\
&= \mp 2 A_B^2 r_B \sin\delta_B \sin\gamma. \qquad (1.22)
\end{aligned}
$$

In the last step we have used the trigonometric identity $\cos(\alpha\pm\beta) = \cos\alpha\cos\beta\mp\sin\alpha\sin\beta$. Equation (1.22) also shows possible limitations of this method pointed out by Atwood, Dunietz and Soni (ADS, [15]):

1. D^0 decays to *CP* final states are either Cabibbo suppressed or color suppressed and the experimentally feasible total is less than 5%.

2. Due to the numerical value of $r_B \approx 0.1$, the *CP* violating asymmetries are only on the order of 10%.

3. The sensitivity to γ depends on the unknown strong phase (difference) δ_B that could severely limit the application of the GLW method.

4. The method requires knowledge of A_B and r_B through the measurement of $\mathcal{B}(B^- \to D^0K^-)$ and $\mathcal{B}(B^- \to \overline{D}^0K^-)$.

The last point is worth considering in more detail because it will lead to the ADS method. In practice, measuring the color suppressed decay $\mathcal{B}(B^- \to \overline{D}^0K^-)$ requires that the $\overline{D}^0$ be identified in a hadronic final state since semileptonic decays suffer from unacceptably high backgrounds due to the invisible neutrino. However, the decay chain $B^- \to \overline{D}^0K^-, \overline{D}^0 \to f$ results in the same final state as $B^- \to D^0K^-, D^0 \to f$, where the D^0 undergoes doubly Cabibbo suppressed decay. In the next section, we will see that this interference is not negligible and prevents the experimental determination of $\mathcal{B}(B^- \to \overline{D}^0K^-)$.

1.3.2 Atwood-Dunietz-Soni method (ADS)

Atwood, Dunietz and Soni [15,16] realized that two of the drawbacks of the GLW method can be used in favor of each other. Ideally, in an interference experiment, the two interfering amplitudes should have comparable magnitudes to allow maximum sensitivity to the phase difference. If one allows for non-CP eigenstates in the decay of the neutral D meson one can offset the r_B-suppression in the B decay by a doubly Cabibbo suppressed (DCS) decay of the D meson. For example, let us consider the final state $f = K^+\pi^-$ and the two decay chains $B^- \to D^0 K^-$, $D^0 \to K^+\pi^-$ and $B^- \to \overline{D}^0 K^-$, $\overline{D}^0 \to K^+\pi^-$. Figure 1.4 shows the corresponding Feynman diagrams for the D decay. $\overline{D}^0 \to K^+\pi^-$ (left) only involves diagonal terms of the CKM matrix (Cabibbo allowed) whereas $D^0 \to K^+\pi^-$ (right) is suppressed by $|V_{cd}V_{us}|^2/|V_{cs}V_{ud}|^2 \approx \lambda^4 \approx 0.003$. The experimental value for this ratio is $\mathcal{B}(D^0 \to K^+\pi^-)/\mathcal{B}(\overline{D}^0 \to K^+\pi^-) = 0.0036 \pm 0.0011$ [8]. Using $r_B \approx 0.1$ (1.16) the amplitude ratio for the decay chain is

$$\frac{A(B^- \to D^0_{K^+\pi^-}K^-)}{A(B^- \to \overline{D}^0_{K^+\pi^-}K^-)} \approx \sqrt{0.0036}/r_B \approx 0.6, \tag{1.23}$$

which is of order unity and therefore ideal for an interference experiment. The extraction of γ can now be done with the help of Eq. (1.20) and (1.21) using at least two different final states.

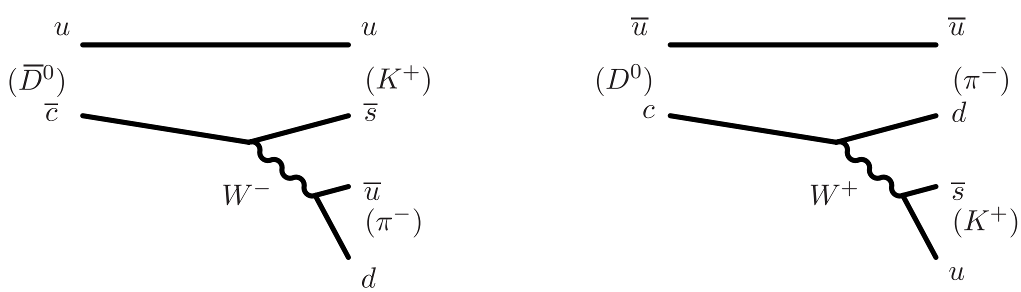


Figure 1.4: Cabibbo allowed $\overline{D}^0 \to K^+\pi^-$ (left) and doubly Cabibbo suppressed (DCS) $D^0 \to K^+\pi^-$ decay (right). The suppression is of order $|V_{cd}V_{us}|^2/|V_{cs}V_{ud}|^2 \approx \lambda^4$.

1.3.3 Giri-Grossman-Soffer-Zupan method (GGSZ)

So far, we have only considered two-body decays of the D meson. Giri, Grossman, Soffer and Zupan (GGSZ, [17,18]) suggested to use multi-body D decays, such as $D \to K^0_S\pi^+\pi^-$, $D \to K^0_S K^+K^-$, $D \to K^0_S\pi^-\pi^+\pi^0$ or $D \to \pi^+\pi^-\pi^0$. Except the last decay all are Cabibbo allowed and due to the expected presence of resonances large strong phases are expected. In order to extract r_B, γ and δ one has to perform a Dalitz analysis of the D decay (see section 2.2 for an introduction to the Dalitz plot).

The amplitudes of the 3-body D-decay (here we use the example $D \to \pi^+\pi^-\pi^0$) is parametrized in terms of its Dalitz variables:

$$\begin{aligned} A(D^0 \to \pi^0(p_1)\pi^+(p_2)\pi^-(p_3)) &\equiv f_{D^0}(s_{12}, s_{13}) = f_{D^0}(s_{+0}, s_{-0}) \\ A(\overline{D}^0 \to \pi^0(p_1)\pi^-(p_2)\pi^+(p_3)) &\equiv f_{D^0}(s_{13}, s_{12}) = f_{D^0}(s_{-0}, s_{+0}), \end{aligned} \tag{1.24}$$

where we have used the CP symmetry of the strong interactions and the fact that the final state is a spin zero state. p_i is the 4-momentum of the pion and $s_{ij} = (p_i + p_j)^2$ the invariant mass squared of the pion pair. We will discuss the functional form of f_{D^0} below. With the above definitions and (1.17) the amplitude of the cascade decay is

$$A(B^- \to D_{\pi^+\pi^-\pi^0}K^-) = A_B \left(f_{D^0}(s_{12}, s_{13}) + r_B e^{i(\delta_B - \gamma)} f_{D^0}(s_{13}, s_{12}) \right). \tag{1.25}$$

The amplitude for the B^+ decay is obtained by exchanging $s_{12} \leftrightarrow s_{13}$ and replacing $\gamma \mapsto -\gamma$:

$$A(B^+ \to D_{\pi^+\pi^-\pi^0}K^+) = A_B \left(f_{D^0}(s_{13}, s_{12}) + r_B e^{i(\delta_B + \gamma)} f_{D^0}(s_{12}, s_{13}) \right). \tag{1.26}$$

The differential partial decay width is given by the complex square of (1.25)

$$\begin{aligned} & d\Gamma(B^- \to D_{\pi^+\pi^-\pi^0}K^-) = \\ & A_B^2 \left(|f_{D^0}(s_{12}, s_{13})|^2 + {r_B}^2 |f_{D^0}(s_{13}, s_{12})|^2 + 2 r_B \Re \left[f_{D^0}(s_{12}, s_{13}) {f_{D^0}}^*(s_{13}, s_{12})\, e^{-i(\delta_B - \gamma)} \right] \right) d\phi\,, \end{aligned} \tag{1.27}$$

where $d\phi$ denotes the phase space variable, which is proportional to $ds_{12}\, ds_{13}$. It seems that in order to measure γ, we need knowledge of the D decay amplitude f_{D^0}. The main new idea by GGSZ is, that all the information necessary to measure γ is already contained in the Dalitz plot and no further knowledge of f_{D^0} is required. This is called the model-independent approach. Since the analysis presented in Chapter 3 of this dissertation does not make use of this approach, we refer the interested reader to [17] for the details.

If the functional form of f_{D^0} were known, the analysis would be simplified since only the three variables r_B, δ_B and γ needed to be fit (A_B is an irrelevant overall normalization factor). CLEO has analyzed $D \to \pi^+\pi^-\pi^0$ decays and found that the three-body decay is dominated by $\rho \to \pi\pi$ decays and a non-resonant component [19]. Thus, f_{D^0} can be modelled by the sum of Breit-Wigner shapes. The theoretical error introduced by this assumption is expected to be much smaller than the statistical error in the measurement of γ. Using this knowledge we can write

$$f_{D^0}(s_{12}, s_{13}) = a_0\, e^{i\delta_0} + \sum_r a_r\, e^{i\delta_r} A_r(s_{12}, s_{13})\,, \tag{1.28}$$

where the first term corresponds to the non-resonant and the second to the resonant contributions. A_r is a Breit-Wigner shape for the resonance r (e.g. ρ^+, ρ^-, ρ^0) with the appropriate spin factors (details can be found in section 3.6.6). The amplitudes a_i and phases δ_i are obtained from an analysis equivalent to the one CLEO did. Using (1.28) together with (1.27) in a Dalitz fit to $B^- \to D_{\pi^+\pi^-\pi^0} K^-$ decays will give a direct measurement of the *CP* violating phase γ. This will be the topic of Chapter 3 of this dissertation.

1.4 The $X(3872)$ state

This section gives an overview over both experimental and theoretical results regarding the "mystery meson" $X(3872)$. By the time of my graduation it has been almost exactly three years since Belle announced the discovery of a new state at 3872 MeV/c^2. Since then, the original Belle publication [20] received over 200 citations and more than 60 articles alone with the word "$X(3872)$" in their title appeared [21]. A vast amount of models and interpretations for this new state have been put forward, some already excluded by experiment, others still being tested. In the following, a summary of the experimental results together with two possible models for the $X(3872)$ are being presented. This also includes a summary of results of this dissertation that will be presented in detail in later chapters.

1.4.1 Experimental results

Belle at KEKB delivered the first evidence [22] for a possible new state at 3.872 GeV/c^2 in B-decays in Sept. 2003. The Belle detector is located at the KEKB asymmetric e^+e^- collider in Japan, which operates at a center-of-mass (CM) energy of $\sqrt{s} = 10.58\,\mathrm{GeV}$ corresponding to the mass of the $\Upsilon(4S)$ resonance. The signal was observed in the decay $B^+ \to J/\psi\,\pi^+\pi^- K^+$ in a total data sample of 152 million $B\overline{B}$ events. The mass of the new state is measured to be $(3872.0 \pm 0.6 \pm 0.5)\,\mathrm{MeV}/c^2$ and the 90% confidence level (C.L.) upper limit on the width is $\Gamma < 2.3\,\mathrm{MeV}$. In May 2005, the Belle Collaboration has updated their results using 275 million $B\overline{B}$ events [23].

In Belle's analysis the J/ψ candidate was reconstructed by a pair of well identified electrons or muons with an invariant mass range $3.077 < m(l^+l^-) < 3.117\,\mathrm{GeV}/c^2$. Additionally, a pair of oppositely charged pions with an invariant mass greater than 575 MeV/c^2 and a loosely identified charged kaon is required. To suppress continuum background only events with a normalized Fox-Wolfram moment of $R_2 < 0.4$ and $|\cos\theta_B| < 0.8$ are selected where θ_B is the polar angle of the B-meson direction in the center-of-mass (CM) frame.

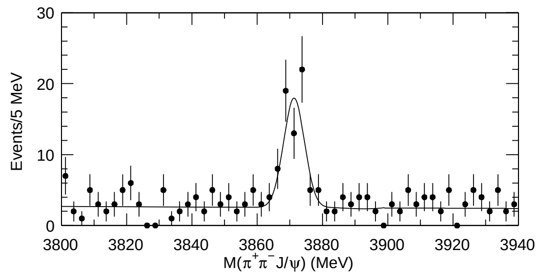


Figure 1.5: Distribution for $J/\psi\,\pi^+\pi^-$ invariant mass for the $X(3872)$ region. The signal yield is 49.1 ± 8.4 events

True B mesons are considered to fall into a 3σ signal box defined as $|\Delta E| < 0.034\,\mathrm{GeV}$ and $5.2725 < m_{ES} < 5.2875\,\mathrm{GeV}/c^2$.[3] Figure 1.5 shows the invariant mass $m(J/\psi\,\pi^+\pi^-)$ near $3872\,\mathrm{MeV}/c^2$ for the selected events. The distribution is fitted with a first-order polynomial for the background and a Gaussian for the signal. The width of the Gaussian is fixed to $3.2\,\mathrm{MeV}/c^2$, the resolution obtained from the $\psi(2S) \to J/\psi\,\pi^+\pi^-$ control sample. The signal yield is 49.1 ± 8.4 events resulting in a product branching fraction of

$$\mathcal{B}(B \to X(3872)K, X(3872) \to J/\psi\,\pi^+\pi^-) = (13.1 \pm 2.4 \pm 1.3) \times 10^{-5}. \tag{1.29}$$

To investigate whether the dipion originates from a ρ meson or not, the $\pi^+\pi^-$ invariant mass in a $\pm 5\,\mathrm{MeV}/c^2$ window around the $X(3872)$ peak is shown in Fig. 1.6. Clearly, this is consistent with a peak at the nominal ρ mass. Belle fits the $m(\pi^+\pi^-)$ invariant mass with the hypothesis of the J/ψ and ρ being in a relative S-wave or P-wave and concludes that the S-wave fits the data much better than the P-wave ($\chi^2/d.o.f. = 43.1/39$ versus $71.0/39$) indicating that J^{++} is strongly favored over J^{-+}.

Since the number of signal events is too low for a full angular analysis, Belle tries to find, for a given J^{PC} hypothesis for the $X(3872)$, angular quantities that have distributions with a zero in some location. In the bins near the zero point, any observed events would have to be accounted for by an upward fluctuation of the background. Going through all possibilities with $J \leq 2$, Belle concludes that the data only support the quantum numbers

[3] In Belle jargon the energy substituted mass m_{ES} is called beam constrained mass M_{bc}

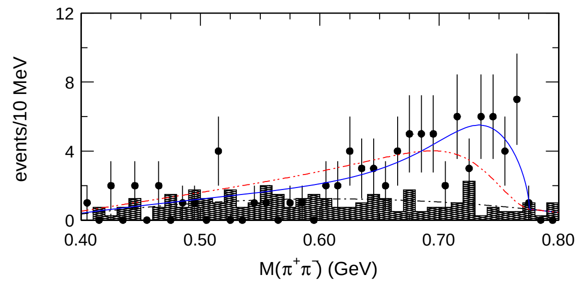


Figure 1.6: $m(\pi^+\pi^-)$ distribution of evens in the $X(3872)$ signal region. The solid (dashed) curve shows the fit that uses a ρ Breit-Wigner line shape with the J/ψ and ρ in a relative S-wave (P-wave). The histogram indicates the side-band determined background and the dot-dashed curve is a fit of this background.

1^{++} and 2^{++} for the $X(3872)$ [23].

Another important experimental test is to search for radiative decays of the $X(3872)$. In case of the $X(3872)$ begin a charmonium state one would expect a considerably large fraction of radiative decays to $\gamma\chi_{c1}$. Belle searches for those decays with $\chi_{c1} \to \gamma J/\psi$ in the $\psi(2S)$ and $X(3872)$ region [22]. Figure 1.7 shows the result of the fit to m_{ES} and the $\gamma\chi_{c1}$ invariant mass. As expected, there is a clear signal for the $\psi(2S)$ but no evidence for this decay in the $X(3872)$ region. A similar analysis has been performed for $X(3872) \to \gamma\chi_{c2}$ with the same negative result [24]. The resulting upper limits on the ratio of partial widths are

$$\frac{\Gamma(X(3872) \to \gamma\chi_{c1})}{\Gamma(X(3872) \to \pi^+\pi^- J/\psi)} < 0.89 \qquad (90\%\ \text{C.L.})$$

$$\frac{\Gamma(X(3872) \to \gamma\chi_{c2})}{\Gamma(X(3872) \to \pi^+\pi^- J/\psi)} < 1.1 \qquad (90\%\ \text{C.L.}).$$

Using $256\,\text{fb}^{-1}$ Belle finds evidence for the decays $X(3872) \to J/\psi\,\gamma$ and $X(3872) \to J/\psi\,\pi^+\pi^-\pi^0$ [25]. Figure 1.8 shows the background subtracted $J/\psi\,\gamma$ invariant mass. It is obtained by fitting the m_{ES} and ΔE distributions in bins of $m(J/\psi\,\gamma)$ for events within the signal box $|\Delta E| < 0.035\,\text{GeV}$ and $5.2745 < m_{ES} < 5.2855\,\text{GeV}/c^2$. The signal yield

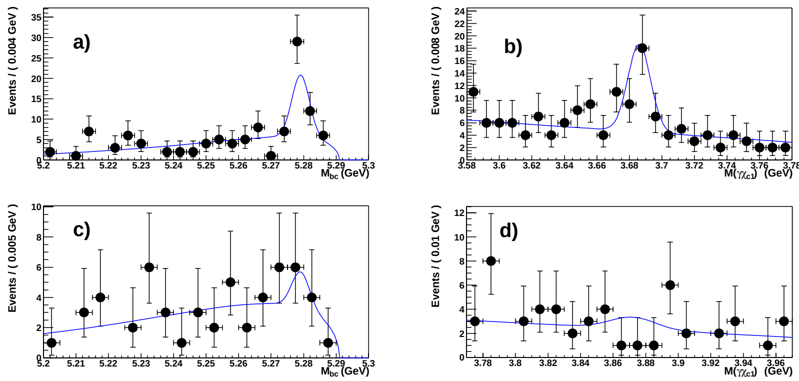


Figure 1.7: Signal-band projections of (a) $m_{ES}$and (b) $m(\gamma\chi_{c1})$ for the $\psi(2S)$ region with the results of the unbinned fit superimposed from Belle. (c) and (d) are the corresponding results for the $X(3872)$ region.

in the $X(3872)$ bin is 13.6 ± 4.4 events with a 4.0σ significance. The product branching ratio is

$$\mathcal{B}(B \to X(3872)K, X(3872) \to J/\psi\gamma) = (1.8 \pm 0.6 \pm 0.1) \times 10^{-6}. \tag{1.30}$$

This result unambiguously establishes the charge conjugation parity of the $X(3872)$ as $C = +1$. This is consistent with the other results reported above.

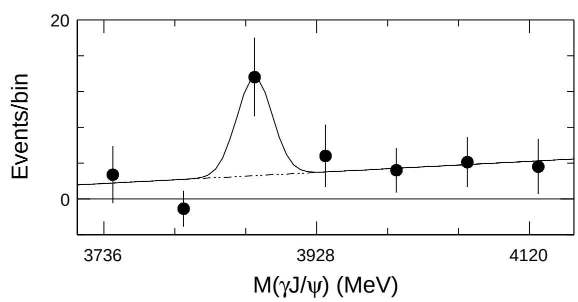


Figure 1.8: The signal yields from a fit to the m_{ES} and ΔE distributions in bins of $m(J/\psi\gamma)$.

A similar analysis is performed for the $X(3872) \to J/\psi\pi^+\pi^-\pi^0$ decay mode. This time the m_{ES} and ΔE fits are done in bins of $m(\pi^+\pi^-\pi^0)$ for events with $|\Delta E| < 0.035\,\mathrm{GeV}$, $5.2725 < m_{ES} < 5.2875\,\mathrm{GeV}/c^2$ and $|m(J/\psi\,3\pi) - 3.872| < 0.0165\,\mathrm{GeV}/c^2$. Figure 1.9 shows the resulting signal yields. The branching ratio is comparable to the $J/\psi\,\pi^+\pi^-$ mode:

$$\frac{\mathcal{B}(X \to J/\psi\,\pi^+\pi^-\pi^0)}{\mathcal{B}(X \to J/\psi\,\pi^+\pi^-)} = 1.0 \pm 0.4 \pm 0.3\,. \tag{1.31}$$

If these are the same states, this implies that there are large isospin breaking effects that allow the $X(3872)$ both decay into $J/\psi\,\rho$ and $J/\psi\,\omega$. Another lesson about the isospin of

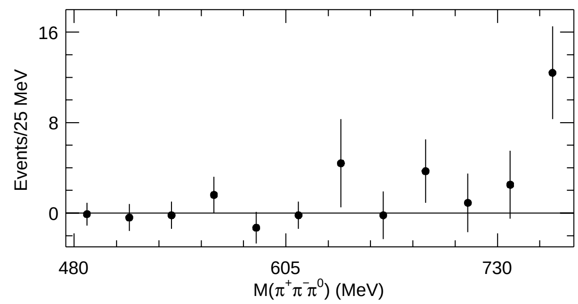


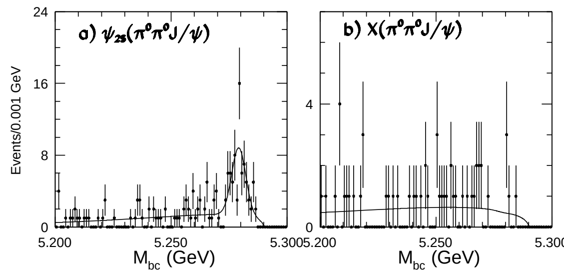
Figure 1.9: The signal yields from a fit to the m_{ES} and ΔE distributions in bins of $m(\pi^+\pi^-\pi^0)$.

the $X(3872)$ can be learned from searching for $X(3872) \to J/\psi\,\pi^0\pi^0$ decays. The ratio $R = \Gamma(X \to J/\psi\,\pi^0\pi^0)/\Gamma(X \to J/\psi\,\pi^+\pi^-)$ is expected to be $R = 1/2$ for $I = 0$ and $R = 0$ for $I = 1$, if isospin is conserved. Belle searches for this decay using $253\,\mathrm{fb}^{-1}$ and selecting events within $-0.06\,\mathrm{GeV} < \Delta E < 0.03\,\mathrm{GeV}$ and $\pm 15\,\mathrm{MeV}/c^2$ of $3872\,\mathrm{MeV}/c^2$. A Gaussian fit to the m_{ES} distribution is shown in Fig. 1.10 that yields 0.2 ± 2.6 events. The ratio R is compared to the $\psi(2S)$ case where R is known to be 0.60 ± 0.05:

$$\frac{\Gamma(X \to J/\psi\,\pi^0\pi^0)}{\Gamma(X \to J/\psi\,\pi^+\pi^-)} < 1.3\frac{\Gamma(\psi(2S) \to J/\psi\,\pi^0\pi^0)}{\Gamma(\psi(2S) \to J/\psi\,\pi^+\pi^-)}\,. \tag{1.32}$$

Unfortunately, with the amount of data used for this analysis it is not possible to distinguish between the $I = 0$ and $I = 1$ hypothesis.

Finally, Belle has preliminary results of the $X(3872)$ decaying to $D^0\overline{D}^0\pi^0$, where the 11.3 ± 3.6 signal events concentrate close to the threshold for the final state, which would strongly disfavor a 2^{++} assignment for the $X(3872)$ [26].



htbp

Figure 1.10: The m_{ES} distribution for (a) $B \rightarrow \psi(2S)K, \psi(2S) \rightarrow J/\psi\pi^0\pi^0$ and (b) $B \rightarrow X(3872)K, X(3872) \rightarrow J/\psi\pi^0\pi^0$ with a superimposed fit. The signal yield is 55 ± 10 events for the $\psi(2S)$ and 0.2 ± 2.6 events for the $X(3872)$.

***BABAR* at PEP-II** was the third collaboration after CDF II to confirm the existence of the $X(3872)$ in the decay mode $X(3872) \rightarrow J/\psi\pi^+\pi^-$ [27]. The most recent result uses 232 million $B\overline{B}$ events collected at the $\Upsilon(4S)$ resonance [28]. This analysis is described in full detail in Chapter 4 of this dissertation. Figure 1.11 shows the $J/\psi\pi^+\pi^-$ invariant mass for (a) $B^- \rightarrow X(3872)K^-$ and (b) $B^0 \rightarrow X(3872)K^0_S$. 61.2 ± 4.5 signal events at 7.5σ statistical significance and 8.3 ± 4.5 signal events at 2.6σ significance are found in the charged and neutral B-decay mode, respectively. The branching fraction and confidence intervals are,

$$\begin{aligned}
\mathcal{B}(B^- \rightarrow XK^-, X \rightarrow J/\psi\pi^+\pi^-) &= (10.1 \pm 2.5 \pm 1.0) \times 10^{-6} \\
1.34 \times 10^{-6} < \mathcal{B}(B^0 \rightarrow XK^0_S, X \rightarrow J/\psi\pi^+\pi^-) &< 10.3 \times 10^{-6} \qquad (90\%\ \text{C.L.}) \\
0.13 < R \equiv \mathcal{B}^0/\mathcal{B}^- &< 1.10 \qquad (90\%\ \text{C.L.}). \qquad (1.33)
\end{aligned}$$

Under the assumption that the excess of events in the B^0 mode is really due to the $X(3872)$, we can measure the mass difference between and the ratio of branching ratios for the $X(3872)$ in charged and neutral B decays:

$$\begin{aligned}
m_{B^- \rightarrow XK^-} &= (3871.3 \pm 0.6 \pm 0.1)\,\text{MeV}/c^2 \\
m_{B^0 \rightarrow XK^0_S} &= (3868.6 \pm 1.2 \pm 0.2)\,\text{MeV}/c^2 \\
\Delta m &= (2.7 \pm 1.3 \pm 0.2)\,\text{MeV}/c^2
\end{aligned}$$

$$R = 0.50 \pm 0.30 \pm 0.05\,. \tag{1.34}$$

The important question whether there exists a charged partner of the $X(3872)$ was

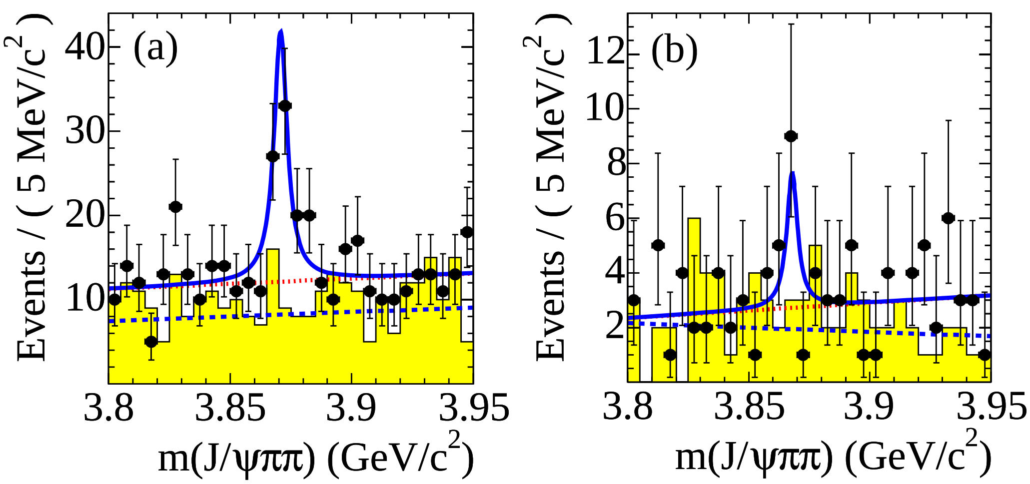


Figure 1.11: $J/\psi\,\pi^+\pi^-$ invariant mass for (a) $B^- \to X(3872)K^-$ and (b) $B^0 \to X(3872)K^0_S$. The dashed line represents the combinatorial background, the dotted lien represents the sum of the combinatorial and peaking background and the solid line the sum of all background plus the signal. The shaded area shows events in the m_{ES} sideband region $|m_{ES} - 5260\,\mathrm{MeV}/c^2| < 6\,\mathrm{MeV}/c^2$.

addressed by *BABAR* [1, 29] in searching for an excess of events in the $J/\psi\,\pi^-\pi^0$ mass spectrum that is shown in Fig. 1.12. The left plot shows events from $B^0 \to J/\psi\,\pi^-\pi^0 K^+$ and the right plot shows events from $B^- \to J/\psi\,\pi^-\pi^0 K^0_S$ decays, respectively. Using 234 million $B\overline{B}$ events no evidence for a charged partner of the $X(3872)$ was found. The limits on the branching ratios are

$$\begin{aligned} \mathcal{B}(B^0 \to X^- K^+, X^- \to J/\psi\,\pi^-\pi^0) &< 5.4 \times 10^{-6} \qquad (90\%\ \text{C.L.}) \\ \mathcal{B}(B^- \to X^- \overline{K}^0, X^- \to J/\psi\,\pi^-\pi^0) &< 22 \times 10^{-6} \qquad (90\%\ \text{C.L.}), \end{aligned} \tag{1.35}$$

and the hypothesis of a charged isospin partner of the $X(3872)$ is ruled out with a likelihood ratio test.

An inclusive search for the production of $X(3872)$ in $B^- \to X_{c\overline{c}}K^-$, where the X

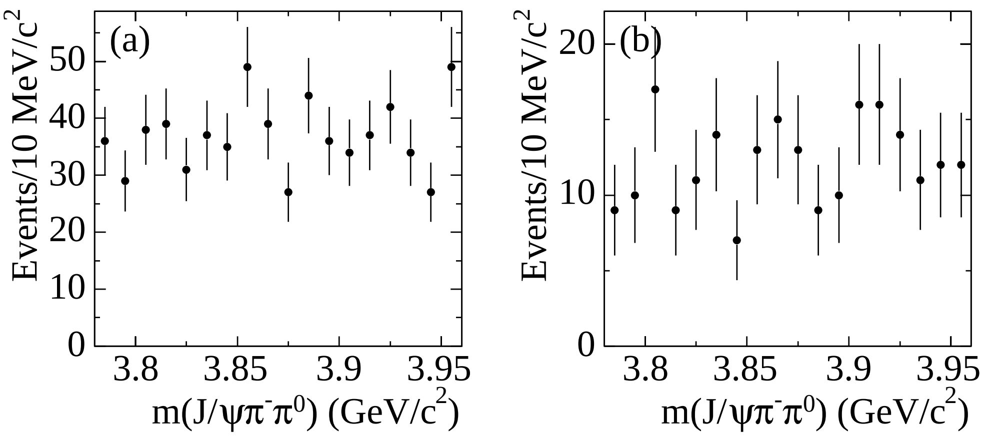


Figure 1.12: $J/\psi\,\pi^-\pi^0$ invariant mass for (a) $B^0 \rightarrow J/\psi\,\pi^-\pi^0 K^+$ (b) and $B^- \rightarrow J/\psi\,\pi^-\pi^0 K^0_S$. No evidence for a charged partner of the $X(3872)$ is found.

is a $c\overline{c}$ state, has been addressed by *BABAR* with 232 million $B\overline{B}$ events [30]. By fully reconstructing one of the two B mesons and measuring the kaon momentum spectrum in the B center-of-mass frame, any charmonia produced with the kaon can be identified in the kaon momentum spectrum. Figure 1.13 shows the known charmonium resonances but not evidence for the $X(3872)$ is found. The upper limit on the absolute branching ratio obtained by this analysis is $\mathcal{B}(B^- \rightarrow X(3872)K^-) < 3.2 \times 10^{-4}$. Together with the branching ratio measurements from Belle and *BABAR*, the lower limit $\mathcal{B}(X(3872) \rightarrow J/\psi\,\pi^+\pi^-) > 4.2\%$ at 90% C.L. can be set.

CDF II at Tevatron confirmed Belle's measurement by searching for a $J/\psi\,\pi^+\pi^-$ resonance produced inclusively in $p\overline{p}$ collisions at $\sqrt{s} = 1.96\,\mathrm{TeV}$ using $220\,\mathrm{pb}^{-1}$ of data [20]. $J/\psi \rightarrow \mu^+\mu^-$ candidates are selected by requiring $m(\mu^+\mu^-)$ within $60\,\mathrm{MeV}/c^2$ of the PDG value. In addition to that, constraints on both the vertex fit for the dimuon and the $J/\psi\,\pi^+\pi^-$ vertex are used to suppress backgrounds with the same final states. Figure 1.14 (left) shows the $J/\psi\,\pi^+\pi^-$ invariant mass distribution. Besides the $\psi(2S)$ at $3.686\,\mathrm{GeV}/c^2$ a small bump at $3.872\,\mathrm{GeV}/c^2$ is observed. The significance of the signal is reported as 11.6 standard deviations. The mass and width are obtained by modelling each peak by

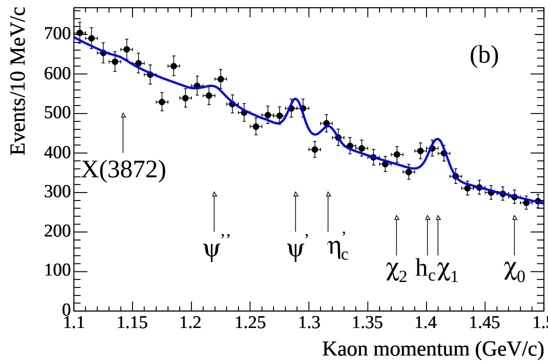


Figure 1.13: Kaon momentum spectrum for $B^- \to X_{c\bar{c}}K^-$ in the $X(3872)$ region. The arrows indicate the position of known charmonium states. No evidence for the $X(3872)$ is found.

a Gaussian and the background shape by a second order polynomial. This results in the following mean and width for the $X(3872)$

$$
\begin{aligned}
m_X &= (3871.3 \pm 0.7 \pm 0.4)\,\mathrm{MeV}/c^2 \\
\sigma_X &= (4.9 \pm 0.7)\,\mathrm{MeV}/c^2\,,
\end{aligned}
$$

where the width is consistent with the detector resolution. To investigate the dipion mass distribution, which was reported by Belle to peak near the ρ mass, the same plot is made again with the requirement $m(\pi^+\pi^-) > 500\,\mathrm{MeV}/c^2$. This is shown in Fig. 1.14 (right). The background is reduced by almost a factor of two and the fit shows that there is no change in the $X(3872)$ signal yield within statistics. This leads to the conclusion that there is little signal with dipion masses below $500\,\mathrm{MeV}/c^2$ that supports evidence that the dipion is originating from $\rho^0 \to \pi^+\pi^-$ decays.

A more detailed study of the dipion mass spectrum has been carried out by CDF II using $360\,\mathrm{pb}^{-1}$ [31]. Fig. 1.15 shows the $\pi^+\pi^-$ invariant mass spectrum for $\psi(2S)$ (left) and $X(3872)$ (right) events together with fits to various hypothesis. The spectrum is inconsistent with calculations for 1_1^P and 3_J^D charmonia. A good fit is obtained for

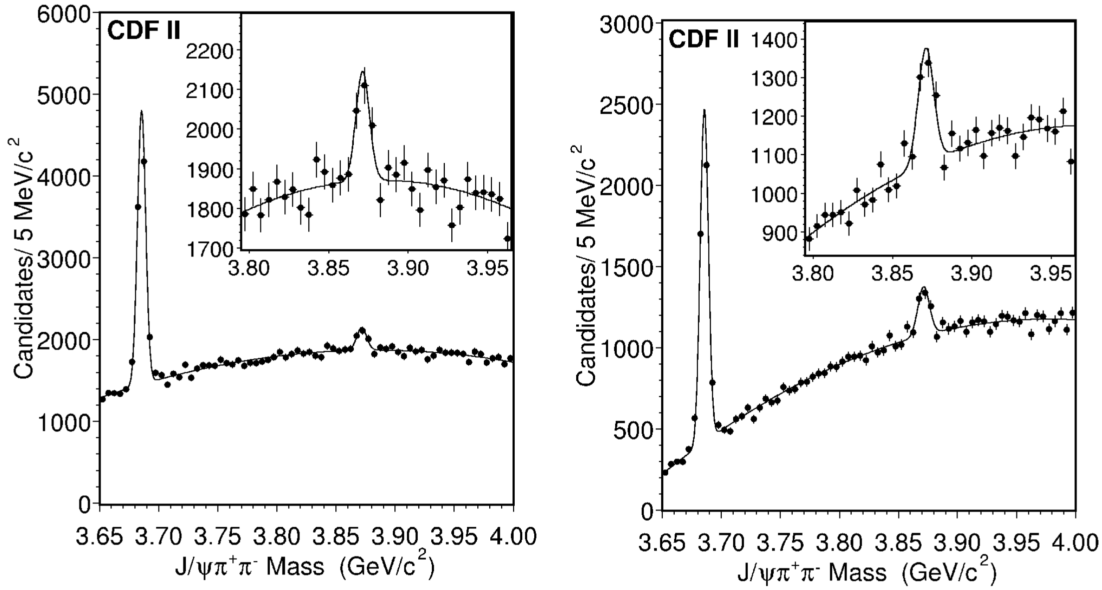


Figure 1.14: (left) The mass distribution of $J/\psi\ \pi^+\ \pi^-$ observed by CDF II in inclusive $p\ \overline{p}$ collisions. A large peak for the $\psi(2S)$ is seen and a signal near a mass of $3872\,\mathrm{MeV}/c^2$ is visible (enlargement shown in the inset). (right) shows the same requiring $m(\pi^+\pi^-) > 500\,\mathrm{MeV}/c^2$.

$X \to J/\psi\,\rho$, an interpretation supported by the C-even decay $X \to J/\psi\,\gamma$ that was shown above. The data is compatible with both S- and P-wave $J/\psi\,\rho$ decays.

DØ at Tevatron performs a similar analysis to CDF II using $230\,\mathrm{pb}^{-1}$ of data collected at the Tevatron between April 2002 and January 2004. Figure 1.16 (left) shows the $\mu^+\mu^-\pi^+\pi^-$ mass spectrum with the $\psi(2S)$ and $X(3872)$. The superimposed fit extracts 522 ± 100 $X(3872)$ candidates from the data. The mass difference between the $X(3872)$ and the J/ψ is measured to be $774.9 \pm 3.1 \pm 3.0\,\mathrm{MeV}/c^2$ and the with of the peak is $17 \pm 3\,\mathrm{MeV}/c^2$, which is consistent with the detector resolution. To investigate the characteristics of the $X(3872)$, they compare the production and decay properties of the $X(3872)$ to the $\psi(2S)$. Figure 1.16 (right) shows the signal yields for the two particles with different selections on kinematic and angular variables applied. Without going into the details at this point, no significant difference between the $X(3872)$ and the $c\overline{c}$ state $\psi(2S)$ is found.

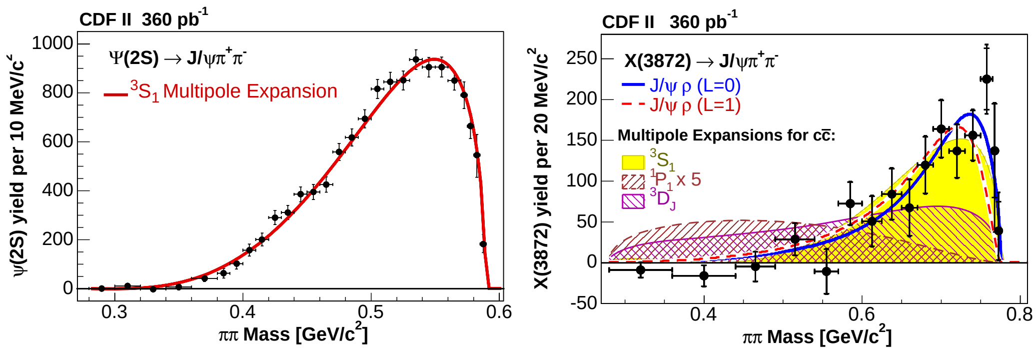


Figure 1.15: The dipion mass mass spectrum as observed by CDF II for the $\psi(2S)$ (left) and $X(3872)$ (right) together with fits to various hypothesis.

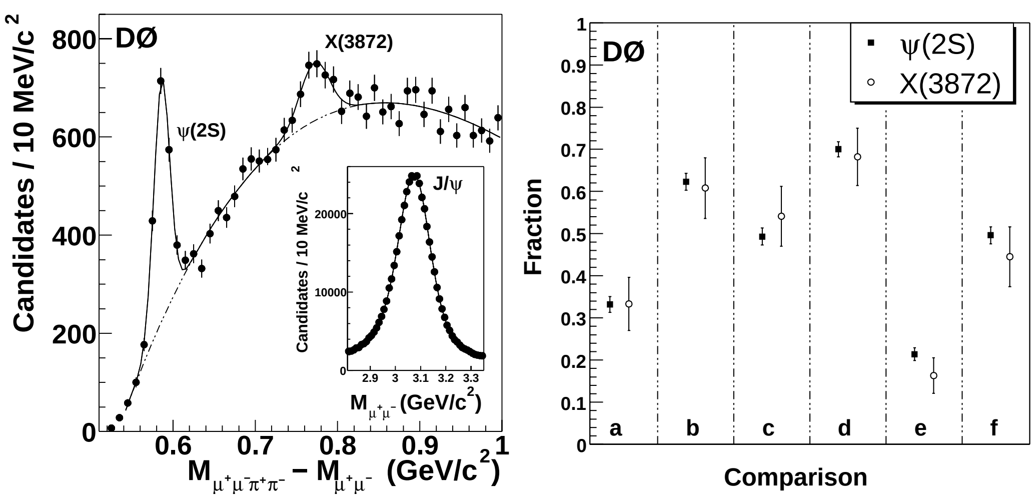


Figure 1.16: (left) The $\mu^+\mu^-\pi^+\pi^-$ mass spectrum with the $\psi(2S)$ and $X(3872)$. The insert shows the mass distribution of the J/ψ candidates used in the analysis. (right) Comparison of event-yield fractions for $X(3872)$ and $\psi(2S)$ in the regions: (a) $p_T > 15\,\mathrm{GeV}/c$, (b) rapidity of $J/\psi\pi^+\pi^- < 1$, (c) $\cos\theta_\pi < 0.4$, (d) effective proper decay length $< 0.01cm$, (e) isolation $= 1$, (f) $\cos\theta_\mu < 0.4$.

1.4.2 Diquark-Antidiquark model

The possibility of diquarks (qq) in addition to mesons ($q\overline{q}$) and baryons (qqq) as building blocks for hadrons is almost as old as QCD itself [32, 33]. Recently, diquarks have been revived as possible explanations for some of the more exotic particles observed in experiments (for a recent review see [34] and references therein). Since each quark is a $SU(3)$ color triplet, the possible diquark color multiplets are $\mathbf{3} \otimes \mathbf{3} = \mathbf{6} \oplus \overline{\mathbf{3}}$. Inevitably, a diquark is a colored object and therefore cannot exists on its own. Whether the two quarks attract or repel is determined by the color factor of the multiplet. Using the proper quark gluon interaction vertices from QCD, one obtains $f_{\mathbf{6}} = -1/3$ and $f_{\overline{\mathbf{3}}} = 2/3$ where positive (negative) numbers indicate attraction (repulsion) [35]. Including flavor $SU(3)$ symmetry the four possible diquark configurations are $(\overline{\mathbf{3}}_{\mathbf{c}}, \overline{\mathbf{3}}_{\mathbf{f}})$ and $(\overline{\mathbf{3}}_{\mathbf{c}}, \mathbf{6}_{\mathbf{f}})$ with each spin 0 and 1. It turns out that $(\overline{\mathbf{3}}_{\mathbf{c}}, \overline{\mathbf{3}}_{\mathbf{f}})$ with spin 0 is the most promising candidate for a diquark ("good" diquark in [34]). Combining a diquark and an antidiquark results in a $\mathbf{3} \otimes \overline{\mathbf{3}} = \mathbf{8} \oplus \mathbf{1}$ multiplet both in flavor and color giving rise to an observable flavor nonet, which is a color singlet.

L. Maiani *et al.* use this idea in their 2005 diquark-antidiquark model [36] to accommodate some recently discovered new states ($X(3872)$, $Y(4260)$, $D_s(2317)$, ...) into the quark model. As most other quark models, the quark masses are determined from a spin-dependent interaction Hamiltonian

$$\mathcal{H} = \sum_i m_i + \sum_{i<j} 2\kappa_{ij}(S_i \cdot S_j)\,, \tag{1.36}$$

where m_i are the constituent quark masses, κ_{ij} depend on the flavor and color state of the quark pair and S_i is the spin of the quark. A diquark-antidiquark pair is defined as $[cq][\overline{cq}]$ with $q = u, d$ (q and $\overline{q}$ do not need to be of the same flavor). Applying (1.36) yields

$$\mathcal{H} = 2m_{[cq]} + 2(\kappa_{cq})_{\overline{\mathbf{3}}}(S_c \cdot S_q + S_{\overline{c}} \cdot S_{\overline{q}}) + 2\kappa_{q\overline{q}} S_q \cdot S_{\overline{q}} + 2\kappa_{c\overline{q}}(S_c \cdot S_{\overline{q}} + S_{\overline{c}} \cdot S_{\overline{q}}) + 2\kappa_{c\overline{c}} S_c \cdot S_{\overline{c}}\,, \tag{1.37}$$

where $(\kappa_{cq})_{\bar{\mathbf{3}}}$ indicates that the two quarks are in a color triplet whereas all the other quark pairs are in a color singlet configuration likewise to mesons. The latter are obtained by fitting the model to the observed hyperfine splittings of $L = 0$ mesons (K, K^*) and the color triplet coefficients are determined likewise from Baryon splittings (Λ, Σ, Y^*). The $X(3872)$ is used as the $J^{PC} = 1^{++}$ state with the symmetric spin distribution $[cq]_{S=1}[\overline{cq}]_{S=0} + [cq]_{S=0}[\overline{cq}]_{S=1}$ and a mass $M(1^{++})$ fixed to the experimentally observed value. Diagonalizing (1.37) and using $2S_1 \cdot S_2 = (S_1 + S_2)^2 - S_1^2 - S_2^2$ the constituent diquark mass $m_{[cq]}$ can be calculated from

$$M(1^{++}) = 2m_{[cq]} - (\kappa_{cq})_{\bar{\mathbf{3}}} + \frac{1}{2}\kappa_{q\bar{q}} - \kappa_{c\bar{q}} + \frac{1}{2}\kappa_{c\bar{c}}. \tag{1.38}$$

as $m_{[cq]} = 1933\,\mathrm{MeV}/c^2$. Using this as input for the other eigenvalues of the Hamiltonian the spectrum of X particles shown in Fig. 1.17 is obtained.

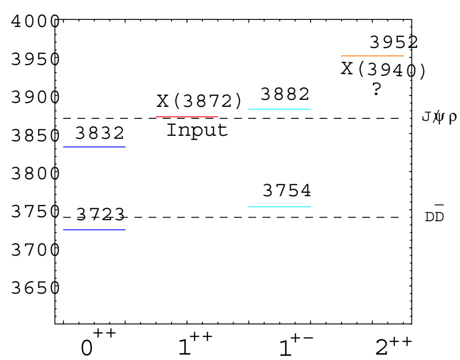


Figure 1.17: The full spectrum of X particles. The $X(3872)$ $J^{PC} = 1^{++}$ state is used as the input for the calculation of the other masses. The dashed lines indicate possible decay channel thresholds [36].

Each of these X particles comes with four different quark contents: two neutral states $X_u = [cu][\bar{c}\bar{u}]$, $X_d = [cd][\bar{c}\bar{d}]$ and two charged states $X^+ = [cu][\bar{c}\bar{d}]$, $X^- = [cd][\bar{c}\bar{u}]$. Allowing for mixing between the neutral mass eigenstates results in a low-mass and a high-mass state as described by

$$\begin{pmatrix} X_l \\ X_h \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} X_u \\ X_d \end{pmatrix}. \tag{1.39}$$

For the difference in the mass of the low and high mass state, the prediction is

$$m(X_h) - m(X_l) = 2\frac{m_d - m_u}{\cos(2\theta)} = \frac{7 \pm 2}{\cos(2\theta)}\,\mathrm{MeV}/c^2\,, \tag{1.40}$$

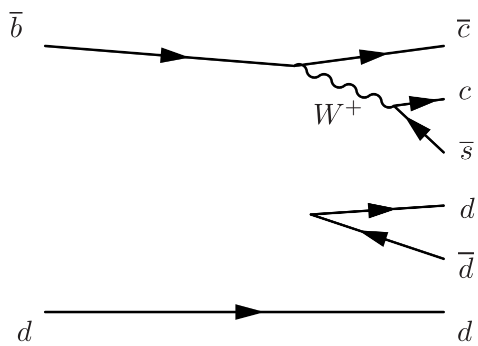
where the mixing angle $\theta \approx \pm 20^o$ is calculated from the ratio of $X \to J/\psi\,3\pi$ and $X \to J/\psi\,2\pi$ events reported by Belle [24]. This results a mass difference of

$$m(X_h) - m(X_l) = (8 \pm 3)\,\mathrm{MeV}/c^2. \tag{1.41}$$

Figure 1.18 shows the Feynman diagram for $B^0 \to XK^0$ decays. Note that both X_u and X_d can be produced in neutral/charged B-decays. The same is true for X^+ and X^-. This is one of the most important predictions of the diquark-antidiquark model: since the narrow width of the observed $X(3872)$ does not allow for two nearby states produced at the same time, one of the neutral X states has to dominate B^0 decays and one has to dominate B^- decays. If this is true, the two states should have a mass difference given by (1.40). Moreover, since the amplitudes of weak B^0 and B^- decays are related by isospin symmetry, the branching ratios for neutral and charged X particles can be related to each other. Assuming that indeed one of the X particles dominates B^0 decays, the following lower bounds can be obtained:

$$R^- = \frac{\mathcal{B}(B^- \to X^- K^0_S, X^- \to J/\psi\,\pi^-\pi^0)}{\mathcal{B}(B^- \to X_{l/h}K^-, X_{l/h} \to J/\psi\,\pi^+\pi^-)} > 0.2 \tag{1.42}$$

$$R^0 = \frac{\mathcal{B}(B^0 \to X^- K^+, X^- \to J/\psi\,\pi^-\pi^0)}{\mathcal{B}(B^- \to X_{l/h}K^-, X_{l/h} \to J/\psi\,\pi^+\pi^-)} > 0.53. \tag{1.43}$$



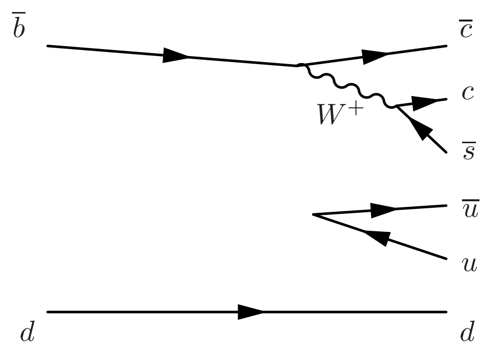


Figure 1.18: Feynman diagrams for the production of $X_u([cu][\overline{cu}])$ (left) and $X_d([cd][\overline{c}\overline{d}])$ (right) in $B^0 \rightarrow J/\psi\pi^+\pi^-K^0$ decays. The $K^0(d\bar{s})$ is either produced with the d quark from the vacuum (left) or with the spectator-d (right). Similar diagrams hold for charged B decays as well as for the production of X^+ and X^-.

In summary, the diquark-antidiquark model predicts a complete spectrum of new X particles. In particular, it predicts four distinct states for the $X(3872)$, two neutral and two charged states. Finding the charged states and measuring the mass difference between the two neutral states will determine if the model is correct or not.

1.4.3 Meson-Meson bound states

Similar to the binding of a proton and a neutron in the deuterium (^2H) nucleus (deuteron) one could imagine that two mesons can form a bound state. As for the deuteron the binding force would be mediated by pion exchange. A detailed analysis of possible bound meson states is presented by N. Törnqvist in [37, 38], which he calls “deusons”. Unfortunately, the nomenclature is not consistent throughout the literature and other authors refer to the same states as meson “molecules”.

Restricting ourselves to ground-state mesons ($l = 0$), we need to consider possible bound states of spin-0 pseudoscalars (P) with [4] $J^{PC} = 0^{-+}$ and spin-1 vectors (V) with $J^{PC} = 1^{--}$. Since the pion itself is a vector ($J^P = 1^-$), parity conservation forbids a bound state of two pseudoscalars (PP). Therefore only PV and VV deusons are possible. Furthermore, pion exchange is in general much weaker (or repulsive) for isovectors than

[4] $P = (-1)^{l+1}$ and $C = (-1)^{l+s}$ for a $q\overline{q}'$ meson with angular momentum l and spin s.

isoscalars, reducing the possible meson combinations even more. Considering only K, D and B mesons as candidates for P or V the possible combinations (all isoscalars) with sufficient binding energy are $P\bar{V}^*$ with $J^{PC} = 0^{-+}, 1^{++}$ and $V\bar{V}^*$ with $J^{PC} = 0^{++}, 0^{-+}, 1^{+-}, 2^{++}$. A detailed calculation and an extensive list can be found in [38]. In the following we only consider bound states of D mesons due to their recent experimental relevance in light of the $X(3872)$.

The most intriguing feature of the $X(3872)$ is that its mass is very close to the $D^0 D^{*0}$ threshold ($m(D^0) + m(D^{*0}) = 3871.2 \pm 0.7\,\mathrm{MeV}/c^2$). This suggests that it is a good candidate for a $P\bar{V}^*$ (i.e. $D^0\overline{D}^{*0}$) deuson and several authors recently investigated this possibility in great detail [39–49]. As explained above, possible quantum numbers are 1^{++} if the two D mesons are in an relative S-wave or 0^{-+} for a P-wave. Higher total angular momentum states would result in a higher mass and are thus not considered as candidates for the $X(3872)$. Other properties are:

- Decays via DD^* annihilation are expected to be small because of the small binding energy (deuteron $\sim 2.2\,\mathrm{MeV}$) and thus large spatial size resulting in a narrow width. OZI allowed decays to J/ψ will be favored compared to states with only light hadrons.

- The D^+D^{*-} channel is about $8\,\mathrm{MeV}/c^2$ higher than the observed peak and therefore closed by phase space. However, see below for isospin breaking effects.

- For exact isospin and pion exchange, only the isoscalar channel ($I = 0$) results in an attractive potential. Thus in the deuson model, the $X(3872)$ is a isoscalar.

- Charged combinations like D^0D^{*+} with vanishing total charm are necessarily isovectors and therefore not predicted by the model [50].

If the $X(3872)$ would really be an isosinglet then the decay to $J/\psi\,\rho^0$ would be forbidden by isospin conservation unless there is isospin mixing in the initial state, which we will discus now in more detail.

Isospin breaking in bound D mesons [41,47] could be produced by mixing of different isospin doublets in the final state. In the case of the $X(3872)$, the nearby D^+D^{*-} threshold at $3879.5\,\mathrm{MeV}/c^2$ suggests that a DD^* molecule at $3872\,\mathrm{MeV}/c^2$ would consist of a mixture of $D^0\overline{D}^{*0}$ and D^+D^{*-}. Assuming the mixture is described by $\alpha \in [0,1]$ and $|X\rangle$ is the wave function of the $X(3872)$ deuson we get [5]

$$|X\rangle = \frac{(1+\alpha)D^0\overline{D}^0 + (1-\alpha)D^+D^-}{\sqrt{2}\sqrt{\alpha^2+1}}\,, \tag{1.44}$$

which can be written in terms of quark content using $D^0(c\overline{u})$ and $D^+(c\overline{d})$ as

$$\begin{aligned} |X\rangle &= \frac{c\overline{c}}{\sqrt{2}\sqrt{\alpha^2+1}}\left((1+\alpha)u\overline{u} + (1-\alpha)d\overline{d}\right) \\ &= \frac{c\overline{c}}{\sqrt{\alpha^2+1}}\left(\frac{u\overline{u}+d\overline{d}}{\sqrt{2}} + \alpha\frac{u\overline{u}-d\overline{d}}{\sqrt{2}}\right) \\ &= \frac{c\overline{c}}{\sqrt{\alpha^2+1}}\left(|I=0\rangle + \alpha\,|I=1\rangle\right)\,, \end{aligned}$$

where $|I=0,1\rangle$ is the usual isosinglet and isotriplet respectively, obtained from combining a u and d quark. We obtain a pure $I=0$ state for $\alpha=0$ equivalent to equal contributions from $D^0\overline{D}^0$ and D^+D^- in the initial state (1.44). Since the mass of the $X(3872)$ is closer to the $D^0\overline{D}^{*0}$ threshold than to the D^+D^{*-} we expect less contribution from the latter ($\alpha > 0$) resulting in a final state $|X\rangle$ of undefined isospin and G-parity.[6] Due to these isospin breaking effects both the decays $X(3872) \to J/\psi\,\rho^0(\pi^+\pi^-)$ and $X(3872) \to J/\psi\,\omega(\pi^+\pi^-\pi^0)$ are permitted to certain extends. Radiative decays to $\chi_{c1}\gamma$ will be forbidden by C-parity if the $X(3872)$ is either the 1^{++} or 0^{-+} deuson, which agrees with Belle's non-observation of this mode. Finally, it is worth noting that using a specific decay model for the $X(3872)$ as a deuson, E. Braaten and M. Kusunoki [40] predict that the branching fraction for $B^0 \to X(3872)K^0$ is likely to be suppressed by at least an order of magnitude compared to that for $B^- \to X(3872)K^-$. However, other models do not make such a strong prediction.

[5] To simplify the notation, one of the D mesons is always to be considered a D^* and $D\bar{D} = \left(|D\bar{D}\rangle + |\bar{D}D\rangle\right)/\sqrt{2}$.

[6] G is the conserved quantum number resulting from isospin symmetry with $G = (-1)^I C$

Chapter 2

Analysis techniques

This section describes common concepts and variables used in the following analyses. In fact, most of these techniques will be used in any data analysis in high energy physics. We start by introducing variables that can be used to discriminate between certain event types. In most cases these types consists of signal and background events. The goal is to classify each event by exploiting certain properties of the kinematics, angular distribution, spatial shape, etc. Usually, once the events have been classified, a unknown physical parameter (e.g. mass and width of a resonance, branching ratio, *CP* parameter, ...) needs to be extracted from the signal events. Depending on the parameter this can be done by simply counting events or by fitting the data to an analytic function derived from a physical model of the process. The latter is performed with the method of maximum likelihood estimators that is described further below.

2.1 Discriminating variables

2.1.1 Kinematic variables ΔE and m_{ES}

B mesons that are produced in $e^+e^- \rightarrow B\overline{B}$ can be identified by their well constrained kinematics due to the known initial momentum of the e^+e^- system. We follow the notation in [51, 52] to derive two kinematic variables (ΔE and m_{ES}) that can be used to separate true B mesons from background events.

In the reaction of $e^+e^- \to B\overline{B}$, we can write energy-momentum conservation as

$$q_0 = q_1 + q_2 \tag{2.1}$$

where q_0 is the 4-momentum of the e^+e^- system and $q_{1,2}$ are the 4-momenta of the two B mesons. Since $q_1^2 = q_2^2 = m_B^2$ for true B mesons, we can test for any reconstructed $B\overline{B}$ pair if (I) q_1^2 and q_2^2 equal to each other and (II) their common value equals m_B^2.

Test (II) is represented by the invariant $\Delta q^2 \equiv q_1^2 - q_2^2$. With the CM energy squared $s = q_0^2 = (q_1 + q_2)^2$ we can write this as

$$\begin{aligned} \Delta q^2 &\equiv q_1^2 - q_2^2 \\ &= q_1^2 - (q_0 - q_1)^2 \\ &= 2q_0 q_1 - s. \end{aligned} \tag{2.2}$$

In the CM frame where $q_1 = (E_1^*, 0)$ and $q_0 = (\sqrt{s}, 0)$ this reads

$$\Delta q^2 = 2\sqrt{s}E_1^* - s\,. \tag{2.3}$$

Dividing both sides by $2\sqrt{s}$ we can define

$$\Delta E \equiv \Delta q^2/2\sqrt{s} = E_1^* - \sqrt{s}/2, \tag{2.4}$$

which will be peaked at zero for $e^+e^- \to B\overline{B}$ events since it is the energy of the B candidate minus its expected CM energy.

For the second independent variable we evaluate q_1^2 with the constraint $\Delta q^2 = 2q_0q_1 - s = 0$ (equivalent to $\Delta E = 0$) and define $m_{ES} \equiv q_1^2|_{\Delta E=0}$. Evaluated in the CM frame the constraint is $E_1^* = \sqrt{s}/2$ and therefore

$$\begin{aligned} m_{ES}^2 &\equiv q_1^2|_{\Delta E=0} = E_1^{*2} - \mathbf{p}_1^{*2} = (\sqrt{s}/2)^2 - \mathbf{p}_1^{*2}\,, \\ m_{ES} &= \sqrt{s/4 - \mathbf{p}_1^{*2}}\,, \end{aligned} \tag{2.5}$$

which corresponds to the mass of the B candidate using its measured momentum and fixing its expected CM energy. If we instead evaluate q_1^2 in the laboratory frame the $\Delta E = 0$ constraint yields $E_0 E_1 - \mathbf{p}_0 \cdot \mathbf{p}_1 = s/2$ and

$$m_{ES}^2 \equiv q_1^2|_{\Delta E=0} = E_1^2 - \mathbf{p}_1^2 = (s/2 + \mathbf{p}_0 \cdot \mathbf{p}_1)^2/E_0^2 - \mathbf{p}_1^2. \tag{2.6}$$

m_{ES} is called the energy-substituted mass for obvious reasons as can be seen from (2.5). The use of m_{ES} in the CM frame (2.5) is complicated at an asymmetric e^+e^- collider because one requires particles masses for the boost to the CM frame. With a slight change in notation we summarize the two kinematic variables to be used in identifying B mesons in $e^+e^- \to B\overline{B}$ collisions:

$$\Delta E = E_B^* - \sqrt{s}/2 \tag{2.7}$$

$$m_{ES} = \sqrt{(s/2 + \mathbf{p}_0 \cdot \mathbf{p}_B)^2/E_0^2 - \mathbf{p}_B^2} \tag{2.8}$$

where $E_0 = E_{e^+} + E_{e^-}$ and $\mathbf{p}_0 = \mathbf{p}_{e^+} + \mathbf{p}_{e^-}$ is the total energy and momentum of the e^+e^--system, $s = E_0^2 - \mathbf{p}_0^2$ the center of mass energy squared, E_B^* the energy of the reconstructed B meson in the e^+e^--rest frame and $\mathbf{p}_B$ the 3-momentum of the B meson in the laboratory frame. For a true B meson we expect ΔE to peak at zero. The energy substituted mass m_{ES} should, of course, peak at the nominal value of the B mass around $5.279\,\mathrm{GeV}/c^2$. Due to the detector resolution effect, it turns out that ΔE and m_{ES} represent an almost uncorrelated pair of variables in the two dimensional space of momentum and energy. Other pairs of variables like mass and energy or mass and momentum are strongly correlated and less useful. Figure 2.1 shows an example for $B^- \to J/\psi\,\pi^+\pi^- K^-$ signal Monte Carlo events.

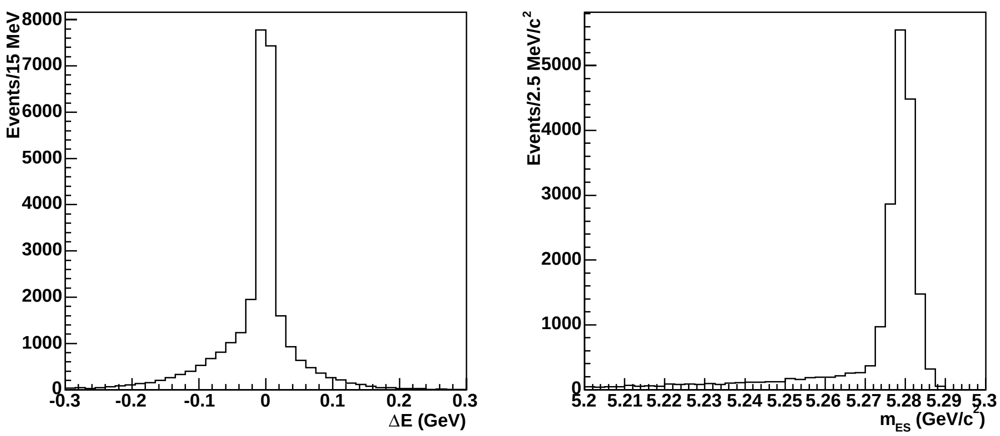


Figure 2.1: ΔE (left) and m_{ES} (right) for $B^- \to J/\psi\,\pi^+\pi^- K^-$ signal Monte Carlo events.

2.1.2 Thrust angle

To discriminate between certain types of background it is useful to have a measure for the spherical symmetry of an event. A particle that decays into two particles with opposite momenta or a particle decaying into two jets has less spherical symmetry than a decay into many particles with random momenta. For this reason we define the thrust of an event as follows:

$$T = \max_{|\hat{\mathbf{T}}|=1} \frac{\sum_i |\hat{\mathbf{T}} \cdot \mathbf{p}_i|}{\sum_i |\mathbf{p}_i|} \tag{2.9}$$

$\hat{\mathbf{T}}$ is called the (normalized) thrust axis of the event and points in the direction that maximizes the sum of the longitudinal momenta of the particles. The summation goes over all the particles in the event with three-momentum $\mathbf{p}_i$. The thrust T can have values between 0.5 and 1, where the latter one describes highly directional events and the former one highly isotropic events.

The thrust angle θ_{Thrust} is now defined as the angle between the thrust axis of the B

candidate and the thrust axis of the rest of the event where all the calculations are done in the CM system, which is the $\Upsilon(4S)$ rest frame. For a typical continuum background event with a two-jet structure the B thrust axis will be collinear with the thrust axis for the rest of the event. For a true signal event those two thrust axis will be uncorrelated. Thus, we expect $\cos\theta_{Thrust}$ to peak near 1 for continuum background and a flat distribution for signal events.

2.1.3 Fox-Wolfram moment

The Fox-Wolfram moments H_l are defined by [53]

$$H_l = \sum_{i,j} \frac{|\mathbf{p}_i||\mathbf{p}_j|}{E_{vis}^2} P_l(\cos\theta_{ij}) \tag{2.10}$$

where θ_{ij} is the angle between the particle momenta $\mathbf{p}_i$ and $\mathbf{p}_j$ and E_{vis} is the total visible energy of the event. The $P_l(x)$ are the Legendre polynomials, $P_0(x) = 1, P_1(x) = x, \ldots$ Energy-momentum conservation requires that $H_0 \simeq 1$ and $H_1 = 0$ if we assume negligible contributions from the particle masses. It is therefore customary to normalize the results to H_0 and we define the second Fox-Wolfram moment as

$$R2 = \frac{H_2}{H_0} \tag{2.11}$$

The highly directional continuum events tend to have high $R2$-values whereas the more spherical B events have lower values of $R2$. See Fig. 4.1 in section 4.2.5 for an example.

2.1.4 Helicity angles

Particle decays have a certain angular distribution depending on the spin structure of initial and final states. The calculation of the matrix element is done within the helicity formalism that was developed in [54]. More accessible descriptions can be found in [55–57]. The main idea is that the helicity operator $\hat{h} = \hat{\mathbf{S}} \cdot \hat{\mathbf{p}}$ with spin $\mathbf{S}$ and linear momentum

p is invariant under rotations and boosts along the momentum direction resulting in well defined angular distributions. From the experimentalists point of view, we compare the angular distributions of particle decays for different types of backgrounds and signal and separate them by imposing selections on the helicity angle. The definition of the helicity angle is pictured in Fig. 2.2. For a two-body decay $A \to B + X, B \to C + D$ the helicity angle is measured in the rest-frame of B between the direction of its parent (A) and daughter (D). Any sequential two-body decay can be described using these angles in the particles respective rest frame.

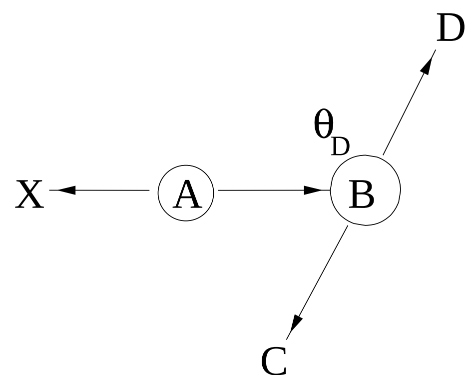


Figure 2.2: Definition of the helicity angle θ_D in the two-body decay $A \to B + X, B \to C + D$. The angle is measured between the direction of D and A in the rest-frame of B.

2.2 Dalitz analysis technique

The Dalitz plot technique was first introduced by Richard Dalitz in 1953 for the analysis of τ-meson data. It allows to represent the entire phase space of any three-body final state in a two-dimensional scatter plot.

A three-body decay $M \to m_1+m_2+m_3$ (see Fig. 2.3 left) of a mother particle with mass M into three daughter particles of masses m_i has 12 degrees of freedom. Four-momentum conservation and the knowledge of the three daughter masses puts seven constraints on the kinematics. In the rest frame of M where $\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3 = 0$ the momenta of all three daughter particles lie in a plane. The orientation of this plane described by the three Euler angles (α, β, γ), and the energy of two of the daughters can be used to fully describe the kinematics. The partial decay rate is [8]

$$d\Gamma = \frac{1}{(2\pi)^5}\frac{1}{16M}\left|\mathcal{M}\right|^2\, dE_1\, dE_2\, d\alpha\, d\cos(\beta)\, d\gamma \tag{2.12}$$

where $\mathcal{M}$ is the Lorentz-invariant matrix element. If M is a scalar (spin 0) particle or if we average over the spin states, integration over the angles in (2.12) yields:

$$d\Gamma = \frac{1}{(2\pi)^3}\frac{1}{32M^3}\overline{\left|\mathcal{M}\right|^2}\, ds_{12}\, ds_{23}\,, \tag{2.13}$$

where $s_{ij} = m_{ij}^2 = (p_i + p_j)^2$ is the invariant mass squared of particles i and j. The equation gives the partial decay rate of M for the differential phase space $ds_{12}\, ds_{23}$. A two-dimensional scatter plot of the decay in s_{12} and s_{23} is called a Dalitz plot. For a decay with constant matrix element $\mathcal{M}$, an important feature is that the Dalitz plot has a uniform density. Any additional structure or inhomogeneity must result from a non-constant matrix element, i.e. the decay is not a pure three-body phase space decay. Information about $\mathcal{M}$ and intermediate resonances can be extracted directly from the Dalitz plot. The kinematic allowed region for a three-body decay is illustrated in Fig. 2.3 (right) on the Dalitz plane. The allowed values for s_{23} for given s_{12} are conveniently evaluated in the rest frame of m_{12}:

$$(s_{23})_{\max} = (E_2^* + E_3^*)^2 - (|\mathbf{p}_2^*| - |\mathbf{p}_3^*|)^2 \tag{2.14}$$

$$(s_{23})_{\min} = (E_2^* + E_3^*)^2 - (|\mathbf{p}_2^*| + |\mathbf{p}_3^*|)^2, \tag{2.15}$$

where * denotes quantities in the m_{12} rest frame with $\mathbf{p}_i^{*2} = E_i^{*2} - m_i^2$, $E_2^* = (s_{12} - m_1^2 - m_2^2)/2m_{12}$ and $E_3^* = (M^2 - s_{12} - m_3^2)/2m_{12}$. Overall four-momentum conservation yields $s_{12} + s_{13} + s_{23} = M^2 + m_1^2 + m_2^2 + m_3^2$, which can be used to derive similar relations for the other particle pairs.

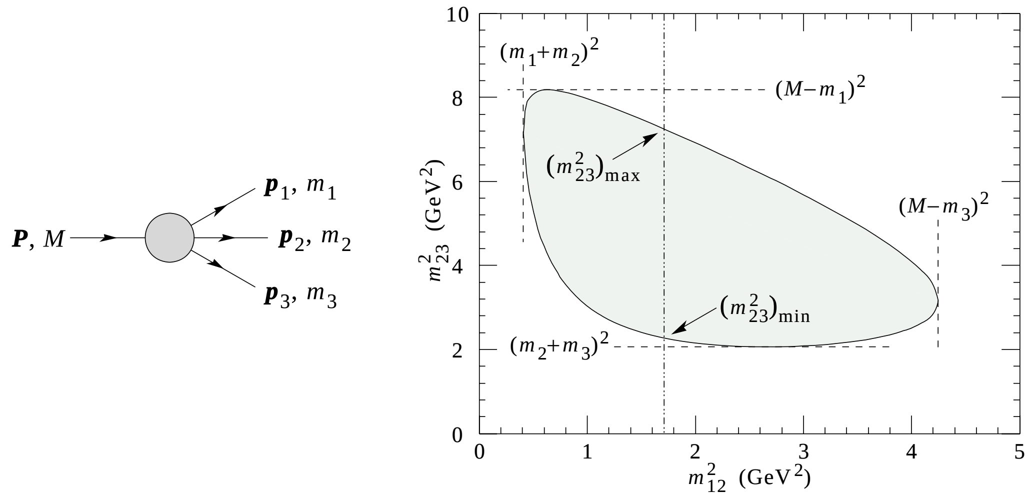


Figure 2.3: Left: A three-body decay of a parent particle with mass M and momentum P into three daughter particles with masses m_i and momenta p_i. Right: Kinematic allowed region for the three-body decay illustrated on the Dalitz plane.

2.3 Maximum likelihood estimators

The analyses described in this dissertation make extensive use of maximum likelihood (ML) estimators [58] to measure unknown parameters $\theta = (\theta_1 \ldots \theta_m)$ of a probability density function (PDF) that is used to describe a physical quantity $\tilde{x}$ (random variable). Assuming θ describes the data correctly, the probability to observe x is $\mathcal{P}(x;\theta)$. For n measurements the likelihood function

$$L(\theta) = \prod_{i=1}^{n} \mathcal{P}(x_i;\theta) \tag{2.16}$$

is the probability for n independent measurements with values x_i. Under the hypothesis that $\mathcal{P}$ is the correct model to describe the data, the ML estimators $\hat{\theta}_i$ are given by the solutions of

$$\frac{\partial L}{\partial \theta_i} = 0. \tag{2.17}$$

Since $\log(y)$ is a monotonically increasing function of y the same estimators can be obtained by solving

$$\frac{\partial \log L}{\partial \theta_i} = \frac{\partial}{\partial \theta_i} \sum_{i=1}^{n} \mathcal{P}(x_i;\theta) = 0 \tag{2.18}$$

instead. In High Energy Physics research, the program MINUIT [59] from the CERN Program Library is used to minimize the negative log-likelihood (NLL) function. In the case of a sufficient large data sample, the Hesse matrix of the NLL

$$(\widehat{V^{-1}})_{ij} = -\frac{\log L}{\partial \theta_i \partial \theta_j}\bigg|_{\theta=\hat{\theta}} \tag{2.19}$$

is an estimator for the covariance matrix of the ML estimator $\hat{\theta}$.

If the number of observations N itself is a random variable distributed according to a Poisson distribution with mean value ν the extended likelihood function is

$$L(\nu,\theta) = \frac{\nu^N}{N!} e^{-N} \prod_{i=1}^{N} \mathcal{P}(x_i;\theta). \tag{2.20}$$

This form of the likelihood function is particularly useful if $\mathcal{P}$ can be written as the sum of PDFs for different event types $t \in T$

$$\mathcal{P}(\mathbf{x};\mu) = \sum_{t\in T} \mu_t \mathcal{P}_t(\mathbf{x})\,, \tag{2.21}$$

where μ_t is the fraction of events of type t and $\sum_t \mu_t = 1$. Inserting (2.21) into (2.20) the log-likelihood function is

$$\log L(\nu,\mu) = -\nu + \sum_{i=1}^{N} \log\left(\sum_{t\in T} \nu\mu_t \mathcal{P}_t(\mathbf{x})\right)\,, \tag{2.22}$$

where we have dropped terms independent of the parameters. By defining $n_t \equiv \mu_t N$ as the expected number of events of type t, the log-likelihood function can be written as a function of the parameters $\mathbf{n}$ only:

$$\log L(\mathbf{n}) = -\sum_{t \in T} n_t + \sum_{i=1}^{N} \log \left(\sum_{t \in T} n_t \mathcal{P}_t(\mathbf{x}_i) \right). \tag{2.23}$$

With the normalization $\sum_t \mu_t = 1$ we obtain the total number of events $N = \sum_t n_t$ as a sum of independent Poisson variables with mean n_t. The estimators $\hat{n}_t$ give directly the estimated mean numbers of events of type t.

2.4 Artificial neural networks

An artificial neural network [58] is a special case of a non-linear test statistics to distinguish between two hypothesis H_0 (null hypothesis) and H_1. According to the Newman-Pearson lemma, a one-dimensional test statistics $t(\mathbf{x})$ for a vector of data $\mathbf{x} = (x_1, \ldots, x_n)$ with the maximum power is simply the likelihood ratio

$$t(\mathbf{x}) = \frac{f(\mathbf{x}|H_0)}{f(\mathbf{x}|H_1)}. \tag{2.24}$$

Of course, in practice we do not know $f(\mathbf{x}|H_{0,1})$ but we can still make a simpler *Ansatz* for the functional form of $t(\mathbf{x})$. Suppose we take $t(\mathbf{x})$ to be of the form

$$t(\mathbf{x}) = K\left(a_0 + \sum_{i=1}^{n} a_i x_i\right). \tag{2.25}$$

The function $K(z)$ is called the activation function and usually taken to be the sigmoid function

$$K(z) = \frac{1}{1 + e^{-z}}. \tag{2.26}$$

The test statistics $t(\mathbf{x})$ can be seen as the output node of a network with n inputs x_i and is called a single-layer perceptron. If the activation function is monotonic (i.e. the sigmoid function) the single-layer perceptron is equivalent to a linear test statistics of which the

Fisher discriminant is one commonly used example. However, the single-layer perceptron can easily be extended by adding additional (hidden) layers h_j with m nodes between the input and output layer. Instead of (2.25) we write

$$t(\mathbf{x}) = K\left(a_0 + \sum_{j=1}^{m} a_j h_j(\mathbf{x})\right) \tag{2.27}$$

and the h_j themselves are given as functions of the nodes in the previous layer (here, the input layer)

$$h_j(\mathbf{x}) = K\left(w_{j0} + \sum_{i=1}^{n} w_{ij} x_i\right). \tag{2.28}$$

In general, this can be extended to multiple hidden layers and is called a multilayer perceptron (MLP). Of course, the task remaining is to determine the weights a_j and w_{ij}. This is called training of the neural network and can be accomplished by minimizing an error function like

$$\mathcal{E} = E_0[(t(\mathbf{x}) - t_0)^2] + E_1[(t(\mathbf{x}) - t_1)^2] \tag{2.29}$$

where E_α is the expectation value and t_α is the preassigned target value for the hypothesis H_α. In practice, if the weights are determined form a training sample of size N, the expectation values in (2.29) can be replaced by the sum of the means:

$$\mathcal{E} = \frac{1}{2N} \sum_{p=1}^{N} (t(\mathbf{x}^p) - t^p)^2 \tag{2.30}$$

where t^p is the desired output value for the pattern (event) p. For example, $t^p = 0$ if $\mathbf{x}^p$ is a background event and $t^p = 1$ if $\mathbf{x}^p$ is a signal event. Choosing the input variables x_i, the number of hidden nodes h_j, the number of total layers and constructing a proper training sample are crucial parameters that determine the separation power of the neural network.

Chapter 3

Measurement of CP parameters in $\mathbf{B^- \to DK^-,\ D \to \pi^+\pi^-\pi^0}$

This chapter presents the first measurement of *CP* parameters in $B^- \to DK^-$, $D \to \pi^+\pi^-\pi^0$ decays. The analysis uses the model-dependent GGSZ Dalitz analysis method that was presented in section 1.3.3 of this dissertation. It is based on the previous branching ratio measurement by *BABAR* [60, 61]. Using 229 million $B\overline{B}$ events, the results for the branching ratio and asymmetry from 133 ± 23 signal events are

$$\begin{aligned}
\mathcal{B}(B^- \to D_{\pi^+\pi^-\pi^0}K^-) &= (5.5 \pm 1.0 \pm 0.7) \times 10^{-6} \\
A(B^- \to D_{\pi^+\pi^-\pi^0}K^-) &= -0.02 \pm 0.16 \pm 0.03\,,
\end{aligned} \tag{3.1}$$

where the asymmetry is defined as $A = (N^- - N^+)/(N^- + N^+)$ with $N^\pm$ as the number of $B^\pm$ signal events. The analysis presented in the following uses the same signal extraction technique as the previous analysis and performs a Dalitz fit of the D decay to extract the *CP* parameters related to the CKM phase γ. Where necessary we will give a brief description of the techniques used in the branching ratio analysis. Otherwise we will concentrate on the *CP*-fit.

3.1 Introduction

We briefly summarize the important equations for a measurement of γ in $B^- \to D^0 K^-$. The reader is referred to section 1.3 for detailed information. The sensitivity to γ arises from the interference of the Cabibbo allowed $B^- \to D^0 K^-$ decay and the doubly Cabibbo suppressed decay $B^- \to \overline{D}^0 K^-$ with $D \to \pi^+\pi^-\pi^0$ in both cases.[7] We define the amplitudes

$$\begin{aligned} A(B^- \to D^0 K^-) &\equiv A_B & (3.2)\\ A(B^- \to \overline{D}^0 K^-) &\equiv A_B r_B e^{i(\delta_B - \gamma)} & (3.3)\end{aligned}$$

with δ_B being the difference in the strong phase between the two decays and γ being the weak phase. Using $f_{D^0}(s_{+0}, s_{-0})$ and $f_{D^0}(s_{-0}, s_{+0})$ from Eq. (1.18) for the amplitudes of the 3-body D^0 and $\overline{D}^0$ decay we obtain the amplitude for the cascade decays as

$$\begin{aligned} A_{B^- \to D_{\pi^+\pi^-\pi^0} K^-}(s_{+0}, s_{-0}) &= A_B \left(f_{D^0}(s_{+0}, s_{-0}) + z_- f_{D^0}(s_{-0}, s_{+0}) \right) \\ A_{B^+ \to D_{\pi^+\pi^-\pi^0} K^+}(s_{+0}, s_{-0}) &= A_B \left(f_{D^0}(s_{-0}, s_{+0}) + z_+ f_{D^0}(s_{+0}, s_{-0}) \right), & (3.4)\end{aligned}$$

where $z_\pm = r_B\, e^{i(\delta_B \pm \gamma)}$ is the parameter that describes the *CP* violation. With r_B being a positive small number of order 0.1 it is experimentally difficult to perform the analysis in the $z(r_B, \delta_B, \gamma)$ parameterization due to inevitable fit biases in r_B. Hence, both *BABAR* and Belle [11, 62] have performed their previous analyses, using different final states, in the Cartesian coordinates

$$\begin{aligned} x_\pm = \Re(z_\pm) = r_B \cos(\delta_B \pm \gamma) \\ y_\pm = \Im(z_\pm) = r_B \sin(\delta_B \pm \gamma)\,, & \quad (3.5)\end{aligned}$$

[7] D denotes either D^0 or $\overline{D}^0$ throughout this document

which are unbounded and hence do not have this problem. The current best measurement of these parameters are listed in Table 3.1. Theses values can be used in a frequentist or

	$B^- \to DK^-$	$B^- \to D^*K^-$ $D^* \to D\pi^0, D\gamma$	$B^- \to DK^{*-}$ $K^{*-} \to K^0_S\pi^-$
x_-	$0.05 \pm 0.05 \pm 0.02$	$-0.13 \pm 0.08 \pm 0.02$	$-0.46 \pm 0.17 \pm 0.03$
y_-	$0.11 \pm 0.07 \pm 0.04$	$-0.20 \pm 0.09 \pm 0.03$	$0.05 \pm 0.27 \pm 0.03$
x_+	$-0.14 \pm 0.05 \pm 0.03$	$0.10 \pm 0.07 \pm 0.03$	$-0.10 \pm 0.15 \pm 0.03$
y_+	$-0.03 \pm 0.06 \pm 0.02$	$0.01 \pm 0.09 \pm 0.06$	$0.00 \pm 0.15 \pm 0.05$

Table 3.1: Average measurements of the Cartesian coordinates in different modes [10]. For each mode the D meson is reconstructed from $K^0_S\pi^+\pi^-$ final states.

Bayesian analysis to obtain a measurement of the CKM phase γ.

The plan of this analysis is as follows. First we describe the data sample and event selection. Then we establish a set of uncorrelated analysis variables and describe the probability density function for each of these variables and all the signal and background event types. The analytic parameterization of f_{D^0} is determined in a separate analysis and described briefly. We continue with a description of the maximum likelihood fit used to obtain a measurement of the CP parameters $z_\pm$. After validating the fitting procedure using Monte Carlo simulations, we present the fit result on data and studies of the systematic errors.

3.2 Data sample and event selection

The Monte Carlo (MC) and data sets used in this analysis are shown in Table 3.2. We use both SP5/6 (processed with `analysis-26`) and SP8 (processed with `analysis-30`) generic MC. Two different simulations of the D Dalitz plot are used. A phase-space (PHSP) only simulation (SP-4854) of $D^0 \to \pi^+\pi^-\pi^0$ decays that is used for the efficiency shape determination. And a simulation according to the previous CLEO measurement (SP-6795) [19]. The on- and off-peak data samples are both processed with `analysis-30`.

	# events (10^6)		Luminosity (fb^{-1})
Event type	SP5/6	SP8	total
B^+B^-	584.0	399.9	1785.6
$B^0\overline{B}^0$	540.7	354.9	1625.6
$c\overline{c}$	425.8	232.7	487.8
uds	677.2	331.8	482.8
Signal (PHSP)	0.347	0.542	
Signal (CLEO)	0.351	0.542	
On peak			288.48

Table 3.2: Data and MC samples used in the B and D analyses.

The event selection in this analysis is based on the branching ratio analysis of this mode [60, 61], and the reader is referred to that documentation for detailed information. In the following, we describe only the additional selection criteria used in this analysis.

We exclude events for which the invariant mass of the bachelor kaon and the oppositely-charged pion daughter of the D is in the range $1.84 < m_{K\pi} < 1.89$ GeV/c^2. This removes the background $B^- \to D^0_{K^-\pi^+}\rho^-$, for which $m_{\pi^+\pi^-\pi^0}$ happens to fall within our D mass cuts. Such events peak in m_{ES} and ΔE and constituted a significant background in the previous $B^- \to D_{\pi^+\pi^-\pi^0}K^-$ analysis, where they were classified as part of the $D\pi X$ background. The signal efficiency of this cut is 97.8% and the $B\overline{B}$ cut efficiency is 69%.

As in the branching fraction analysis, we have decided to veto events in which the $\pi^+\pi^-$ daughters of the D candidate originate from a K^0_S decay. The reason for doing this is as follows: First, the treatment of the decay $D \to K_S^0\pi^0$ as a signal mode has been done in a separate *BABAR* analysis [11] using runs 1-4 data, which was optimized for the selection of those events. Second, removing K_S^0 decays from the event sample simplifies the treatment of the background. Without the K_S^0 veto, one needs to either trust the MC to give the correct fraction of K_S^0 decays in the background, or introduce several new fit parameters for describing this fraction in each background type.

By studying the efficiency of signal and background decays, we decided to use the following cuts to reject K_S^0 decays. We exclude $\pi^+\pi^-$ candidate pairs whose invariant mass is between $0.489\,\mathrm{GeV}/c^2$ and $0.508\,\mathrm{GeV}/c^2$. The size of this mass veto window in relation to the kinematically allowed Dalitz region is pictured in Fig. 3.1. We instantly see that the effect on the Dalitz plot is very small.

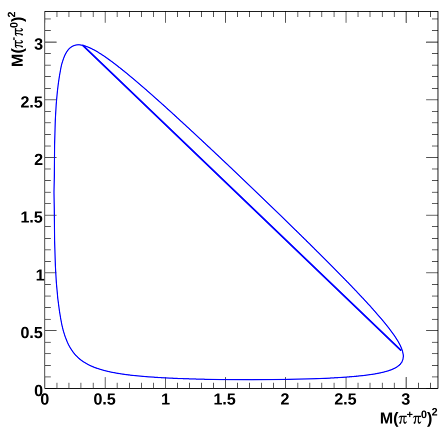


Figure 3.1: Schematic representation of the K_S^0 mass veto window applied to the Dalitz plot. Events within the diagonal line are rejected. The area within the boundary is the kinematic allowed region for $D \to \pi^+\pi^-\pi^0$ decays.

Moreover, we remove events with a D flight distance ($D_{\Delta d}$) greater than 1.5 cm. $D_{\Delta d}$ is calculated as the spatial distance between the D vertex and the K^- vertex. Table 3.3 shows the number of true K_S^0 candidates found on generic $B\overline{B}$ MC with different selections on $D_{\Delta d}$ as well as with and without the K_S^0 mass veto cut. The right two columns show the relative changes in signal efficiency. The chosen K_S^0 rejection cuts remove 94.3% of the K_S^0 events and are 97.4% efficient for $B^- \to D_{\pi^+\pi^-\pi^0}K^-$ signal events.

$D_{\Delta d}$/ cm	true K^0_S in generic $B\overline{B}$ MC				$B^- \to D_{\pi^+\pi^-\pi^0}K^-$ events			
	cut on $D_{\Delta d}$		$+K^0_S$ veto		cut on $D_{\Delta d}$		$+K^0_S$ veto	
	386	100%	153	39.6%	31518	100%	30929	98.1%
< 3.0	221	57.3%	51	13.2%	31442	99.8%	30855	97.9%
< 1.5	141	36.5%	22	5.7%	31285	99.3%	30701	97.4%
< 1.0	106	27.5%	17	4.4%	31128	98.8%	30549	96.9%
< 0.7	80	20.7%	11	2.8%	30938	98.2%	30363	96.3%
< 0.5	62	16.1%	10	2.6%	30707	97.4%	30138	95.6%
< 0.4	53	13.7%	9	2.3%	30476	96.7%	29918	94.9%
< 0.3	45	11.7%	9	2.3%	30110	95.5%	29566	93.8%

Table 3.3: Study of the K^0_S rejection cuts. Columns 2 and 3 list the number of truth-matched K^0_S events on generic $B\overline{B}$ MC with and without the K^0_S mass veto for different values of the D flight distance cut. The last two columns list the relative signal efficiency for those cuts. The highlighted line represents the final cut.

Two neural network variables are use in this analysis to provide separation between signal and background events. The first neural network variable q is computed from input variables that provide separation between continuum and $B\overline{B}$ events. The second variable d combines input variables that separate correctly reconstructed π^0 and D^0 candidates from misreconstructed ones. It provides separation between signal and all misreconstructed D-background. A detailed description of the neural networks can be found in [61]. Figure 3.2 shows the distributions of the neural network variables together with the efficiencies for signal and background when cutting on the variable. Since the neural network variable d cannot be used in the Dalitz fit due to correlations (see section 3.4) we tighten the cut from $d > 0.1$ to $d > 0.25$ compared to the branching ratio analysis.

In events that have multiple candidates, we select the candidate with the value of m_{ES} closest to 5.279 GeV/c^2, the nominal B^- mass.

The absolute signal efficiency is obtained from the CLEO signal simulation and corrected using the official correction tables and recommendations from the Particle ID (PID), Tracking and Neutrals groups for release 18. Only DK_{sig} events (see section 3.3) are considered signal. We use the PID tables provided by the PID group to weight each track by its data efficiency and calculate an event weight. For the π^0 efficiency we use the

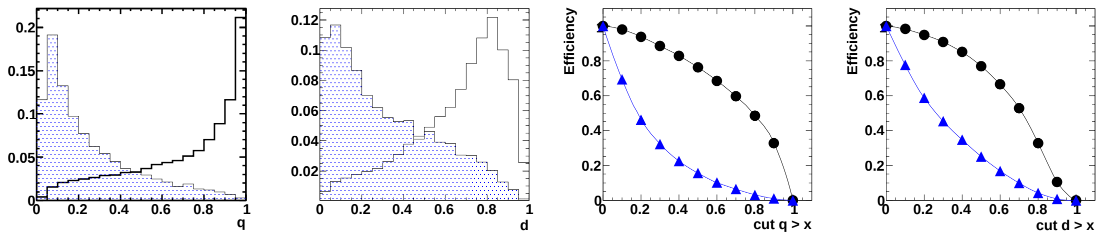


Figure 3.2: (left two plots) Distribution of the neural network variables q and d for background (hatched) and signal. (right two plots) The cut efficiencies to remove background (triangles) and signal (dots) when applying the selection $q, d > x$.

recommended correction of 0.968311. The same procedure is done with the phase space simulation as a cross-check. Table 3.4 lists the efficiencies before and after the corrections. The final efficiency used in this analysis is thus 11.41%.

Efficiency (%)	CLEO	PHSP
no correction	12.14	13.23
PID correction	11.78	12.84
π^0 correction	11.41	12.43

Table 3.4: Signal efficiency after different corrections for the CLEO and phase-space simulation.

3.3 Event types

The following ten event types (one signal and nine background types) are used throughout this analysis. We use the notation X_D for an event type with a correctly reconstructed D candidate, and $X_{\not{D}}$ for an event type with a misreconstructed D candidate:

- $DK_{\rm sig}$: $B^- \to D_{\pi^+\pi^-\pi^0}K^-$ events that were correctly reconstructed. These are the only events considered as signal. Note that this is somewhat different from the DK_D category of [60], where the requirement was only that the D be correctly reconstructed. Here we require that the entire B candidate is correctly reconstructed, since a wrong-sign kaon candidate can bias the CP measurement. In practice, the fraction of $B^- \to D_{\pi^+\pi^-\pi^0}K^-$ events with a correctly-reconstructed D and a fake K is small. The old and new definition of "signal" are therefore very similar.

- $DK_{\rm bgd}$: $B^- \to D_{\pi^+\pi^-\pi^0}K^-$ events that are misreconstructed. Note that this is somewhat different from the $DK_{\not{D}}$ category of [60], which includes only events in which the D is misreconstructed.

- $D\pi_D$: $B^- \to D^0\pi^-$, $D^0 \to \pi^+\pi^-\pi^0$ decays, where the decay $D^0 \to \pi^+\pi^-\pi^0$ is correctly reconstructed and the remaining π^- is mistaken to be the kaon.

- $D\pi_{\not{D}}$: $B^- \to D^0\pi^-$, $D^0 \to \pi^+\pi^-\pi^0$ decays, where the D candidate is misreconstructed. The kaon candidate may be either the remaining π^- or a particle from the other B meson in the event.

- DKX : $B \to D^{(*)}K^{(*)-}$, excluding $D \to \pi^+\pi^-\pi^0$ decays, with a misreconstructed D candidate.

- $D\pi X$: $B \to D^{(*)}\pi^-$ and $B \to D^{(*)}\rho^-$, excluding $D^0 \to \pi^+\pi^-\pi^0$ decays, with a misreconstructed D candidate.

- $BBC_{\not{D}}$: All other $B\overline{B}$ events with a misreconstructed D candidate.

- BBC_D : Other $B\overline{B}$ events with a correctly reconstructed $D \to \pi^+\pi^-\pi^0$ decay.
- $qq_{\not{D}}$: Continuum $e^+e^- \to q\overline{q}$ events with a misreconstructed D candidate.
- qq_D : Continuum $e^+e^- \to q\overline{q}$ events with a correctly reconstructed $D \to \pi^+\pi^-\pi^0$ decay.

3.4 Variable correlation studies

In order to decide which analysis variables can be used in a Dalitz plot fit, we conducted studies of correlations between possible variables. In the branching ratio analysis of this mode, it was found that small, non-linear correlations between the background distributions of fit variables result in a significant bias. This was solved by not using m_{ES} and m_D as fit variables.

The neural net variable d and its input parameters are highly correlated with the Dalitz plot variables. This is because most of the d input variables depend strongly on the momentum of the π^0. Therefore this variable is not used in the Dalitz plot fit.

As shown in this section, the Dalitz variable distributions have small, yet significant correlations with the m_{ES} and m_D distributions. Therefore, we are cutting on m_{ES} and m_D rather than using them in the fit. We show in this section that the variables ΔE and q are uncorrelated with s_{+0} and s_{-0} and will be used together with the Dalitz variables in the fit to measure the CP parameters.

3.4.1 1-D Correlations of analysis variables

To check for possible correlations between our analysis variables, we divide the MC samples in each of the variables ΔE, q, m_{ES}, and m_D with the binning shown in Table 3.5. For each bin of variable i, we made histograms of the other variables $j \neq i$ and performed a Kolmogorov-Smirnov (KS) test [63] between the histogram in the different bins. Tables 3.6 to 3.10 list the results. A low KS probability indicates correlations between variables i and j. Figures 3.3 to 3.6 show distributions with particularly low values.

	Bin 1	Bin 2	Bin 3	Bin 4
ΔE (MeV)	$(-70, -25)$	$(-25, 0)$	$(0, 25)$	$(25, 60)$
q	$(0.1, 0.2)$	$(0.2, 0.4)$	$(0.4, 0.7)$	$(0.7, 1.0)$
m_{ES}(GeV/c^2)	$(5.20, 5.24)$	$(5.24, 5.26)$	$(5.26, 5.28)$	$(5.28, 5.30)$
m_D(GeV/c^2)	$(1.805, 1.84)$	$(1.84, 1.86)$	$(1.86, 1.88)$	$(1.88, 1.925)$

Table 3.5: Binnings used for the KS comparisons.

	Bins 1,2	Bins 1,3	Bins 1,4	Bins 2,3	Bins 2,4	Bins 3,4
$m_{ES}(\Delta E)$	0.85009	0.73107	0.93885	0.38350	0.70753	0.95120
$m_D(\Delta E)$	0.62565	0.82801	0.58600	0.78014	0.94076	0.31087
$q(\Delta E)$	0.10397	0.69409	0.96303	0.56791	0.46216	0.99908
$s_{-0}(\Delta E)$	0.79043	0.45685	0.71429	0.81399	0.66677	0.65186
$s_{+0}(\Delta E)$	0.08032	0.05592	0.29049	0.98409	0.26768	0.81920
$m_{ES}(q)$	0.44222	0.37835	0.04888	0.87244	0.09863	0.01279
$m_D(q)$	0.80627	0.67542	0.64751	0.60491	0.84932	0.91112
$s_{-0}(q)$	0.74894	0.99280	0.95268	0.14221	0.18733	0.34757
$s_{+0}(q)$	0.77493	0.38147	0.35849	0.68479	0.33700	0.92544
$m_D(m_{ES})$	0.64959	0.97761	0.73920	0.92023	0.51343	0.89923
$s_{-0}(m_{ES})$	0.80018	0.88319	0.46723	0.96282	0.63195	0.50504
$s_{+0}(m_{ES})$	0.07161	0.01451	0.04428	0.94910	0.74056	0.64685
$s_{-0}(m_D)$	0.85561	0.65314	0.60934	0.37380	0.99953	0.27898
$s_{+0}(m_D)$	0.50661	0.30152	0.88817	0.66741	0.71759	0.53655

Table 3.6: KS test for correlations in $BBC_{\not{D}}$. In each row, $x(y)$ indicates that the KS probabilities compare the histograms of variable x in bins of variable y. Probabilities less than 0.01 are in italics. See Table 3.5 for the bin ranges.

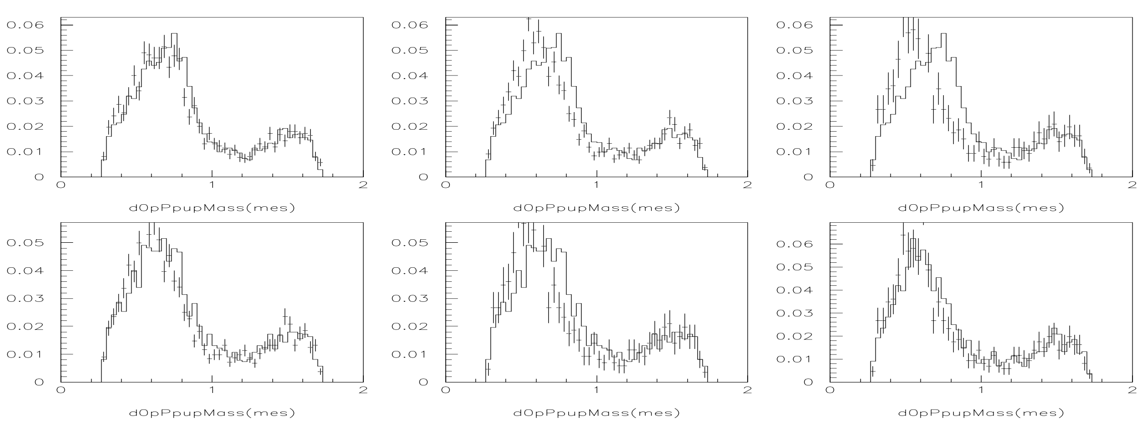


Figure 3.3: Comparison of s_{-0} in bins of m_{ES} for $D\pi X$ events corresponding to the KS probabilities of Table 3.9 (from left to right, top to bottom).

	Bins 1,2	Bins 1,3	Bins 1,4	Bins 2,3	Bins 2,4	Bins 3,4
$m_{ES}(\Delta E)$	0.28395	0.62840	0.45293	0.99846	0.72791	0.93610
$m_D(\Delta E)$	0.99161	0.85667	0.94378	0.55168	0.81582	0.46429
$q(\Delta E)$	0.66880	0.68085	0.47289	0.98864	0.88583	0.84802
$s_{-0}(\Delta E)$	0.45790	0.63849	0.92200	0.24200	0.95943	0.87477
$s_{+0}(\Delta E)$	0.49215	0.22806	0.20413	0.99820	0.80507	0.97716
$m_{ES}(q)$	0.89412	0.12506	*0.00227*	0.38719	0.03010	0.33143
$m_D(q)$	0.87311	0.06921	0.09345	0.29561	0.45084	0.78220
$s_{-0}(q)$	0.32656	0.82834	0.46593	0.58815	0.33722	0.82488
$s_{+0}(q)$	0.99663	0.62460	0.39228	0.99701	0.67679	0.80969
$m_D(m_{ES})$	0.10919	0.34694	0.99671	0.82946	0.93795	0.99120
$s_{-0}(m_{ES})$	0.69502	0.89079	0.88188	0.99813	0.93158	0.93861
$s_{+0}(m_{ES})$	0.70925	0.92879	0.85514	0.59944	0.99992	0.94279
$s_{-0}(m_D)$	0.12292	0.49611	0.29906	0.47361	0.93576	0.67997
$s_{+0}(m_D)$	0.78541	0.57557	0.11712	0.99545	0.21293	0.25807

Table 3.7: KS test for correlations in $qq_{\not D}$. In each row, $x(y)$ indicates that the KS probabilities compare the histograms of variable x in bins of variable y. Probabilities less than 0.01 are in italics. See Table 3.5 for the bin ranges.

	Bins 1,2	Bins 1,3	Bins 1,4	Bins 2,3	Bins 2,4	Bins 3,4
$m_{ES}(\Delta E)$	0.60199	0.84207	0.22632	0.91552	0.62885	0.45079
$m_D(\Delta E)$	0.96144	0.45095	0.98135	0.78327	0.99897	0.64991
$q(\Delta E)$	0.91009	0.23767	0.61756	0.10852	0.74274	0.31324
$s_{-0}(\Delta E)$	0.83659	0.92609	0.09428	0.63933	0.63363	0.39361
$s_{+0}(\Delta E)$	0.38185	0.64747	0.49310	0.79212	0.82248	0.42574
$m_{ES}(q)$	0.95701	0.70660	0.42206	0.90395	0.60795	0.68795
$m_D(q)$	0.71770	0.38342	0.45221	0.11639	0.77951	0.97165
$s_{-0}(q)$	0.06657	0.25413	0.75196	0.84852	0.23763	0.50730
$s_{+0}(q)$	0.42054	0.45182	0.94930	0.76732	0.45072	0.84810
$m_D(m_{ES})$	0.70348	0.91141	0.24713	0.45401	0.71686	0.32906
$s_{-0}(m_{ES})$	0.48828	0.64797	0.55364	0.67328	0.13752	0.29764
$s_{+0}(m_{ES})$	0.82629	0.52138	0.78886	0.90014	0.64186	0.70178
$s_{-0}(m_D)$	*0.00640*	*0.00329*	0.11214	0.98493	0.68523	0.62572
$s_{+0}(m_D)$	0.34202	0.78095	0.43126	0.69949	0.75910	0.69264

Table 3.8: KS test for correlations in qq_D. In each row, $x(y)$ indicates that the KS probabilities compare the histograms of variable x in bins of variable y. Probabilities less than 0.01 are in italics. See Table 3.5 for the bin ranges.

	Bins 1,2	Bins 1,3	Bins 1,4	Bins 2,3	Bins 2,4	Bins 3,4
$m_{ES}(\Delta E)$	0.23212	0.97048	0.30068	0.69550	0.97993	0.71033
$m_D(\Delta E)$	0.07001	0.91724	0.36380	0.12105	0.99036	0.10972
$q(\Delta E)$	0.87135	0.97401	0.07209	0.89851	0.33114	0.26918
$s_{-0}(\Delta E)$	0.26197	0.78646	0.90335	0.98009	0.38209	0.95142
$s_{+0}(\Delta E)$	0.79952	0.35878	0.40727	0.48684	0.85884	0.44531
$m_{ES}(q)$	0.67897	0.87794	0.29283	0.99440	0.39936	0.38617
$m_D(q)$	0.69629	0.99240	0.58685	0.96654	0.10499	0.20911
$s_{-0}(q)$	0.59794	0.12840	0.63841	0.65257	0.87183	0.27060
$s_{+0}(q)$	0.96332	0.31967	*0.00485*	0.23770	*0.00034*	0.03412
$m_D(m_{ES})$	0.97417	0.15612	0.96218	0.52693	0.85493	0.81117
$s_{-0}(m_{ES})$	0.51052	0.79206	0.42789	0.71484	0.84089	0.77075
$s_{+0}(m_{ES})$	0.02331	*0.00000*	*0.00000*	*0.00003*	*0.00000*	*0.00464*
$s_{-0}(m_D)$	0.08562	0.02694	0.01373	0.28299	0.28445	0.98791
$s_{+0}(m_D)$	0.88521	0.02533	*0.00000*	0.14921	*0.00004*	*0.00256*

Table 3.9: KS test for correlations in $D\pi X$. In each row, $x(y)$ indicates that the KS probabilities compare the histograms of variable x in bins of variable y. Probabilities less than 0.01 are in italics. See Figs. 3.3 and 3.4. See Table 3.5 for the bin ranges.

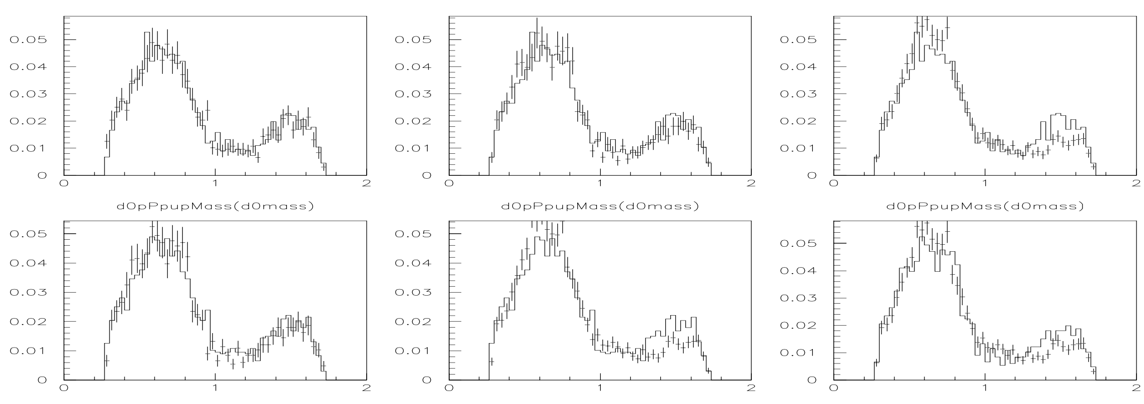


Figure 3.4: Comparison of s_{+0} in bins of m_D for $D\pi X$ events corresponding to the KS probabilities of Table 3.9 (from left to right, top to bottom).

	Bins 1,2	Bins 1,3	Bins 1,4	Bins 2,3	Bins 2,4	Bins 3,4
$m_{ES}(\Delta E)$	0.50372	0.97961	0.29342	0.57620	0.91531	0.29877
$m_D(\Delta E)$	0.08617	0.98832	0.28366	0.39926	*0.00474*	0.18294
$q(\Delta E)$	0.29336	0.82335	0.96075	0.58739	0.66002	0.95231
$s_{-0}(\Delta E)$	0.94220	0.87434	0.76969	0.66894	0.55288	0.98887
$s_{+0}(\Delta E)$	0.81799	0.85517	0.39840	0.79181	0.91457	0.89001
$m_{ES}(q)$	0.85553	0.81694	*0.00880*	0.86322	0.01057	*0.00544*
$m_D(q)$	0.51720	0.79002	0.44661	0.07877	*0.00548*	0.75020
$s_{-0}(q)$	0.20447	0.47838	0.72828	0.66474	0.02687	0.11399
$s_{+0}(q)$	0.10513	0.34842	0.29851	0.98101	0.61993	0.91044
$m_D(m_{ES})$	0.89057	0.45996	0.49679	0.81437	0.51006	0.55743
$s_{-0}(m_{ES})$	0.09971	*0.00023*	*0.00026*	0.02748	*0.00418*	0.29616
$s_{+0}(m_{ES})$	0.99260	*0.00000*	*0.00000*	*0.00005*	*0.00000*	0.24407
$s_{-0}(m_D)$	0.97862	0.22565	0.01610	0.84595	0.27140	0.86307
$s_{+0}(m_D)$	0.94996	0.43259	0.94171	0.95786	0.89141	0.26122

Table 3.10: KS test for correlations in DKX. In each row, $x(y)$ indicates that the KS probabilities compare the histograms of variable x in bins of variable y. Probabilities less than 0.01 are in italics. See also Figs. 3.5 and 3.6. See Table 3.5 for the bin ranges.

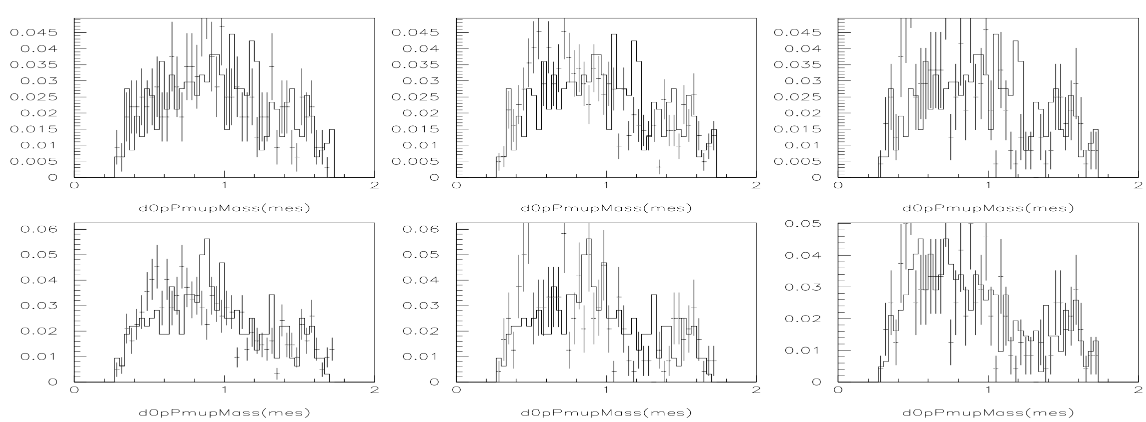


Figure 3.5: Comparison of s_{-0} in bins of m_{ES} for DKX events corresponding to the KS probabilities of Table 3.10 (from left to right, top to bottom).

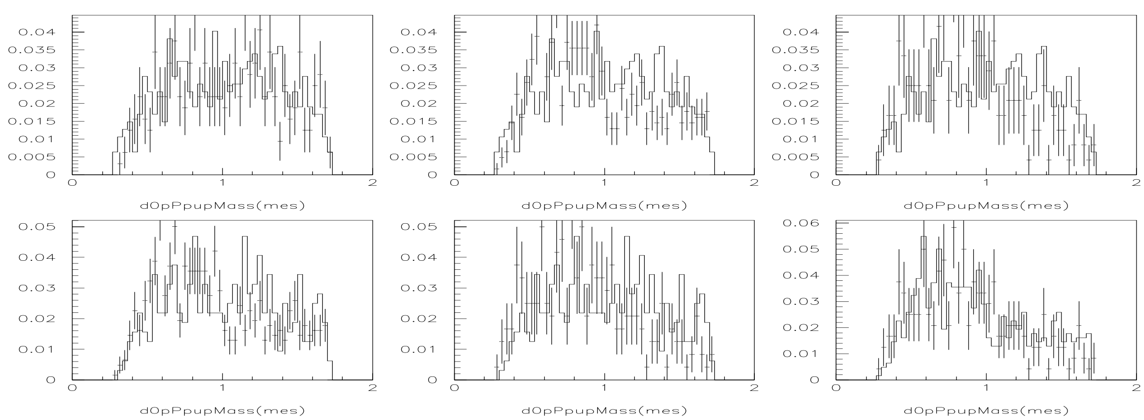


Figure 3.6: Comparison of s_{+0} in bins of m_{ES} for DKX events corresponding to the KS probabilities of Table 3.10 (from left to right, top to bottom).

3.4.2 Correlation studies with 2-D Dalitz distributions

In addition to the 1-D correlation studies of the previous section, we performed KS tests between 2-D Dalitz distributions in the bins listed in Table 3.11 for the different event types. Tables 3.12 to 3.17 list the KS probabilities of this test. Low probabilities indicate possible correlations between the 2D Dalitz distributions and the variable used in the binning.

	Bin 1	Bin 2	Bin 3
ΔE (MeV)	$(-70, -25)$	$(-25, 25)$	$(25, 60)$
q	$(0.1, 0.2)$	$(0.2, 0.7)$	$(0.7, 1.0)$
m_{ES}(GeV/c^2)	$(5.20, 5.24)$	$(5.24, 5.26)$	$(5.26, 5.30)$
m_D(GeV/c^2)	$(1.805, 1.845)$	$(1.845, 1.875)$	$(1.875, 1.925)$

Table 3.11: Binnings used for the 2D-KS comparisons of the Dalitz distributions.

	Bins 1,2	Bins 1,3	Bins 2,3
m_{ES}	0.80528	0.02165	0.07048
ΔE	0.01614	0.13611	0.22403
q	0.20087	0.05519	0.12012
m_D	0.52057	0.08130	0.10963

Table 3.12: KS probabilities for the two-dimensional comparison of the Dalitz distributions for $BBC_{\not{D}}$ events. The bin definitions are listed in Table 3.11.

	Bins 1,2	Bins 1,3	Bins 2,3
m_{ES}	0.76663	0.44127	0.81126
ΔE	0.38409	0.25026	0.51625
q	0.47830	0.04475	0.03021
m_D	0.32201	0.00964	0.01852

Table 3.13: KS probabilities for the two-dimensional comparison of the Dalitz distributions for $qq_{\not{D}}$ events. The bin definitions are listed in Table 3.11.

	Bins 1,2	Bins 1,3	Bins 2,3
m_{ES}	0.55411	0.88925	0.36111
ΔE	0.38744	0.11640	0.74168
q	0.58916	0.90730	0.56165
m_D	0.10490	0.38350	0.51012

Table 3.14: KS probabilities for the two-dimensional comparison of the Dalitz distributions for qq_D events. The bin definitions are listed in Table 3.11.

	Bins 1,2	Bins 1,3	Bins 2,3
m_{ES}	0.00095	0.00000	0.00000
ΔE	0.20996	0.18657	0.72411
q	0.78332	0.31413	0.25159
m_D	0.21288	0.00033	0.07307

Table 3.15: KS probabilities for the two-dimensional comparison of the Dalitz distributions for $D\pi X$ events. The bin definitions are listed in Table 3.11. The Dalitz distributions for m_{ES} and m_D are shown in Figs. 3.7 and 3.8.

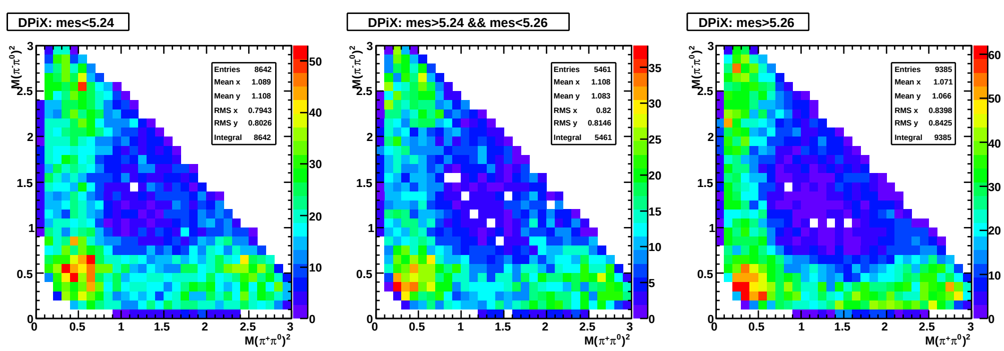


Figure 3.7: Dalitz distributions of m_{ES} for $D\pi X$ events in the three bins used for the KS comparison shown in Table 3.15.

	Bins 1,2	Bins 1,3	Bins 2,3
m_{ES}	0.24916	0.00000	0.00000
ΔE	0.53465	0.34538	0.70650
q	0.91606	0.47534	0.14354
m_D	0.74505	0.03849	0.03367

Table 3.16: KS probabilities for the two-dimensional comparison of the Dalitz distributions for DKX events. The bin definitions are listed in Table 3.11. The Dalitz distributions for m_{ES} and m_D are shown in Figs. 3.9 and 3.10.

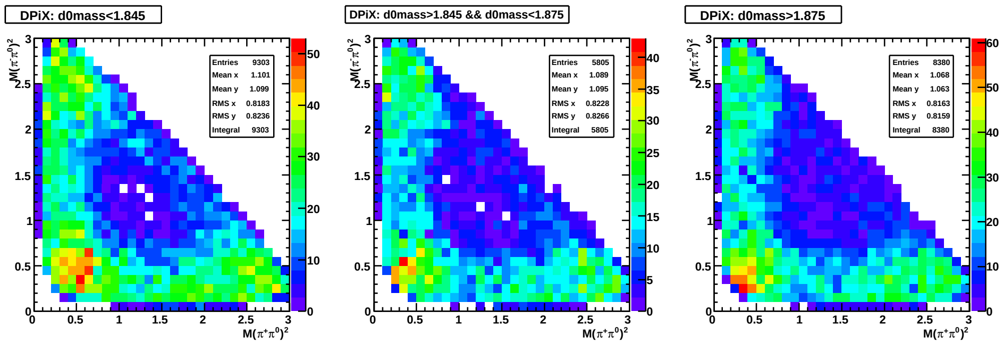


Figure 3.8: Dalitz distributions of m_D for $D\pi X$ events in the three bins used for the KS comparison shown in Table 3.15.

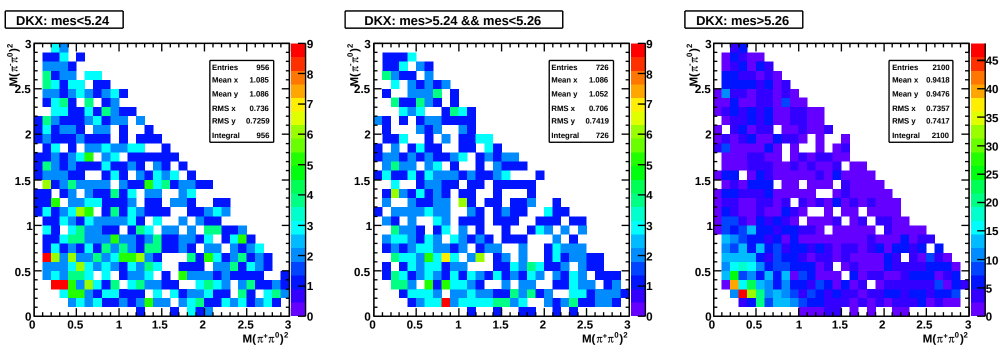


Figure 3.9: Dalitz distributions of m_{ES} for DKX events in the three bins used for the KS comparison shown in Table 3.16.

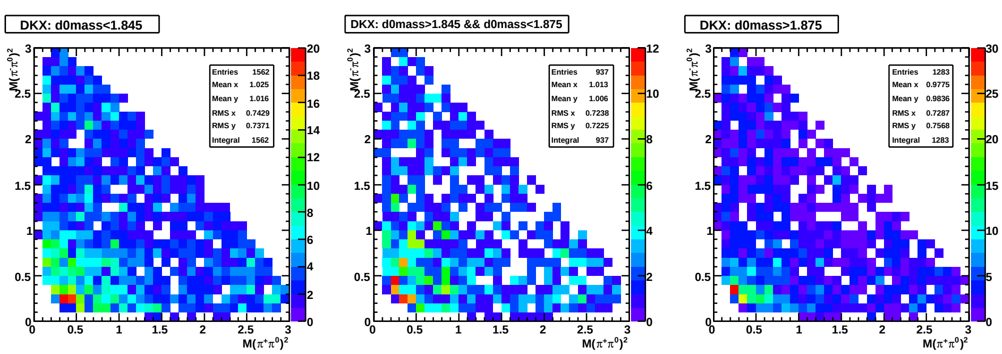


Figure 3.10: Dalitz distributions of m_D for DKX events in the three bins used for the KS comparison shown in Table 3.16.

	Bins 1,2	Bins 1,3	Bins 2,3
m_{ES}	0.00000	0.00000	0.00000
ΔE	0.84778	0.78585	0.89935
q	0.30894	0.48435	0.20788
m_D	0.00087	0.21046	0.30470

Table 3.17: KS probabilities for the two-dimensional comparison of the Dalitz distributions for $D\pi_D$ events. The bin definitions are listed in Table 3.11. The Dalitz distributions for m_D are shown in Fig. 3.11.

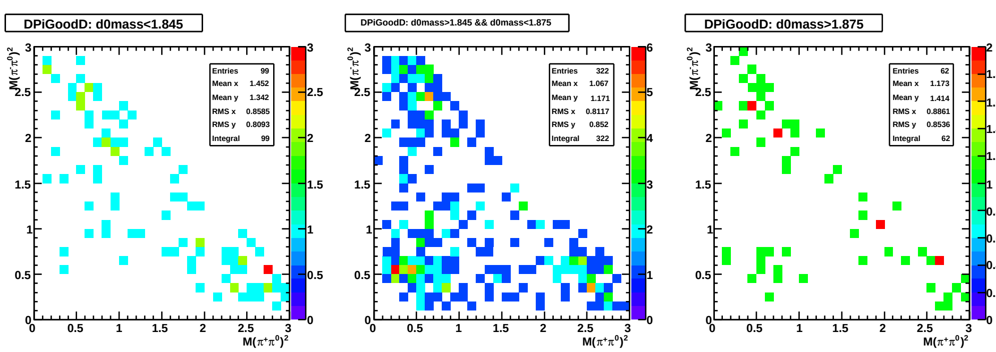


Figure 3.11: Dalitz distributions of m_D for $D\pi_D$ events in the three bins used for the KS comparison shown in Table 3.17.

3.5 Validation of event distributions

Since all of the PDF shapes in this analysis are obtained from MC samples, it is necessary to validate them against the data and check the level of agreement. The general idea is to first establish the agreement between signal and sideband region on MC. Once this is confirmed, the MC and data sidebands can be used to quantify the agreement between data and MC.

3.5.1 Comparison of MC Dalitz plot distributions in signal region and sidebands

To determine whether the background Dalitz plot distributions can be validated with sideband data, we compared these distributions in MC between the signal region and five different sidebands. The sidebands are:

- Upper ΔE: $0.06 < \Delta E < 0.140$ GeV
- Lower ΔE: $-0.140 < \Delta E < -0.07$ GeV
- m_{ES}: $5.2 < m_{ES} < 5.272\,\mathrm{GeV}/c^2$
- Upper m_D: $m_D > 1.9$ GeV/c^2
- Lower m_D: $m_D < 1.82$ GeV/c^2

The KS probabilities comparing the Dalitz plot distribution of the different backgrounds in the signal region and the sidebands are shown in Table 3.18. The MC statistics used in this study are listed in Table 3.19. We find good to very good agreement for the Dalitz plot distributions between the MC signal and sideband region. Hence, we can assume that a comparison of data and MC sidebands will also reveal any possible problems in the signal region. This test is performed in the next section.

Event type	Upper ΔE	Lower ΔE	m_{ES}	Upper m_D	Lower m_D
DKX	0.89	0.44	0.06	0.42	0.61
$D\pi X$	0.94	0.80	0.31	0.29	0.44
$BBC_{\not D}$	0.98	0.77	0.17	0.30	0.26
$qq_{\not D}$	0.77	0.90	0.43	0.55	0.68
qq_D	0.87	0.95	0.74	0.91	0.91

Table 3.18: Kolmogorov-Smirnov probabilities comparing the Dalitz plot distribution of background event types in the signal region and sidebands. The MC statistics used in this study are listed in Table 3.19.

Event type	Signal region	Upper ΔE	Lower ΔE	m_{ES}	Upper m_D	Lower m_D
DKX	171	144	202	530	403	426
$D\pi X$	628	514	821	3156	2095	2188
$BBC_{\not D}$	401	344	531	2620	1625	1700
$qq_{\not D}$	2307	2412	2849	19235	11565	11792
qq_D	34	33	37	267	153	154

Table 3.19: Numbers of MC events used to calculate Table 3.18.

3.5.2 Comparison of data and MC Dalitz plot distributions in sidebands

To check the data-MC agreement, we compare the data and MC Dalitz plot distributions in the same sidebands as in the previous section. The upper and lower m_D sideband are merged together to obtain larger statistics. The two-dimensional data Dalitz distributions are compared to properly weighted MC samples and the total $\chi^2/ndof$ between the two histograms together with the p-value is calculated. The test is performed for all values of q, $q < 0.25$ (enhances continuum events), and $q > 0.25$ (suppresses continuum events). Table 3.20 and Table 3.21 show the results for the different sidebands and a minimum average number of events in each Dalitz plot bin of 15 and 30 events, respectively.

Again, we find good agreement between the data and MC sidebands. Together with the studies of the previous section, we conclude that the background distributions obtained from MC give a good description of the distributions found in data.

Sideband	$0.1 < q < 1$	$0.1 < q < 0.25$	$q > 0.25$	$q > 0.6$
Upper ΔE	1.082 (31.31%)	0.842 (60.65%)	1.009 (42.97%)	1.139 (28.58%)
Lower ΔE	1.134 (22.89%)	1.203 (25.21%)	0.927 (58.31%)	1.191 (26.15%)
m_{ES}	1.034 (32.55%)	1.011 (43.79%)	0.982 (56.83%)	1.122 (13.24%)
m_D	0.980 (52.85%)	1.167 (22.44%)	1.302 (7.01%)	1.301 (16.13%)

Table 3.20: Data and MC comparison of the Dalitz shapes in different sidebands and for different bins of q. The values given are the total $\chi^2/ndof$ and the resulting p-value in percent. The average minimum number of events per Dalitz bin is 15 events. See Table 3.21 for the same test with 30 as the minimum number of events per bin.

Sideband	$0.1 < q < 1$	$0.1 < q < 0.25$	$q > 0.25$	$q > 0.6$
Upper ΔE	1.342 (9.10%)	0.967 (46.54%)	1.166 (28.34%)	1.348 (23.18%)
Lower ΔE	1.018 (41.21%)	1.230 (27.07%)	1.161 (28.79%)	1.007 (43.16%)
m_{ES}	1.034 (32.55%)	0.956 (62.17%)	1.003 (47.56%)	1.150 (15.14%)
m_D	1.209 (16.79%)	1.327 (16.58%)	0.813 (69.50%)	2.174 (2.08%)

Table 3.21: Data and MC comparison of the Dalitz shapes in different sidebands and for different bins of q. The values given are the total $\chi^2/ndof$ and the resulting p-value in percent. The average minimum number of events per Dalitz bin is 30 events. See Table 3.20 for the same test with 15 as the minimum number of events per bin.

3.5.3 Comparison of 1D fit variables in data and MC sidebands

Although this has been validated already in the previous branching ratio analysis, we repeat the comparison of data and MC sidebands for the remaining fit variables ΔE and q. Figure 3.12 shows the ΔE distribution on data and MC for the different sidebands. The p-values of a χ^2 test for this set of plots and for other ranges of q are listed in Table 3.22. The comparison of the q distribution can be seen in Fig. 3.13 and Table 3.23. Both from the p-values and the plots we conclude that there is good agreement between data and MC.

Sideband	$0.1 < q < 1$	$0.1 < q < 0.25$	$q > 0.25$	$q > 0.6$
m_{ES}	0.803	0.315	0.908	0.653
Lower m_D	0.692	0.299	0.849	0.386
Upper m_D	0.179	0.526	0.131	0.079

Table 3.22: P-values of a χ^2 test between data and MC for the ΔE distribution in different sidebands and using different cuts on q.

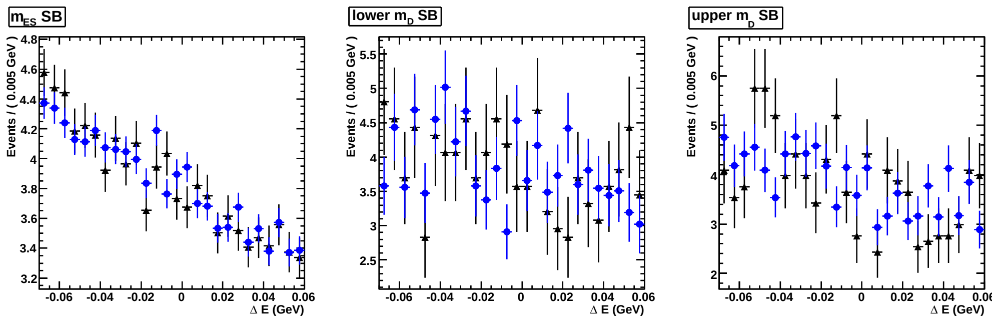


Figure 3.12: Comparison of ΔE distributions in data (black triangles) and MC (blue circles) in the m_{ES} (left), lower m_D (middle) and upper m_D (right) sideband for $0.1 < q < 1.0$. The p-values of a χ^2-test are listed in Table 3.22.

Sideband	$0.1 < q < 1$
m_{ES}	0.645
Lower m_D	0.140
Upper m_D	0.506
Upper ΔE	0.404
Lower ΔE	0.392

Table 3.23: P-values of a χ^2 test between data and MC for the q distribution in different sidebands.

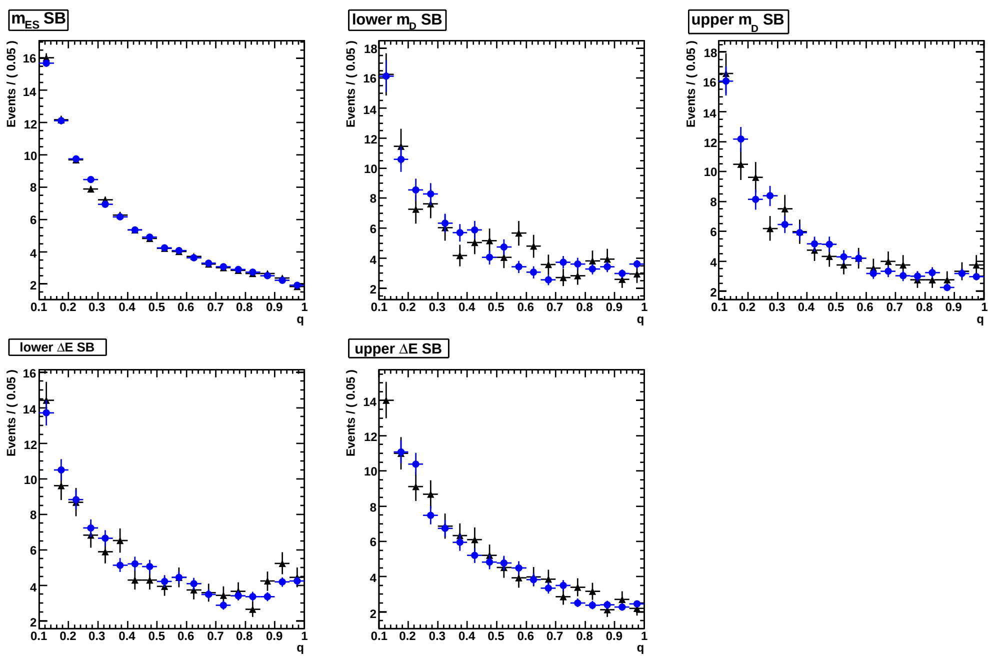


Figure 3.13: Comparison of q distributions in data (black triangles) and MC (blue circles) for different sidebands (see plot titles). The p-values of a χ^2-test between data and MC are listed in Table 3.23.

3.6 Probability density function

This section describes the probability density function (PDF) used in the fit. Separate functions for the 1-D analysis variables ΔE, d, q and the 2-D Dalitz plot, as well as for the ten different event types are presented. Eventually, the full PDF for this analysis has about 400 parameters, that are determined by fitting suitable MC samples.

3.6.1 Overview

The total PDF is a sum over the PDFs of all event types t:

$$\begin{aligned}\mathcal{P}^i = \sum_t f_t\,\mathcal{P}^i_t &= f_{DK_{\rm sig}}\,\mathcal{P}^i_{DK_{\rm sig}} + f_{DK_{\rm bgd}}\,\mathcal{P}^i_{DK_{\rm bgd}} \\ &+ f_{D\pi_D}\,\mathcal{P}^i_{D\pi_D} + f_{D\pi_{\not D}}\,\mathcal{P}^i_{D\pi_{\not D}} \\ &+ f_{DKX}\,\mathcal{P}^i_{DKX} + f_{D\pi X}\,\mathcal{P}^i_{D\pi X} \\ &+ f_{BBC_{\not D}}\,\mathcal{P}^i_{BBC_{\not D}} + f_{BBC_D}\,\mathcal{P}^i_{BBC_D} \\ &+ f_{qq_D}\,\mathcal{P}^i_{qq_D} + f_{qq_{\not D}}\,\mathcal{P}^i_{qq_{\not D}}\,, \end{aligned} \tag{3.6}$$

where the subscripts correspond to the event types of section 3.3, f_t is the expected fraction of events of type t and $\mathcal{P}^i_t$ is the PDF for these events. The superscript $i = 1, 2$ indicates the two types of fits that are used in the two-step fitting procedure of this analysis (see section 3.8.5). In the first fit, we use the variables

1. ΔE
2. q (neural net separating $B\overline{B}$ from continuum)
3. d (neural net separating good D candidates from fake ones)

Under the assumption of no correlations between the distributions of the various event types in the fit variables (this assumption is justified by the studies in the branching ratio analysis [61]), the PDF for events of type t used in the first fit is

$$\mathcal{P}_t^1(\Delta E, q, d) = \mathcal{E}_t(\Delta E)\,\mathcal{Q}_t(q)\,\mathcal{C}_t(d). \tag{3.7}$$

The variables used in the second fit are the uncorrelated (see sec. 3.4) set

1. ΔE
2. q
3. s_{+0} and s_{-0}

hence the single event type PDF can be expressed as the product

$$\mathcal{P}_t^2(\Delta E, q, s_{+0}, s_{-0}) = \mathcal{E}_t(\Delta E)\,\mathcal{Q}_t(q)\,\mathcal{D}_t(s_{+0}, s_{-0}). \tag{3.8}$$

The following subsections discuss the parameterizations of these PDFs.

3.6.2 Parametrization of $\mathcal{E}_t(\Delta E)$

The ΔE PDFs are parameterized using Gaussians, asymmetric Gaussians, and 2nd-order polynomials. Table 3.24 lists the functional form used for each event type, and the fits to MC samples from which these functions and their shape parameters were obtained are shown in Figs. 3.14 through 3.18.

Event type t	$\mathcal{E}_t(\Delta E)$	$\mathcal{Q}_t(q'), \mathcal{C}_t(d')$
DK_{sig}	$G + P$	$G + AG$
DK_{bgd}	$G + P$	$G + AG$
$D\pi_D$	G	$G + AG$
$D\pi_{\not{D}}$	P	G
$D\pi X$	P	$G + AG$
DKX	P	$G + AG$
BBC_D	G	G
$BBC_{\not{D}}$	P	$G + AG$
qq_D	P	G
$qq_{\not{D}}$	P	$G + AG$

Table 3.24: Functional forms of the ΔE and q PDFs of each event type, indicated with G = Gaussian, AG = asymmetric Gaussian and P = 2nd order polynomial.

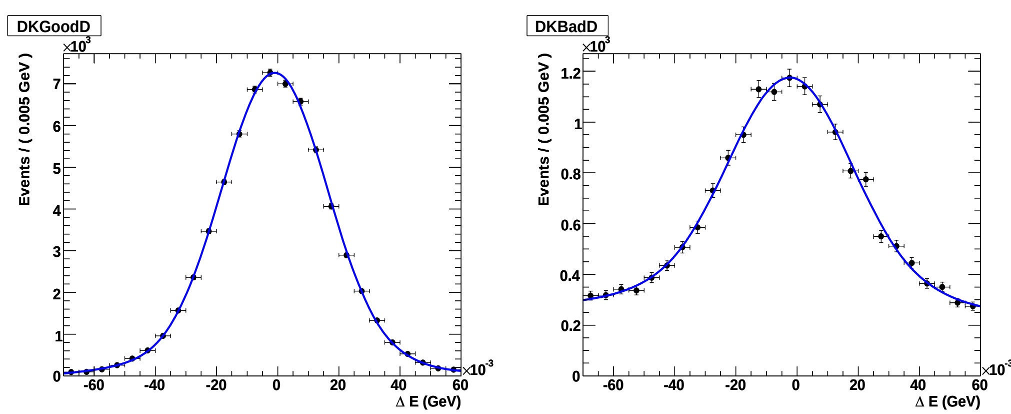


Figure 3.14: The distributions of ΔE for DK_{sig} (left) and DK_{bgd} (right) obtained from MC.

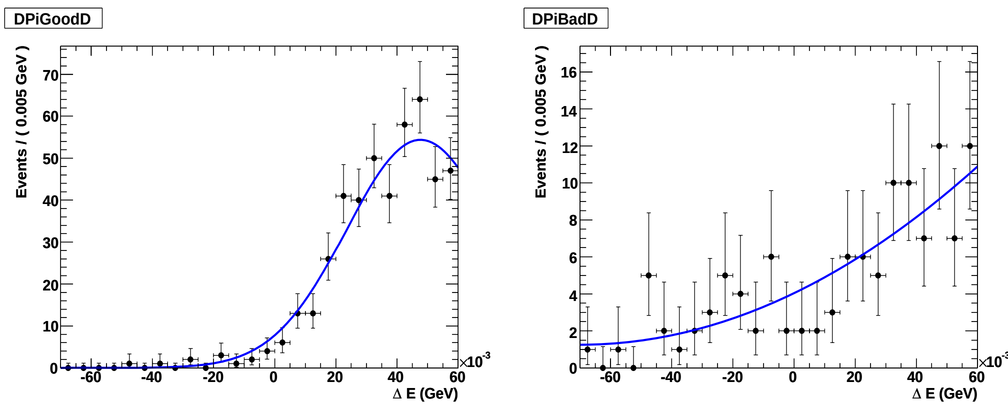


Figure 3.15: The distributions of ΔE for $D\pi_D$ (left) and $D\pi_{\not{D}}$ (right) obtained from MC.

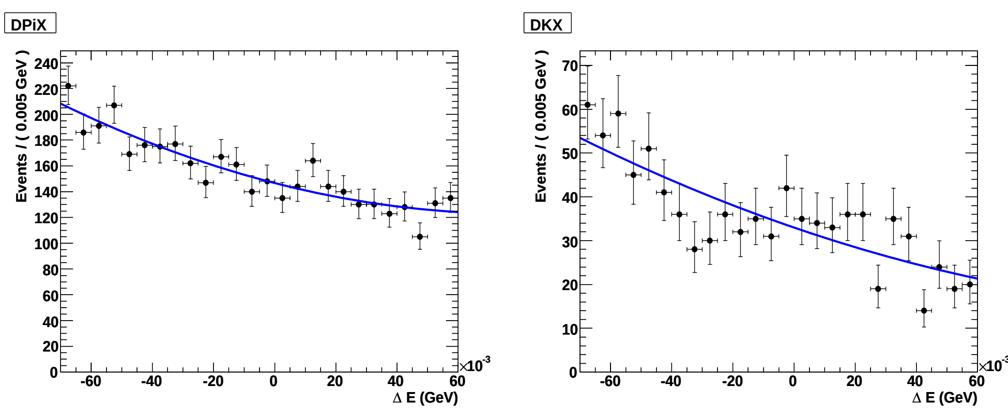


Figure 3.16: The distributions of ΔE for $D\pi X$ (left) and DKX (right) obtained from MC.

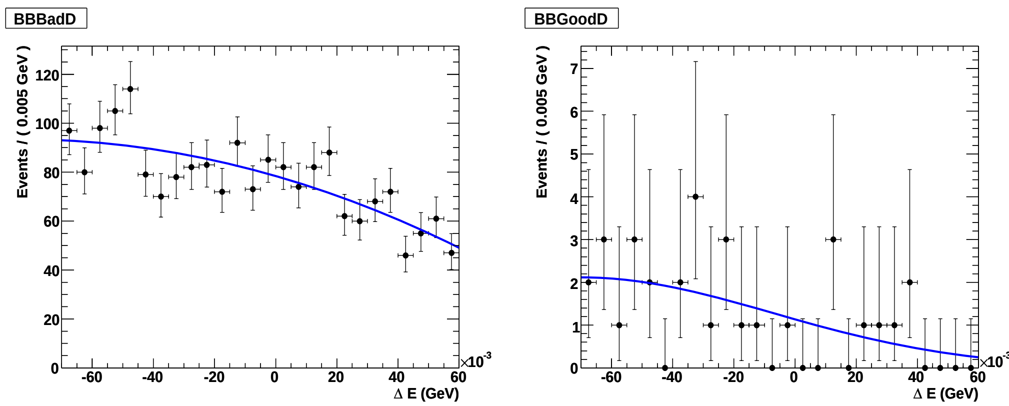


Figure 3.17: The distributions of ΔE for $BBC_{\not D}$ (left) and BBC_D (right) obtained from MC.

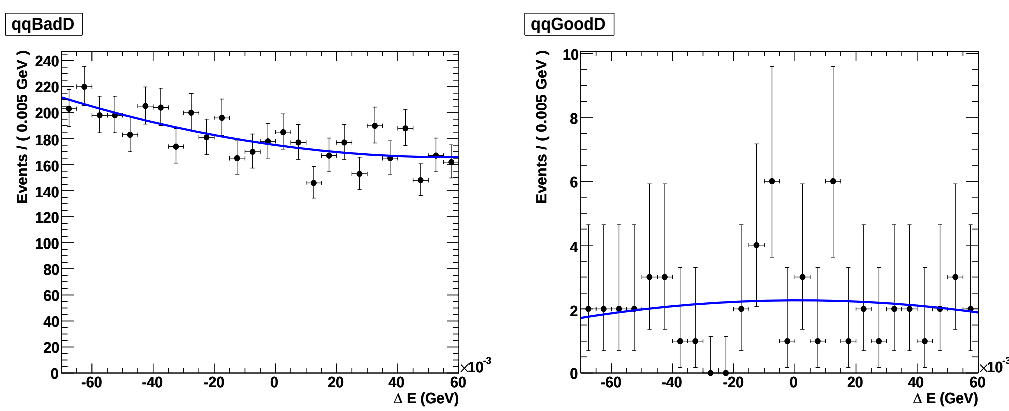


Figure 3.18: The distributions of ΔE for $qq_{\not D}$ (left) and qq_D (right) obtained from MC.

3.6.3 Parametrization of $\mathcal{Q}_t(q)$

Figure 3.2 (p. 54) shows the distribution of the neural network variable q for signal and background. Finding an analytical parameterization for this shape is rather difficult and most analyses resort to using histogram based PDFs to describe them. We use a different approach and define the transformation

$$q \mapsto q' = \operatorname{arctanh}\left(\frac{q - 0.55}{0.45}\right) , \tag{3.9}$$

which maps the selection interval $q \in (0.1, 1.0)$ to $q' \in (-\infty, \infty)$. It turns out that q' has a Gaussian-like shape and can easily be parametrized analytically. We use the sum of an asymmetric Gaussian and a Gaussian to fit the event types DK_{sig}, DK_{bgd}, $D\pi X$, DKX, $D\pi_D$, $BBC_{\not{D}}$ and $qq_{\not{D}}$. A single Gaussian is used for the event types BBC_D, $D\pi_{\not{D}}$ and qq_D. Table 3.24 summarizes the different shapes used in each fit and Figs. 3.19 through 3.23 show the fitted MC distributions.

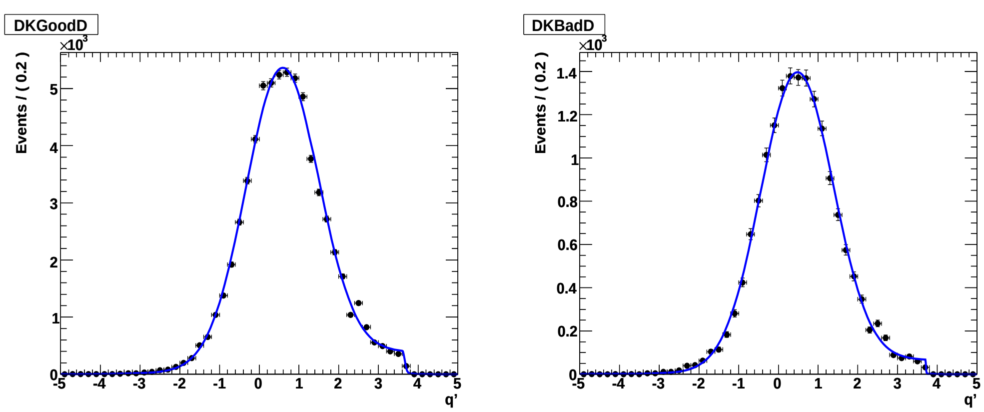


Figure 3.19: The distributions of q' for DK_{sig} (left) and DK_{bgd} (right) obtained from MC.

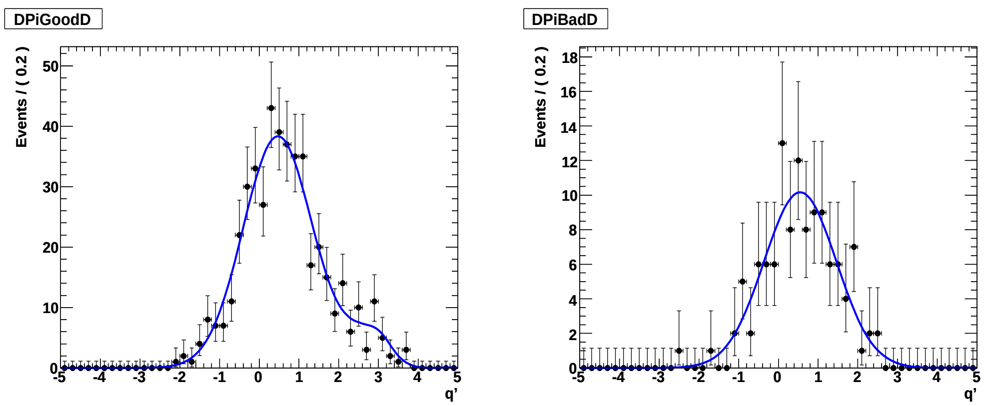


Figure 3.20: The distributions of q' for $D\pi_D$ (left) and $D\pi_{\not D}$ (right) obtained from MC.

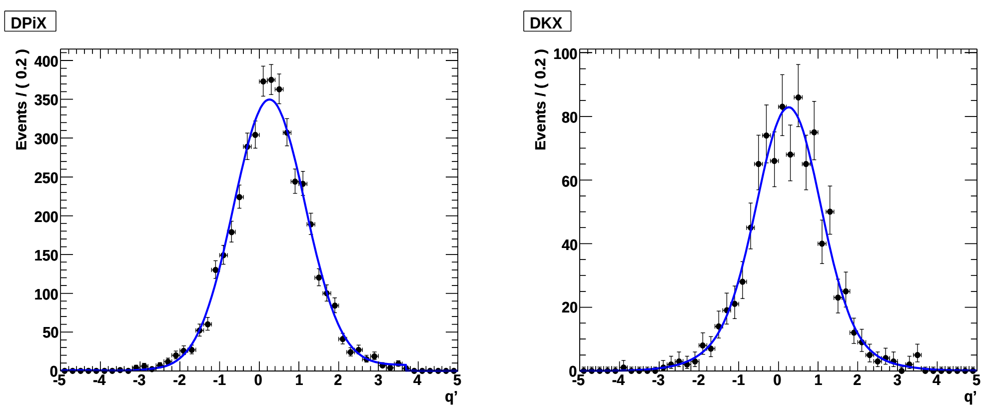


Figure 3.21: The distributions of q' for $D\pi X$ (left) and DKX (right) obtained from MC.

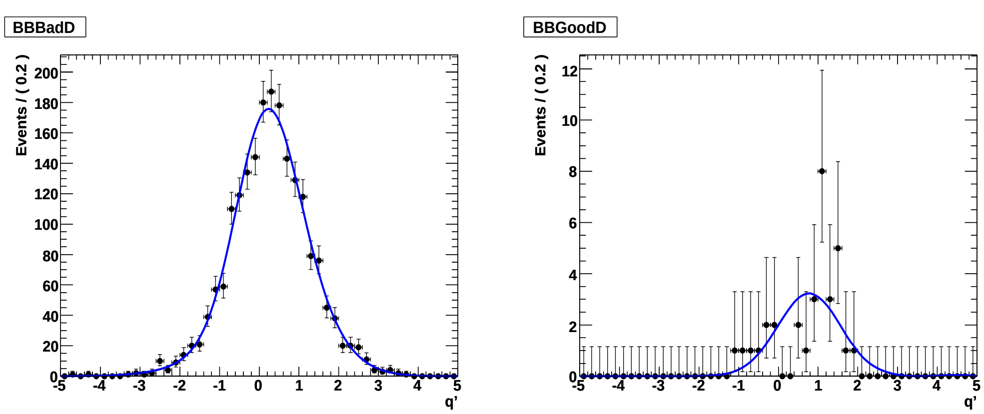


Figure 3.22: The distributions of q' for $BBC_{\not D}$ (left) and BBC_D (right) obtained from MC.

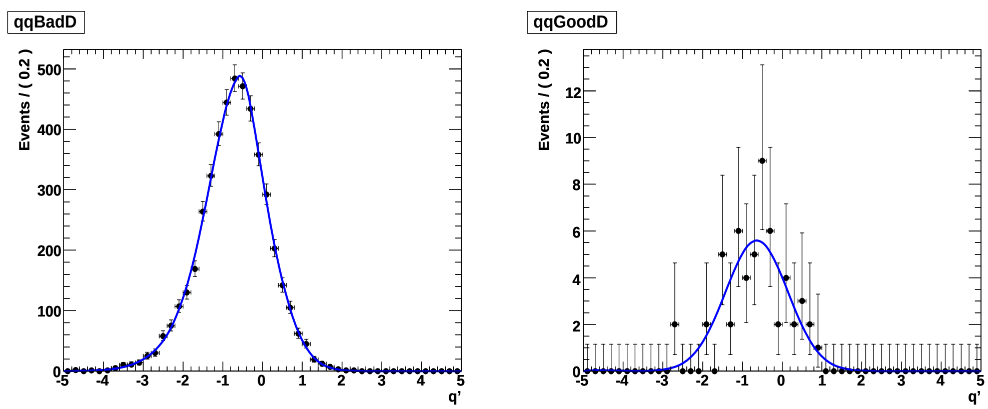


Figure 3.23: The distributions of q' for $qq_{\not{D}}$ (left) and qq_D (right) obtained from MC.

3.6.4 Parametrization of $\mathcal{C}_t(d)$

The same transformation as described in the previous section is used for d but with changed numerical values due to the different selection interval:

$$d \mapsto d' = \operatorname{arctanh}\left(\frac{d - 0.625}{0.375}\right), \tag{3.10}$$

which maps the selection interval $d \in (0.25, 1.0)$ into $d'(-\infty, \infty)$. Figs. 3.24 through 3.28 show the fitted MC sample to the shapes listed in Table 3.24 (p. 74).

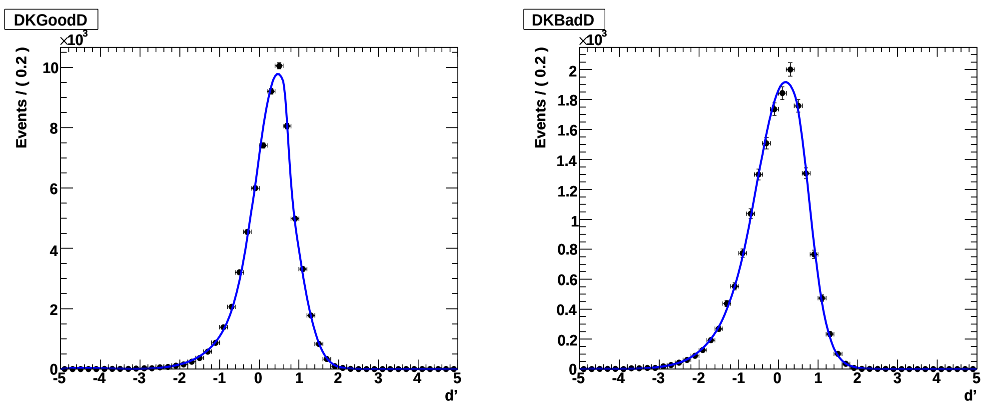


Figure 3.24: The distributions of d' for DK_{sig} (left) and DK_{bgd} (right) obtained from MC.

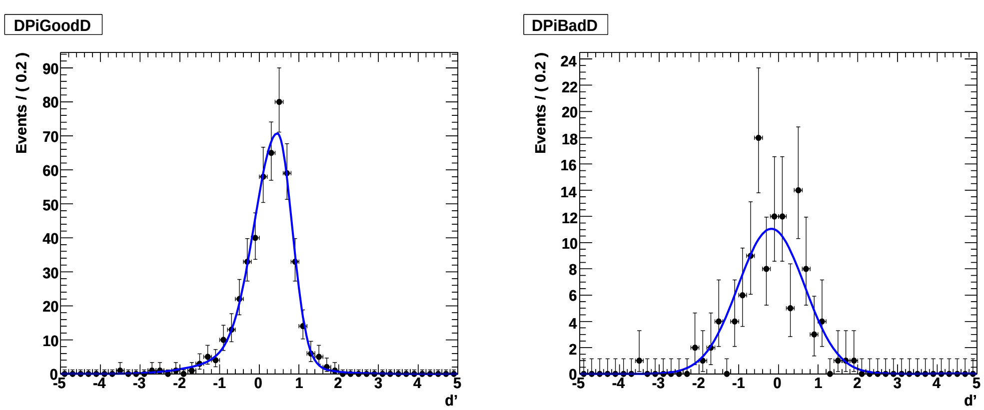


Figure 3.25: The distributions of d' for $D\pi_D$ (left) and $D\pi_{\not{D}}$ (right) obtained from MC.

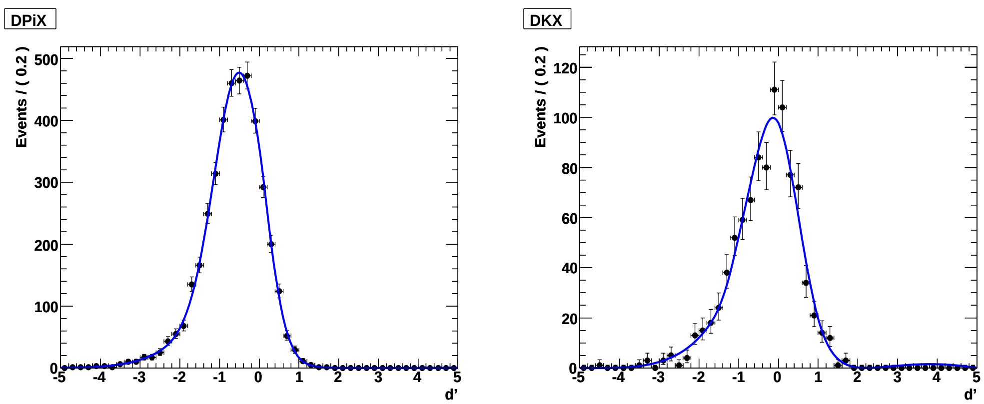


Figure 3.26: The distributions of d' for $D\pi X$ (left) and DKX (right) obtained from MC.

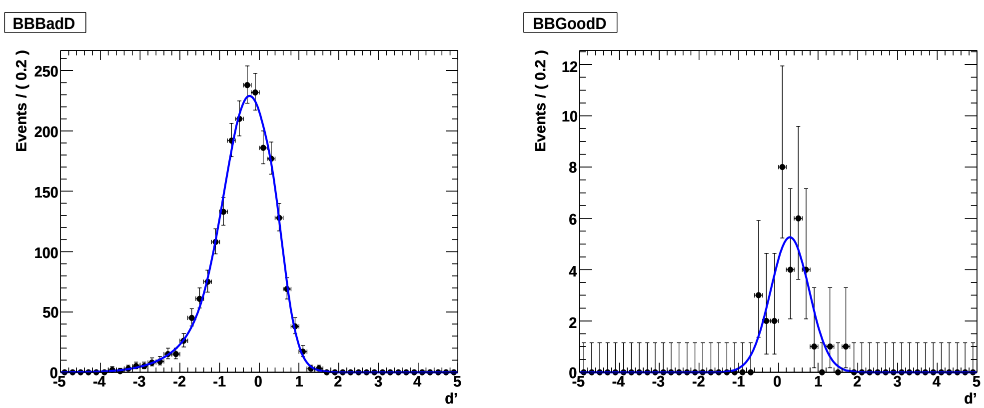


Figure 3.27: The distributions of d' for $BBC_{\not D}$ (left) and BBC_D (right) obtained from MC.

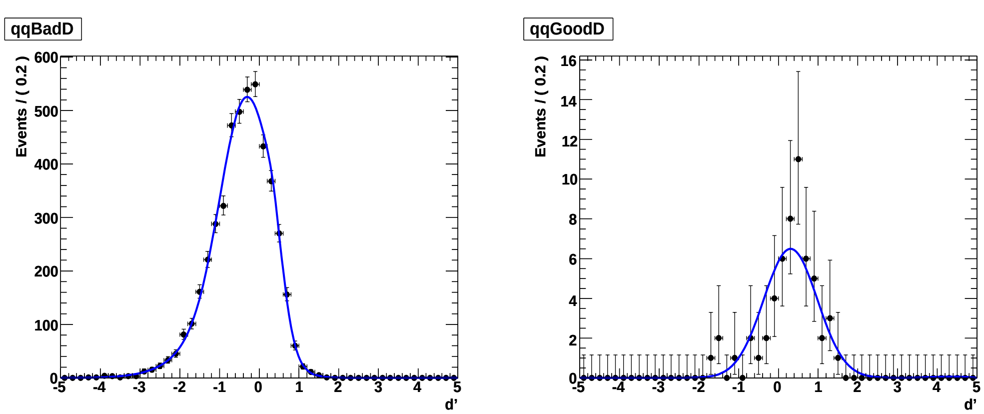


Figure 3.28: The distributions of d' for $qq_{\not D}$ (left) and qq_D (right) obtained from MC.

3.6.5 Dalitz Plot Efficiency PDF

The efficiency as a function of the Dalitz plot variables s_{+0} and s_{-0} is determined from phase space signal MC including the proper particle ID corrections for the charged pions and kaon in the event. We fit the efficiency with a cubic polynomial:

$$\begin{aligned}
\mathrm{Eff}(\mathrm{s}_{+0}, \mathrm{s}_{-0}) = 1 &+ s_1(s_{+0} + s_{-0}) \\
&+ s_2(s_{+0}^2 + s_{-0}^2) + s_3(s_{+0}^3 + s_{-0}^3) \\
&+ s_4(s_{-0}s_{+0}^2 + s_{+0}s_{-0}^2) + s_5\, s_{+0}s_{-0} \\
&+ a_1(s_{+0} - s_{-0}) + a_2(s_{+0}^2 - s_{-0}^2) \\
&+ a_3(s_{+0}^3 - s_{-0}^3) + a_4(s_{-0}s_{+0}^2 - s_{+0}s_{-0}^2) \qquad (3.11)
\end{aligned}$$

The parametrization is split into symmetric (s_i) and asymmetric (a_i) coefficients that are mostly uncorrelated, helping the fit converge. In addition to the relation (3.11), $\mathrm{Eff}(\mathrm{s}_{+0}, \mathrm{s}_{-0}) \equiv 0$ for all points outside the physical Dalitz boundary or within the K_S^0 mass veto window described in section 3.2. This veto window shows up as a diagonal void line at the high mass region of the Dalitz plot.

The function $\mathrm{Eff}(\mathrm{s}_{+0}, \mathrm{s}_{-0})$ gives only the reconstruction efficiency. Truth-matched DK_{sig} signal events satisfy an additional requirement, namely, that they were correctly reconstructed, and hence have different efficiency function parameters. The efficiency function for these events is $\mathrm{Eff_D}(\mathrm{s}_{+0}, \mathrm{s}_{-0})$. Its parameters are obtained from phase-space signal MC events that passed all the cuts and were truth-matched.

The results of the efficiency fits are listed in Table 3.25. Fig. 3.29 shows the efficiency (upper left), data generated from the fitted efficiency $\mathrm{Eff_D}$ (upper right) and the projections onto the Dalitz variables (lower plots). We notice that asymmetric coefficients for both efficiencies are consistent with zero. To simplify the fitting code, we therefore neglect the asymmetry in the efficiency functions.

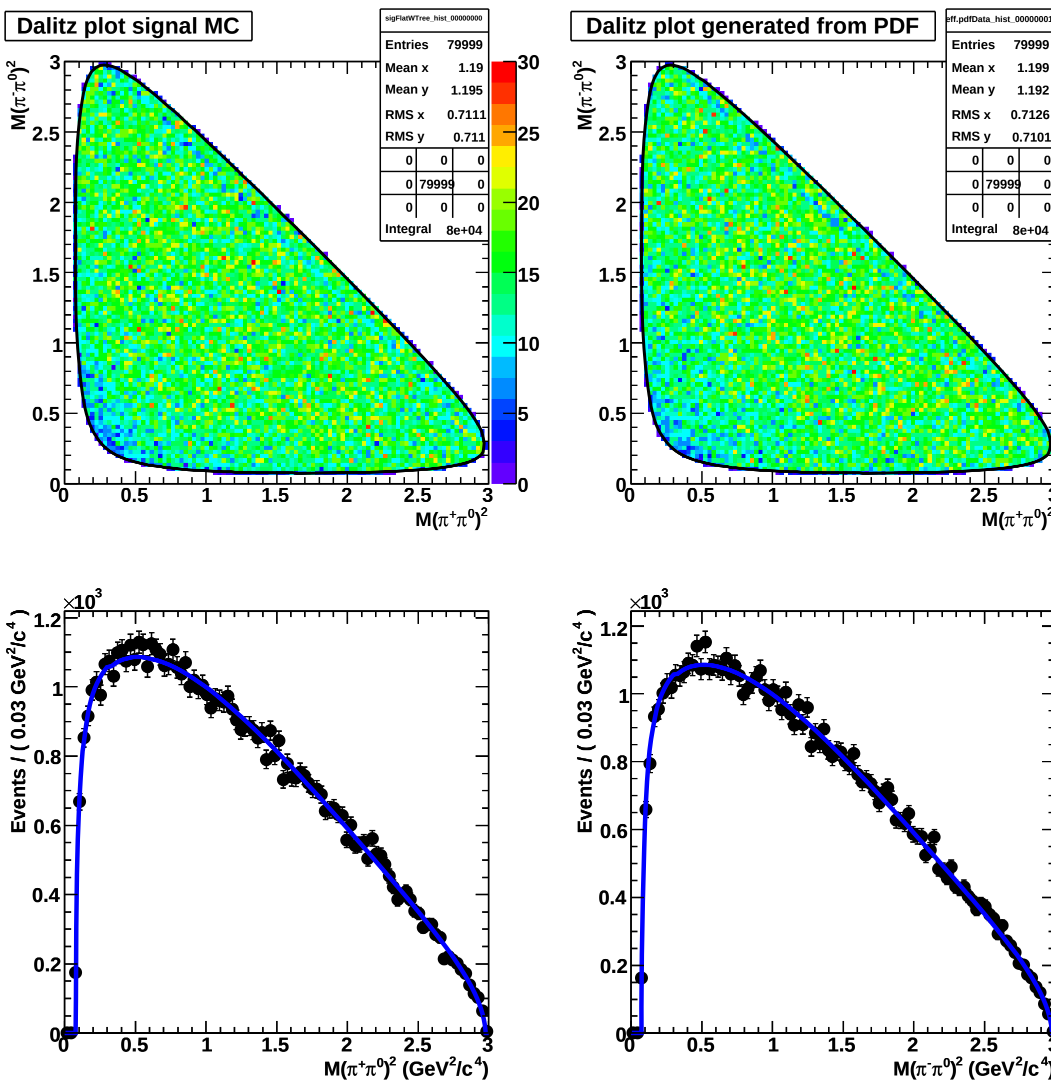


Figure 3.29: Fit of $\mathrm{Eff_D}$ on phase-space signal MC. Upper left: Efficiency map for correctly reconstructed MC $B^- \to D_{\pi^+\pi^-\pi^0} K^-$events. Upper right: The Dalitz plot distribution of events generated using the PDF used to fit the MC efficiency. Bottom: Projections onto the Dalitz plot axes of phase-space MC (data points) and the efficiency fit function (line).

	$\mathrm{Eff_D}$	Eff
s_1	276 ± 8.10	0.56 ± 0.035
s_2	-113 ± 3.29	-0.27 ± 0.017
s_3	15.7 ± 0.84	0.042 ± 0.0044
s_4	29.8 ± 2.19	0.043 ± 0.012
s_5	-166 ± 8.24	-0.27 ± 0.044
a_1	1.65 ± 12.2	-0.013 ± 0.069
a_2	-3.13 ± 10.3	-0.0043 ± 0.058
a_3	0.72 ± 2.19	0.0021 ± 0.012
a_4	1.45 ± 2.33	0.0065 ± 0.013

Table 3.25: Fit results for the efficiency coefficients of Eq. (3.11) for the signal efficiency $\mathrm{Eff_D}$ and the reconstruction efficiency Eff. Fig. 3.29 shows the fit for $\mathrm{Eff_D}$.

3.6.6 Dalitz plot signal PDF

The PDF for signal events is a product of the "physical" PDF and the efficiency function:

$$\mathcal{D}_{DK_{\mathrm{sig}}}(s_{+0}, s_{-0}) = \mathcal{D}^{\mathrm{phys}}_{DK_{\mathrm{sig}}}(s_{+0}, s_{-0})\,\mathrm{Eff_D}(\mathrm{s_{+0}, s_{-0}}). \tag{3.12}$$

For B^- decays, the physical PDF is

$$\mathcal{D}^{\mathrm{phys}}_{DK_{\mathrm{sig}}}(s_{+0}, s_{-0}) = \left|f_{D^0}(s_{+0}, s_{-0}) + (x_- + iy_-)f_{D^0}(s_{-0}, s_{+0}),\right|^2. \tag{3.13}$$

where f_{D^0} is the complex Dalitz plot amplitude for D^0 decays, and we have written the complex coefficient $r_B e^{i(\delta-\gamma)}$ using the two real variables x_- and y_- with $z_- = x_- + iy_-$. The physical PDF for B^+ decays is obtained by the exchange $s_{+0} \leftrightarrow s_{-0}$ and $z_- \leftrightarrow z_+$. In Section 3.8 we will show that the parameterization $z(x, y)$ is not suitable for this analysis and we will use a polar parameterization $z(\rho, \theta)$ instead.

The D decay Dalitz distributions are calculated with the isobar model,

$$f_{D^0}(s_{+0}, s_{-0}) = a_{\mathrm{NR}} e^{i\phi_{\mathrm{NR}}} + \sum_r a_r e^{i\phi_r} A_r(s_{+0}, s_{-0}), \tag{3.14}$$

where the first term is a flat, non-resonant contribution and the sum is over all 2-body resonances. The 2-body resonance amplitude describes the decay of the D^0 to particle C and resonance r, followed by the decay of the resonance r to particles A and B ($D^0 \to (r \to AB)C$). Following the notation of [64] we write:

$$A_r = F_D F_r \frac{F_s}{m_r^2 - q^2 - i m_r \Gamma(q)}, \tag{3.15}$$

where we take the D form factor to be $F_D = 1$, the expressions for the resonance form factor F_r are listed in Table 3.26, $q^2 = m_{AB}^2 = (p_A + p_B)^2$ is the reconstructed mass squared of the resonance candidate, m_r is the nominal mass of the resonance, $\Gamma(q)$ a mass-dependent width and F_s the spin-factor for a resonance of spin s:

$$\begin{aligned}
F_0 &= 1 && (3.16)\\
F_1 &= m_{AC}^2 - m_{BC}^2 + \frac{(m_{D^0}^2 - m_C^2)(m_B^2 - m_A^2)}{q^2} && (3.17)\\
F_2 &= \left(m_{BC}^2 - m_{AC}^2 + \frac{(m_{D^0}^2 - m_C^2)(m_A^2 - m_B^2)}{q^2}\right)^2 \\
&- \frac{1}{3}\left(m_{AB}^2 - 2m_{D^0}^2 - 2m_C^2 + \frac{(m_{D^0}^2 - m_C^2)^2}{q^2}\right) \\
&\times \left(m_{AB}^2 - 2m_A^2 - 2m_B^2 + \frac{(m_A^2 - m_B^2)^2}{q^2}\right). && (3.18)
\end{aligned}$$

Note that we use the reconstructed mass q in the denominator rather than the resonance mass m_r as was done in [64]. This seems a more reasonable approach, especially for broad resonances like the ρ. The mass-dependent width is expressed as

$$\Gamma(q) = \Gamma_r \left(\frac{p_{AB}}{p_r}\right)^{2s+1} \left(\frac{m_r}{q}\right) F_r^2 \tag{3.19}$$

where Γ_r is the width of the resonance and p_r is the momentum of either daughter in the resonance rest frame. p_{AB} is the same but with the two-track invariant mass assigned to the parent instead of the nominal resonance mass.

The parameters of the resonances used in the signal PDF are listed in Table 3.27. We

Spin	Form factor F_r
0	1
1	$\sqrt{\frac{1+R^2p_r^2}{1+R^2p_{AB}^2}}$
2	$\sqrt{\frac{9+3R^2p_r^2+R^4p_r^4}{9+3R^2p_{AB}^2+R^4p_{AB}^4}}$

Table 3.26: Blatt-Weisskopf penetration form factors. R is the meson radial parameter, whose value we take to be $1.5\,\text{GeV}^{-1}$, and p_r is the momentum of either daughter in the resonance rest frame. p_{AB} is the two-particle invariant mass.

take the mass and width of the $f_0(1370)$ from the E791 Dalitz plot analysis of $D_s \to \pi^+\pi^-\pi^+$ [65]. The width of the $f_0(980)$ is also taken from that analysis. These three parameters have large uncertainties in the PDG, whose input comes mostly from scattering experiments. Therefore, it seems more relevant to our decay to take them from their relatively precise measurement in the D_s decay. The ω resonance is included without any $\rho - \omega$ mixing.

Furthermore, the PDF includes three non-resonant P-wave amplitudes (one for each pair of pions) with $A_r = F_1$. Figures 3.30 through 3.36 show the Dalitz distributions of individual components used in our signal PDF.

State	Mass (MeV)	Width (MeV)	Source
$\rho \to \pi\pi$	775.8 ± 0.5	150.3 ± 1.6	PDG
$\rho(1450) \to \pi\pi$	1465 ± 25	400 ± 60	PDG
$\rho(1700) \to \pi\pi$	1720 ± 20	250 ± 100	PDG
$f_0(980) \to \pi^+\pi^-$	980 ± 10	44 ± 3	PDG (mass), [65] (width)
$f_0(1370) \to \pi^+\pi^-$	1434 ± 18	173 ± 32	[65]
$f_0(1500) \to \pi^+\pi^-$	1507 ± 5	109 ± 7	PDG
$f_0(1710) \to \pi^+\pi^-$	1714 ± 5	140 ± 10	PDG
$f_2(1270) \to \pi^+\pi^-$	1275.4 ± 1.2	$185.1^{+3.5}_{-2.6}$	PDG
$\sigma \to \pi^+\pi^-$	500	400	
Phase space			
The following are used only for systematics			
$f_2'(1525) \to \pi^+\pi^-$	1525 ± 5	73 ± 6	PDG
$\omega \to \pi^+\pi^-$	782.59	8.49	PDG
Phase space P-wave			

Table 3.27: Components of the signal Dalitz PDF, and the sources of their parameters.

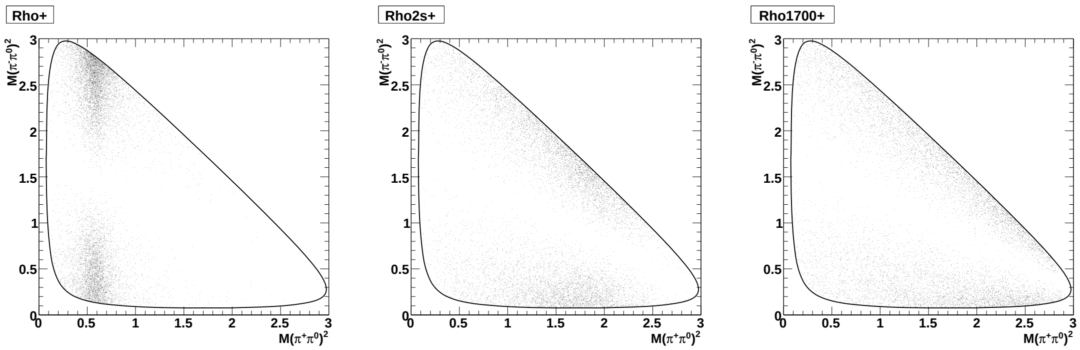


Figure 3.30: Dalitz plot distributions of $\rho^+\pi-$, $\rho(1450)^+\pi^-$, and $\rho(1700)^+\pi^-$ toy MC events.

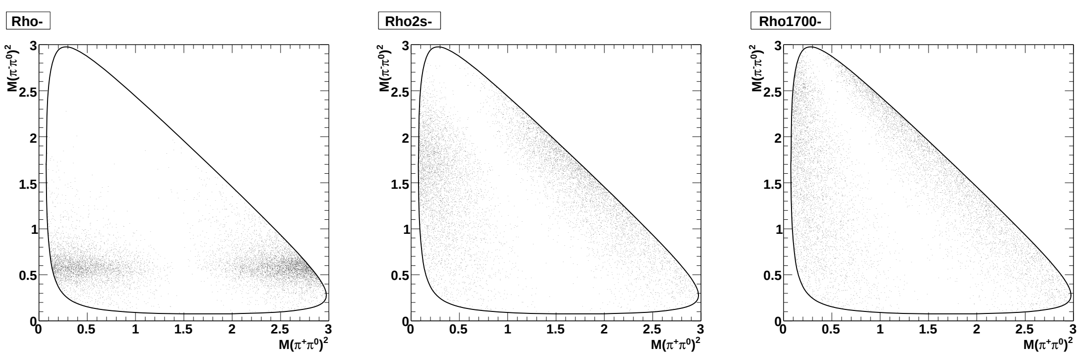


Figure 3.31: Dalitz plot distributions of $\rho^-\pi^+$, $\rho(1450)^-\pi^+$, and $\rho(1700)^-\pi^+$ MC events.

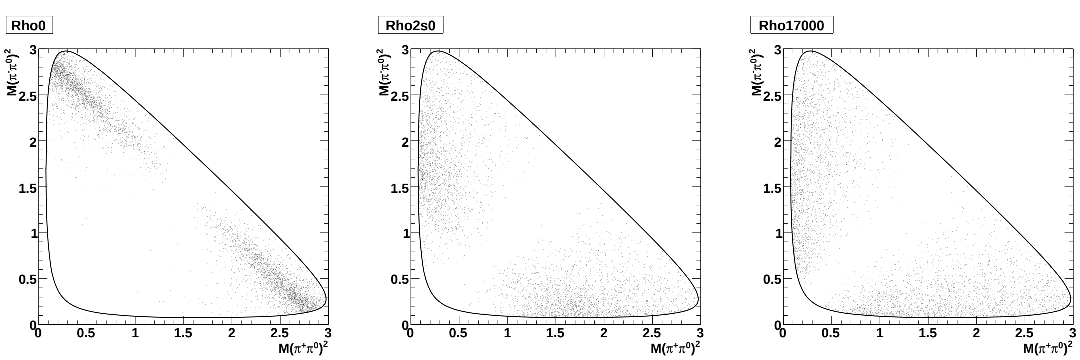


Figure 3.32: Dalitz plot distributions of $\rho^0\pi^0$, $\rho(1450)^0\pi^0$, and $\rho(1700)^0\pi^0$ toy MC events.

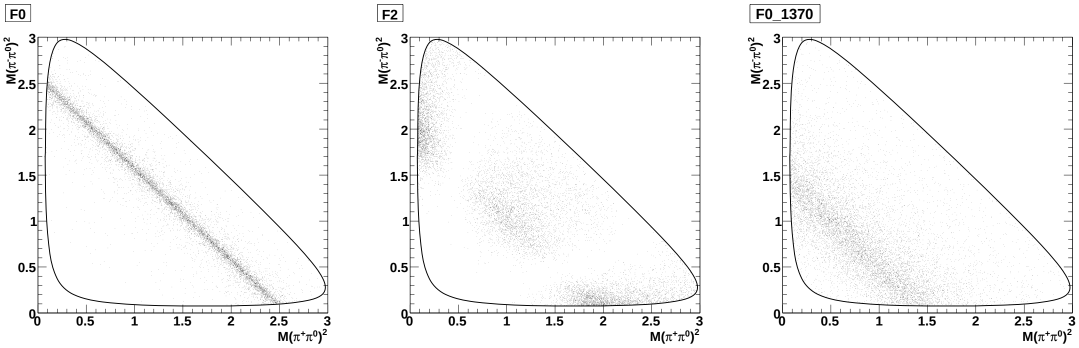


Figure 3.33: Dalitz plot distributions of $f_0\pi^0$, $f_2\pi^0$, and $f_0(1370)\pi^0$ toy MC events.

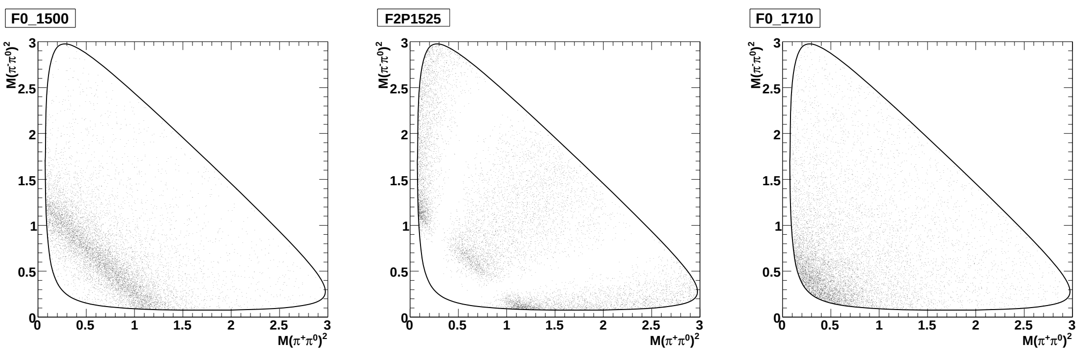


Figure 3.34: Dalitz plot distributions of $f_0(1500)\pi^0$, $f_2'(1525)\pi^0$, and $f_0(1710)\pi^0$ toy MC events.

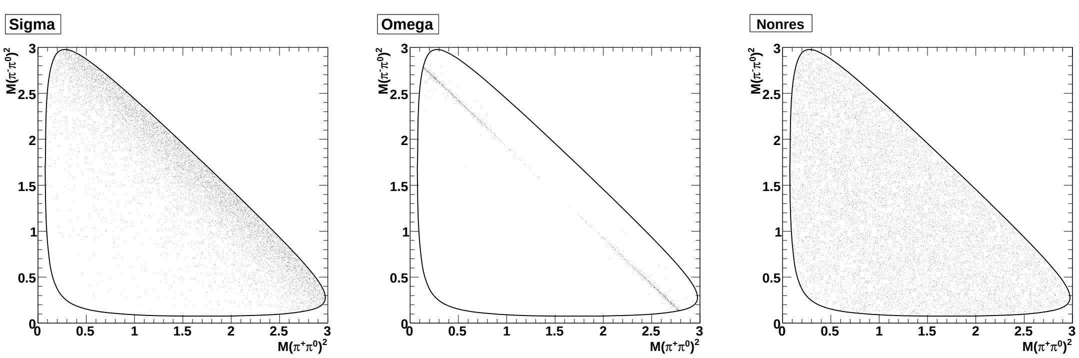


Figure 3.35: Dalitz plot distributions of $\sigma\pi^0$, $\omega\pi^0$, and nonresonant toy MC events.

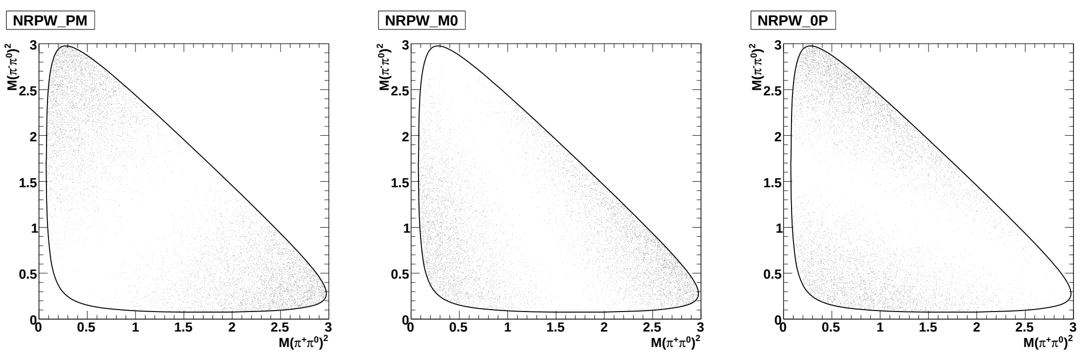


Figure 3.36: Dalitz plot distributions of nonresonant P-wave toy events, with the P-wave particles being the $\pi^+\pi^-$ (left), $\pi^-\pi^0$ (center), or $\pi^0\pi+$ (right).

3.6.7 Dalitz plot background PDF

The Dalitz distributions for background events with a correctly reconstructed D-candidate ($D\pi_D$, BBC_D and qq_D) are the same as for signal events with $x = y = 0$:

$$\mathcal{D}_{D\pi_D} \equiv \mathcal{D}_{BBC_D} \equiv \mathcal{D}_{qq_D} \equiv |f_{D^0}(s_{+0}, s_{-0})|^2, \tag{3.20}$$

where $f_{D^0}(s_{+0}, s_{-0})$ was defined in Eq. (3.14).

The Dalitz background distributions for $D\pi_{\not{D}}$, DKX, $BBC_{\not{D}}$ and $qq_{\not{D}}$ events are obtained from the generic $B\overline{B}$ and $q\overline{q}$ MC samples. We model the Dalitz distributions as the incoherent sum of three unpolarized ρ resonances and a non-resonant (NR) component:

$$\mathcal{D}_{\rm inc}(s_{+0}, s_{-0}) = f_{\rm NR}\mathcal{P}_{\rm NR} + (1 - f_{\rm NR})\left[(1 - f_{\rho^0})(f_{\rho^+}\mathcal{P}^0_{\rho^+} + (1 - f_{\rho^+})\mathcal{P}^0_{\rho^-}) + f_{\rho^0}\mathcal{P}^0_{\rho^0}\right]. \tag{3.21}$$

The non-resonant component $\mathcal{P}_{\rm NR}$ is parametrized by a cubic polynomial (Eq. (3.11)). Each resonant component $\mathcal{P}^0_\rho$ is the product of the efficiency function and a relativistic, spin zero, Breit-Wigner amplitude at the ρ mass A^0_ρ:

$$\mathcal{P}^0_\rho = \rm Eff(s_{+0}, s_{-0})|A^0_\rho(s_{+0}, s_{-0})|^2. \tag{3.22}$$

This gives a good description of ρ resonances from background events that do not have an angular correlation with any of the other particles in the decay. Table 3.28 lists the relative fractions of the different components found by the fit. To quantify the fit quality, we generate events from the fitted PDF, with 100 times the statistics of the input MC sample, and perform a KS test between the two Dalitz distributions. The KS probabilities of this test can be found in Table 3.28. The fits for each event type are shown in Figs. 3.37 through 3.40.

The $D\pi X$ Dalitz shape has an accumulation of events at low s_{+0} masses (see Fig. 3.41). This is due to $B \to D^*\pi^-, D^* \to D^0\pi^+$ decays, with the D^0 final state often containing

Event type	f_{NR}	f_{ρ^+}	f_{ρ^0}	KS prob
$D\pi_{\not D}$	0.94 ± 0.09	0.30 ± 0.60	0.00 ± 0.97	0.564
DKX	0.97 ± 0.02	1.00 ± 0.62	0.00 ± 0.11	0.120
$BBC_{\not D}$	0.76 ± 0.03	0.70 ± 0.07	0.13 ± 0.07	0.759
$qq_{\not D}$	0.89 ± 0.02	0.53 ± 0.10	0.22 ± 0.08	0.396

Table 3.28: Relative fractions of the non-resonant (NR), ρ^+ and ρ^0 component in the background Dalitz distributions together with the KS probability of the fit results.

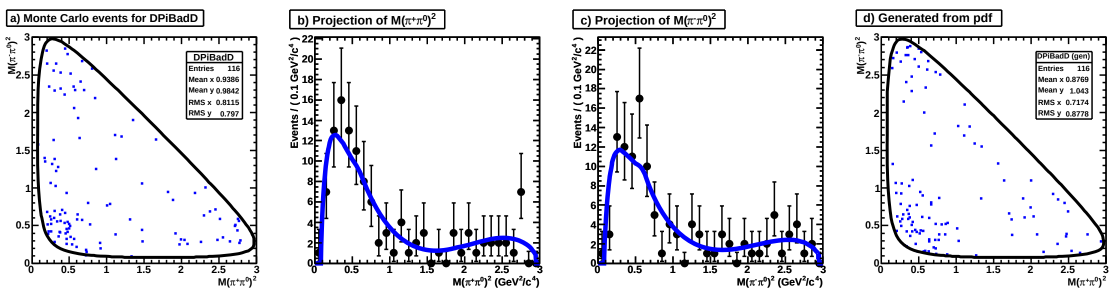


Figure 3.37: Fit to the Dalitz distribution of $D\pi_{\not D}$ events. a) shows the MC events used in the fit, b) and c) show the projections on s_{+0} and s_{-0} and d) shows the Dalitz distribution of events generated from the fitted PDF.

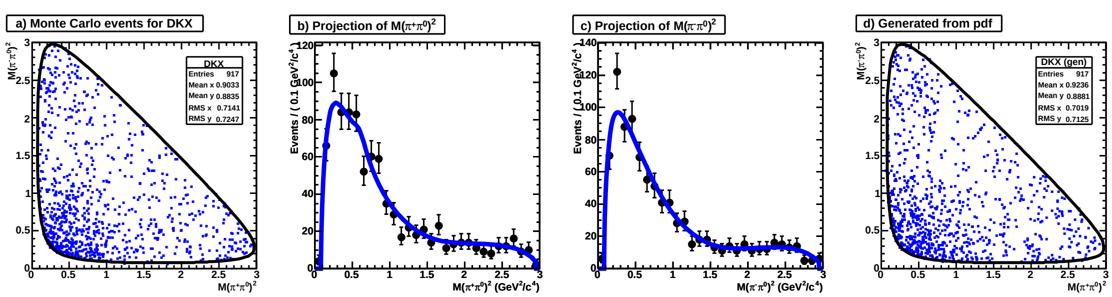


Figure 3.38: Fit to the Dalitz distribution of DKX events. a) shows the MC events used in the fit, b) and c) show the projections on s_{+0} and s_{-0} and d) shows the Dalitz distribution of events generated from the fitted PDF.

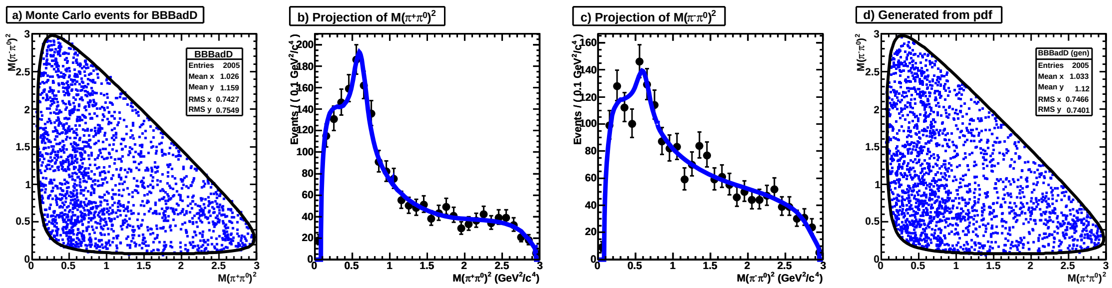


Figure 3.39: Fit to the Dalitz distribution of $BBC_{\not{D}}$ events. a) shows the MC events used in the fit, b) and c) show the projections on s_{+0} and s_{-0} and d) shows the Dalitz distribution of events generated from the fitted PDF.

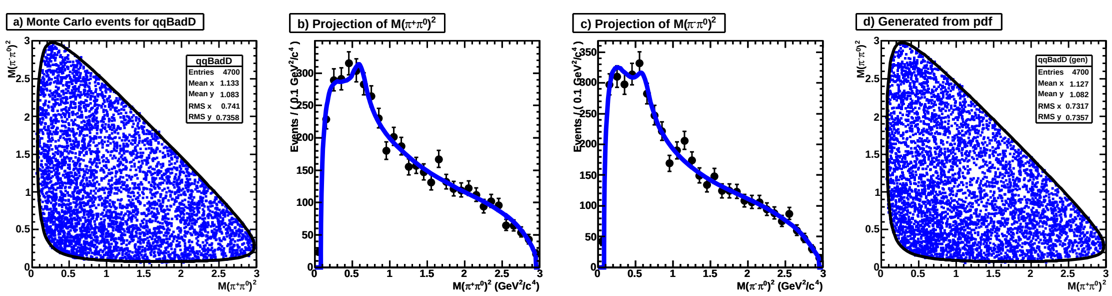


Figure 3.40: Fit to the Dalitz distribution of $qq_{\not{D}}$ events. a) shows the MC events used in the fit, b) and c) show the projections on s_{+0} and s_{-0} and d) shows the Dalitz distribution of events generated from the fitted PDF.

a K^- that is taken to be the bachelor kaon. Combining the soft pion from the D^* decay with a π^0 candidate that is often also soft, results in this low-mass accumulation.

We use a histogram-based PDF to parameterize this background type. Fig. 3.41 (left) shows the binning of the histogram used for the PDF. Each bin is weighted according to its area inside the Dalitz boundaries. The two plots on the right show the PDF projections on s_{+0} and s_{-0}, respectively.

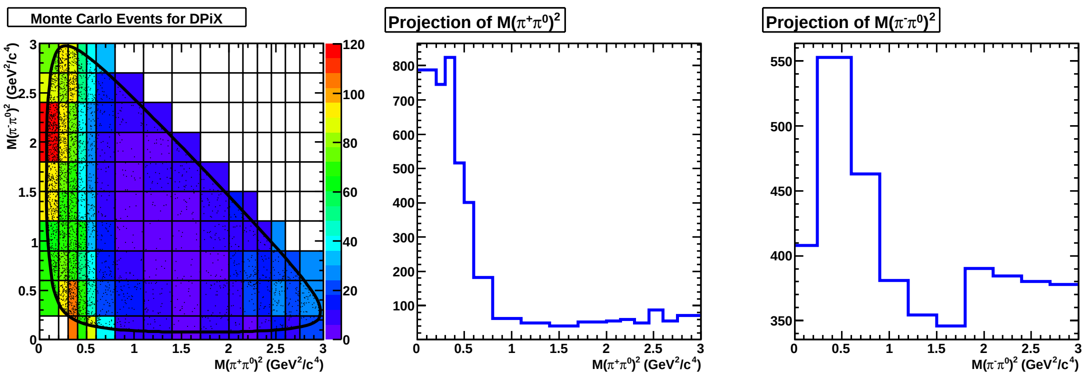


Figure 3.41: Dalitz distribution of $D\pi X$ events including the binning used for the histogram PDF (left). The two plots on the right show the projections of the PDF on s_{+0} and s_{-0}, respectively.

We use the same approach for signal events with a badly reconstructed D-candidate (DK_{bgd}). Fig. 3.42 shows the histogram used for the PDF. Overlaid is a plot of the DK_{bgd} events (with reduced statistics).

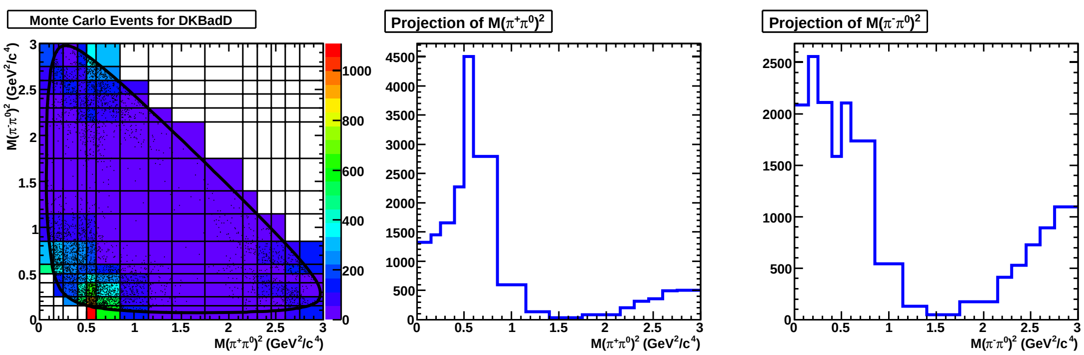


Figure 3.42: Dalitz distribution of DK_{bgd} events including the binning used for the histogram PDF (left). The two plots on the right show the projections of the PDF on s_{+0} and s_{-0}, respectively.

3.7 Dalitz plot fit of $D \to \pi^+\pi^-\pi^0$

The resonance parameters of $f_{D^0}(s_{+0}, s_{-0})$ (Eq. (3.14) in section 3.6.6) are determined in a separate analysis [66, 67]. Using $232\,\mathrm{fb}^{-1}$ of data, a clean $D_{\pi^+\pi^-\pi^0}$ sample is obtained from $D^{*+} \to D^0\pi_s^+$ decays where the charge of the soft π_s^+ tags the flavor of the D^0. The resolution in $\Delta m = m_D^* - m_D^0$ is approximately $0.3\,\mathrm{MeV}/c^2$ and only D^0 candidates with Δm within $0.6\,\mathrm{MeV}/c^2$ of the central value are retained. Figure 3.43 shows the $\pi^+\pi^-\pi^0$ invariant mass of these D^0 candidates.

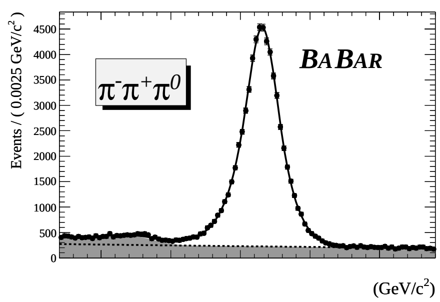


Figure 3.43: $\pi^+\pi^-\pi^0$ invariant mass of D^0 candidates obtained from $D^{*+} \to D^0\pi_s^+$ decays. The dots represent the data, the dashed line shows the combinatorial background and the shaded region represents the total background.

For the Dalitz plot analysis, only D^0 candidates within the 1σ signal region ($1.848 < m(\pi^+\pi^-\pi^0) < 1.880$) are used. This window also rejects all background events from misreconstructed $\pi^+K^-\pi^0$ decays that can be seen in the left tail of Fig. 3.43. The Dalitz fit uses the resonance masses and width listed in Table 3.27 (p. 86). The amplitude of the ρ^+ component is fixed to one and its phase is fixed to zero. All other amplitudes and phases are floating in the fit. The *fit fraction* for each PDF component in terms of Eq. (3.14) is defined as

$$F_r = \frac{\int |a_r A_r(s_{+0}, s_{-0})|^2 \, ds_{-0} ds_{+0}}{\int |f_{D^0}(s_{+0}, s_{-0})|^2 \, ds_{-0} ds_{+0}} \,. \tag{3.23}$$

Due to interference, the fit fraction is not the relative contribution of the component. Table 3.29 lists the result of the Dalitz fit and Fig. 3.44 shows the projections of the PDF on the three two-pion invariant masses.

Component	Amplitude a_r	Phase ϕ_r	Fraction $F_r(\%)$
ρ^+	1	0	67.8
ρ^0	0.588 ± 0.006	16.2 ± 0.6	26.2 ± 0.5
ρ^-	0.714 ± 0.008	-2.0 ± 0.6	34.6 ± 0.8
$\rho^+(1450)$	-0.21 ± 0.06	34 ± 18	0.11 ± 0.07
$\rho^0(1450)$	0.33 ± 0.06	10 ± 8	0.30 ± 0.11
$\rho^-(1450)$	0.82 ± 0.05	15.9 ± 2.9	1.79 ± 0.22
$\rho^+(1700)$	2.25 ± 0.18	-16.7 ± 2.4	4.1 ± 0.7
$\rho^0(1700)$	2.51 ± 0.15	-17.1 ± 1.7	5.0 ± 0.6
$\rho^-(1700)$	2.00 ± 0.11	-50.2 ± 3.3	3.25 ± 0.36
$f_0(980)$	0.0525 ± 0.0039	120.6 ± 4.7	0.250 ± 0.037
$f_0(1370)$	0.222 ± 0.034	-24 ± 9	0.37 ± 0.11
$f_0(1500)$	0.203 ± 0.022	192 ± 9	0.39 ± 0.08
$f_0(1710)$	0.391 ± 0.046	231 ± 8	0.31 ± 0.07
$f_2(1270)$	0.303 ± 0.009	9.0 ± 3.4	1.32 ± 0.08
σ	-0.238 ± 0.015	7.9 ± 4.3	0.82 ± 0.10
Nonresonant	0.57 ± 0.07	168.6 ± 3.7	0.84 ± 0.21

Table 3.29: Result of the D^0 Dalitz fit. The amplitudes (a_r) and phases (ϕ_r) are defined relative to the ρ^+. The last column shows the fit fraction F_r according to Eq. (3.23).

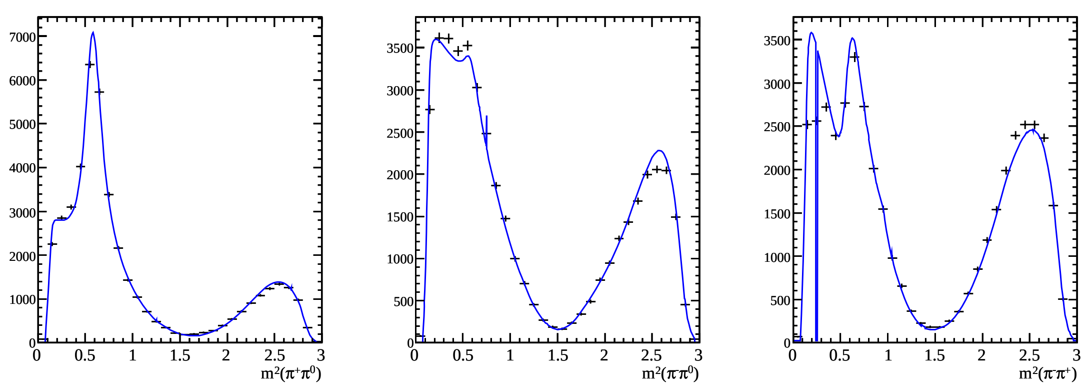


Figure 3.44: Projections of the data and PDF for the Dalitz fit to $D^0 \to \pi^+\pi^-\pi^0$ decays. From left to right, the invariant masses squared of $\pi^+\pi^0$, $\pi^-\pi^0$ and $\pi^+\pi^-$ are shown, respectively.

3.8 Maximum likelihood fit

This section describes how the PDF described in section 3.6 is used to extract the *CP* parameters from the data sample.

3.8.1 Combining Dalitz shape, signal yield and asymmetry

In order to measure the *CP* parameters using the Dalitz shape one minimizes the negative log-likelihood (NLL)

$$\mathcal{L}_{\mathrm{DP}} = -\sum \log \mathcal{P}^2 \tag{3.24}$$

with respect to the *CP* parameters. $\mathcal{P}^2$ is the PDF given by equations (3.6) and (3.8) and the sum goes over all events.

We note, however, that the signal branching fraction and decay rate asymmetry also depend on the *CP* parameters. This information is used in the GLW (section 1.3.1) and ADS (section 1.3.2) methods but has not been exploited yet by any of the previous Dalitz analysis and is not captured by this NLL. From very general arguments [17], one can conclude that the branching ratio and asymmetry have a sensitivity to the *CP* parameters that is similar to that of the Dalitz shape distribution. To incorporate all the available information in the data, we minimize

$$\mathcal{L} = \mathcal{L}_{\mathrm{DP}} + \mathcal{L}_{\mathrm{BA}}\,, \tag{3.25}$$

where $\mathcal{L}_{\mathrm{BA}}$ is an additional term to the log-likelihood that represents the information contained in the branching ratio and asymmetry. This term can be written in form of a χ^2

$$\mathcal{L}_{\mathrm{BA}} \equiv \frac{1}{2} Y_i V_{ij}^{-1} Y_j, \tag{3.26}$$

where we have defined Y_1 (Y_2) to be the difference between the measured and expected signal yield (asymmetry), and V is the covariance matrix for these two observables including some systematic errors described below. The expected signal yield and asymmetry are

$$\begin{aligned} N_{\rm exp} &= N^+_{\rm exp} + N^-_{\rm exp} \\ A_{\rm exp} &= \frac{N^-_{\rm exp} - N^+_{\rm exp}}{N^-_{\rm exp} + N^+_{\rm exp}}, \end{aligned} \tag{3.27}$$

where $N^{\pm}_{\rm exp}$ is the expected number of $B^{\pm}$ signal events for a given value of the *CP* parameters $z_{\pm}$

$$\begin{aligned} N^-_{\rm exp} &= \eta \frac{\int |f_{D^0}(s_{+0}, s_{-0}) + z_- f_{D^0}(s_{-0}, s_{+0})|^2 \, ds_{+0} ds_{-0}}{\int |f_{D^0}(s_{+0}, s_{-0})|^2 \, ds_{+0} ds_{-0}} \\ N^+_{\rm exp} &= \eta \frac{\int |f_{D^0}(s_{-0}, s_{+0}) + z_+ f_{D^0}(s_{+0}, s_{-0})|^2 \, ds_{+0} ds_{-0}}{\int |f_{D^0}(s_{-0}, s_{+0})|^2 \, ds_{+0} ds_{-0}}, \end{aligned} \tag{3.28}$$

and we have used the quantities defined in Eq. (3.4). η is a normalizing factor depending on the number of $B\overline{B}$ events $N_{B\overline{B}}$, the no-*CP* branching fractions, and the absolute efficiency ϵ:

$$\eta \equiv \frac{1}{2} N_{B\overline{B}} \, \mathcal{B}(B^- \to D^0 K^-) \, \mathcal{B}(D^0 \to \pi^+\pi^-\pi^0) \, \epsilon. \tag{3.29}$$

The statistical and systematic uncertainties in η are included in the error matrix V.

In the subsections below we explore the advantages of minimizing $\mathcal{L}$ instead of $\mathcal{L}_{\rm DP}$. We then describe the fit procedure to realize this in section 3.8.5

3.8.2 Behavior of the Dalitz shape NLL $\mathcal{L}_{\rm DP}$

Figure 3.45 shows the dependence of $\mathcal{L}_{\rm DP}$ (3.24) on the parameters $x_{\pm}$ and $y_{\pm}$, calculated from a data luminosity equivalent toy MC experiment. The lines indicate 1σ contours with the line surrounding the white area being the 1σ contour line. It is evident that at negative values of x, $\mathcal{L}_{\rm DP}$ is almost constant, resulting in a reduced sensitivity in this region.

The asymmetry with respect to x and the flatness of $\mathcal{L}_{\rm DP}$ at low x values are a conse-

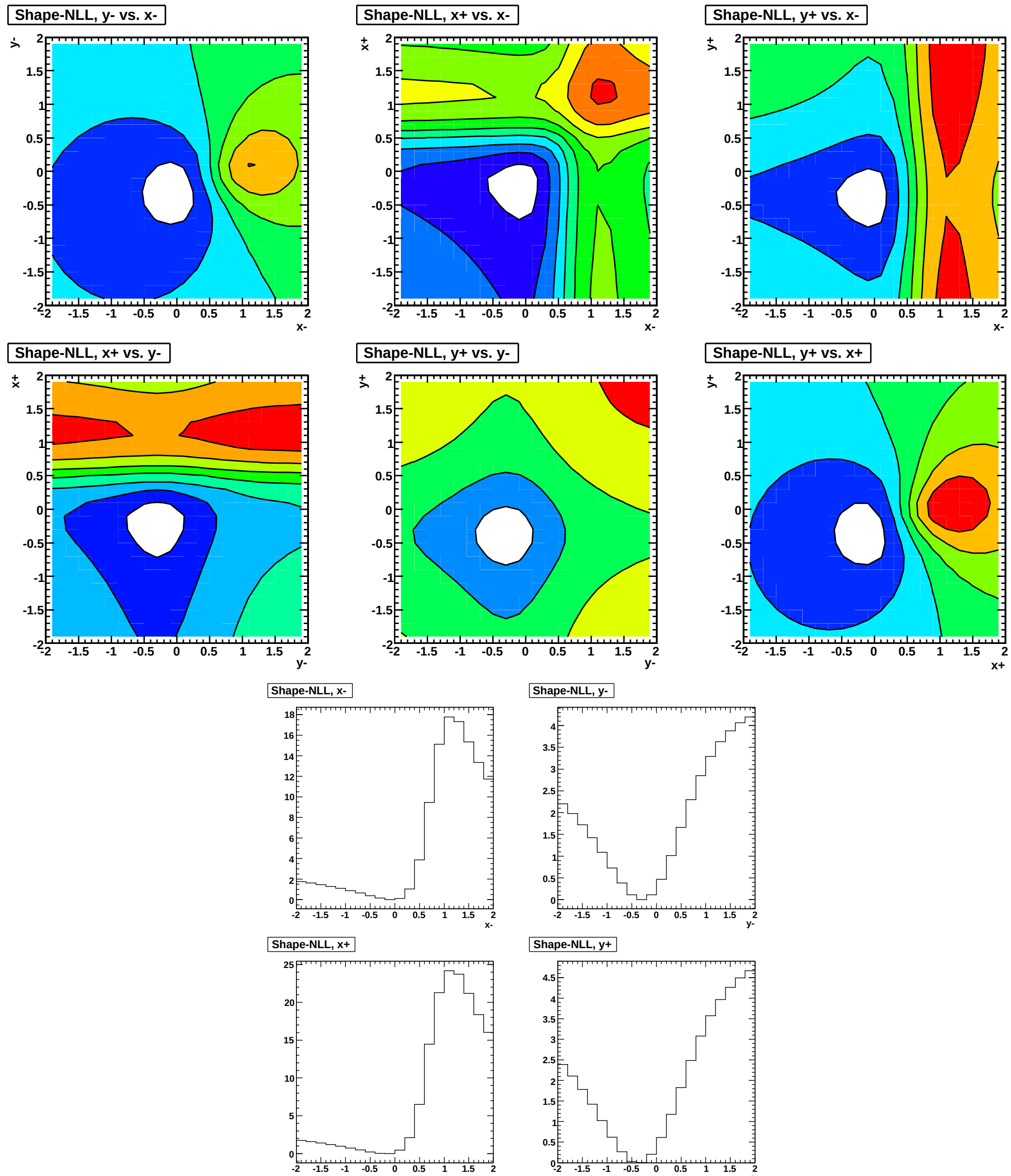


Figure 3.45: Dependence of the negative log likelihood $\mathcal{L}_{\mathrm{DP}}$, computed with a single toy MC experiment containing both signal and background, on all pair combinations of the parameters $x_\pm, y_\pm$ (2D plots, 1σ contours) and on the individual parameters (1D plots). The true values of $x_\pm, y_\pm$ are 0. The values of the parameters not shown in each plot are set to their true values.

quence of the small value of the phase difference $\Delta\phi = \phi_{\rho^-} - \phi_{\rho^+} \approx 2^\circ$ between the ρ^- and ρ^+ resonances, as well as the relatively large ratio $a_{\rho^-}/a_{\rho^+} \approx 0.7$ between the magnitudes of their amplitudes (see Table 3.29). To illustrate the effect of $\Delta\phi$ on the NLL shape we show in Fig. 3.46 $\mathcal{L}_{\rm DP}$ for B^- events as a function of x_- and y_- for varying values of $\Delta\phi$. One sees that as $\Delta\phi$ increases, the shape of $\mathcal{L}_{\rm DP}$ rotates in the $x - y$ plane. Moreover, the sensitivity (density of contour lines) is maximal around $\Delta\phi \sim 90^\circ$. From the upper left plot, which corresponds to the physical situation in $D^0 \to \pi^+\pi^-\pi^0$, we observe that the NLL is highly asymmetric, that the sensitivity greatly depends on the true values of x and y and that there are non-linear correlations between x and y. These properties of the NLL make it very difficult to obtain unbiased results for x and y in the maximum likelihood fit.

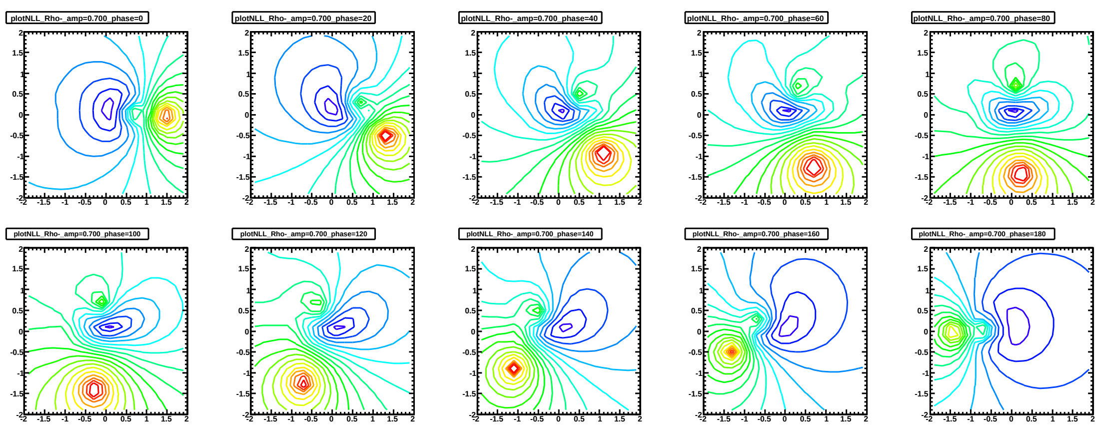


Figure 3.46: $\mathcal{L}_{\rm DP}$ of a signal-only B^- toy sample as a function of x_- and y_- if the Dalitz plot contains only ρ^+ and ρ^-. The phase difference between them varies from 0° (top left) to 180° (bottom right) in 20° steps.

In Figure 3.46, one can see a very tall peak in $\mathcal{L}_{\rm DP}$ at, for example $x \approx 1.4$ for the top left plot. This peak appears approximately where the *CP* parameters are such that the ρ^+ totally destructively interferes with the ρ^-, making the Dalitz plot highly asymmetric. With the Dalitz plot distribution thus being very different from the highly symmetric distribution in the no-*CP* violation case (a symmetry that is due to approximately equal ρ^+ and ρ^- amplitudes), this results in very large values for $\mathcal{L}_{\rm DP}$.

3.8.3 Behavior of the yields term $\mathcal{L}_{\rm BA}$

Figure 3.47 shows the expected signal yields $N^{\pm}_{\rm exp}$ of (3.28) as a function of the *CP* parameters. A significant dependence is observed, which can be used to increase the sensitivity of the measurement to the values of these parameters. The source of the circular shape of regions of constant yield vs. x and y is as follows. The yield (either N^+ or N^-) is proportional to

$$N \propto \int \left|A + (x+iy)\overline{A}\right|^2 \, d\phi\,, \tag{3.30}$$

where the integral is over the Dalitz plot and we use the shorthand notation

$$\begin{aligned} A &\equiv f_{D^0}(s_{+0}, s_{-0}) \\ \overline{A} &\equiv f_{D^0}(s_{-0}, s_{+0}). \end{aligned} \tag{3.31}$$

Squaring and dividing (3.30) by $\int |A|^2$ gives

$$N \propto 1 + (x^2+y^2) + 2x\frac{\int \Re(A^*\overline{A})}{\int |A|^2} - 2y\frac{\int \Im(A^*\overline{A})}{\int |A|^2}. \tag{3.32}$$

For constant N, this gives a circle whose center (minimum value of N) is at

$$\begin{aligned} x^0 &= -\frac{\int \Re(A^*\overline{A})}{\int |A|^2} \\ y^0 &= \frac{\int \Im(A^*\overline{A})}{\int |A|^2}. \end{aligned} \tag{3.33}$$

It is easy to show that $y^0 = 0$ as a result of the symmetry of the *boundary* of the Dalitz plot. Dividing the Dalitz plot into the region above and below the symmetry line $s_{+0} = s_{-0}$, we can write

$$\begin{aligned} y^0 \int |A|^2 &= \int \Im(A^*\overline{A}) \\ &= \int_{s_{+0}>s_{-0}} \Im(A^*\overline{A}) + \int_{s_{+0}<s_{-0}} \Im(A^*\overline{A}). \end{aligned} \tag{3.34}$$

Since changing from the "above" region to the "below" region is equivalent to exchanging $s_{+0} \leftrightarrow s_{-0}$ and hence $A \leftrightarrow \overline{A}$ according to (3.31) we obtain

$$\begin{aligned} y^0 \int |A|^2 &= \int_{s_{+0}>s_{-0}} \Im(A^*\overline{A}) + \int_{s_{+0}>s_{-0}} \Im(\overline{A}^*A) \\ &= \int_{s_{+0}>s_{-0}} \Im\left(A^*\overline{A} - (\overline{A}^*A)^*\right). \end{aligned} \tag{3.35}$$

The integrand of the last line vanishes, resulting in

$$y^0 = 0. \tag{3.36}$$

Furthermore, we expect the magnitude of $|x^0|$ to be of order $\int |A|^2 = 1$. Both these expectations are seen in Fig. 3.47.

In Fig. 3.48 we make use of the dependence of the yields on the *CP* parameters by showing the dependence of $\mathcal{L}_{\mathrm{BA}}$ on these parameters. One can see that $\mathcal{L}_{\mathrm{BA}}$ does not give a unique solution for $x_\pm$ and $y_\pm$, since there are two observables and four unknowns. For example, the y versus x plot shows that while one can find a decent solution in the radial direction (i.e., the x-axis when one is at $y = 0$), the solution in the tangential direction (y in this example) will have a large error, since infinitesimal variations in this direction result in a negligible change in the NLL. However, $\mathcal{L}_{\mathrm{BA}}$ still puts a significant constraint on the *CP* parameters, given additional information to resolve this ambiguity.

Finally, Fig. 3.49 shows the dependence of the combined NLL function $\mathcal{L}$. This function depends more strongly on the *CP* parameters and is better behaved than either $\mathcal{L}_{\mathrm{DP}}$ or $\mathcal{L}_{\mathrm{BA}}$. It is worthwhile to note some features of $\mathcal{L}$. Fig. 3.49 still shows traces of the ambiguity seen in Fig. 3.48. In addition, the circular shape clearly seen in the y_+ vs. x_+ plot and the nonlinear $x - y$ correlation that it indicates, means that there is on average a bias in the radial direction. For example, with the true values being $x = y = 0$, both upward and downward fluctuations in y lead to an upward fluctuation in x.

It is now obvious that the Cartesian parameterization of the *CP* parameter $z(x, y)$

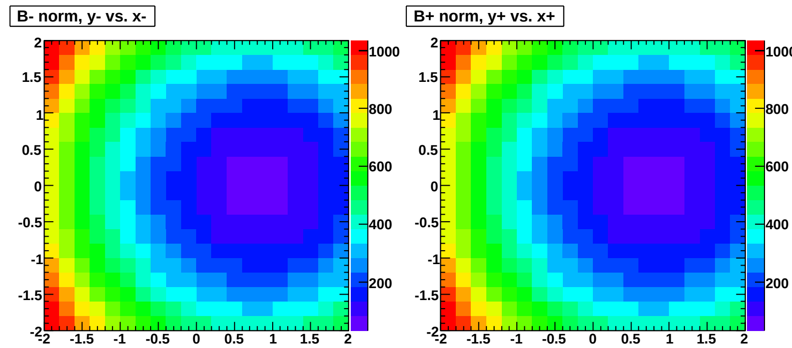


Figure 3.47: Dependence of the expected number of events $N^{\pm}_{\text{exp}}$ on the true values of the CP parameters $x_{\pm}, y_{\pm}$ in a toy experiment.

chosen by previous analyses is inappropriate for this analysis. Due to the circular shape of $\mathcal{L}$ polar coordinates $z(\rho, \theta)$ are a much better suited pair of parameters.

3.8.4 Polar coordinates for the CP parameters

The circular shape of $\mathcal{L}_{\text{BA}}$ in the $x - y$ plane (Fig. 3.49) results in non-linear correlations between the fit parameters x and y. This causes various problems that are illustrated in Appendix 3.A. Therefore, instead of using the Cartesian coordinates x and y, we define the following polar parameterization $z(\rho, \theta)$ reflecting the symmetry of $\mathcal{L}_{\text{BA}}$:

$$\begin{aligned} \rho_{\pm} &= \sqrt{(x_{\pm} - x^0)^2 + (y_{\pm} - y^0)^2} \\ \theta_{\pm} &= \tan^{-1}\left(\frac{y_{\pm} - y^0}{x_{\pm} - x^0}\right) , \end{aligned} \tag{3.37}$$

where (x^0, y^0) is the origin of the polar coordinate system, defined in Eq. (3.33) with

$$\begin{aligned} x^0 &= 0.8496 \\ y^0 &= 0. \end{aligned} \tag{3.38}$$

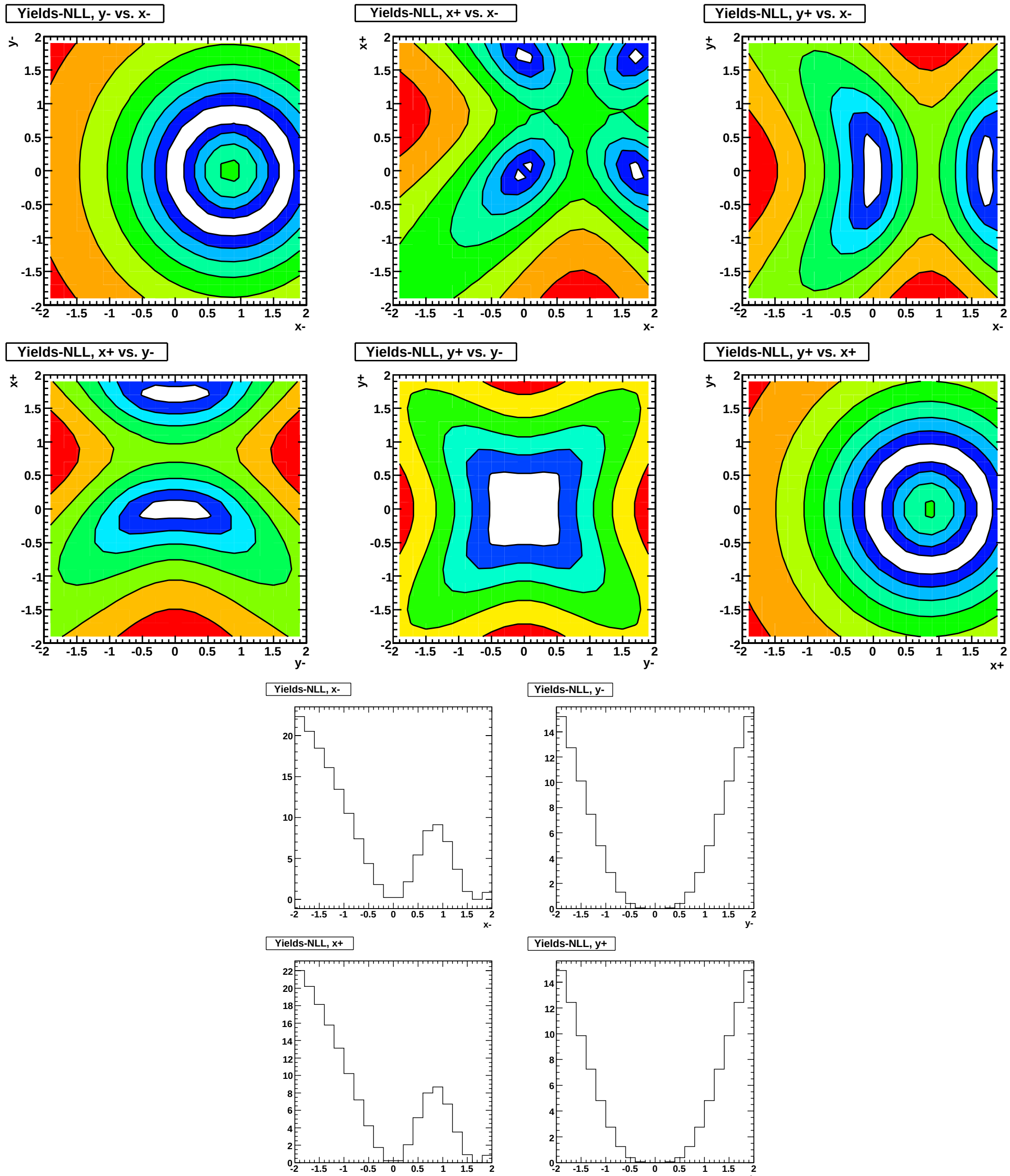


Figure 3.48: Dependence of $\mathcal{L}_{\mathrm{BA}}$ on all pair combinations of the parameters $x_{\pm}, y_{\pm}$ (2D plots, 1σ contours) and on the individual parameters (1D plots) in a toy experiment. The true values of $x_{\pm}, y_{\pm}$ are 0. The values of the parameters not shown in each plot are set to their true values.

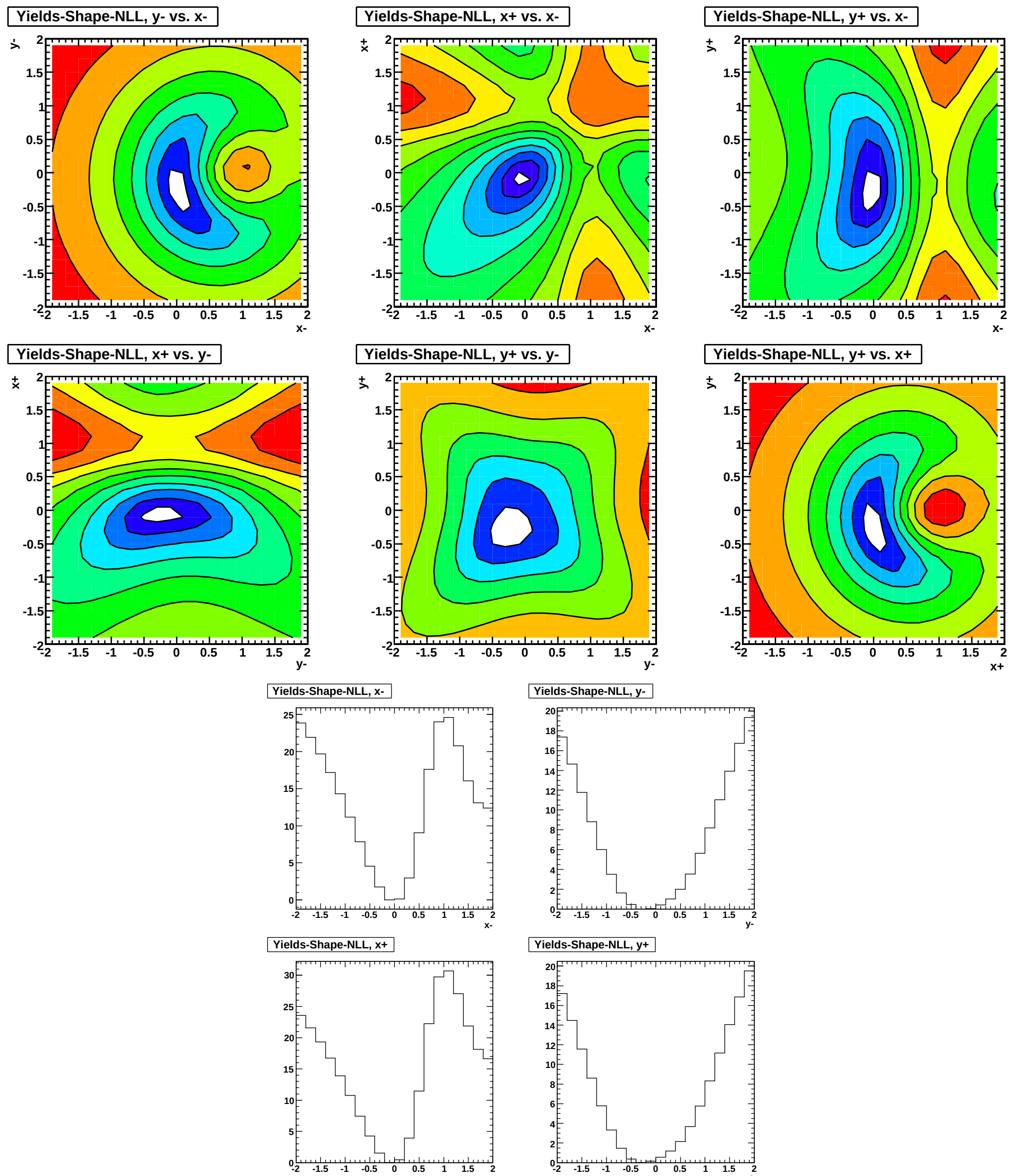


Figure 3.49: Dependence of $\mathcal{L}$ on all pair combinations of the parameters $x_\pm, y_\pm$ (2D plots, 1σ contours) and on the individual parameters (1D plots) in a toy experiment. The true values of $x_\pm, y_\pm$ are 0. The values of the parameters not shown in each plot are set to their true values.

The no-CP violation case in these coordinates is $\rho_{\pm} = x^0$ and $\theta_{\pm} = 180°$. This set of variables represents an (almost) uncorrelated pair of variables. Some correlation remains due to the fact that $\mathcal{L}_{\mathrm{DP}}$ does not respect the circular symmetry of $\mathcal{L}_{\mathrm{BA}}$, implied by these coordinates.

3.8.5 Fit Procedure

In order to make use of all the CP-relevant information provided in the data, namely, the Dalitz shape, the signal branching fraction, and the asymmetry, we carry out the following 2-step fitting procedure:

1. We fit the data with the PDF $\mathcal{P}^1$ of Eqs. (3.6) and (3.7), which depends on the variables ΔE, q', and d'. From this fit we obtain the parameters

   (a) $N_{DK_{\rm sig}}$ = the number of $DK_{\rm sig}$ events.

   (b) $A_{DK_{\rm sig}}$ = the $DK_{\rm sig}$ decay rate asymmetry

   (c) $N_{D\pi_D}$ = the number of $D\pi_D$ events

   (d) $N_{BB_{\not D}}$ = the total number of events of types DKX, $D\pi X$, and $BBC_{\not D}$.

   (e) $R_{D\pi X} \equiv N_{D\pi X}/N_{BB_{\not D}}$ = the ratio between the number of $D\pi X$ events and $N_{BB_{\not D}}$.

   (f) $N_{qq_{\not D}}$ = the number of $qq_{\not D}$ events.

   All shape parameters (parameters describing the shapes of $\mathcal{P}_t$) are fixed to the values obtained earlier on MC or data (Sections 3.6.2, 3.6.3 and 3.7). In addition, we use the MC to fix several ratios between certain numbers of events that, together with the floating parameters listed above, give the number of events for all event types. The values of these ratios after all cuts are listed here, with N always referring to a number of events:

   (a) $R_{DKX} \equiv N_{DKX}/N_{D\pi X} = 0.228 \pm 0.057$

   (b) $R_{D\pi_{\not D}} \equiv N_{D\pi_{\not D}}/N_{D\pi_D} = 0.253 \pm 0.026$

   (c) $R_{BBC_D} \equiv N_{BBC_D}/N_{BB_{\not D}} = 0.00450 \pm 0.00080$

   (d) $R_{qq_D} \equiv N_{qq_D}/N_{qq_{\not D}} = 0.0116 \pm 0.0016$

   (e) $R_{DK_{\rm bgd}} \equiv N_{DK_{\rm bgd}}/N_{DK_{\rm sig}} = 0.2540 \pm 0.0022$

The error of R_{DKX} comes from the measured branching fractions in the PDG. All other errors are due to MC statistics only, and are not used in this analysis. The values of these parameters are later varied for systematic error evaluation (section 3.11.1).

2. We compute $\mathcal{L}_{\rm BA}$ using $N_{DK_{\rm sig}}$, $A_{DK_{\rm sig}}$, and their error matrix obtained in the previous step. The NLL $\mathcal{L}_{\rm DP}$ is calculated from the data using the PDF $\mathcal{P}^2$, which depends on the variables ΔE, d', and the Dalitz variables s_{+0} and s_{-0}. We then minimize $\mathcal{L} = \mathcal{L}_{\rm BA} + \mathcal{L}_{\rm DP}$, floating only the CP parameters

   (a) ρ_-

   (b) θ_-

   (c) ρ_+

   (d) θ_+

3.9 Monte Carlo fit studies

All the MC studies in this section are performed with the D decay parameters found in our fit to the D^* data. Unless specified otherwise, the *CP* violating parameters $z_\pm$ are set to the values found in the $B^- \to D^0_{K^0_S\pi^+\pi^-}K^-$ analysis [11] converted to our set of polar coordinates centered at (x^0, y^0). Table 3.30 shows these default parameters. The number of B^- and B^+ signal events used in the toys is calculated from (3.28), each separately fluctuated by their Poisson errors for every toy experiment. The other parameters are set to the values found in the data fit listed in Table 3.34 (p. 125). Each toy MC fit uses the 2-step fitting procedure described in section 3.8.5.

Parameter	Value
ρ_-	0.775
θ_-	175.3
ρ_+	0.979
θ_+	178.9
$< N_{DK_{\rm sig}} >$	202
$< A_{DK_{\rm sig}} >$	−0.1687

Table 3.30: *CP* violating parameters used for all the MC studies in this section. The origin of the polar coordinates is at $x^0 = 0.850$ and $y^0 = 0$.

3.9.1 Toy MC studies

We generate 2000 toy MC experiments including signal and background and fit them with the 2-step fitting procedure. Since the systematic errors incorporated in $\mathcal{L}_{\rm BA}$ are not taken into account in the toy generation we set them to zero for this set of experiments. The results for each floating variable are shown in Figs. 3.50 through 3.56. Each set of plots shows the distribution of the pull, error and fitted value. The pull for a variable ζ with true value $\zeta^{\rm true}$, measured value $\zeta^{\rm meas}$ and measurement error $\sigma^{\rm meas}$ is defined as

$$\mathrm{pull}(\zeta) = \frac{\zeta^{\rm meas} - \zeta^{\rm true}}{\sigma^{\rm meas}}. \tag{3.39}$$

For a large number of measurements, pull(ζ) is distributed according to a standard normal distribution (Gaussian centered around zero with unit width) if the measurement is unbiased and the measured error on average represents the true measurement error. All pull distributions in Figs. 3.50 through 3.56 are standard normal distributions except for $\theta_{\pm}$. The width of the pull in $\theta_{\pm}$ is about 1.2. Although it is not apparent from the NLL projections, which appear to be very Gaussian, the cause is most likely due to a bifurcation of the θ fit results for different values of ρ. In fact, fitting the pull distribution for experiments with $\rho_{\pm} < 0.9$ results in a unit width Gaussian. This is similar to what was observed using Cartesian coordinates (see Appendix 3.A). Figure 3.57 shows the minimum values of the NLL found by the branching ratio (step-1) fit and CP (setp-2) fit.

We repeat the same experiments but this time including the systematic errors in $\mathcal{L}_{\mathrm{BA}}$ to obtain a correct estimate for the expected fit errors on data. Figures 3.58 and 3.59 show the results for the CP parameters. As expected the width of the $\rho_{\pm}$ pulls are too narrow.

Finally, Fig. 3.60 through 3.62 show the NLL projections of $\mathcal{L}_{\mathrm{DP}}$, $\mathcal{L}_{\mathrm{BA}}$ and $\mathcal{L}$ for each pair of CP parameters and each CP parameter individually. By construction there is no sensitivity to $\theta_{\pm}$ in $\mathcal{L}_{\mathrm{BA}}$. Overall the polar coordinates result in much better behaved fit parameters than the Cartesian coordinates.

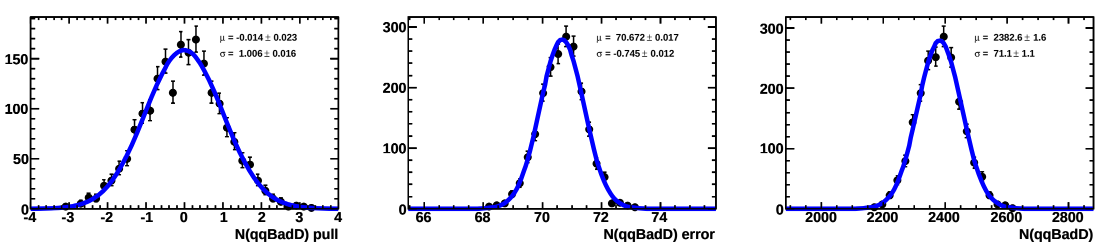


Figure 3.50: Results of 2000 toy MC experiments (signal and background, no systematic error in $\mathcal{L}_{\mathrm{BA}}$) for the pull (left), error (center) and fitted values (right) of $N_{qq\not{D}}$.

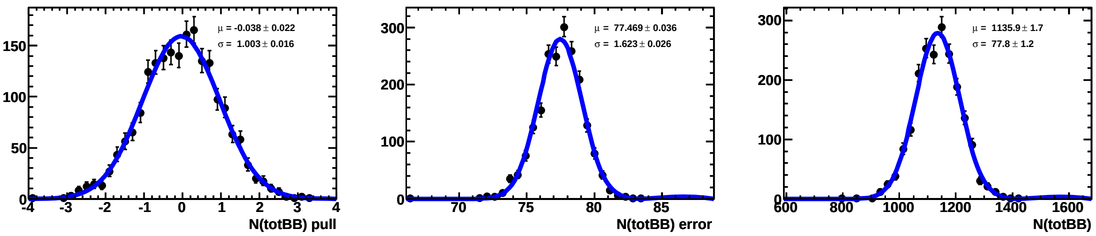


Figure 3.51: Results of 2000 toy MC experiments (signal and background, no systematic error in $\mathcal{L}_{\mathrm{BA}}$) for the pull (left), error (center) and fitted values (right) of $N_{BB\not{p}}$.

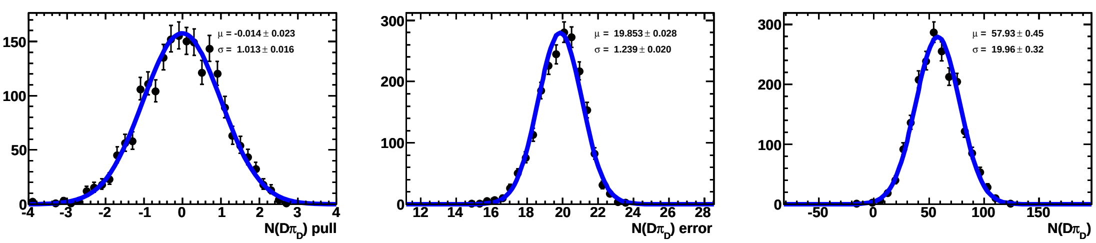


Figure 3.52: Results of 2000 toy MC experiments (signal and background, no systematic error in $\mathcal{L}_{\mathrm{BA}}$) for the pull (left), error (center) and fitted values (right) of $N_{D\pi_D}$.

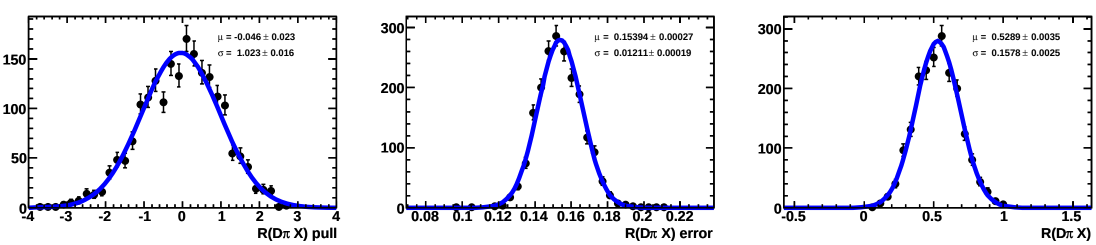


Figure 3.53: Results of 2000 toy MC experiments (signal and background, no systematic error in $\mathcal{L}_{\mathrm{BA}}$) for the pull (left), error (center) and fitted values (right) of $R_{D\pi}$.

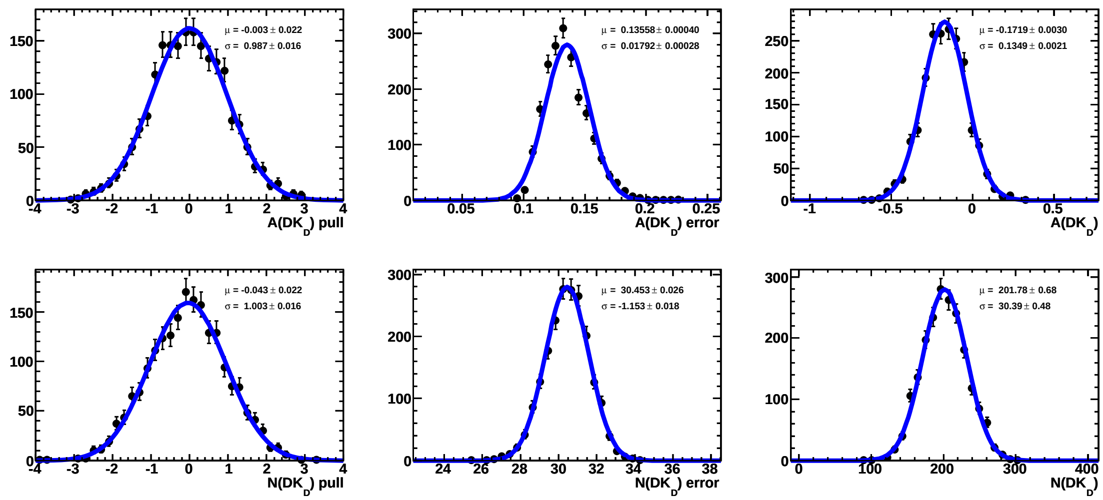


Figure 3.54: Results of 2000 toy MC experiments (signal and background, no systematic error in $\mathcal{L}_{\mathrm{BA}}$) for the pull (left), error (center) and fitted values (right) of $A_{DK_{\mathrm{sig}}}$ (top) and $N_{DK_{\mathrm{sig}}}$ (bottom).

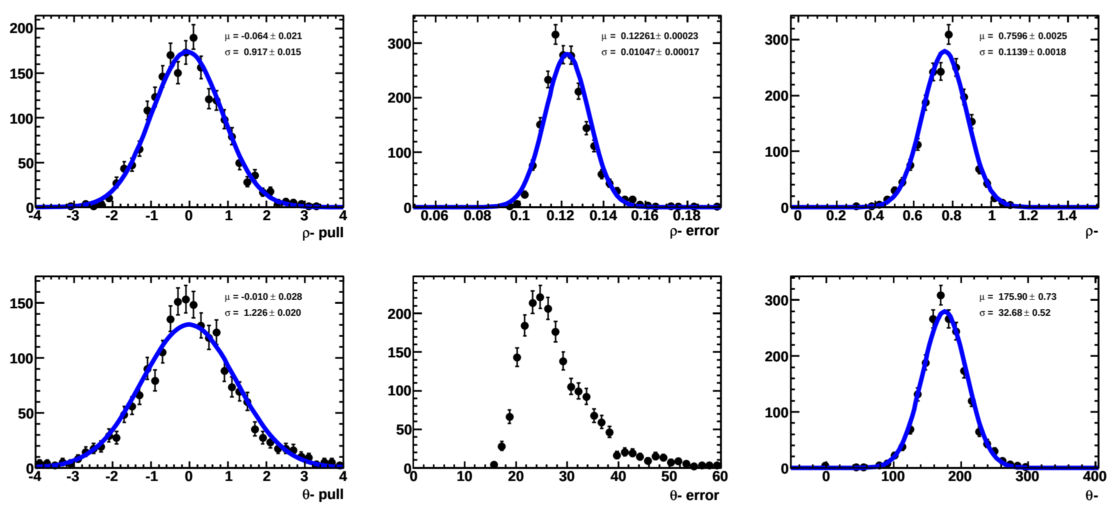


Figure 3.55: Results of 2000 toy MC experiments (signal and background, no systematic error in $\mathcal{L}_{\mathrm{BA}}$) for the pull (left), error (center) and fitted values (right) of ρ_- (top) and θ_- (bottom). See Fig. 3.58 for the corresponding toy MC including systematic errors.

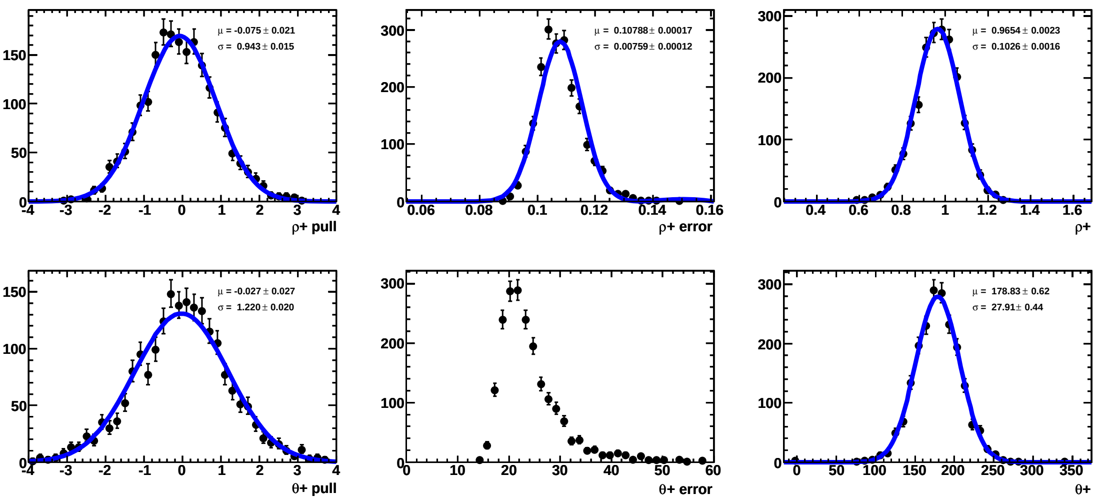


Figure 3.56: Results of 2000 toy MC experiments (signal and background, no systematic error in $\mathcal{L}_{\mathrm{BA}}$) for the pull (left), error (center) and fitted values (right) of ρ_+ (top) and θ_+ (bottom). See Fig. 3.59 for the corresponding toy MC including systematic errors.

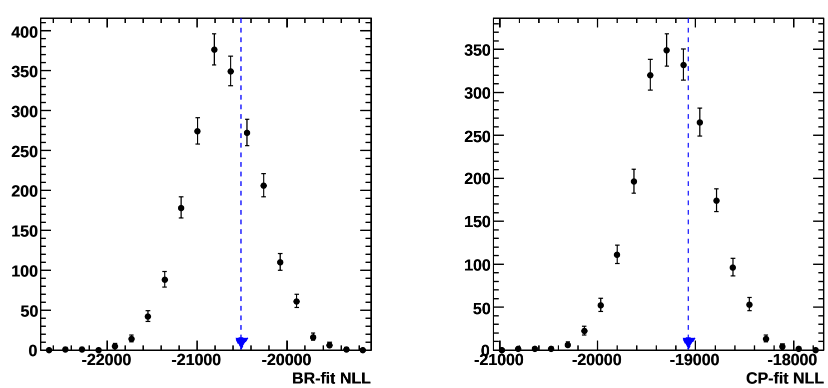


Figure 3.57: Minimum value of the NLL from the 2000 toy MC experiments for the yields *step-1* fit (left) and the *CP step-2* fit (right). The arrow indicates the NLL found in the data fit (section 3.10).

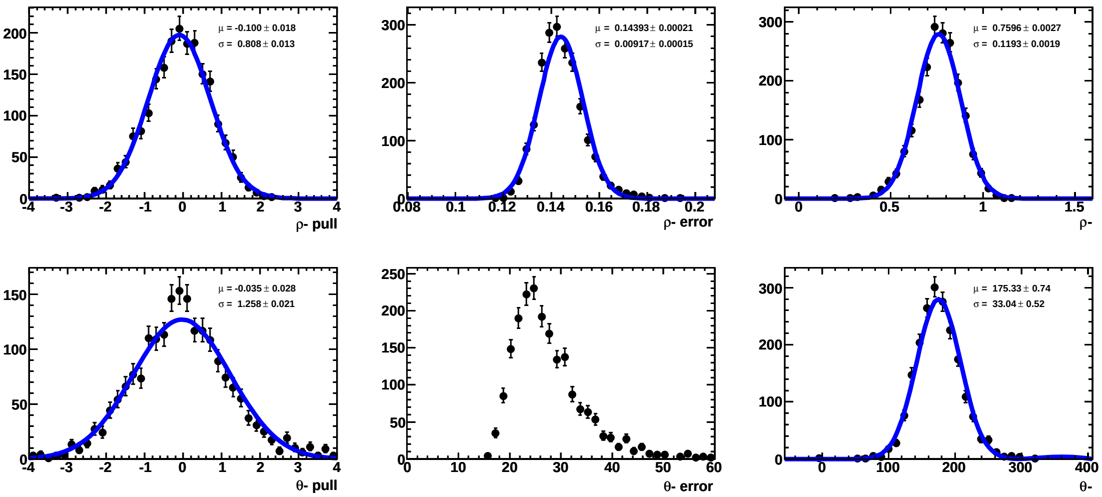


Figure 3.58: Results of 2000 toy MC experiments (signal and background, including systematic error in $\mathcal{L}_{\mathrm{BA}}$) for the pull (left), error (center) and fitted values (right) of ρ_- (top) and θ_- (bottom). See Fig. 3.55 for the corresponding toy MC without systematic errors.

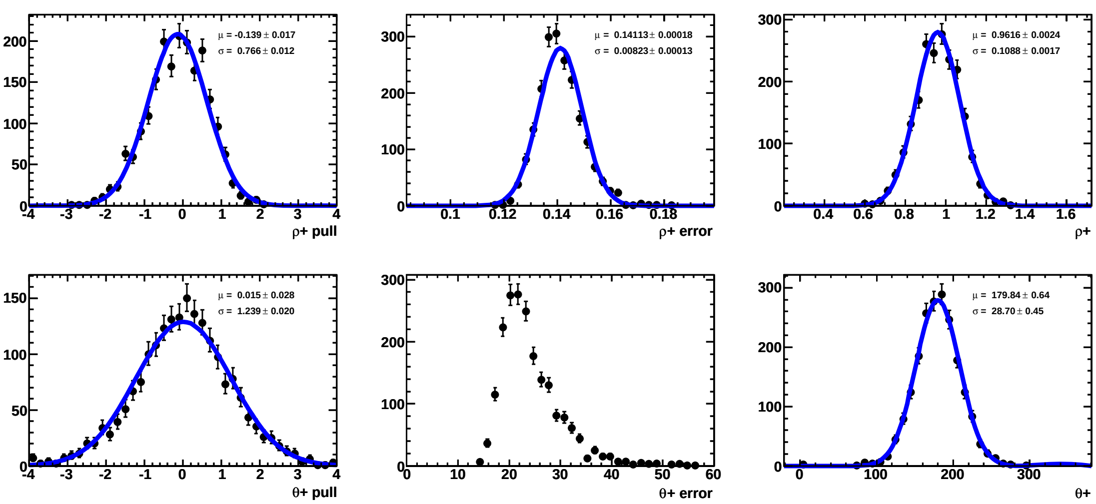


Figure 3.59: Results of 2000 toy MC experiments (signal and background, including systematic error in $\mathcal{L}_{\mathrm{BA}}$) for the pull (left), error (center) and fitted values (right) of ρ_+ (top) and θ_+ (bottom). See Fig. 3.56 for the corresponding toy MC without systematic errors.

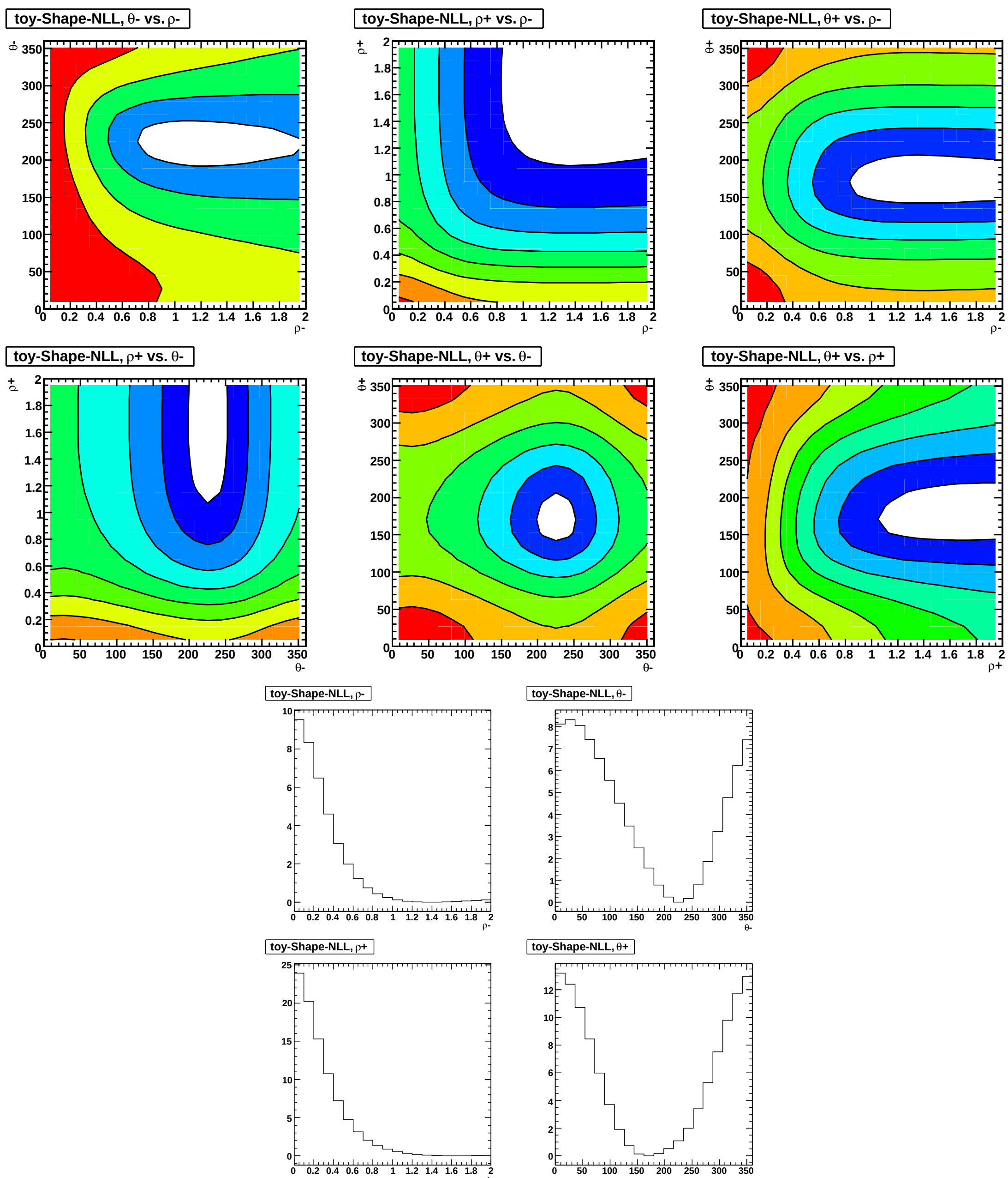


Figure 3.60: Dependence of $\mathcal{L}_{\rm DP}$ on all pair combinations of the parameters $\rho_\pm$, $\theta_\pm$ (2D plots) with 1σ contours and the individual parameters in a toy experiment (1D plots).

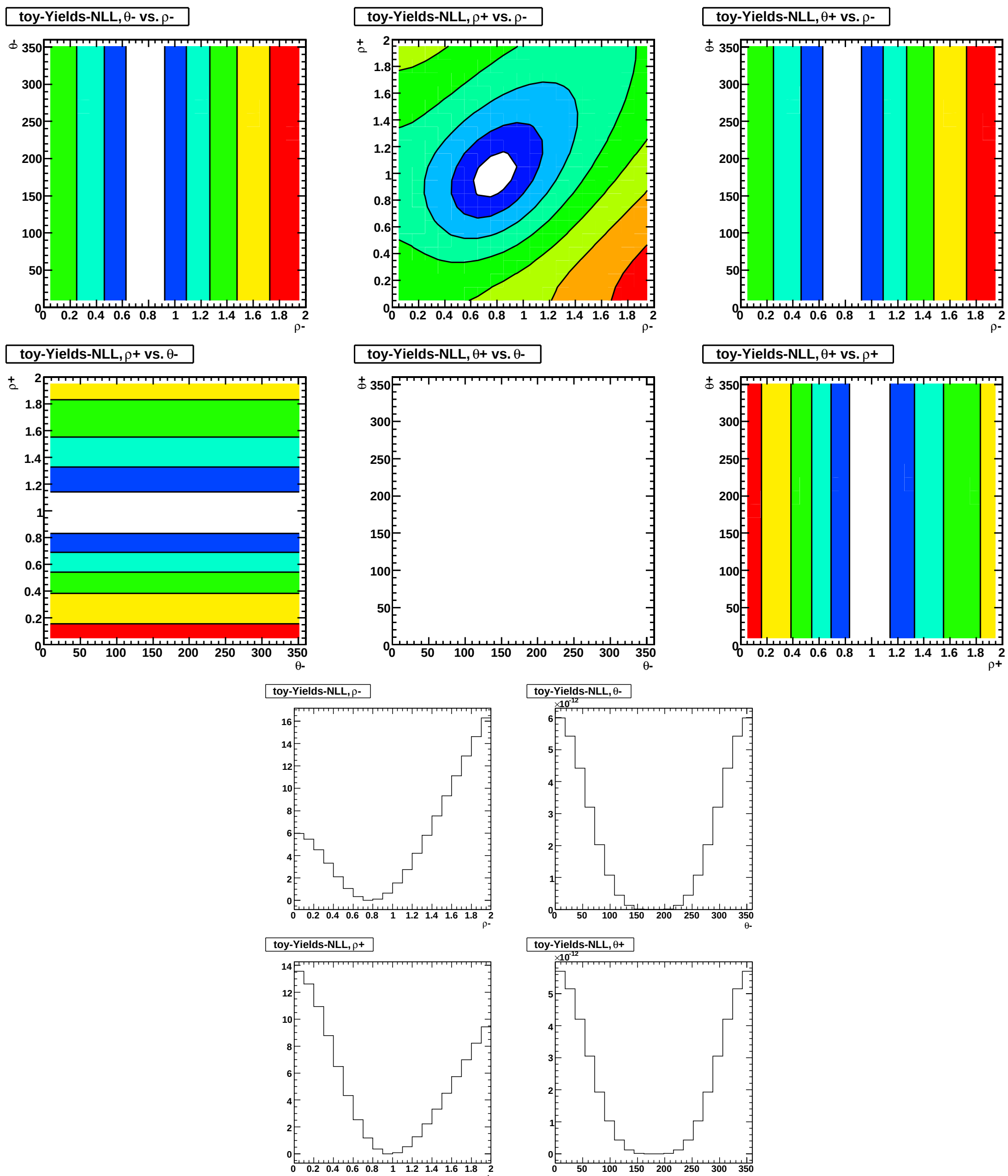


Figure 3.61: Dependence of $\mathcal{L}_{\mathrm{BA}}$ on all pair combinations of the parameters $\rho_{\pm}$, $\theta_{\pm}$ (2D plots) with 1σ contours and the individual parameters in a toy experiment (1D plots).

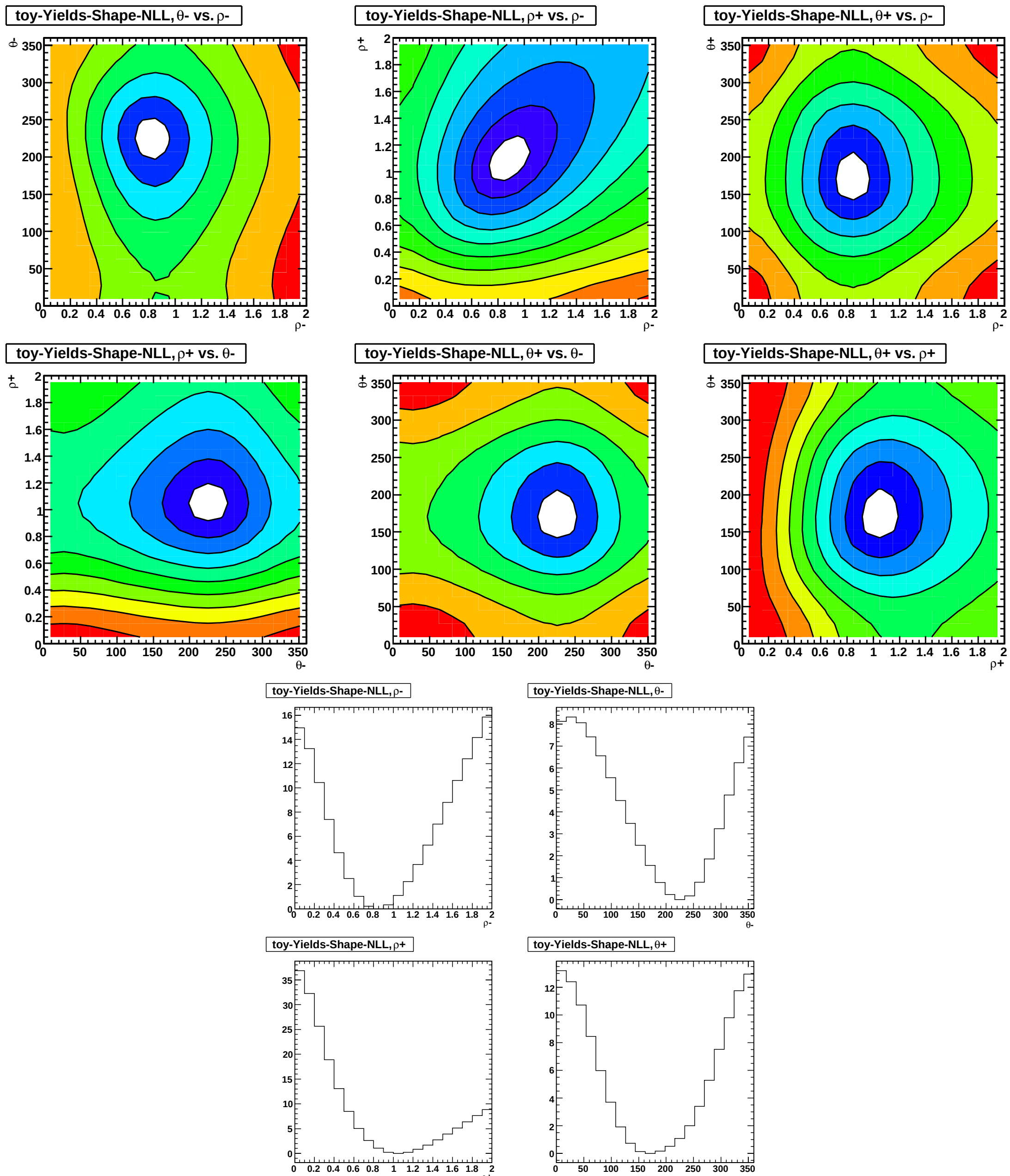


Figure 3.62: Dependence of $\mathcal{L} = \mathcal{L}_{\mathrm{DP}} + \mathcal{L}_{\mathrm{BA}}$ on all pair combinations of the parameters $\rho_{\pm}$, $\theta_{\pm}$ (2D plots) with 1σ contours and the individual parameters in a toy experiment (1D plots).

3.9.2 Selection bias and Dalitz plot variable measurement resolution

The resolution of the Dalitz variables due to reconstruction effects is calculated as the difference between the generated and reconstructed values of s_{+0} and s_{-0} for $DK_{\rm sig}$ events. Figure 3.63 shows the average resolution for s_{+0} (left) and s_{-0} (right). However, it should be noted that the resolutions for these two variables are in general correlated and depend on s^{tr}_{+0} and s^{tr}_{-0}, the true values of s_{+0} and s_{-0}.

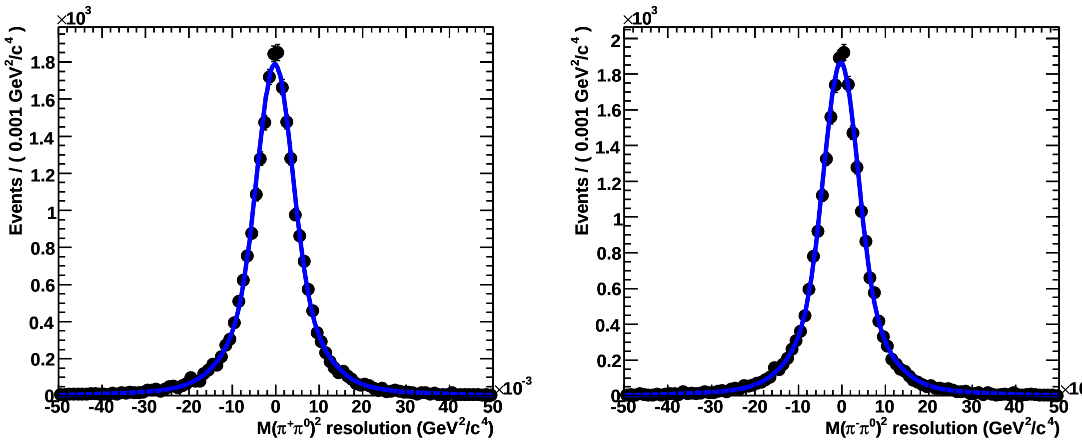


Figure 3.63: Resolution of the Dalitz variables s_{+0} (left) and s_{-0} (right) for $DK_{\rm sig}$ events. The superimposed curve shows a fit using the sum of three Gaussians with common mean.

We use the flat signal MC sample to check whether the finite resolution of the measurement of the Dalitz plot variables affects the CP fit. The sample consists of 41800 events that pass the selection cuts. For each event, we calculate the physical signal likelihood $\mathcal{D}^{\rm phys}_{DK_{\rm sig}}(s^{tr}_{+0}, s^{tr}_{-0})$, given values of x and y between -0.3 and 0.3 taken in 0.1-wide steps.[8] This likelihood is used to randomly reject events, so that the true Dalitz plot variables of the remaining ~ 3000 events (the actual number of events depends on the values of x and y) have the signal PDF distribution, up to effects due to the Dalitz plot dependence of the efficiency.

We first check for biases in the resulting "shaped" samples by fitting their *true* Dalitz variables using the signal PDF including the efficiency, $\mathcal{D}_{DK_{\rm sig}}(s^{tr}_{+0}, s^{tr}_{-0})$. The pull distri-

[8] This study was performed using Cartesian coordinates. The results still hold for the polar coodinates.

butions and differences between measured and generated values of the *CP* parameters are shown in Fig 3.64. Note that the different entries in each plot are highly correlated, as they are selected from the same flat sample and contain a large fraction of identical events. Nonetheless, these plots demonstrate that there is no significant bias in these fits, which contain more than ten times the number of signal events than we expect in the data.

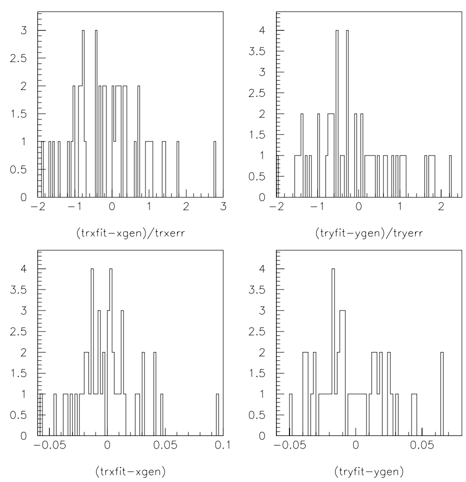


Figure 3.64: Top: Pull distributions for x (left) and y (right) when fitting the *true* Dalitz plot variables of "shaped" signal samples. Bottom: Differences between the generated and measured values of x (left) and y (right). The different entries in each plot, corresponding to different generated values of x and y, are highly correlated.

Next, we study the effect of the measurement resolution, by fitting the shaped samples' *measured* Dalitz variables, i.e., using the PDF $\mathcal{D}_{DK_{\rm sig}}(s_{+0}, s_{-0})$. The difference between each fit to the measured and true Dalitz variables is shown in Fig 3.65. Again, the entries in each plot are highly correlated. However, one can see that the difference between the fit to the true and the fit to the measured Dalitz variables is much smaller than the expected statistical error given our data.

From these plots we conclude that the effect of a selection bias on the measurement of x and y is consistent with the MC statistical error of order 0.02, and that Dalitz variable measurement resolution effects are of order 0.01.

As an additional test, we generate toy MC including a model for the Dalitz resolu-

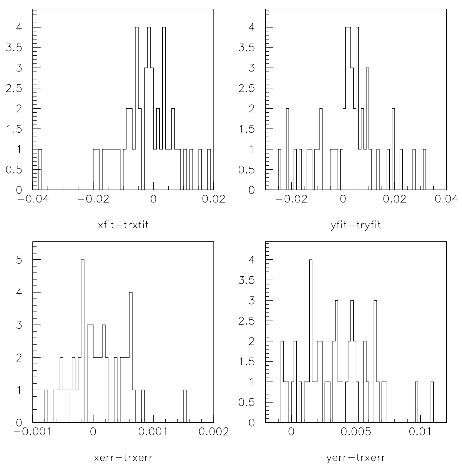


Figure 3.65: Top: Differences between the values of x (left) and y (right) when fitting the *measured* and the *true* Dalitz plot variables of “shaped” signal samples. Bottom: Differences between the fit errors on x (left) and y (right). The different entries in each plot, corresponding to different generated values of x and y, are highly correlated.

tion. For each toy MC experiment, the generated events are smeared over the Dalitz plot according to the resolution model of Fig. 3.63 and then fitted with the original (unsmeared) signal PDF. Fig. 3.66 shows the result of 5000 toy MC experiments (100 events per experiment, $x = y = 0.0$). The squares represent toy MC experiments without the Dalitz resolution model. The triangles show the result of the toy MC including the Dalitz resolution. We conclude that the resolution on the Dalitz variables does not have any measurable effect on either the pull nor the error distribution of x and y.

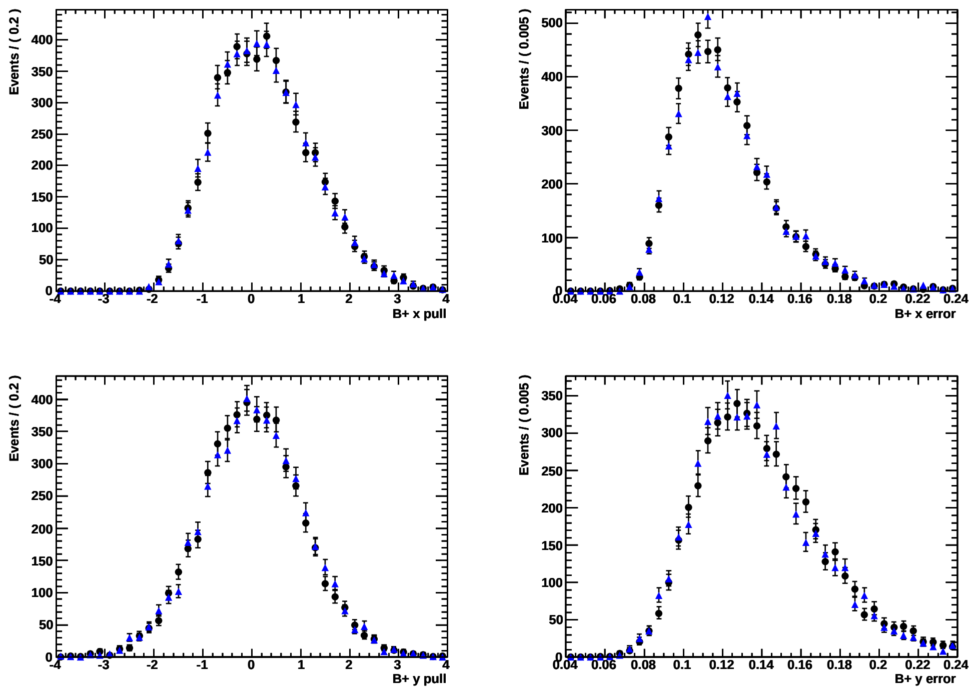


Figure 3.66: Comparison of 5000 toy MC experiments (100 B^+ events, $x, y = 0.0$) with (triangles) and without (squared) the Dalitz resolution model.

3.9.3 Full MC fit

In this section, we describe the fit performance on a MC cocktail consisting of appropriately weighted signal, $B\overline{B}$ generic and $q\overline{q}$ generic events. The generic B^+B^- and $B^0\overline{B}^0$ samples are about five times larger than the on-peak data luminosity whereas the continuum samples are only 1.5 times the on-peak data luminosity. The $DK_{\rm sig}$ signal events are obtained by re-weighting the phase-space signal MC sample according to the $\mathcal{D}_{DK_{\rm sig}}$ Dalitz-PDF. This ensures the correct Dalitz distributions for signal events. After re-weighting, the $DK_{\rm sig}$ sample is about 26 times larger than the expected number of signal events on data. $DK_{\rm bgd}$ events are obtained directly from the phase-space signal MC sample.

The first study consists of 26 fits, each using an independent signal sample. The qq_D and $qq_{\not D}$ events are randomly selected out of the continuum MC giving rise to a non-negligible overlap due to the small size of this sample. Finally, 5-6 of the 26 signal samples are assigned the same $B\overline{B}$ sample from which the other event types are extracted. Before each fit we count the true number of events for each type. $D \to \pi^+\pi^-\pi^0$ decays are incorrectly simulated in the generic MC by a incoherent sum of ρ-resonances. Therefore, we use $D\pi_D$ events generated from the $\mathcal{D}_{D\pi_D}$ PDF described in section 3.6.7 rather than events from the generic MC (BBC_D, qq_D and $D\pi_D$ are removed entirely due to their negligible yields).

Table 3.32 lists the true and fitted parameters with their error averaged over the 5-6 fits that share the same $B\overline{B}$ sample. A summary of the deviations of the fitted value from the true value in standard deviations can be found in Table 3.31. To separate statistical and systematic effects in the fit, we perform a second study making use of the entire $B\overline{B}$ generic MC sample. Due to its small size we do not use the continuum sample in this test. After scaling all event yields by a factor of 5.0 we are left with five independent signal samples. Table 3.33 lists the results averaged over the five fits. We do not observe any large biases and all fitted parameters are within about one standard deviation of their true values. The same study has been repeated without replacing $D\pi_D$ events in the generic

MC with comparable results.

Set	1	2	3	4	5
ρ_-	-0.3	-0.5	0.9	0.2	-1.1
θ_-	0.4	0.3	-0.4	-0.8	-1.0
ρ_+	-0.1	0.4	0.3	0.3	0.8
θ_-	-1.4	-0.8	1.7	-0.5	-0.2
$R_{D\pi X}$	1.5	-0.3	0.7	0.6	0.7
$N_{D\pi_D}$	0.8	0.7	0.8	0.5	-0.3
$A_{DK_{\text{sig}}}$	0.2	-0.5	0.9	0.5	-1.8
N_{qq_D}	0.4	0.4	0.1	-0.3	0.0
$N_{DK_{\text{sig}}}$	1.0	0.1	1.5	0.8	0.9
$N_{BB_{\not D}}$	-1.1	-0.7	-1.2	-0.3	-0.4

Table 3.31: Deviation of the fitted average value from the true average value in numbers of σ. $D\pi_D$ Dalitz distributions are replaced by data generated from $\mathcal{D}_{D\pi_D}$. Compare to the vales in Table 3.32.

	<true>	<fit> $\pm < \sigma >$	<true>	<fit> $\pm < \sigma >$	<true>	<fit> $\pm < \sigma >$
	Set 1		Set 2		Set 3	
ρ_-	0.779	0.733 ± 0.139	0.779	0.703 ± 0.146	0.779	0.910 ± 0.142
θ_-	175	191.9 ± 37.0	175	182.1 ± 26.9	175	164.0 ± 25.3
ρ_+	0.976	0.959 ± 0.131	0.976	1.038 ± 0.142	0.976	1.014 ± 0.135
θ_-	179	153.8 ± 18.2	179	158.9 ± 24.6	179	213.2 ± 20.5
$R_{D\pi X}$	0.576	0.835 ± 0.169	0.587	0.534 ± 0.162	0.548	0.672 ± 0.174
$N_{D\pi_D}$	82.0	100.7 ± 23.0	75.0	90.2 ± 22.4	77.0	94.7 ± 21.9
$A_{DK_{\rm sig}}$	-0.189	-0.163 ± 0.120	-0.170	-0.235 ± 0.135	-0.183	-0.080 ± 0.114
N_{qq_D}	2854	2883.6 ± 76.3	2854	2887.8 ± 76.3	2854	2863.1 ± 76.3
$N_{DK_{\rm sig}}$	202	232.7 ± 30.9	202	203.9 ± 29.8	202	252.0 ± 32.3
$N_{BB_{\not D}}$	1148	1057.8 ± 78.5	1156	1101.1 ± 79.4	1157	1062.0 ± 80.7
	Set 4		Set 5			
ρ_-	0.779	0.812 ± 0.141	0.779	0.637 ± 0.133		
θ_-	175	154.7 ± 27.4	175	152.6 ± 23.3		
ρ_+	0.976	1.015 ± 0.137	0.976	1.086 ± 0.137		
θ_-	179	166.1 ± 23.2	179	174.7 ± 20.3		
$R_{D\pi X}$	0.570	0.673 ± 0.166	0.603	0.727 ± 0.169		
$N_{D\pi_D}$	67.0	78.4 ± 21.1	74.0	67.6 ± 21.7		
$A_{DK_{\rm sig}}$	-0.187	-0.129 ± 0.121	-0.159	-0.373 ± 0.122		
N_{qq_D}	2854	2831.4 ± 76.2	2854	2857.1 ± 76.6		
$N_{DK_{\rm sig}}$	202	226.4 ± 30.7	202	230.7 ± 31.5		
$N_{BB_{\not D}}$	1164	1139.7 ± 81.0	1155	1123.5 ± 82.3		

Table 3.32: Results of the full MC fit. Each of the five sets has an independent signal and $B\overline{B}$ sample but all share the same $q\overline{q}$ sample. The numbers are averages over three or four fits. The deviations from the true values in number of σ are listed in Table 3.31.

	<true>	<fit> $\pm < \sigma >$	$\Delta(\sigma)$
ρ_-	0.779	0.8095 ± 0.0699	0.4
θ_-	175	173.27 ± 9.67	-0.2
ρ_+	0.976	1.0026 ± 0.0760	0.3
θ_-	179	172.30 ± 8.68	-0.8
$R_{D\pi X}$	0.578	0.5838 ± 0.0481	0.1
$N_{D\pi_D}$	379	389.8 ± 39.0	0.3
$A_{DK_{\rm sig}}$	-0.167	-0.1637 ± 0.0501	0.1
$N_{DK_{\rm sig}}$	1011	1050.9 ± 57.6	0.7
$N_{BB_{\not D}}$	5804	5739 ± 104	-0.6

Table 3.33: Average over five fits using the full generic $B\overline{B}$ MC sample without continuum events. Each fit has an independent signal sample but all use the same $B\overline{B}$ sample. The last column (Δ) shows the difference of the fitted value to the true value in numbers of σ.

3.10 Data fit

The results of the step-1 fit to obtain the yields and the signal asymmetry are given in Table 3.34. Table 3.35 summarizes the results of the step-2 fit in two configurations: using only the shape of the Dalitz plot (using the NLL $\mathcal{L}_{\mathrm{DP}}$), and using both the shapes and the signal yield and asymmetry (using the NLL $\mathcal{L} = \mathcal{L}_{\mathrm{DP}} + \mathcal{L}_{\mathrm{BA}}$). It can be seen that the $\mathcal{L}$-fit is significantly more sensitive to $\rho_{\pm}$.

The correlation matrices of the two fits are shown in Table 3.36. The relatively large correlations between ρ_- and ρ_+ are due to the systematic errors in $\mathcal{L}_{\mathrm{BA}}$.

Projections of the data and the PDF onto the fit variables are shown in Figs. 3.67 through 3.68. The dependences of the NLL's on the *CP* parameters are shown in Fig. 3.69 for $\mathcal{L}_{\mathrm{DP}}$, Fig. 3.70 for $\mathcal{L}_{\mathrm{BA}}$and Fig. 3.71 for $\mathcal{L}$.

Parameter	Value	Correlation matrix $R_{D\pi X}$	$N_{D\pi_D}$	$A_{DK_{\mathrm{sig}}}$	$N_{qq\not{D}}$	$N_{DK_{\mathrm{sig}}}$	$N_{BB\not{D}}$
$R_{D\pi X}$	0.53 ± 0.15	1	0.223	-0.037	0.096	0.478	-0.388
$N_{D\pi_D}$	57.2 ± 19.6	0.223	1	-0.028	0.019	-0.092	-0.257
$A_{DK_{\mathrm{sig}}}$	-0.024 ± 0.148	-0.037	-0.028	1	-0.006	-0.038	0.032
$N_{qq\not{D}}$	2383 ± 71	0.096	0.019	-0.006	1	0.090	-0.544
$N_{DK_{\mathrm{sig}}}$	170.1 ± 29.0	0.478	-0.092	-0.038	0.090	1	-0.452
$N_{BB\not{D}}$	1138 ± 76	-0.388	-0.257	0.032	-0.544	-0.452	1
$-\log\mathcal{L}_{\min}$	-20505.9						

Table 3.34: Results of the step-1 fit on the run 1-5 data sample.

	$\mathcal{L}_{\mathrm{DP}}$ fit			$\mathcal{L}_{\mathrm{DP}}$+$\mathcal{L}_{\mathrm{BA}}$ fit		
Parameter	Value	−MINOS	+MINOS	Value	−MINOS	+MINOS
ρ_-	0.968 ± 0.557	-0.368	$+1.24$	0.804 ± 0.148	-0.140	$+0.159$
θ_-	174.5 ± 46.3	-39.9	$+58.1$	173.1 ± 43.3	-37.8	$+55.0$
ρ_+	0.919 ± 0.359	-0.274	$+0.540$	0.833 ± 0.145	-0.138	$+0.155$
θ_+	147.3 ± 23.9	-23.9	$+24.8$	147.2 ± 23.3	-23.2	$+23.8$
$-\log\mathcal{L}_{\min}$	-19068.9			-19068.8		

Table 3.35: Results of the step-2 fit on the run 1-5 data sample. The first fit is done with only the Dalitz shape NLL $\mathcal{L}_{\mathrm{DP}}$, and the second incorporates also the signal yield and asymmetry.

Parameter	$\mathcal{L}_{\rm DP}$ fit					$\mathcal{L}_{\rm DP}+\mathcal{L}_{\rm BA}$ fit				
	Global	ρ_-	θ_-	ρ_+	θ_+	Global	ρ_-	θ_-	ρ_+	θ_+
ρ_-	0.137	1	0.137	0.000	0.000	0.400	1	0.020	0.400	0.005
θ_-	0.137	0.137	1	0.000	0.000	0.020	0.020	1	0.008	0.000
ρ_+	0.028	0.000	0.000	1	0.028	0.400	0.400	0.008	1	0.013
θ_+	0.028	0.000	0.000	0.028	1	0.013	0.005	0.000	0.013	1

Table 3.36: Correlation matrix of the data fit done with the NLLs $\mathcal{L}_{\rm DP}$(left) and $\mathcal{L} = \mathcal{L}_{\rm DP} + \mathcal{L}_{\rm BA}$ (right).

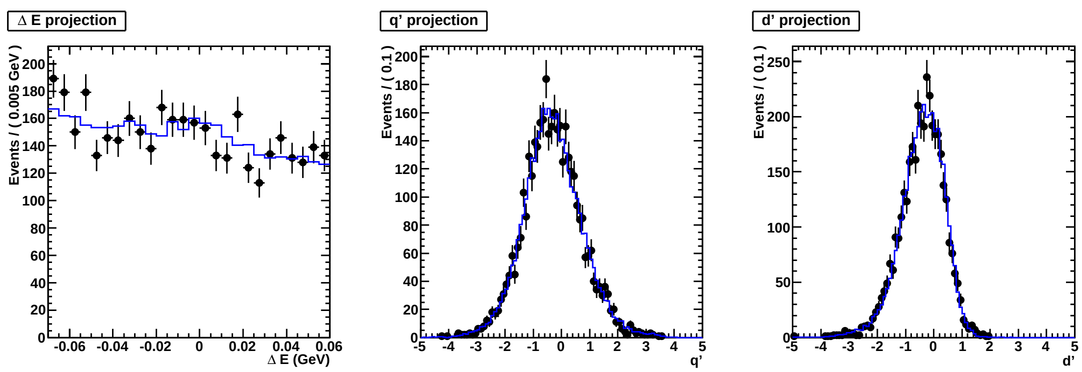


Figure 3.67: ΔE (left), q' (center) and d' projection (right) of the data fit. The solid line shows the fitted PDF and the dots show the data.

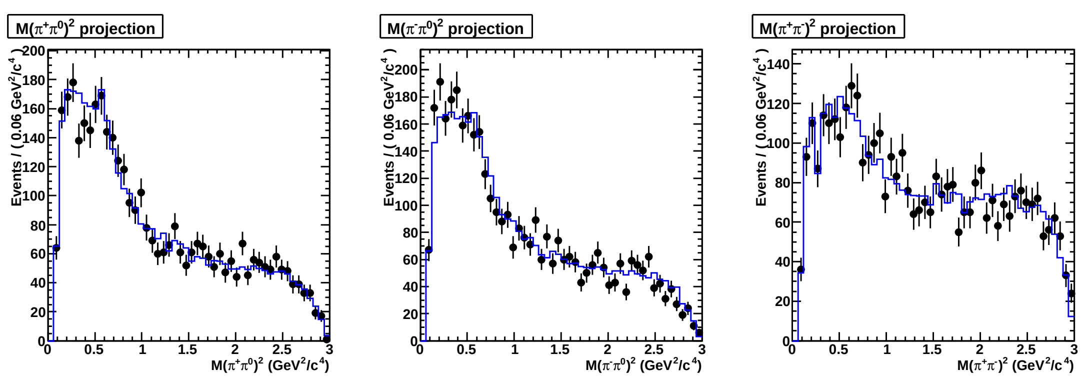


Figure 3.68: s_{+0} (left), s_{-0} (center) and s_{+-} projection (right) of the data fit. The solid line shows the fitted PDF and the dots show the data.

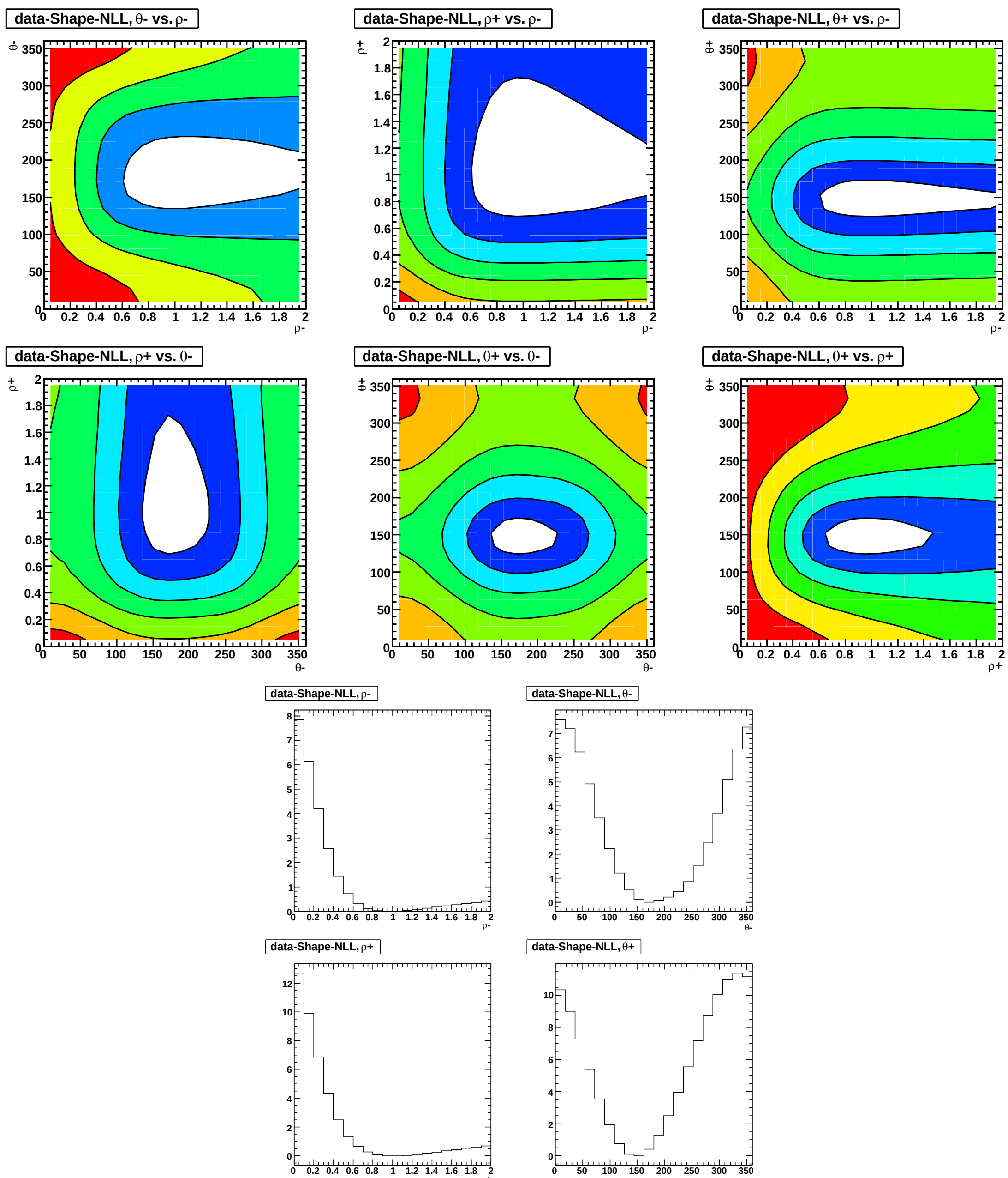


Figure 3.69: Dependence of $\mathcal{L}_{\mathrm{DP}}$ on all pair combinations of the parameters $\rho_{\pm}$, $\theta_{\pm}$ (2D plots) with 1σ contours and the individual parameters (1D plots) for the data.

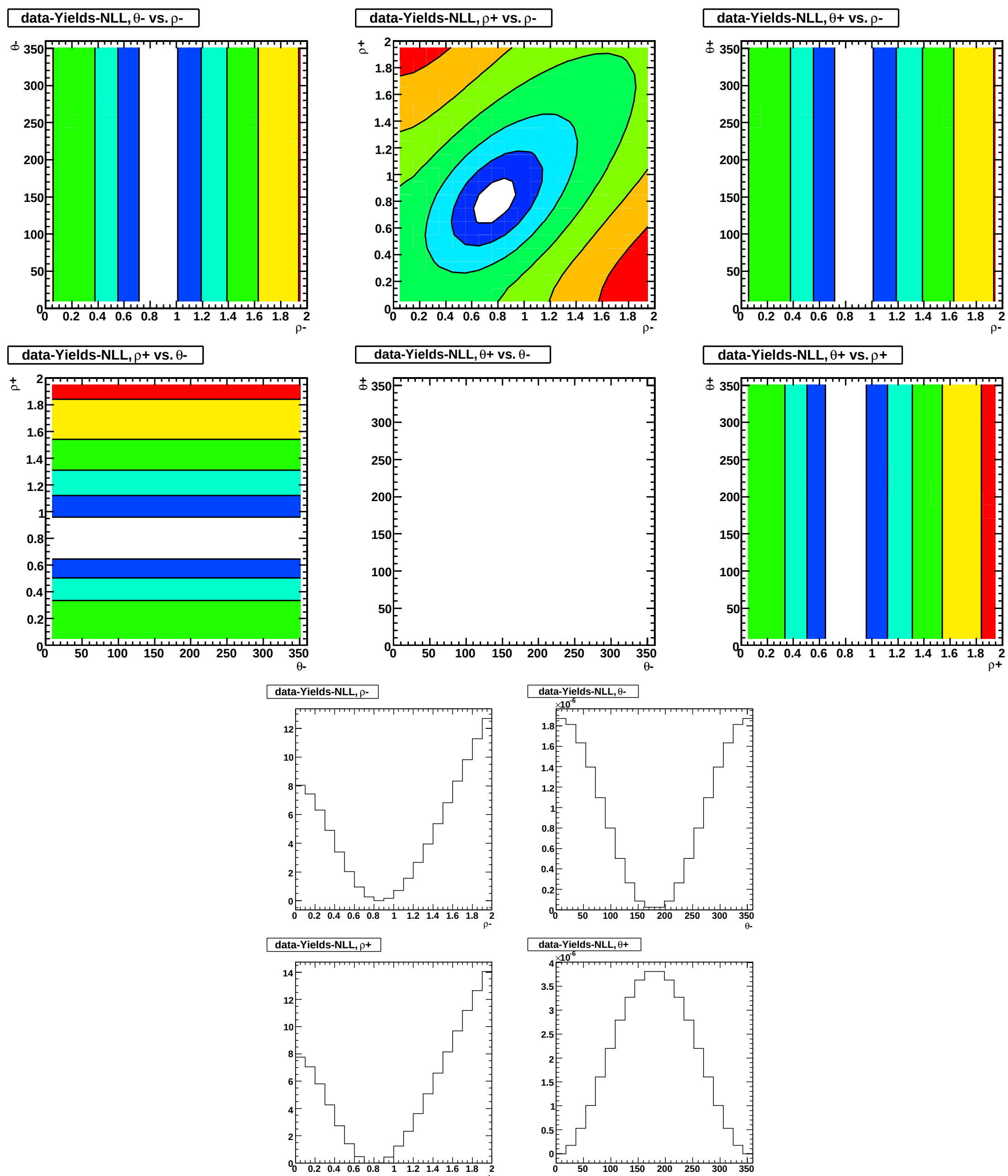


Figure 3.70: Dependence of $\mathcal{L}_{\mathrm{BA}}$ on all pair combinations of the parameters $\rho_{\pm}$, $\theta_{\pm}$ (2D plots) with 1σ contours and the individual parameters (1D plots) for the data.

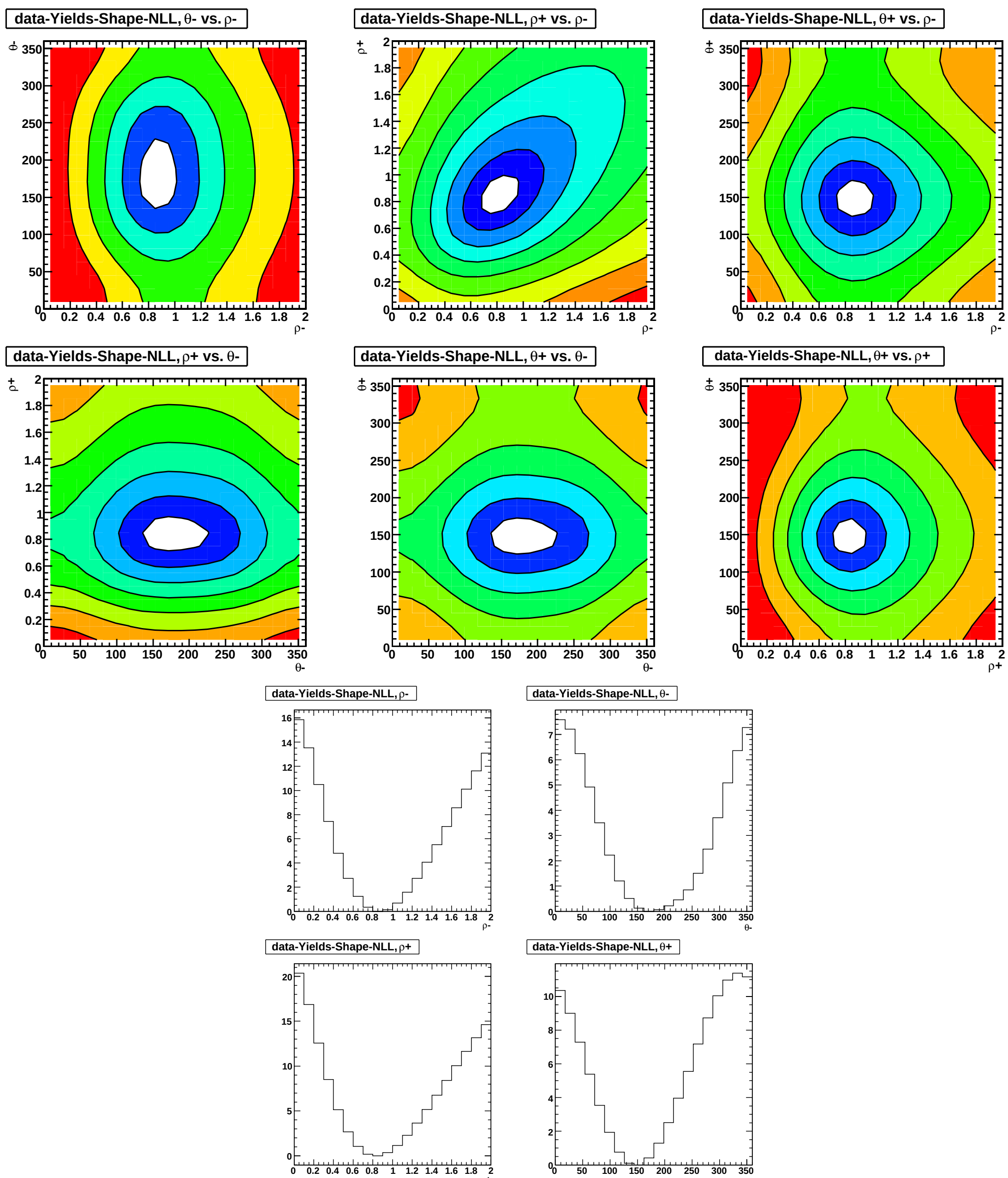


Figure 3.71: Dependence of $\mathcal{L} = \mathcal{L}_{\rm DP} + \mathcal{L}_{\rm BA}$ on all pair combinations of the parameters $\rho_\pm$, $\theta_\pm$ (2D plots) with 1σ contours and the individual parameters (1D plots) for the data.

3.11 Systematic errors

Table 3.37 lists the systematic uncertainties to be added to the error from the data fit. Some of the systematic uncertainties in this analysis are already incorporated in the main fit through the $\mathcal{L}_{\mathrm{BA}}$ term in the likelihood fit. In order to separate these errors from the statistical error, Table 3.38 lists the individual contributions to the total error of the fit. Finally, Table 3.39 summarizes the systematic uncertainties on the signal branching fraction. More details can be found in the following sections.

Source	$A_{DK_{\mathrm{sig}}}$	$N_{DK_{\mathrm{sig}}}$	ρ_-	θ_-	ρ_+	θ_+	Section
Fixed fractions	0.0023	7.93	0.0214	1.38	0.0190	1.40	3.11.1
MC statistics	0.0058	8.55	0.0173	10.48	0.0171	3.18	3.11.2
Sig. Dalitz model	n/a	n/a	0.05	10	0.05	10	3.11.3
DK_{bgd} Dalitz shape	n/a	n/a	0.0007	1.09	0.0007	0.48	3.11.4
Bgd. Dalitz shapes	n/a	n/a	0.004	3.51	0.004	1.00	3.11.5
Bkd. shapes of ΔE, q', d'	0.0020	5.18	0.0127	2.69	0.0121	1.26	3.11.6
Asym. in DK_{bgd}, DKX	0.0083	0.23	0.0046	0.10	0.0034	0.12	3.11.7
Detector charge asym.	0.0251	0.09	0.0121	0.60	0.0107	0.03	3.11.8
Kaon charge in qq_D	n/a	n/a	0.0031	0.94	0.0033	1.04	3.11.9
PID efficiency	n/a	n/a	0.0023	0.19	0.0023	0.32	3.11.10
D^* Bgd. shape	n/a	n/a	0.0087	2.23	0.0085	1.61	3.11.11
Dalitz variable resolution	n/a	n/a	0.01	0.7	0.01	0.7	3.11.12
Total	0.0272	12.76	0.0616	15.47	0.0603	10.92	

Table 3.37: Summary of *additive* systematic uncertainties in addition to the errors used in $\mathcal{L}_{\mathrm{BA}}$.

Source	ρ_-	θ_-	ρ_+	θ_+	Section
$\mathcal{B}(B^- \to D^0 K^-)$	0.0804	3.88	0.0774	2.13	3.11.16
$\mathcal{B}(D^0 \to K^-\pi^+\pi^0)$	0.0368	1.62	0.0355	0.98	3.11.16
$\frac{\mathcal{B}(D^0\to\pi^+\pi^-\pi^0)}{\mathcal{B}(D^0\to K^-\pi^+\pi^0)}$	0.0056	0.03	0.0053	0.02	3.11.16
Signal efficiency	0.0141	0.08	0.0135	0.06	3.11.16
$N_{B\overline{B}}$	0.0046	0.03	0.0044	0.02	3.11.16
Total	0.0898	4.21	0.0865	2.35	

Table 3.38: Summary of *subtractive* systematic uncertainties. These errors are already incorporated in the main fit through the $\mathcal{L}_{\mathrm{BA}}$ term of the likelihood.

Source	BF error (%)	Section
PID efficiency	3.1	3.11.10
π^0 efficiency	3.0	3.11.13
Tracking efficiency	1.5	3.11.14
B counting	1.1	3.11.15
Total	4.70	

Table 3.39: Systematic errors on the signal branching fraction.

3.11.1 Variation of fixed event yield fractions

The five event yield fractions that are fixed in the fits are varied, and the two-fit analysis is repeated to evaluate the resulting systematic error. The fractions are varied conservatively by $\pm 50\%$, except for R_{DKX}, which is varied by its PDG uncertainty of $\pm 25\%$. For each pair of fits we take

$$\sigma_\alpha^2 = \frac{(\alpha^+ - \alpha)^2 + (\alpha^- - \alpha)^2}{2} \tag{3.40}$$

as the systematic uncertainty for parameter $\alpha \in \{A_{DK_{\rm sig}}, N_{DK_{\rm sig}}, \rho_-, \theta_-, \rho_+, \theta_+\}$ where α is the nominal fit value and $\alpha^\pm$ is the fit value with one of the fixed fractions varied by the above amount. The resulting uncertainties in the parameters of interest are shown in Table 3.40.

Source	$A_{DK_{\rm sig}}$	$N_{DK_{\rm sig}}$	ρ_-	θ_-	ρ_+	θ_+
R_{qq_D}	0.0006	1.97	0.0056	0.38	0.0062	0.76
R_{BBC_D}	0.0006	0.58	0.0034	0.61	0.0027	0.36
$R_{D\pi\not{p}}$	0.0009	1.09	0.0025	0.43	0.0016	0.29
$R_{DK_{\rm bgd}}$	0.0005	7.56	0.0196	0.94	0.0175	1.06
R_{DKX}	0.0019	0.53	0.0048	0.57	0.0027	0.17
Total	0.0023	7.93	0.0214	1.38	0.0190	1.40

Table 3.40: Systematic uncertainties due to variation in the fixed fit fractions.

3.11.2 MC statistics

For all PDFs that are obtained by fitting MC samples the following method is used to evaluate the systematic uncertainty due to finite MC statistics. The MC fit for a PDF

β with N^β floating parameters results in a N^β-dimensional parameter vector p^β and a $N^\beta \times N^\beta$ error matrix V^β. If A is the (orthogonal) matrix of eigenvectors of V^β, we obtain the parameter vector and the error matrix in the diagonal basis as $q = A^T p^\beta$ and $W = A^T V^\beta A$. In this basis, we vary q_i by $\pm\sqrt{W_{ii}}$ and transform it back into the original basis. We repeat the data fit and obtain new values $\alpha_j^\pm$ for each parameter of interest $\alpha \in \{A_{DK_{\rm sig}}, N_{DK_{\rm sig}}, \rho_-, \theta_-, \rho_+, \theta_+\}$. This procedure is repeated for all parameters p_i^β and for all PDFs β resulting in $2N = 2\sum_\beta N^\beta = 338$ data fit results. The total systematic uncertainty due to MC statistics is

$$\sigma^2_{\rm MC\ stat} = \frac{1}{2}\sum_{j=1}^{N}\left[\left(\alpha_j^+ - \alpha\right)^2 + \left(\alpha_j^- - \alpha\right)^2\right]. \tag{3.41}$$

3.11.3 $D^0 \to \pi^+\pi^-\pi^0$ Dalitz model

To evaluate the chages in the CP parameters due to the signal Dalitz model we repeat the data fit with different $D^0 \to \pi^+\pi^-\pi^0$ Dalitz models. Table 3.41 lists the differences to the nominal fit result for different Dalitz models with increasing number of components. In addition we show the difference to the nominal fit result for a fit with the meson radial parameter R of the Blatt-Weisskopf penetration factor set to zero.

Based on these results and considering that the first two models in Table 3.41 are too simplistic and hence unrealistic, we assign a systematic error of 0.05 for $\rho_\pm$ and 10° for $\theta_\pm$.

3.11.4 $DK_{\rm bgd}$ Dalitz shape

The $DK_{\rm bgd}$ Dalitz shape is obtained from high statistics MC with the CLEO parameters. To estimate the uncertainty on the CP parameters, we repeat the data fit with a shape obtained from signal toy MC generated according to the D decay parameters we find and the $K_S\pi^+\pi^-$ values for $\rho_\pm$ and $\theta_\pm$. We then smear the generated MC with the residuals between the true and reconstructed s_{+0} and s_{+0} values found in the phase space signal

Dalitz model	ρ_-	θ_-	ρ_+	θ_+
NR_S, $\rho(770)$	0.0862	16.42	0.0595	−6.82
+ $f_0(980)$	0.0760	21.14	0.0453	5.01
+ $\rho(1450)$	−0.0071	6.27	−0.0215	−7.89
+ $\rho(1700)$	0.0307	3.21	0.0100	−9.91
+ $f_0(1370, 1500, 1710)$, $f_2(1270)$	−0.0434	−9.72	−0.0463	−1.40
+ σ	0	0	0	0
+ NR_P	0.0163	0.06	0.0142	−1.22
+ ω, $f_2'(1525)$	0.0144	3.15	0.0128	−1.82
$R = 0$	0.0014	7.97	0.0007	0.01

Table 3.41: Differences to the nominal fit result for different $D \to \pi^+\pi^-\pi^0$ Dalitz models and with the meson radial parameter R set to zero.

MC simulation. A new shape for the DK_{bgd} events is obtained from this sample and the data fit is repeated. The differences between the results of this fit and the nominal fit are taken as the systematic error.

3.11.5 Uncertainties in simulation of background Dalitz plot shape

We obtain a histogram of the Dalitz distributions on both the MC and data in the m_{ES} sideband. The ratio of the normalized histograms in each bin gives a first-order estimate of the data/MC agreement. In section 3.5.2 we showed that these distributions are in good statistical agreement. Nonetheless, we evaluate the error due to the fact that the agreement is not perfect.

To do this, we apply the full data-MC difference to the Dalitz PDF of the $D\pi X$ background and repeat the data fit. The procedure is repeated with the data-MC difference assigned to the DK_{bgd} PDF. Table 3.42 lists the differences to the nominal data fit. We take the set with the larger differences ($D\pi X$) as the systematic error due to the background Dalitz plot shape.

3.11.6 ΔE, q' and d' shapes

The systematic uncertainties due to MC-data differences in the one-dimensional PDF shapes of ΔE, q' and d' are evaluated as follows. For each of the three PDFs β and each

Source	ρ_-	θ_-	ρ_+	θ_+
Diff. applied to $D\pi X$	0.004	-3.51	0.004	-1.00
Diff. applied to $DK_{\rm bgd}$	-0.0001	-0.55	-0.0002	0.09

Table 3.42: Systematic errors due to the uncertainties in the simulation of background Dalitz plot shape. The data/MC differences observed are applied separately to the $D\pi X$ and $DK_{\rm bgd}$ Dalitz background shape.

event type t, we obtain the PDF parameter vector p_t^β in the m_{ES} sideband on MC. These parameters are used to fit the m_{ES} sideband on data, in which we allow the parameters $p_{qq\not{p}}^\beta$ of the dominant background $qq\not{p}$ to float. The difference $q_{qq\not{p}}^\beta - p_{qq\not{p}}^\beta$ between the new parameters $q_{qq\not{p}}^\beta$ and the original MC sideband parameters is applied to our nominal fit parameters and we repeat the nominal data fit. In other words, we assign the entire MC-data difference to the dominating background type. The changes in the analysis variables with respect to the nominal fit are taken as systematic uncertainty and listed in Table 3.43.

Source	$A_{DK_{\rm sig}}$	$N_{DK_{\rm sig}}$	ρ_-	θ_-	ρ_+	θ_+
Bgd. ΔE shapes	0.0016	3.26	0.0048	1.59	0.0062	0.17
Bgd. d' shapes	0.0011	3.93	0.0115	1.01	0.0103	0.58
Bgd. q' shapes	0.0003	0.88	0.0022	1.92	0.0015	1.11
Total	0.0020	5.18	0.0127	2.69	0.0121	1.26

Table 3.43: Differences to the nominal fit and resulting systematic errors due to the uncertainty in the ΔE, q' and d' shapes

3.11.7 Possible asymmetry in $DK_{\rm bgd}$ and DKX

The nominal fit was performed with zero asymmetry for the $DK_{\rm bgd}$ and DKX event types. Since these are $B \to DK$ decays they can have a Standard Model asymmetry. As in the branching ratio analysis [61], we assume a possible asymmetry in DKX of $A_{DKX} \lesssim 0.022$ and repeat the fit with $A_{DKX} = \pm 0.022$. To evaluate the error due to a possible asymmetry in $DK_{\rm bgd}$ we set $A_{DK_{\rm bgd}}$ equal to $A_{DK_{\rm sig}}$ in the fit under the assumption that nothing in the reconstruction can change the charge asymmetry of signal events. Table 3.44 summarizes the resulting systematic errors.

Source	$A_{DK_{\rm sig}}$	$N_{DK_{\rm sig}}$	ρ_-	θ_-	ρ_+	θ_+
$A_{DKX} = +0.022$	0.0043	−0.01	0.0025	−0.09	−0.0018	0.07
$A_{DKX} = -0.022$	−0.0044	0.23	−0.0021	0.03	0.0022	−0.08
$A_{DK_{\rm bgd}} = A_{DK_{\rm sig}}$	−0.0055	0.01	−0.0032	0.03	0.0019	0.05
Total	0.0083	0.23	0.0046	0.10	0.0034	0.12

Table 3.44: Differences to the nominal fit and resulting systematic errors due to asymmetries in DKX and $DK_{\rm bgd}$ events.

3.11.8 Global detector charge asymmetry

A possible charge asymmetry in the detection efficiency for K^+ versus K^- is evaluated by floating a global charge asymmetry, which affects all event types equally, in the fit. The fitted value for the global asymmetry is -0.007 ± 0.018. We take the difference in the analysis variables to the nominal fit as a systematic error.

3.11.9 Kaon charge correlation in qq_D.

The Cabibbo allowed decay $\overline{D}^0 \to K^+ + X$ can introduce a wrong sign charge correlation in the qq_D background events. The systematic effect is expected to be small since qq_D events contribute less than 1% to the total. Nevertheless, we evaluate the systematic error due to a wrong sign kaon by assuming a 100% correlation between a wrong sign kaon and qq_D events. The difference to the nominal fit is taken as a systematic error.

3.11.10 PID efficiency

Since the corrections due to PID are very small (see Table 3.4), we conservatively take the magnitude of the correction (3.1%) as systematic error to the signal branching fraction. In addition, we repeat the data fit without the PID corrections on the Dalitz efficiency function and take the almost negligible difference to the nominal fit as the systematic error. As a cross-check, we repeat the fit with a flat efficiency function. The result of this fit are within the errors of the "no-PID" fit and we conclude that the efficiency function does not have any appreciable effect on the CP parameters.

3.11.11 Uncertainty due to the D^* background shape

Instead of taking the background Dalitz plot PDF from the data sideband as described in section we repeat the D Dalitz fit with the background PDF extracted from the MC signal region and obtain a new set of Dalitz parameters. With these we repeat the *CP*-fit and take the difference to the nominal fit as the systematic error.

3.11.12 Finite Dalitz variable resolution

Based on the studies of section 3.9.2 we assign an error of 0.01 to the Cartesian *CP* parameters. This translates into an error of 0.01 on $\rho_\pm$ and 0.7° on $\theta_\pm$.

3.11.13 π^0 efficiency

We use the ratio of data and MC efficiencies for π^0 mesons from the official recipe to be 0.968311 and assign a 3.0% systematic error.

3.11.14 Tracking efficiency

No efficiency corrections are needed for tracking in release 18. We assign a systematic error of 0.5% per track resulting in 1.5% total systematic uncertainty due to the three tracks in our decay mode.

3.11.15 B counting

The systematic error on the number of $B\overline{B}$ pairs is 1.1%.

3.11.16 Subtractive systematic errors

To separate the systematic errors included in $\mathcal{L}_{\mathrm{BA}}$ from the total error returned by the fit, we determine the individual contributions separately. The two largest uncertainties are from the secondary branching fractions [68]

$$
\begin{aligned}
\mathcal{B}(B^- \to D^0 K^-) &= (3.7 \pm 0.6) \times 10^{-4} \\
\mathcal{B}(D^0 \to K^- \pi^+ \pi^0) &= (13.2 \pm 1.0) \times 10^{-2}.
\end{aligned}
$$

We repeat the data fit without including the error on $\mathcal{B}(B^- \to D^0 K^-)$ in the error matrix of $\mathcal{L}_{\mathrm{BA}}$ and obtain a new (and smaller) value for the error $\tilde{\sigma}_\alpha$ of the floating parameter $\alpha \in \{\rho_\pm, \theta_\pm\}$. The systematic error due to $\mathcal{B}(B^- \to D^0 K^-)$ is calculated from $\sigma^2 = \sigma_\alpha^2 - \tilde{\sigma}_\alpha^2$ where σ_α is the error from the nominal data fit. The same procedure is repeated for $\mathcal{B}(D^0 \to K^- \pi^+ \pi^0)$. The relative error on the ratio of branching fractions from [69]

$$
\frac{\mathcal{B}(D^0 \to \pi^+ \pi^- \pi^0)}{\mathcal{B}(D^0 \to K^- \pi^+ \pi^0)} = (10.59 \pm 0.06 \pm 0.13) \times 10^{-2} \tag{3.42}
$$

is 1.4%. This small error cannot be reliably determined by the above method. Instead, we change the nominal values in $\mathcal{L}_{\mathrm{BA}}$ by $\pm 1\sigma$, repeat the data fit, and take the average squared difference as the systematic uncertainty for $\rho_\pm$ and $\theta_\pm$, respectively. The same is repeated for the uncertainties on $N_{B\overline{B}}$ (section 3.11.15) and combined error on the efficiency due to tracking and π^0 corrections (section 3.11.13 and 3.11.14).

3.12 Physics results

3.12.1 Branching fraction and asymmetry measurement

The branching fraction is calculated from

$$\mathcal{B}(B^{\pm} \to (\pi^{+}\pi^{-}\pi^{0})_{D}K^{\pm}) = \frac{N_{DK_{\rm sig}}}{N_{B\overline{B}}\,\epsilon}. \tag{3.43}$$

Using an absolute efficiency of $\epsilon = 11.41\%$ and $N_{B\overline{B}} = 324,041,437,$ we measure

$$\mathcal{B}(B^{\pm} \to (\pi^{+}\pi^{-}\pi^{0})_{D}K^{\pm}) = (4.6 \pm 0.8\ (\text{stat.}) \pm 0.7\ (\text{syst.})) \times 10^{-6} \tag{3.44}$$

and a decay rate asymmetry

$$A(B^{\pm} \to (\pi^{+}\pi^{-}\pi^{0})_{D}K^{\pm}) = -0.02 \pm 0.15\ (\text{stat.}) \pm 0.03\ (\text{syst.})\,, \tag{3.45}$$

where the asymmetry is defined as $A = \frac{N^{-}-N^{+}}{N^{-}+N^{+}}$ with $N^{\pm}$ as the number of $DK_{\rm sig}$ signal events from $B^{\pm}$ decays.

3.12.2 *CP* parameter measurement

The *CP* parameters measured in $B^{\pm} \to (\pi^{+}\pi^{-}\pi^{0})_{D}K^{\pm}$ are

$$\begin{aligned}
\rho_{-} &= 0.804 \pm 0.118\ (\text{stat.}) \pm 0.109\ (\text{syst.}) \\
\theta_{-} &= 173.1^{\circ} \pm 43.1^{\circ}\ (\text{stat.}) \pm 16.0^{\circ}\ (\text{syst.}) \\
\rho_{+} &= 0.833 \pm 0.116\ (\text{stat.}) \pm 0.105\ (\text{syst.}) \\
\theta_{+} &= 147.2^{\circ} \pm 23.2^{\circ}\ (\text{stat.}) \pm 11.2^{\circ}\ (\text{syst.})\,.
\end{aligned} \tag{3.46}$$

The statistical error is the error from the fit reduced by the systematic error from Table 3.38. The systematic error is the combined error of Table 3.37 and Table 3.38.

The polar coordinates $\rho_{\pm}$ and $\theta_{\pm}$ are defined with respect to the Cartesian and physical parameters $z_\pm = x_\pm + iy_\pm = r_B e^{i(\delta\pm\gamma)}$ as

$$
\begin{aligned}
{\rho_\pm}^2 &= (x_\pm - x^0)^2 + y_\pm^2 \\
&= (r_B\cos(\delta\pm\gamma) - x^0)^2 + {r_B}^2\sin^2(\delta\pm\gamma) \\
\tan\theta_\pm &= \frac{y_\pm}{x_\pm - x^0} \\
&= \frac{r_B\sin(\delta\pm\gamma)}{r_B\cos(\delta\pm\gamma) - x^0}
\end{aligned}
\tag{3.47}
$$

with the numerical value of the coordinate offset along the real axis being $x^0 = 0.8496$.

3.13 Summary

In summary, this chapter presented the first measurement of the *CP* parameters in $B^{\pm} \rightarrow (\pi^0\pi^+\pi^-)_D K^{\pm}$ decays. After $B^{\pm} \rightarrow (K^0_S\pi^+\pi^-)_D K^{\pm}$ this is the second mode where a GGSZ/Dalitz analysis was performed with the goal to measure the CKM angle γ. It is the first analysis where the *CP* parameters are not exclusively extracted from the Dalitz shape or the rate asymmetry alone, but the information from both sources is used simultaneously in the fit. Due to the additional constraints and resulting likelihood shape, a new set of polar coordinates was introduced. This is different from the previous GGSZ analyses by *BABAR* and Belle where Cartesian coordinates have been used. It was shown that the additional information from the decay rate asymmetry significantly improves the sensitivity of the measurement in this decay mode.

As is true for any γ-analysis with the current available statistics, this analysis is not a high precision measurement on its own. However, adding this mode to the other γ measurements will decrease the overall error. Intentionally, it was not attempted to extract a value of γ based on the results of this analysis. Due to the relatively large errors on two of the four *CP* parameters ($\theta_{\pm}$), the sensitivity on γ would be greatly decreased. It was considered more important to present the results in a way that makes their combination with other measurements straightforward. This is left to the "averaging groups" that have the necessary knowledge and experience in correctly combining measurements.

At the beginning of the B-factory programs it was the word that "Measuring β is easy, measuring α is hard and measuring γ is impossible". With advances in both theory and experiment, we succeeded in the first direct measurements of γ. With the startup of the Large Hadron Collider (LHC) at CERN or possibly the construction of a *SuperB* factory, using the same methods as described in this dissertation will eventually result in a precise measurement of the CKM angle γ.

Appendix 3.A Inadequacy of Cartesian coordinates

The inadequacy of the Cartesian coordinates in the MLL-fit using $\mathcal{L} = \mathcal{L}_{\mathrm{DP}} + \mathcal{L}_{\mathrm{BA}}$ (section 3.8) is best demonstrated with toy MC. We perform the same toy study as in section 3.9.1 but this time using the Cartesian coordinates x and y. The results for each floating variable are shown in Figs. 3.72-3.78. The pull distributions for the event yields, $R_{D\pi}$ fraction and the asymmetry (Figs. 3.72-3.76) are standard normal distributions and need no further discussions. This is not the case for the x and y pull distributions.

The x-pull is shifted to positive values and too narrow, while the y-pull is shifted to negative values and too wide when fitted to a Gaussian. The narrow x-pull is explained by the fact that the toy experiments do not properly simulate the systematic errors included in $\mathcal{L}_{\mathrm{BA}}$ and therefore overestimate the error. This effect mostly manifests itself in x because the curvature of $\mathcal{L}_{\mathrm{BA}}$ in the y-direction is very small compared to the x-direction for true values of x and y close to zero (see Fig. 3.48). The source of the upward bias in x can be seen in the same figure. Due to the non-linear correlations between x and y, a fluctuation of y in either direction will lead to a higher value of x (x moves (counter)clockwise on the circle in Fig. 3.48 for an upward (downward) fluctuation of y).

The reason for the large width of the y-pull is best explained with the help of Fig. 3.79 that shows the y-pull versus the error on y and the fitted value of x. For values of x larger than about 0.4, the distribution of y-pulls becomes bifurcated representing the ambiguity in $\mathcal{L}_{\mathrm{BA}}$for these values of x. At the same time the error on y decreases due to the larger curvature of $\mathcal{L}_{\mathrm{BA}}$in the y-direction. In other words, the y-pull distribution represents two types of experiments. The "physical" experiment with small values of x and y, and the "unphysical" experiment with $x \gtrapprox 0.4$ with an ambiguity in y. In fact, the y-pull distribution of experiments with $x < 0.4$ has indeed unit width. The bias of the y-pull is not understood quantitatively at this point.

The NLL projections of $\mathcal{L}_{\mathrm{DP}}$, $\mathcal{L}_{\mathrm{BA}}$ and $\mathcal{L}$ for the Cartesian coordinates using toy MC can be found in section 3.8. For reference, Figs. 3.80 through 3.82 show the NLL projections for the data.

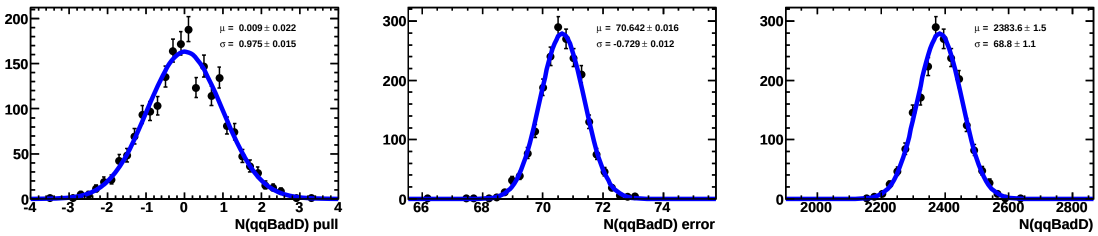


Figure 3.72: Results of 2000 toy MC experiments (signal and background) for the pull (left), error (center) and fitted values (right) of $N_{qq\not{p}}$.

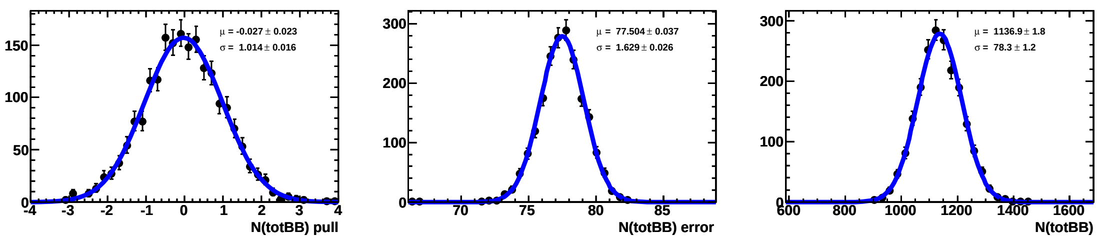


Figure 3.73: Results of 2000 toy MC experiments (signal and background) for the pull (left), error (center) and fitted values (right) of $N_{BB\not{p}}$.

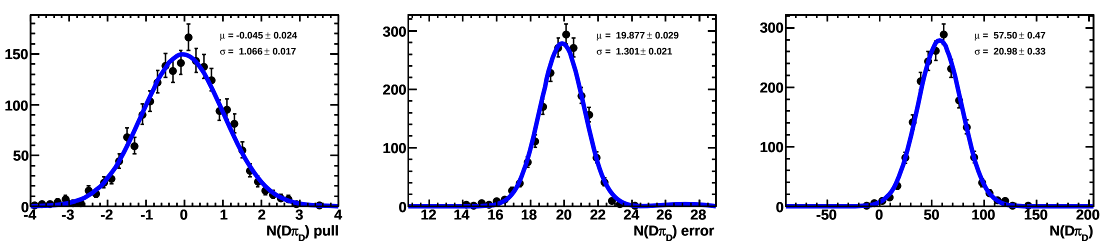


Figure 3.74: Results of 2000 toy MC experiments (signal and background) for the pull (left), error (center) and fitted values (right) of $N_{D\pi_D}$.

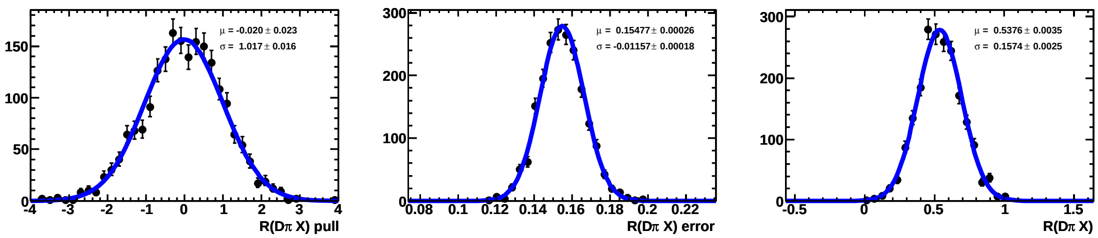


Figure 3.75: Results of 2000 toy MC experiments (signal and background) for the pull (left), error (center) and fitted values (right) of $R_{D\pi}$.

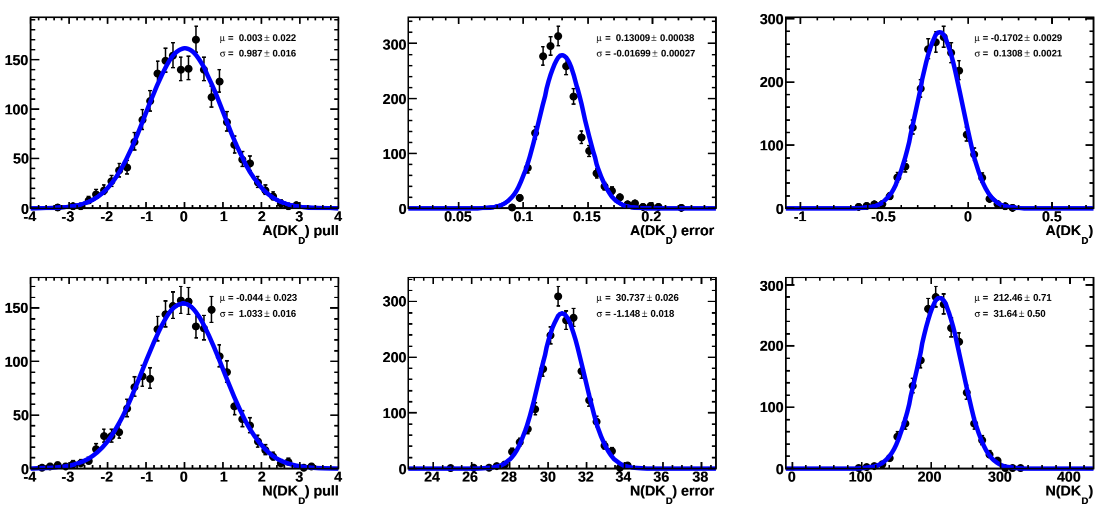


Figure 3.76: Results of 2000 toy MC experiments (signal and background) for the pull (left), error (center) and fitted values (right) of $A_{DK_{\mathrm{sig}}}$ (top) and $N_{DK_{\mathrm{sig}}}$ (bottom).

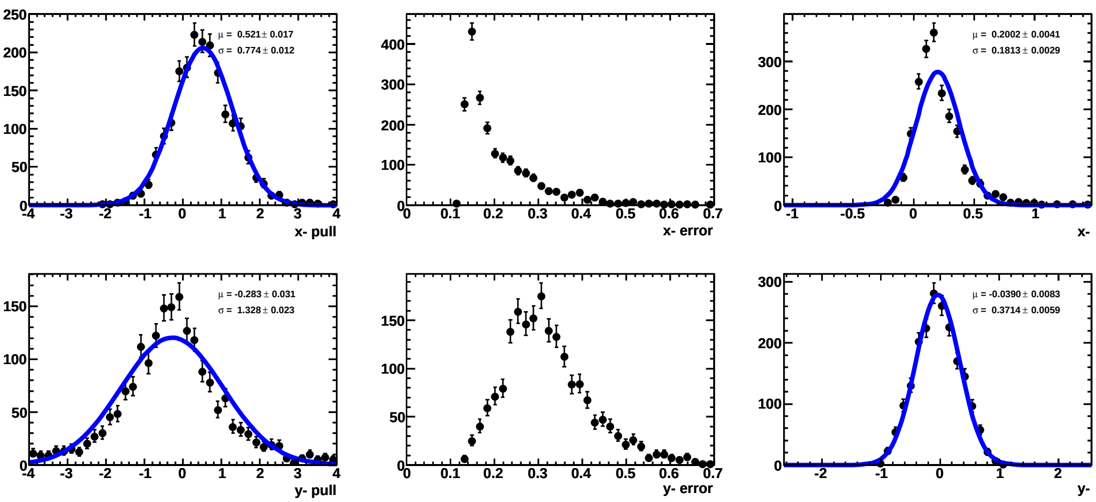


Figure 3.77: Results of 2000 toy MC experiments (signal and background) for the pull (left), error (center) and fitted values (right) of x_- (top) and y_- (bottom).

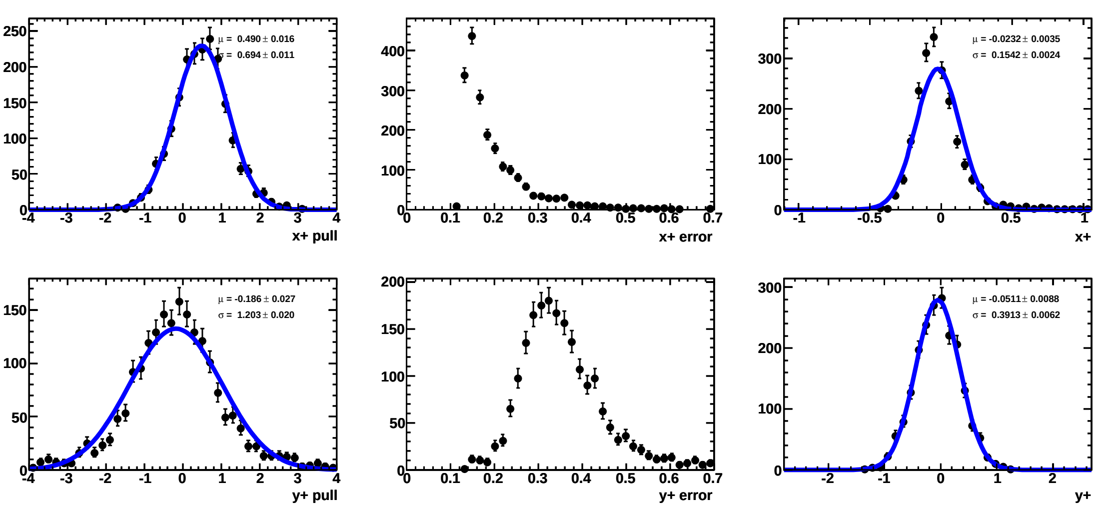


Figure 3.78: Results of 2000 toy MC experiments (signal and background) for the pull (left), error (center) and fitted values (right) of x_+ (top) and y_+ (bottom).

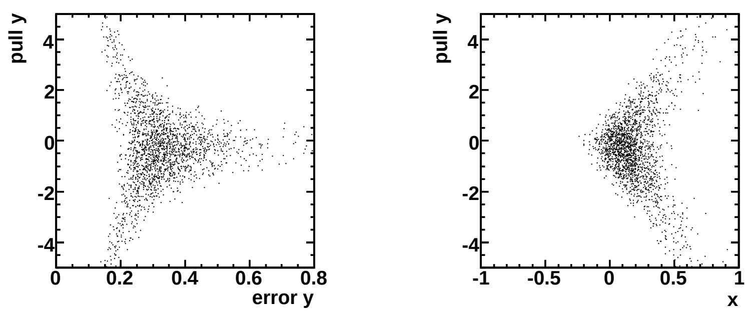


Figure 3.79: Scatter plots of the y-pull versus the error on y (left) and the fitted value of x.

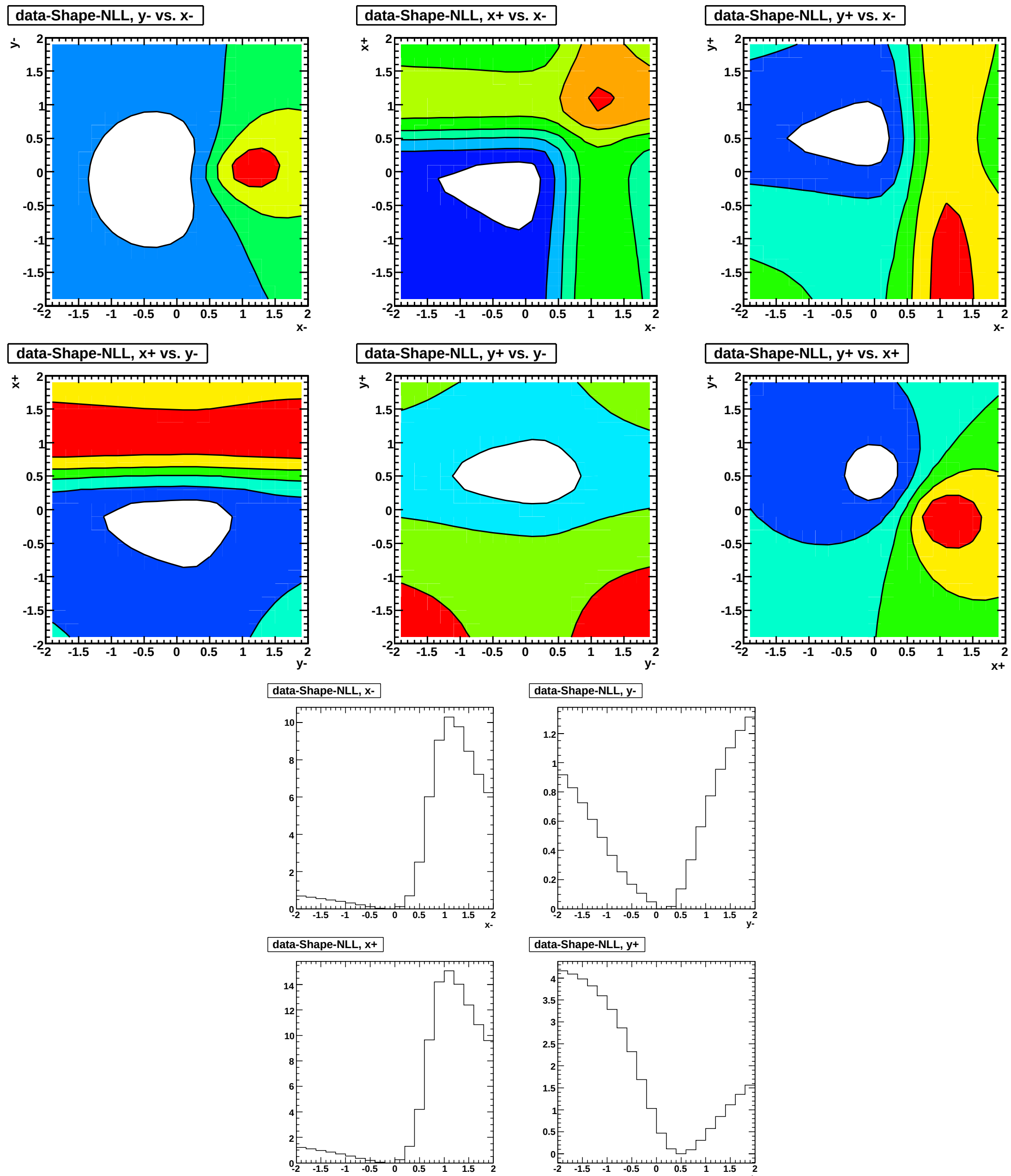


Figure 3.80: Dependence of $\mathcal{L}_{\mathrm{DP}}$ on all pair combinations of the parameters $x_{\pm}, y_{\pm}$ (2D plots, 1σ contours) and on the individual parameters (1D plots) for the data.

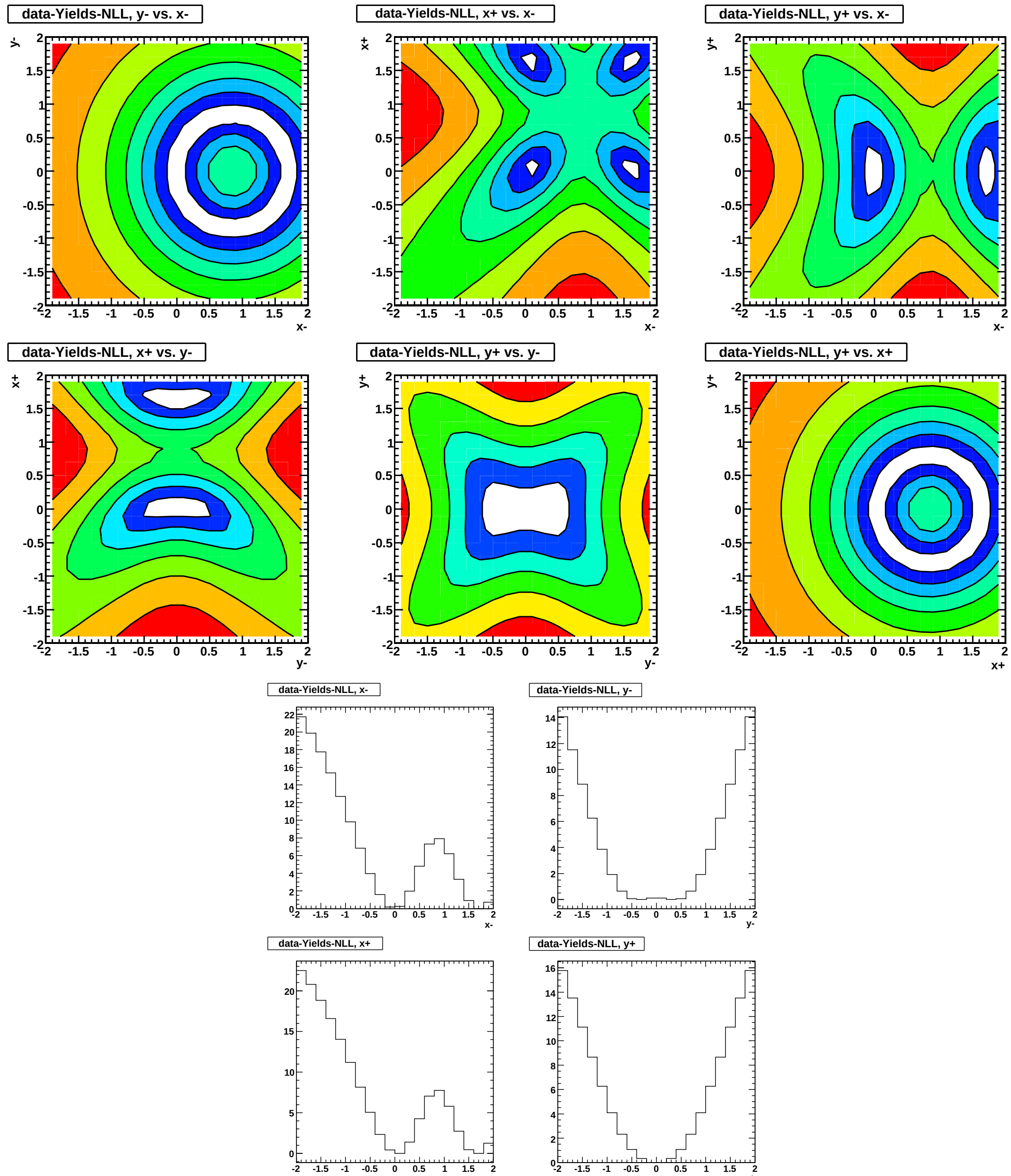


Figure 3.81: Dependence of $\mathcal{L}_{\mathrm{BA}}$ on all pair combinations of the parameters $x_\pm, y_\pm$ (2D plots, 1σ contours) and on the individual parameters (1D plots), using the signal yield and asymmetry from the step-1 data fit.

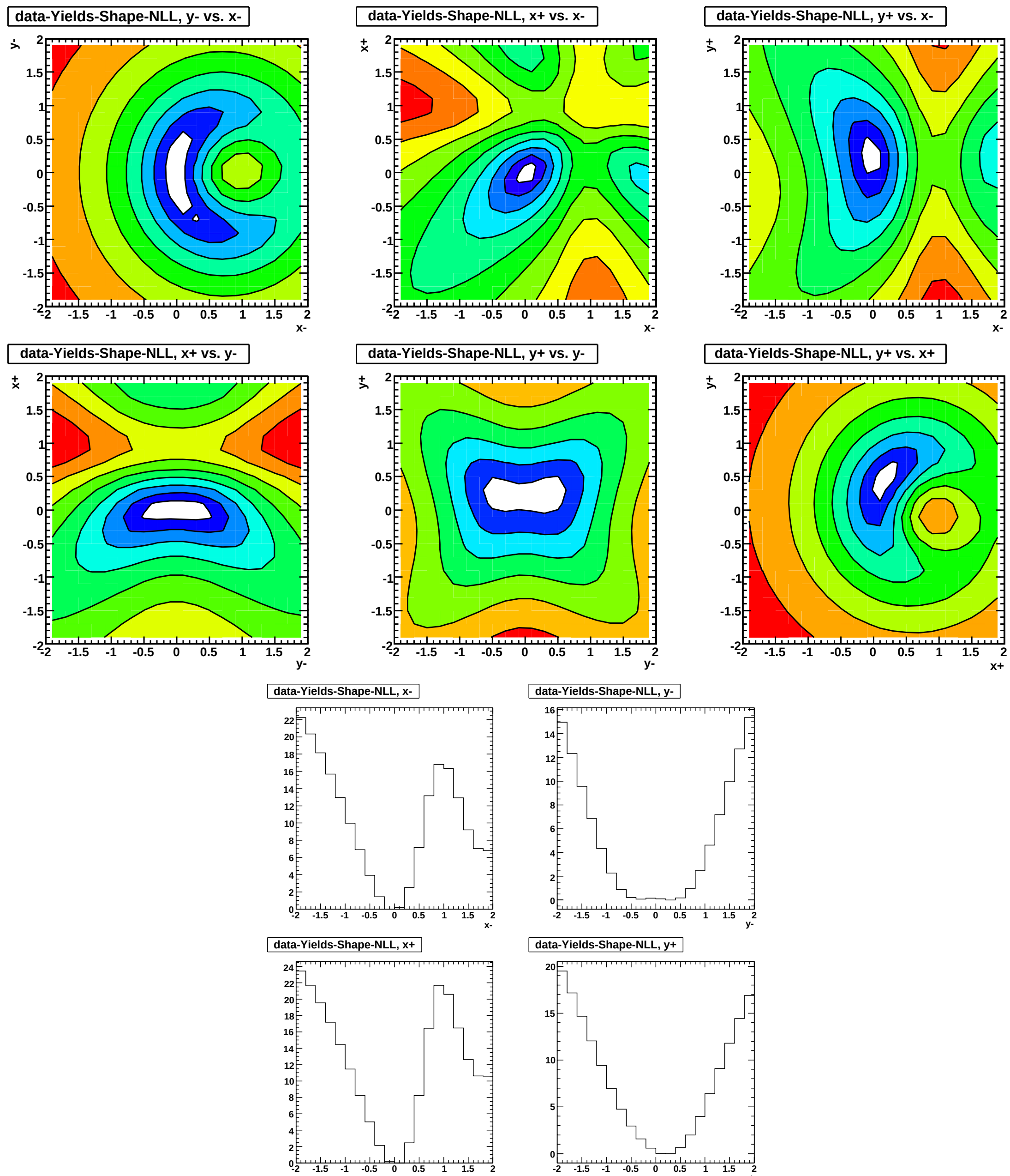


Figure 3.82: Dependence of $\mathcal{L} = \mathcal{L}_{\rm DP} + \mathcal{L}_{\rm BA}$ on all pair combinations of the parameters $x_{\pm}, y_{\pm}$ (2D plots, 1σ contours) and on the individual parameters (1D plots) for the data.

Appendix 3.B Fit validation using $B^- \to D\pi^-$

Negligible *CP*-violation is expected for $B^- \to D_{\pi^+\pi^-\pi^0}\pi^-$ decays. Therefore, a fit to this higher statistics Dalitz plot can be used as a validation of the analysis method. *CP* parameters obtained from the fit should be consistent with no *CP*-violation within the statistical errors. For this validation we perform the same fit as for the $B^- \to DK^-$ sample, minimizing the function $\mathcal{L} = \mathcal{L}_{\mathrm{DP}} + \mathcal{L}_{\mathrm{BA}}$, using a signal PDF plus a single background component. The selections used for this validation are similar to the one in the main analysis, except

1. $0.005 < \Delta E < 0.090\,\mathrm{GeV}/c^2$
2. $q > 0.4$
3. $d > 0.5$,

which have been chosen so as to optimize the signal to background ratio $S/\sqrt{S+B}$. In addition we define a lower ($-0.08 < \Delta E < 0.01\,\mathrm{GeV}/c^2$) and upper ($0.1 < \Delta E < 0.14\,\mathrm{GeV}/c^2$) ΔE sideband region. The ΔE width is about 22 MeV for this mode. Fitting the ΔE distributions to the sum of a Gaussian and a linear function, we find (870 ± 47) $B^- \to D_{\pi^+\pi^-\pi^0}\pi^-$ events and (928 ± 51) $B^+ \to D_{\pi^+\pi^-\pi^0}\pi^+$ events. Figure 3.83 shows the results of these fits.

To determine how best to obtain the background Dalitz shape, we do simple comparisons of the Dalitz distributions in data and MC. Fig. 3.84 shows the Dalitz plot of $B^- \to D_{\pi^+\pi^-\pi^0}\pi^-$ in the ΔE sideband and signal region (defined above). The agreement is reasonable, but not very good. Fig. 3.85 shows the Dalitz distributions in the ΔE sideband on data and MC. The agreement is good. Given the results of these comparisons, we use a histogram-based PDF for the background, taking the shape from the data ΔE sideband. This histogram is shown in Fig. 3.86. For the *CP* fit, we are using the same signal Dalitz PDF as in the $B^- \to D^0K^-$ analysis, and $\mathcal{L}_{\mathrm{BA}}$ is calculated as in section 3.8.1, with the $B^- \to D^0\pi^-$ branching fraction and the above efficiency used in Eq. (3.29). The

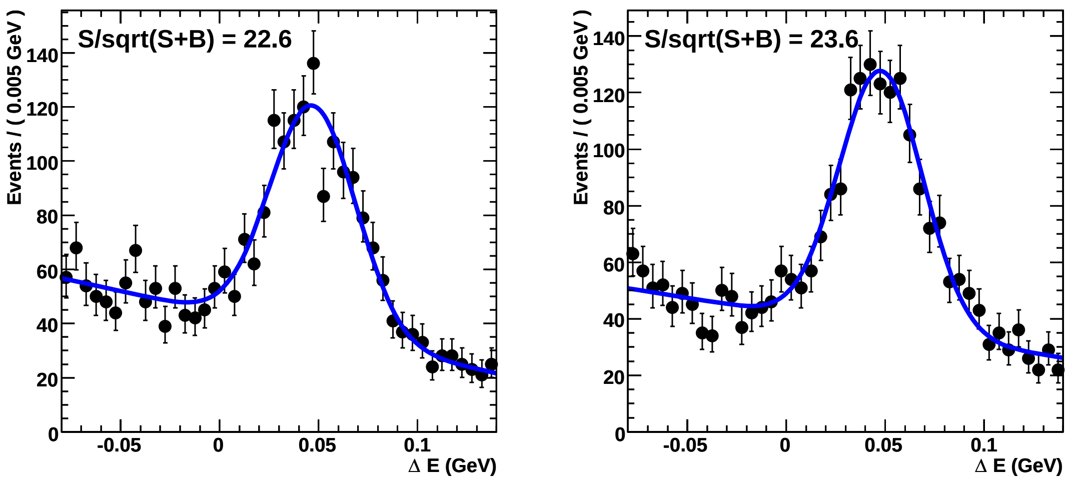


Figure 3.83: Fits to the ΔE distribution of negative (left) and positive (right) $B \to D_{\pi^+\pi^-\pi^0}\pi$ candidates in the data.

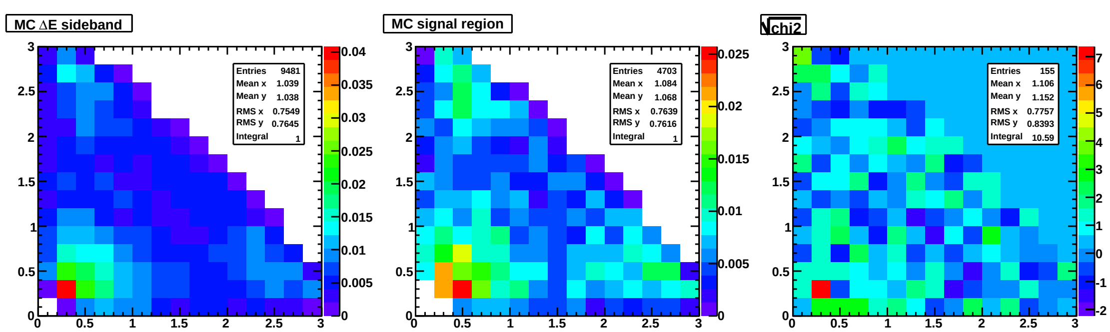


Figure 3.84: Comparison of MC Dalitz distributions for $B^- \to D_{\pi^+\pi^-\pi^0}\pi^-$ in the ΔE sideband (left) and ΔE signal region (center) together with the signed $\sqrt{\chi^2}$ distribution (right). The χ^2 probability of the two histograms to be drawn form the same distribution is 0.4%.

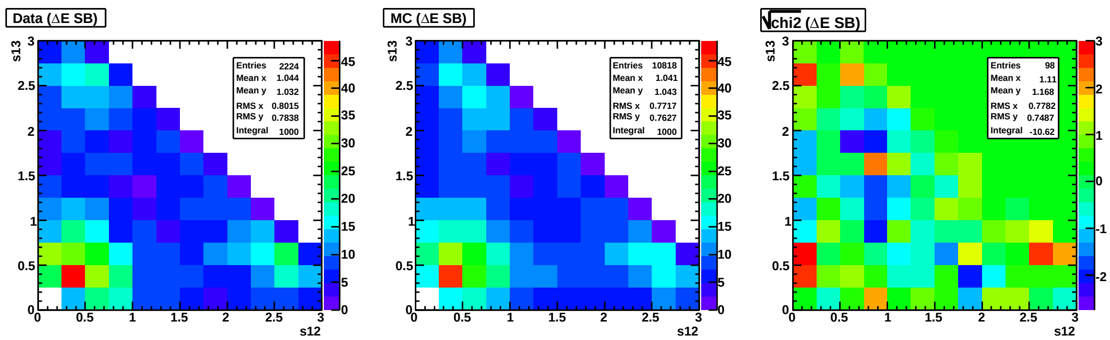


Figure 3.85: Comparison of data (left) and MC (center) Dalitz distributions for $B^- \to D_{\pi^+\pi^-\pi^0}\pi^-$ in the ΔE sideband together with the signed $\sqrt{\chi^2}$ distribution (right). The χ^2 probability of the two histograms to be drawn from the same distribution is 16.0%.

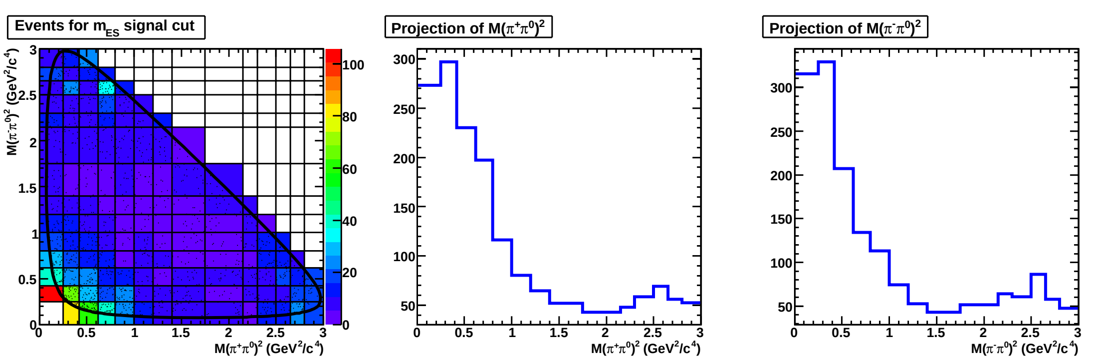


Figure 3.86: The histogram-based PDF used for the background shape (left), and the data (points) in the ΔE sideband, from which the histogram was derived. The 1-D projections are also shown.

results of the fit are summarized in Table 3.45. It is seen that the *CP* parameters obtained from the $B^- \to D^0\pi^-$sample are consistent with no *CP* violation, although the difference is somewhat large (1.7 times the statistical error) for y_+ and θ_+, respectively.

	Value	σ from 0		Value	σ from $x^0/180°$
x_-	0.002 ± 0.045	0.04	ρ_-	0.853 ± 0.044	0.07
y_-	-0.095 ± 0.103	-0.92	θ_-	186.4 ± 7.0	0.91
x_+	-0.023 ± 0.049	-0.47	ρ_+	0.893 ± 0.045	0.96
y_+	-0.188 ± 0.109	-1.72	θ_+	192.2 ± 7.1	1.72

Table 3.45: Fit results for $B^- \to D^0\pi^-$ using Cartesian and polar coordinates. Errors are only due to statistics, as well as the systematics associated with $\mathcal{L}_{\mathrm{BA}}$, namely, the error on the efficiency, the luminosity, the branching fractions $\mathcal{B}(B^- \to D^0\pi^-)$, $\mathcal{B}(D \to K^-\pi^+\pi^0)$, and the ratio $\mathcal{B}(D \to \pi^+\pi^-\pi^0)/\mathcal{B}(D \to K^-\pi^+\pi^0)$.

Chapter 4

Analysis of $\mathbf{B \to X(3872)K,\ X(3872) \to J/\psi\,\pi^+\pi^-}$

This chapter describes the analysis of $B \to X(3872)$, $X(3872) \to J/\psi\,\pi^+\pi^-$ decays, which updates and supersedes the previous *BABAR* measurement [27]. Moreover, it is the first dedicated search for the $X(3872)$ in neutral B decays. The decay modes under investigation are:

- $B^0 \to X(3872)K^0_S, \quad X(3872) \to J/\psi\,\pi^+\pi^-, \quad K^0_S \to \pi^+\pi^-$
- $B^- \to X(3872)K^-, \quad X(3872) \to J/\psi\,\pi^+\pi^-.$

The ratio of branching fractions and the masses of the $X(3872)$ in these two decay modes will help to decide between different models for the $X(3872)$. The initially favored Charmonium assignment [70–72] is considered unlikely in the meantime due to the discrepancy of the predicted mass. A review of the Diquark-Antidiquark and meson molecule model can be found in section 1.4.2 and 1.4.3 of this dissertation.

4.1 Data sample

The data we use in this analysis consists of the following subsamples:

- **On-peak data**

  Run 1-4 data are used with the runs 1-3 as defined in the BlackDiamond dataset and the "final run 4 dataset" for run 4.

- **Signal Monte Carlo**

  We have generated signal Monte Carlo for each of the following modes:

  - $B^0 \to X(3872)K^0_S, \quad X(3872) \to J/\psi\,\rho^0, \quad \rho^0 \to \pi^+\pi^-, \quad K^0_S \to \pi^+\pi^-$
  - $B^- \to X(3872)K^-, \quad X(3872) \to J/\psi\,\rho^0, \quad \rho^0 \to \pi^+\pi^-$

  The decay model is sequential, two-body phase-space and the $X(3872)$ is generated as a zero-width particle. Note that we simulate the dipion as originating from a ρ^0. This seems to match the observed dipion invariant mass better than a pure phase-space model. However, at this point, it is not clear, whether it is really $\rho \to \pi\pi$ or not.

- **Generic Monte Carlo**

  We are using $B^0\overline{B}^0$, B^+B^-, $c\overline{c}$ and uds MC samples. SP5 MC is used to simulate run 1-3 and SP6 for run 4.

Table 4.1 summarizes the data sample. The equivalent integrated luminosity for the MC samples are calculated using $L = N/\sigma$ with the number of events N and the cross section σ. For on-peak data we are showing the official B-counting numbers ($N_{B\overline{B}}$) and the all MC samples are scaled to the total on-peak data luminosity of $210.6\,\mathrm{fb}^{-1}$.

Mode	# events	σ(nb)	L(fb^{-1})	scale factor
$B^0 \to X(3872)K^0_S, X(3872) \to J/\psi\,\pi^+\pi^-$	57,000			
$B^- \to X(3872)K^-, X(3872) \to J/\psi\,\pi^+\pi^-$	57,000			
$B^0\overline{B}^0$ generic MC	453,426,225	0.551	822.92	0.256
B^+B^- generic MC	465,004,433	0.551	843.93	0.249
uds MC	551,116,526	2.090	263.69	0.798
$c\overline{c}$ MC	274,406,199	1.300	211.08	0.997
Run 1	21,181,864		19.459	
Run 2	66,441,247		60.267	
Run 3	34,076,579		31.061	
Run 4	110,107,681		99.763	
Total on-peak ($N_{B\overline{B}}$)	231,807,371		210.550	1

Table 4.1: Data sample summary. Number of events, cross section σ, integrated luminosity L and scale factor to on-peak data luminosity.

4.2 Candidate reconstruction and preselection

The data were reconstructed with `analysis-22/23` at GridKa/SLAC. For all the data and Monte Carlo, except signal MC, we use the `Jpsitoll` skim (release 14 processing). We use two tag bit filters, in order to preselect events on the level of the tag database. `JpsiELoose || JpsiMuLoose` selects only the J/ψ candidates from the skim and `R2All<0.5` is used for continuum suppression.

4.2.1 Reconstruction of $J/\psi \to l^+l^-$

We are only reconstructing J/ψ candidates in the two leptonic modes:

- $J/\psi \to e^+e^-$

  Oppositely-charged tracks from the list `PidLHElectrons` and electron candidates from the same list undergoing the standard bremsstrahlung recovery are used to form a J/ψ candidate. The electrons are fitted with a geometric and a mass constraint and must satisfy $2.9 < m(e^+e^-) < 3.2\,\mathrm{GeV}/c^2$.

- $J/\psi \to \mu^+\mu^-$

  Oppositely charged muon candidates from the list `muNNVeryLoose` fitted in the same way as electrons and satisfying $3.0 < m(\mu^+\mu^-) < 3.2\,\mathrm{GeV}/c^2$ are used to form a J/ψ candidate.

In both cases we constrain the dilepton invariant mass to the J/ψ mass from the PDG.

4.2.2 Reconstruction of $\psi(2S)/X(3872) \to J/\psi\pi^+\pi^-$

The $\psi(2S)$ decays into the same final states we are interested in, and thus we expect to see this decay mode in our spectrum as well. We form a $\psi(2S)$ (or $X(3872)$) candidate by combining the J/ψ candidate with two pion candidates from the list `piLHVeryLoose` using a geometric fit.

4.2.3 Reconstruction of $K^0_S \to \pi^+\pi^-$

Our K^0_S selection is the same as used in the $\mathcal{B} \to K^0_S\pi^0$ analysis (see `AnalTools/611` HN). K^0_S candidates are formed from oppositely `ChargedTracks` requiring $|m(\pi^+\pi^-) - m_{K^0_S}| < 25\,\mathrm{MeV}/c^2$, a χ^2 consistency of the fit > 0.001 and flight length significance $l/\sigma(l) > 3$.

4.2.4 Reconstruction of B meson candidates

The final B meson candidate is formed by combining the $\psi(2S)$ candidate with either a K^0_S candidate or a charged kaon candidate from the `KLHVeryLoose` list by using a geometric fit and requiring $|\Delta E| < 0.3\,\mathrm{GeV}$ and $5.2 < m_{ES} < 5.3\,\mathrm{GeV}/c^2$.

4.2.5 Fox-Wolfram moment

Figure 4.1 shows $R2$ for signal (left), data and generic MC (right). The plot uses the optimized cuts described later and we conclude that there is only a tiny contamination from continuum ($q\overline{q}$) background.

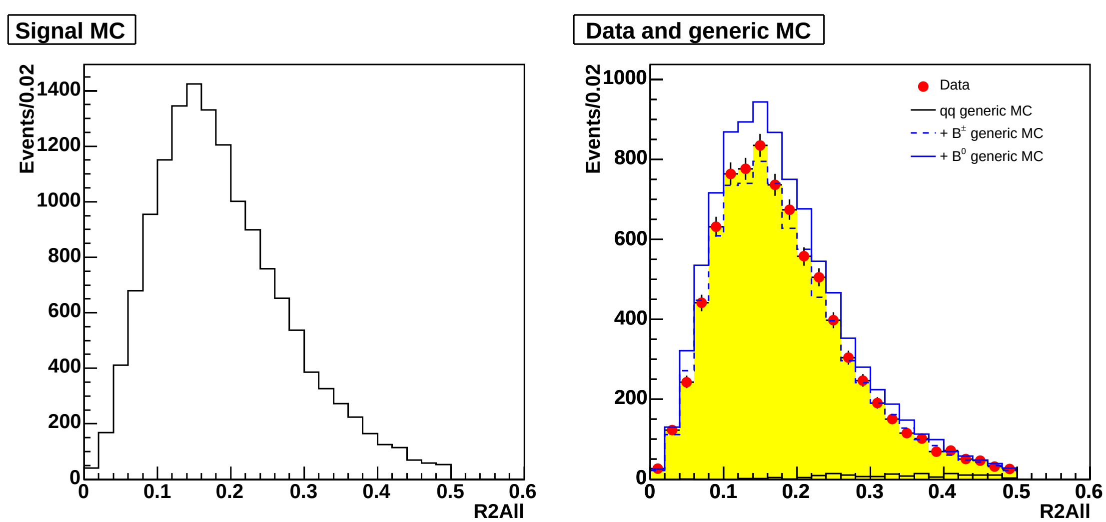


Figure 4.1: Fox-Wolfram moment $R2$ in $B^- \to J/\psi\pi^+\pi^-K^-$for signal MC (left), generic MC and data (right).

4.3 Final candidate selection

4.3.1 Optimization procedure

The final selection criteria are optimized by maximizing the following quantity [73]:

$$\frac{n_S}{a/2 + \sqrt{n_B}} \rightarrow \max \tag{4.1}$$

with n_S as the number of signal events from signal MC, n_B the weighted (and normalized to on-peak data luminosity) sum of background events from different generic MC samples ($B\overline{B}$, $c\overline{c}$ and uds) and a the desired significance of signal to background separation in numbers of sigmas (we choose $a = 3$). [9]

From our very similar previous analysis of $B \rightarrow J/\psi\, \pi^{\pm}\pi^{0}K$ [74] we already had a good idea of possible useful discriminating variables. We did the optimization as a grid search starting with larger binnings and then reoptimizing around the found optimum with smaller bins. The variables, ranges and step sizes used in the optimization for the $\pi^+\pi^-$ modes are as follows:

- $|\Delta E| < 5, 10 \ldots 40\,\mathrm{MeV}/c^2$
- $|m_{ES} - 5.279\,\mathrm{GeV}/c^2| < 4, 6 \ldots 10\,\mathrm{MeV}/c^2$
- $|m(J/\psi\, \pi^+\pi^-) - 3.872\,\mathrm{GeV}/c^2| < 4, 6 \ldots 10\,\mathrm{MeV}/c^2$
- K^0_S mass: $|m(K^0_S) - 497.7\,\mathrm{MeV}/c^2| < 10, 15 \ldots 25\,\mathrm{MeV}/c^2$
- Thrust angle: $|\cos\theta_{Thrust}| < 0.85, 0.9 \ldots 1.0$
- Fox-Wolfram Moment: $R2 < 0.40, 0.45, 0.50$
- Muon Particle ID $\in \{\texttt{muNNVeryLoose}, \texttt{muNNLoose}, \texttt{muNNTight}\}$
- Kaon Particle ID $\in \{\texttt{KLHVeryLoose}, \texttt{KLHLoose}, \texttt{KLHTight}\}$

[9] Notice that it is not crucial for the optimization how we treat multiple candidates as long as we only count one candidate per event. See section 4.3.6 for how we treat multiple candidates.

- Pion Particle ID $\in \{\texttt{piLHVeryLoose}, \texttt{piLHLoose}, \texttt{piLHTight}\}$

The lepton invariant mass cut and electron PID selection is fixed to the values described in section 4.3.3.

4.3.2 Optimization results

Besides the final candidate selection criteria, which are described below, the optimization revealed that neither the thrust angle cut nor an additional cut on the Fox-Wolfram Moment $R2$ is helpful in increasing the sensitivity. The common m_{ES} signal window for all modes is $|m_{ES} - 5.279\,\mathrm{GeV}/c^2| < 6\,\mathrm{MeV}/c^2$. The optimized ΔE signal window is $|\Delta E| < 15\,\mathrm{MeV}$ for the $\pi^+\pi^-$ and the optimized X-particle signal region is $|m(J/\psi\,\pi^+\pi^-) - 3.872\,\mathrm{GeV}/c^2| < 6\,\mathrm{MeV}/c^2$.

4.3.3 Final J/ψ selection

Our final J/ψ selection uses the standard charmonium mass cuts and lepton selectors as follows:

- $2.95 < m(J/\psi \to e^+e^-) < 3.14\,\mathrm{GeV}/c^2$

  Both electrons have to pass the `PidLHElectron` (not optimized) electron selector

- $3.06 < m(J/\psi \to \mu^+\mu^-) < 3.14\,\mathrm{GeV}/c^2$

  One muon has to pass the `muNNVeryLoose` and the other muon has to pass the `muNNLoose` muon selector.

Fig. 4.2 shows the J/ψ mass distributions in the electronic and muonic mode with the arrows indicating the final mass cut on the J/ψ candidate.

4.3.4 Final K_S^0 selection

Since we already had a somewhat optimized preselection for the K_S^0, the signal is very clean. The final cut we choose is $|m(K_S^0 \to \pi^+\pi^-) - 497.7)| < 15\,\mathrm{MeV}/c^2$. Figure 4.3

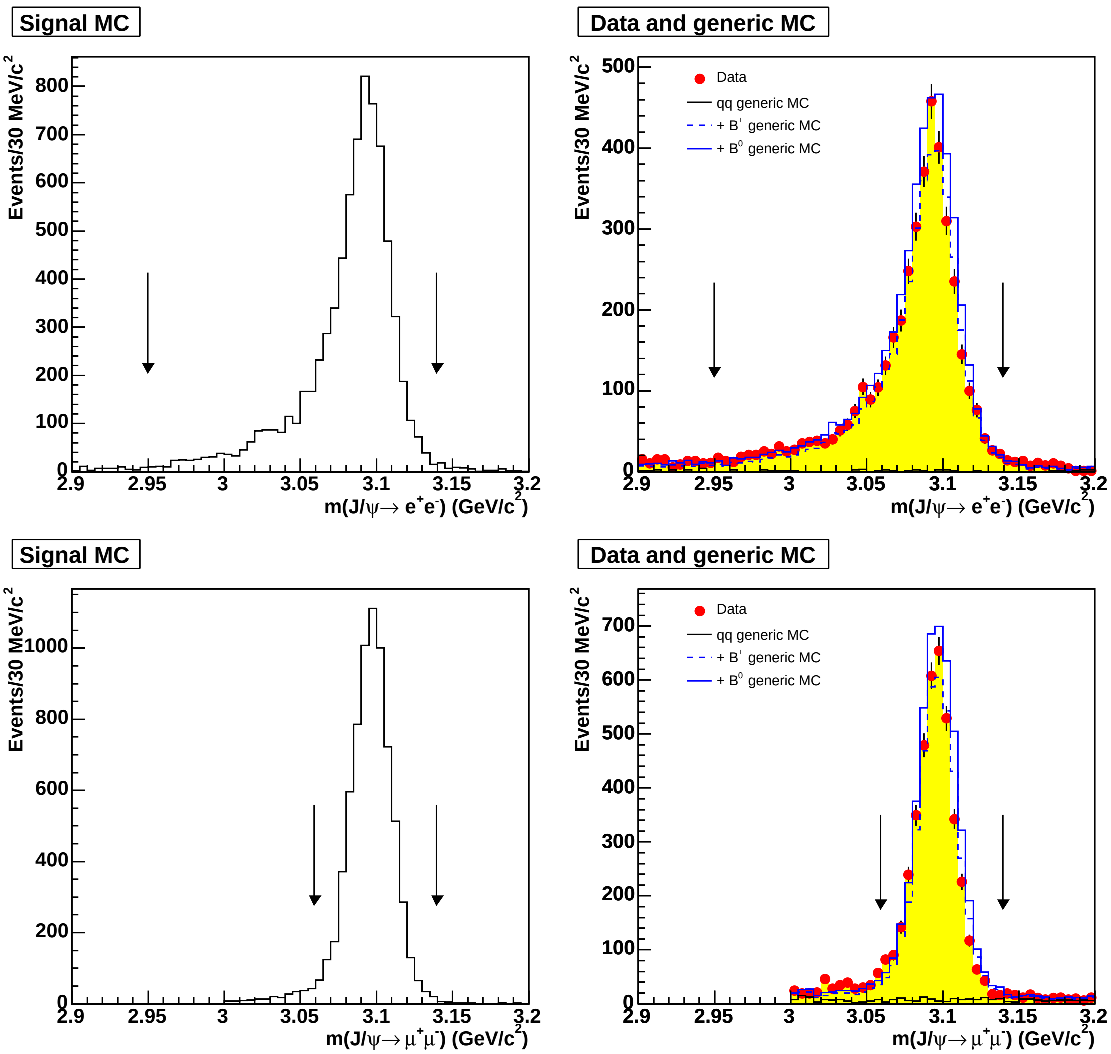


Figure 4.2: $J/\psi \rightarrow e^+e^-$ (top) and $J/\psi \rightarrow \mu^+\mu^-$ (bottom) in $B^- \rightarrow J/\psi\pi^+\pi^-K^-$. One can see the longer tail in the electron mode because of the energy loss due to bremsstrahlung. The small arrows indicate our final cuts.

shows the $K_S^0 \to \pi^+\pi^-$ invariant mass in $B^0 \to J/\psi\,\pi^+\pi^- K_S^0$.

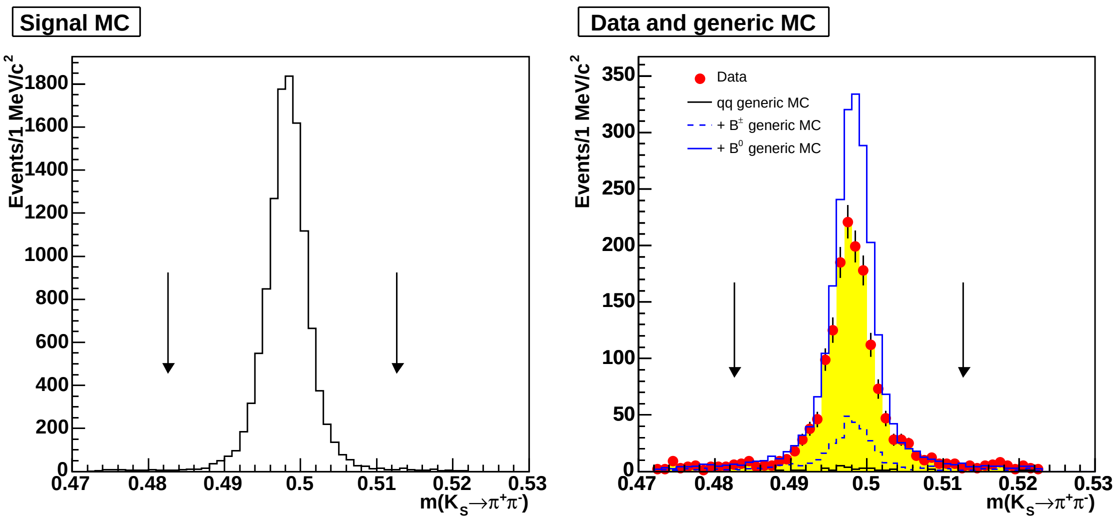


Figure 4.3: $K_S^0 \to \pi^+\pi^-$ invariant mass in $B^0 \to J/\psi\,\pi^+\pi^- K_S^0$with the optimized cuts applied. The left plot shows signal MC and the right plot shows background MC and data (red dots). The solid, blue line is the weighted sum of all MC samples. The arrows indicate our final mass cut.

4.3.5 Final $K^\pm$ and $\pi^\pm$ selection

Kaon candidates have to pass the `KLHVeryLoose` selector. One of the two pions in the $\pi^+\pi^-$ modes has to pass `piLHLoose` and the other pion has to pass the `piLHTight` selector.

4.3.6 Summary of selection cuts and efficiencies

Table 4.2 summarizes the final selection criteria. Applying all of the above cuts to our signal MC samples and counting the remaining events $n_{S,MC}$, we obtain the (cut-)efficiencies

$$\epsilon_{MC} = \frac{n_{S,MC}}{N_{MC}} \tag{4.2}$$

as listed in Table 4.3. The error on ϵ_{MC} is taken as binomial distributed with

$$\sigma_\epsilon = \sqrt{\epsilon(1-\epsilon)/N_{MC}}. \tag{4.3}$$

Variable	Selection
ΔE	$\lvert\Delta E\rvert < 15\,\mathrm{MeV}$ ($\pi^+\pi^-$ modes)
m_{ES}	$\lvert m_{ES} - 5.279\,\mathrm{GeV}/c^2\rvert < 6\,\mathrm{MeV}/c^2$
$X^\pm$ candidate mass	$\lvert m(J/\psi\,\pi^+\pi^-) - 3.872\,\mathrm{GeV}/c^2\rvert < 6\,\mathrm{MeV}/c^2$
$J/\psi \to e^+e^-$ mass	$2.95 < m(J/\psi \to e^+e^-) < 3.14\,\mathrm{GeV}/c^2$
$J/\psi \to \mu^+\mu^-$ mass	$3.06 < m(J/\psi \to \mu^+\mu^-) < 3.14\,\mathrm{GeV}/c^2$
K^0_S mass	$\lvert m(K^0_S \to \pi^+\pi^-) - 497.7)\rvert < 15\,\mathrm{MeV}/c^2$
Electron PID	`PidLHElectron`, `PidLHElectron`
Muon PID	`muNNVeryLoose`, `muNNLoose`
Pion PID	`piLHLoose`, `piLHTight`
Kaon PID	`KLHVeryLoose`

Table 4.2: Final selection cuts besides the preselection criteria mentioned in section 4.2.

Furthermore we list the number of remaining candidates per event. In case there is more than one remaining candidate per event we choose the candidate with the smallest $|\Delta E|$ for all of our plots (and the efficiency calculation) - except when showing ΔE itself.

Mode	N_{MC}	$n_{S,MC}$	ϵ_{MC}	Error ϵ_{MC}	cand/event
$B^0 \to J/\psi\,\pi^+\pi^-K^0_S$	57,000	9,565	16.78%	0.16%	1.01
$B^- \to J/\psi\,\pi^+\pi^-K^-$	57,000	12,007	21.06%	0.17%	1.01

Table 4.3: Efficiencies and number of candidates per event after applying all the optimized cuts on signal MC. These efficiencies are not used in the final analysis. See section 4.4 for the determination of the fit efficiency.

4.4 Signal extraction

In this section, we describe the analysis method for the two modes $B^0 \to J/\psi\,\pi^+\pi^- K^0_S$ and $B^- \to J/\psi\,\pi^+\pi^- K^-$. The method is based on a two-dimensional unbinned extended maximum likelihood (UEML) fit to the m_{ES} and $m(J/\psi\,\pi^+\pi^-)$ (called m_X) invariant mass distribution. The method and the data sample used as input for the fit are described in the following sections.

4.4.1 Candidate selection and fit variables

The obvious fit variables in this analysis include ΔE, m_{ES} and m_X. However, there remain a significant number of multiple candidates if the fit includes all of those three variables and the modelling of the likelihood function gets more complicated. Therefore, we decided to cut on ΔE, select the candidate with the smallest $|\Delta E|$ and perform a two-dimensional UEML fit to m_{ES} and m_X. The detailed analysis method consists of the following steps:

1. Apply all the optimized cuts (except for the fit variables) listed in Tab. 4.2.

2. Apply selection cuts for fit variables (this will be the fit range).

   - $5.2 < m_{ES} < 5.3\,\mathrm{GeV}/c^2$
   - $3.8 < m_X < 4.0\,\mathrm{GeV}/c^2$

3. Select the candidate with the smallest $|\Delta E|$ (best-ΔE).

4. Perform a two-dimensional UEML fit to m_{ES} and m_X.

4.4.2 Probability density function and event types

For each event type $t \in T$ we define a PDF $\mathcal{P}_t(\mathbf{x};\theta)$ evaluated for each event with $\mathbf{x} = (m_{ES}, m_X)$ and dependent on the parameter(s) θ. We further assume that m_{ES} and m_X

are uncorrelated quantities and thus write the PDF as a product of two one-dimensional PDFs

$$\mathcal{P}_t(\mathbf{x};\theta) = g_t(m_{ES})h_t(m_X). \tag{4.4}$$

Based on Monte Carlo studies we define three different event types $T = \{S, P, C\}$:

- **Signal** events are genuine $B \to X(3872)K, X(3872) \to J/\psi\pi^+\pi^-$ events and taken from signal MC.
- **Peaking** background events from non-resonant $B \to J/\psi\pi^+\pi^-K$ or other intermediate resonances like $B \to J/\psi K_1(1273), K_1 \to K\rho$ and $B \to J/\psi K^*\pi, K^* \to K\pi$. Those events are obtained from generic $B\overline{B}$ MC by selecting all events which have the same final states as our signal modes and form a good B-candidate.
- **Combinatorial** background events forming a fake B-candidate taken from $B\overline{B}$ generic MC by removing all peaking background events as described above.

Each of those event types has a different parametrization of its PDF $\mathcal{P}_t(\mathbf{x}) = g_t(m_{ES})h_t(m_X)$, which we obtain by fitting g_t and h_t separately to the m_{ES} and m_X distributions, respectively. All the fits are done with the RooFit [75] package and if not otherwise noted, all parameters are kept floating. The following sections described the PDFs and Tab. 4.4 summarizes the PDF parameters obtained from the fit.

Signal PDF parametrization

Due to the mass-constraint on the J/ψ candidate in $X \to J/\psi\pi^+\pi^-$ the resolution function of the $X(3872)$ (generated with zero natural width in the MC) is not a simple Gaussian as one might expect. In general, the distribution is more peaked around the central value with longer tails. As a first approach, we model this distribution by a Lorentzian[10] convoluted

[10] Note that Lorentzian, Cauchy distribution and Breit-Wigner are just different names for the same distribution. To avoid confusion, we refer to a Lorentzian in case of the resolution model and to Breit-Wigner for the line-shape model.

with a Gaussian in m_X. For the m_{ES} distribution we use the usual Gaussian:

$$g_S(x;\mu,\sigma) \quad = \quad G(x;\mu,\sigma) \sim e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{4.5}$$

$$h_S(x;m,\Gamma,\sigma) \quad \sim \quad \int_{-\infty}^{+\infty} G(x';0,\sigma) L(x-x';m,\Gamma)\, dx' \tag{4.6}$$

where L is the Lorentz function $L(x;m,\Gamma) \sim \left((x-m)^2 + (\Gamma/2)^2\right)^{-1}$ with central value m and Γ as the full width at half maximum (FWHM). Fig. 4.4 shows the 1-dimensional UEML-fit to the m_{ES} and m_X distributions for the $B^0 \to J/\psi\,\pi^+\pi^- K^0_S$ mode. Note, that the ARGUS tail used to model remaining combinatorics in the signal sample, is not part of the signal PDF, but only used to obtain an accurate fit to the Gaussian signal peak. The final parameters from the fit including errors are listed in Tab. 4.4.

In both modes the Gaussian width σ of the fit to the X-mass is consistent with zero. In the following MC experiments used to validate our general fitting procedure and the final fit on data, we therefore use a simple Lorentzian as a model for the $X(3872)$ mass distribution.

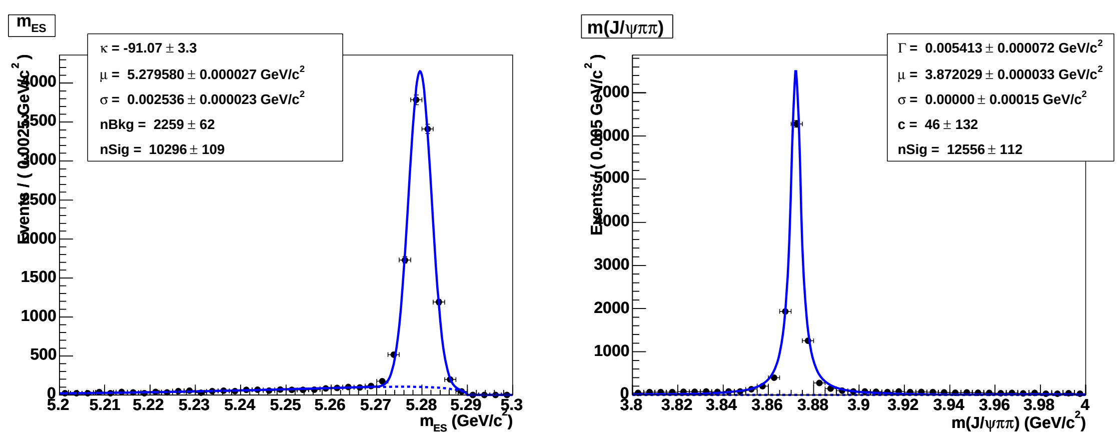


Figure 4.4: 1-dimensional UEML-fit to the m_{ES} (left) and m_X (right) distribution on signal MC for $B^0 \to J/\psi\,\pi^+\pi^- K^0_S$.

Peaking background PDF parametrization

Peaking background events are modelled by a Gaussian in m_{ES} (that's why we call them peaking events) and a linear function in m_X:

$$g_P(x;\mu,\sigma) \sim e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{4.7}$$

$$h_P(x;c) = \frac{1}{N}(1+cx) \tag{4.8}$$

Note, that the slope of the first order polynomial is not c, but rather c/N where N is the normalization of the polynomial which can be negative as well. Fig. 4.5 (left) shows the 1-dimensional UEML-fit to the m_{ES} distribution in $B^0 \to J/\psi\,\pi^+\pi^-K^0_S$. As explained above, the ARGUS tail is not used in the peaking background PDF g_P and only included for technical reasons. The final parameters from the fit for both modes including errors are listed in Tab. 4.4.

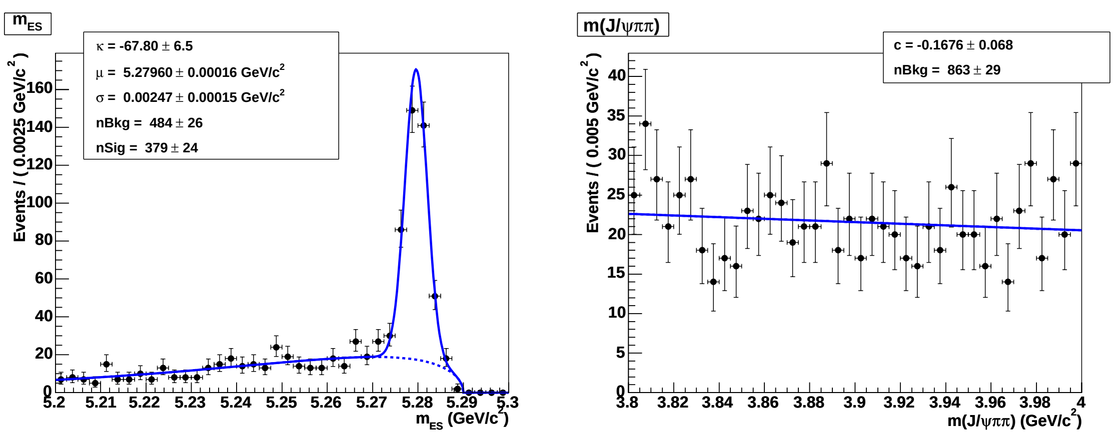


Figure 4.5: 1-dimensional UEML-fit to the m_{ES} (left) and m_X (right) distribution on generic MC for peaking background events in $B^0 \to J/\psi\,\pi^+\pi^-K^0_S$.

Combinatorial background PDF parametrization

Finally, combinatorial background events are expected to follow an ARGUS shape in m_{ES} and observed to be linear in m_X:

$$g_C(x; m_0, \kappa) \sim x\sqrt{1 - x^2/m_0^2}\, e^{\kappa(1-x^2/m_0^2)} \tag{4.9}$$

$$h_P(x; c) = \frac{1}{N}(1 + cx) \tag{4.10}$$

where m_0 is the ARGUS end-point and set to $5.29\,\mathrm{GeV}/c^2$, the kinematic limit for $\Upsilon(4S) \to B\overline{B}$. Fig. 4.6 shows the 1-dimensional UEML-fit to the m_{ES} and m_X distributions for the $B^0 \to J/\psi\,\pi^+\pi^- K^0_S$ mode. The final parameters from the fit for both modes including errors are listed in Tab. 4.4.

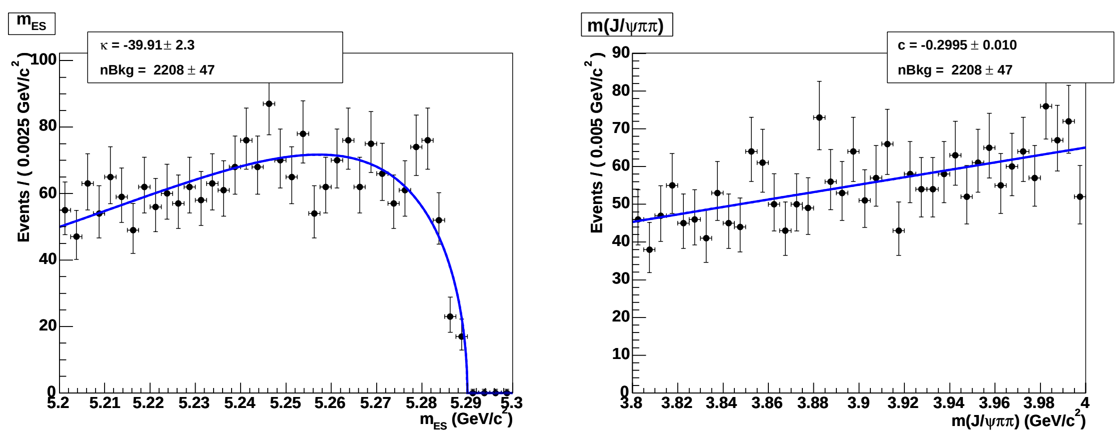


Figure 4.6: 1-dimensional UEML-fit to the m_{ES} (left) and m_X (right) distribution on generic MC for combinatorial background events in $B^0 \to J/\psi\,\pi^+\pi^- K^0_S$.

Summary of PDF parametrization

Tab. 4.4 summarizes the final parameters including errors obtained from the previously described fits for both the modes $B^0 \to X(3872)K^0_S,\ X \to J/\psi\,\pi^+\pi^-$ and $B^- \to X(3872)K^-,\ X \to J/\psi\,\pi^+\pi^-$. Furthermore we give the corresponding parameters which we find from the same fits to the $\psi(2S)$ region ($3.6 < m(J/\psi\pi\pi) < 3.8\,\mathrm{GeV}/c^2$) and using $\psi(2S)$-signal MC.[11]

Parameter	$B^0 \to XK^0_S$	$B^- \to XK^-$	$B^0 \to \psi(2S)K^0_S$	$B^- \to \psi(2S)K^-$
Signal				
m_{ES} mean μ	5.2796 ± 0.0001	5.2793 ± 0.0001	5.2796 ± 0.0001	5.2793 ± 0.0001
m_{ES} width σ	2.54 ± 0.02	2.52 ± 0.02	2.67 ± 0.04	2.76 ± 0.03
m_X mean m	3.8720 ± 0.0001	3.8721 ± 0.0001	3.6861 ± 0.0001	3.6861 ± 0.0001
m_X Gaussian σ	0.00 ± 0.15	0.00 ± 0.25	0	0
m_X Lorentz Γ	5.41 ± 0.07	5.38 ± 0.06	4.09 ± 0.08	4.09 ± 0.08
Peaking bgd.				
m_{ES} mean μ	5.2796 ± 0.0002	5.2794 ± 0.0001		
m_{ES} width σ	2.47 ± 0.15	2.54 ± 0.07		
m_X parameter c	-0.17 ± 0.07	-0.42 ± 0.11		
Comb. bgd.				
ARGUS shape κ	-39.9 ± 2.3	-38.2 ± 1.2	-45.9 ± 3.9	-49.5 ± 1.7
m_X parameter c	-0.30 ± 0.01	-0.30 ± 0.01	-0.30 ± 0.01	-0.30 ± 0.01

Table 4.4: Summary of PDF parameters from separate one-dimensional UEML-fits for the three different event types, two signal modes and the two $\psi(2S)$ benchmark modes. All mean values (μ, m) are in GeV/c^2 and widths (σ, Γ) are in MeV/c^2.

4.4.3 Fit efficiency

We define our efficiency for reconstructed signal events as

$$\epsilon_{MC} = \frac{n_{S,MC}}{N_{MC}} \tag{4.11}$$

where $n_{S,MC}$ is the number of signal events returned by the full 2D-UEML-fit (including all event types) to our signal MC sample consisting of N_{MC} events. Multiplied by the efficiency correction factor described in section 4.5 we obtain the final efficiency ϵ. We use the same correction factor for the $\psi(2S)$-benchmark modes as for our signal modes. Fig. 4.7 shows the projections of the fit and Tab. 4.5 lists the numbers including errors for our two signal modes and the $\psi(2S)$ benchmark modes. It also shows the number of events in the input dataset for the fitter, which is obtained by the procedure described

[11] Since this has been added after unblinding and the final fit on data, we use a pure Lorentzian for the $\psi(2S)$ mass-fit, since this is the shape we decided to use beforehand. We also did not have a peaking background sample at hand for this mode.

in section 4.4.1. Note, that the numbers in this table are different from those shown in Tab. 4.3, which include the optimized cuts on m_{ES} and m_X.

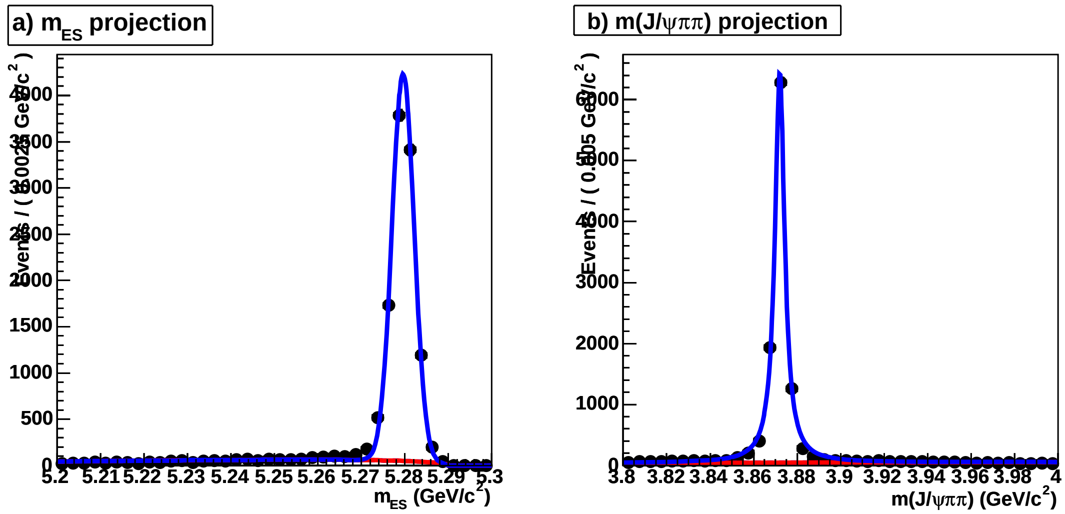


Figure 4.7: Result of the full 2D-UEML-fit (including all event types) for pure signal MC in $B^0 \to X(3872)K^0_S,\ X \to J/\psi\,\pi^+\pi^-$. The fit result is $n_S = 10652 \pm 104$, $n_P = 0.0 \pm 1.5$ and $n_C = 1904 \pm 46$. The number of signal events n_S is used to calculate our efficiency for reconstructing signal events (see Tab. 4.5).

	$B^0 \to XK^0_S$	$B^- \to XK^-$	$B^0 \to \psi(2S)K^0_S$	$B^- \to \psi(2S)K^-$
N_{MC}	57,000	57,000	30,000	30,000
events after cuts	12,556	15,786		
$n_{S,MC}$	$10,652 \pm 104$	$13,338 \pm 117$	$4,922 \pm 72$	$6,375 \pm 83$
efficiency ϵ_{MC} (%)	18.69 ± 0.18	23.40 ± 0.21	16.41 ± 0.24	21.25 ± 0.28
corrected eff. ϵ (%)	17.35 ± 0.17	22.19 ± 0.19	15.23 ± 0.23	20.15 ± 0.28

Table 4.5: Fit efficiency ϵ_{MC} for the two signal and benchmark modes determined on signal MC and the corrected efficiency ϵ.

4.4.4 Fit validation

Monte Carlo experiments

In this section we describe the validation studies done, to ensure that the full (2-dimensional) UEML-fit is returning the correct number of signal events. Those studies have been performed on a mix of signal and background MC with about the same number of events as expected on data.

Since the $B\overline{B}$ generic sample is roughly four times larger than the on-peak data, we divide this sample into four sets. The initial signal sample (57,000 events) is split into 4×50 sets with each 250 events and added to the $B\overline{B}$ sample. Eventually, we conduct 200 experiments including fits, each with an independent set of signal events and experiments within the same set sharing the same $B\overline{B}$ sample. The fits are done with floating event yields and all other parameters fixed to the values in Tab. 4.4.

Using the fit-efficiencies from Tab. 4.5 we calculate the number of input signal events $n_{S,input}$ and compare with the average number of signal events returned by the fit in each set. Fig. 4.8 shows one of the fits in $B^0 \to J/\psi \pi^+\pi^- K^0_S$. Plots a) and b) are m_{ES} and m_X projections whereas in plots c) and d) an additional cut (the optimized signal region cut) on m_X is applied when projecting m_{ES} and vice versa to enhance the visibility of the signal. Finally, we conduct another set of 50 experiments with the full $B\overline{B}$ sample as background and four times the amount of signal events. Scaled to the number of signal events in sets 1-4, we get another average value with smaller errors due to the higher statistics. To validate the fit performance in case of no signal, we repeat the above procedure without the signal MC sample.

Tab. 4.6 summarizes the result for all sets and both modes. We give the average number of events of type t $< n_t >$ and the average error $< \sigma_t >$ on this number over the 50 experiments in each set. Note, that the naive scaling of the error by $1/\sqrt{50}$ cannot be applied in this case, since all the experiments within the same set share the same $B\overline{B}$ sample and are therefore correlated. However, in all cases the average number of signal events in each set is within (the average) errors of the number of input signal events. Repeating the same test without the $B\overline{B}$ sample, and thus uncorrelated samples, the number of signal events is consistent with the expected error of $< \sigma_S > /\sqrt{50}$.

We also repeated part of those experiments, with a floating mass mean value for the

m_X fit, and obtain results compatible within errors to the above and to the number of expected signal events. This is important since we will use a floating mean value in the final fit on data to measure the mass of the $X(3872)$.

$< n_t > \pm < \sigma_t >$	$B^0 \to J/\psi\pi^+\pi^- K^0_S$		$B^- \to J/\psi\pi^+\pi^- K^-$	
$n_{S,input}$	46.7 ± 0.5	0	58.5 ± 0.5	0
Set 1				
signal events n_S	52.0 ± 8.9	3.3 ± 4.1	50.7 ± 11.4	0.0 ± 2.1
peaking events n_P	108.4 ± 14.7	111.9 ± 14.3	448 ± 29.1	441.0 ± 27.4
comb. events n_C	648.6 ± 27.0	638.7 ± 26.8	2713.5 ± 55.1	2702.1 ± 54.9
Set 2				
signal events n_S	46.7 ± 8.8	0.0 ± 1.4	63.8 ± 12.0	3.4 ± 7.9
peaking events n_P	109.8 ± 15.0	112.0 ± 13.9	430.7 ± 28.6	434.9 ± 54.5
comb. events n_C	711.3 ± 28.2	700.9 ± 28.0	2672.1 ± 54.6	2661.0 ± 54.5
Set 3				
signal events n_S	51.6 ± 9.1	0.0 ± 3.1	59.4 ± 11.9	0.0 ± 3.5
peaking events n_P	107.0 ± 14.4	111.8 ± 13.5	444.5 ± 29.0	444.3 ± 27.5
comb. events n_C	633.9 ± 26.6	625.2 ± 26.4	2638.0 ± 54.4	2626.8 ± 54.2
Set 4				
signal events n_S	49.5 ± 8.6	1.4 ± 3.7	77.3 ± 12.9	13.2 ± 9.2
peaking events n_P	106.3 ± 14.2	108.7 ± 13.8	462.5 ± 29.7	469.6 ± 29.5
comb. events n_C	666.7 ± 27.2	656.9 ± 27.0	2778.9 ± 55.8	2767.1 ± 55.7
Full $B\overline{B}$ sample				
signal events n_S	49.7 ± 4.4	0.0 ± 1.8	63.4 ± 6.1	0.0 ± 1.8
peaking events n_P	107.9 ± 7.3	112.3 ± 7.0	446.2 ± 14.6	451.6 ± 14.1
comb. events n_C	665.3 ± 13.6	655.5 ± 13.5	2700.6 ± 27.5	2689.2 ± 27.4

Table 4.6: Number of events n_t and error σ_t of type t returned by the fit averaged over 50 experiments in each set. Each experiment has a different signal sample with $n_{S,input}$ signal events and experiments within the same set share one $B\overline{B}$ background sample. Finally, another set of 50 experiments is conducted with the full background samples and results scaled to the size of sets 1-4.

Toy Monte Carlo

As a further test of our fit, we conduct 1000 toy-MC experiments. Each experiment consists of $N_{toy} = n_{S,toy} + n_{P,toy} + n_{C,toy}$ events sampled from our total PDF $\mathcal{P}$ with all parameters fixed to the values in Tab. 4.4 and $n_{t,toy}$ set to the yields obtained from the MC experiments (full $B\overline{B}$ sample) in the previous section (see Tab. 4.6). Each sample

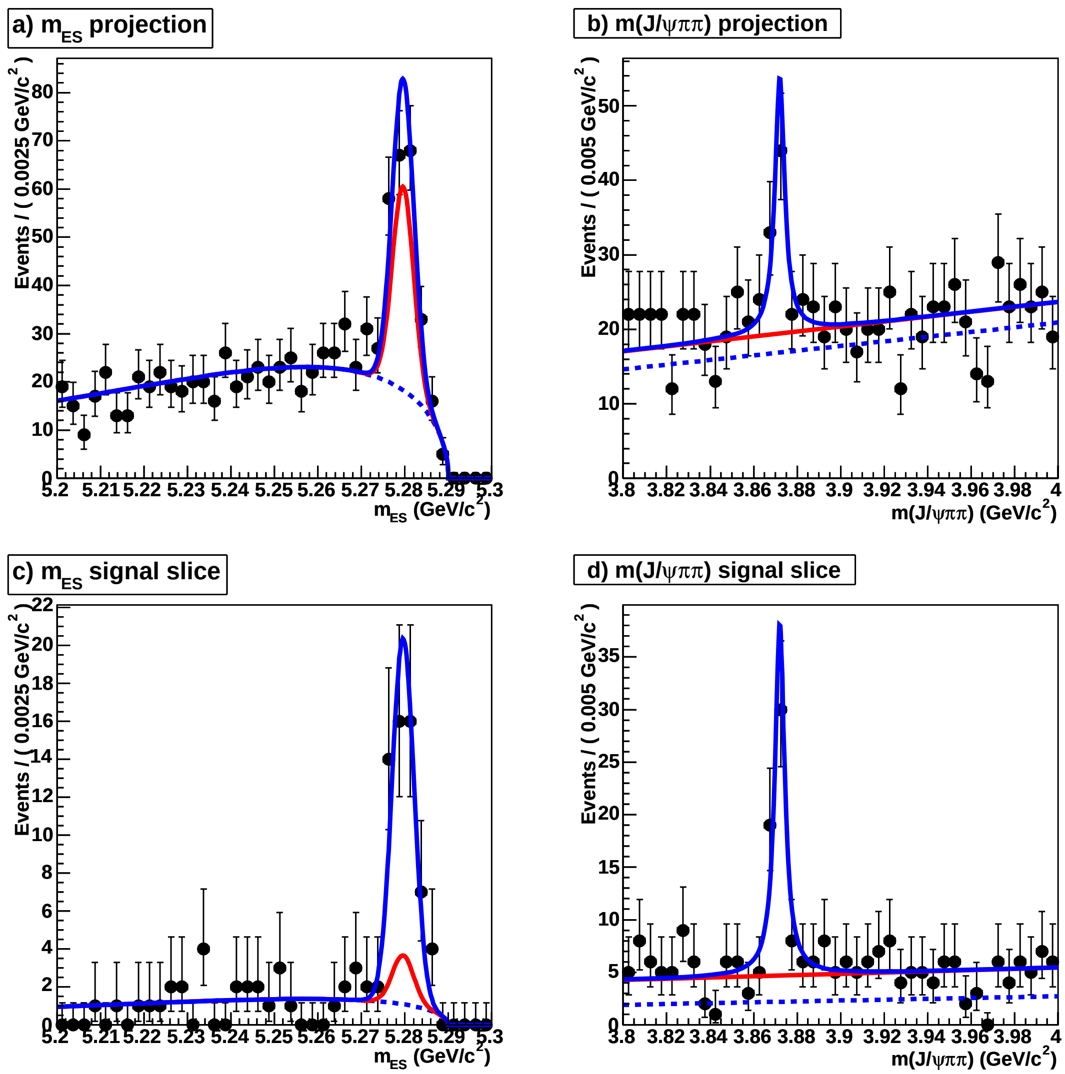


Figure 4.8: The fit result of an arbitrary experiment in $B^0 \to J/\psi\pi^+\pi^-K^0_S$. The lines represent (from bottom to top) combinatorial background (dotted, blue), peaking background (solid, red) and signal events (solid, blue). a) and b) are projections on m_{ES} and m_X, respectively. Plots c) and d) show the signal band projection with an additional cut on the signal region of the variable not shown.

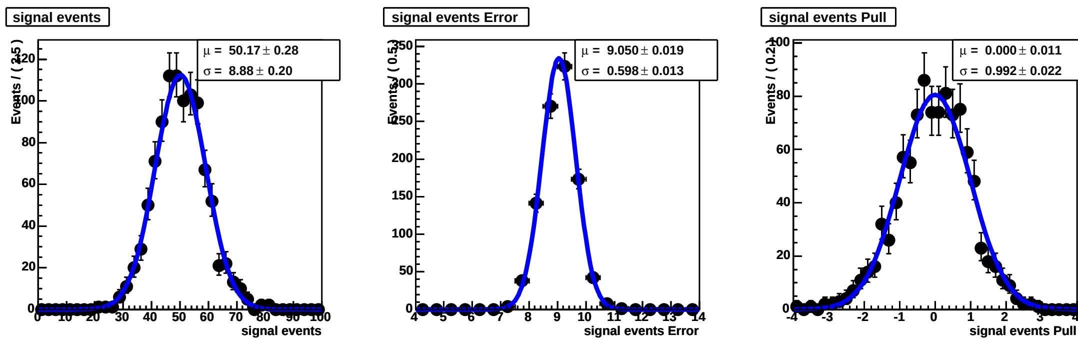


Figure 4.9: Result of 1000 toy-MC experiments for $B^0 \to J/\psi\,\pi^+\pi^-K^0_S$ with $n_{S,toy} = 50$. The left plot shows the number of signal events n_S returned by each fit, in the middle plot we can see the distribution of errors on n_S and the right plot shows the resulting pull distribution.

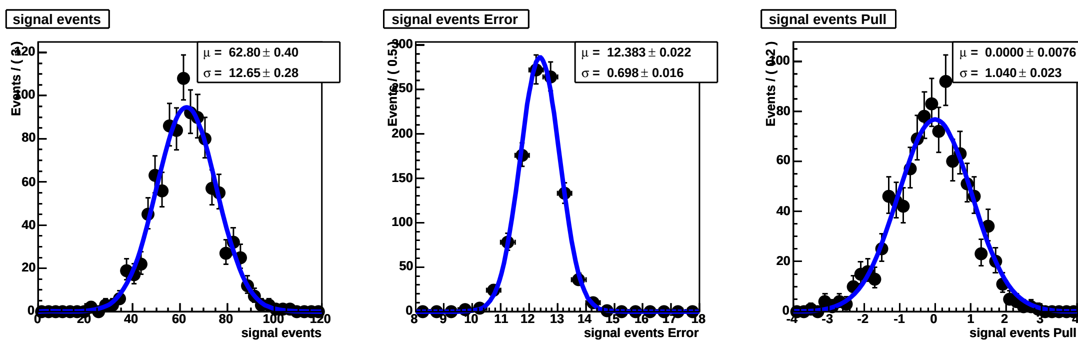


Figure 4.10: Result of 1000 toy-MC experiments for $B^- \to J/\psi\,\pi^+\pi^-K^-$with $n_{S,toy} = 63$. The left plot shows the number of signal events n_S returned by each fit, in the middle plot we can see the distribution of errors on n_S and the right plot shows the resulting pull distribution.

is fitted under the same conditions as in the MC experiments with floating event yields and fixed parameters. Fig. 4.9 shows the number of signal events n_S (left) and error σ_S (middle) returned by each fit for $B^0 \to J/\psi\,\pi^+\pi^-K^0_S$. The right plots shows the pull distribution $pull_S = (n_S - n_{S,toy})/\sigma_S$ of the number of signal events. The fit to a Gaussian shows that all distributions follow a normal distribution and the pull follows a unit normal distribution as expected. Fig. 4.10 shows the same for $B^- \to J/\psi\,\pi^+\pi^-K^-$. As in the previous MC experiments, we can conclude that the fit performs well and that there is no notable bias resulting from the choice of our event shapes.

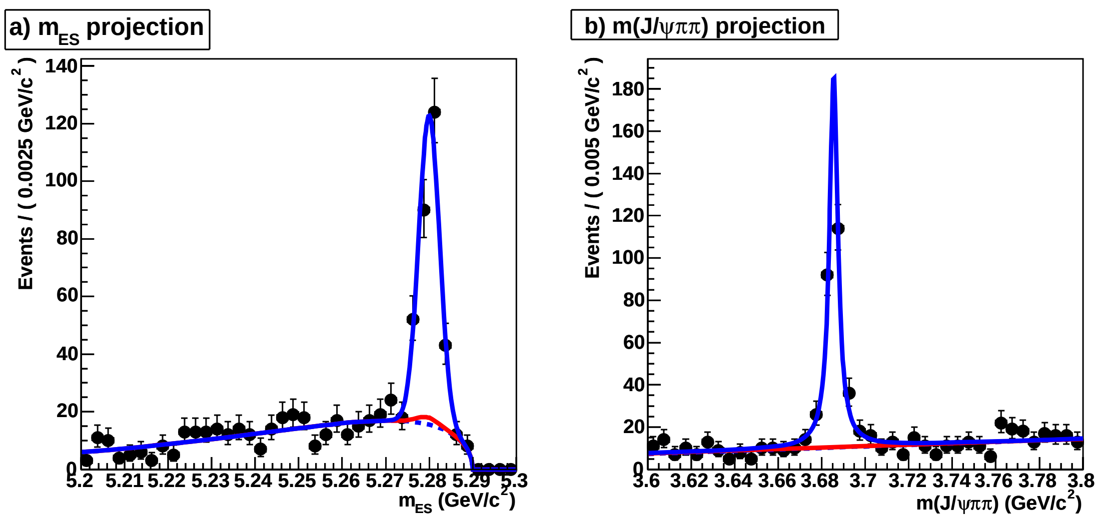


Figure 4.11: The fit result on $210.6\,\mathrm{fb}^{-1}$ of data for the benchmark mode $B^0 \to \psi(2S)K^0_S$. The lines represent (from bottom to top) combinatorial background (dotted, blue), peaking background (solid, red) and signal events (solid, blue).

Benchmark mode - $\psi(2S)$ fit

The decay $\psi(2S) \to J/\psi\,\pi^+\pi^-$ produced in $B \to \psi(2S)K$ can be used as a benchmark for this analysis. It is only $190\,\mathrm{MeV}/c^2$ below the X-mass, very narrow and provides us a well measured benchmark mode with comparably high statistics. Therefore, we perform the exact same 2D UEML-fit as described in the previous sections, but in the mass region $3.6 < m(J/\psi\,\pi^+\pi^-) < 3.8\,\mathrm{GeV}/c^2$. We use a Lorentzian for the $\psi(2S)$ and float all parameters, except the m_{ES} mean and width for the peaking background component, which we fix to the values obtained from our MC studies for the $X(3872)$. Table 4.7 lists the final values of the fit and the event yields (we omit the combinatorial shape parameters).

We realize that the fitted $\psi(2S)$ mass is shifted downwards compared to the world average of $3686.09\,\mathrm{MeV}/c^2$ [8]. We will use this fact, to correct the fitted X-mass. Table 4.8 summarizes the shift of the $\psi(2S)$ mass.

Using the (corrected) fit-efficiencies from Tab. 4.5 we can calculate the number of

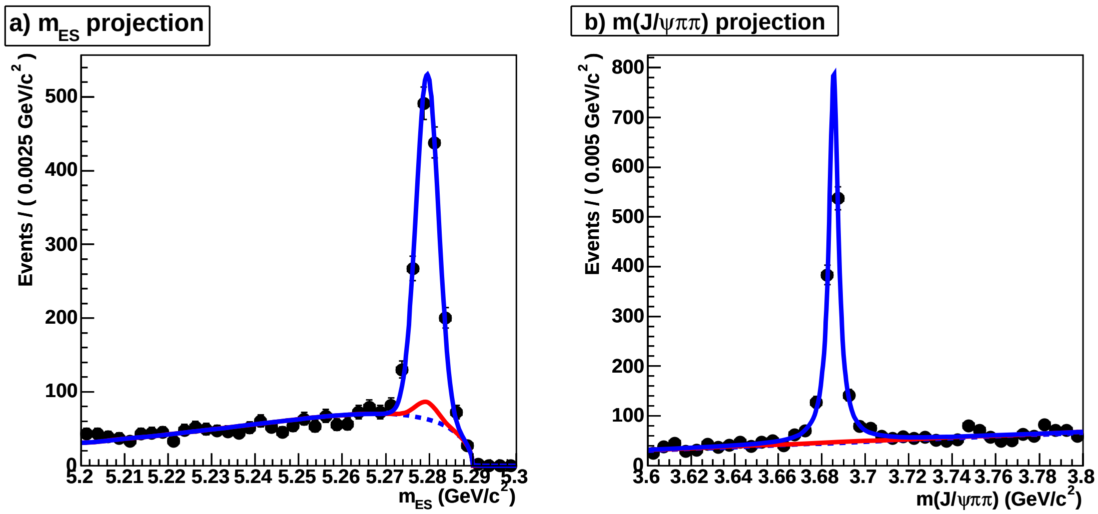


Figure 4.12: The fit result on 210.6 fb^{-1} of data for the benchmark mode $B^- \to \psi(2S)K^-$. The lines represent (from bottom to top) combinatorial background (dotted, blue), peaking background (solid, red) and signal events (solid, blue).

	$B^0 \to \psi(2S)K^0_S$	$B^- \to \psi(2S)K^-$
Parameters		
m_X mean (GeV/c^2)	3.6855 ± 0.0002	3.6855 ± 0.0001
m_X FWHM (MeV/c^2)	4.66 ± 0.46	5.07 ± 0.24
m_{ES} mean (GeV/c^2)	5.2800 ± 0.0001	5.2796 ± 0.0001
m_{ES} width (MeV/c^2)	2.40 ± 0.13	2.60 ± 0.07
m_X parameter c_{comb}	-0.29 ± 0.01	-0.29 ± 0.01
m_X parameter c_{peak}	-0.26 ± 0.02	-0.29 ± 0.02
m_{ES} ARGUS κ	-67.9 ± 6.3	-35.3 ± 2.4
Event yields		
signal n_S	$\mathbf{252.5 \pm 17.1}$	$\mathbf{1159.3 \pm 37.2}$
peaking n_P	6.1 ± 11.2	58.3 ± 24.5
comb. n_C	434.4 ± 23.0	1916.5 ± 48.2
Expected yield		
$n_{S,PDG}$	$\mathbf{282.5 \pm 33.9}$	$\mathbf{1189.3 \pm 84.5}$

Table 4.7: Fit results on 210.6 fb^{-1} of on-peak data for the $\psi(2S)$-benchmark mode with the measured and expected signal yields (bold) calculated from Eq. 4.12

$\psi(2S)$ mass (MeV/c^2)	$B^0 \to \psi(2S)K^0_S$	$B^- \to \psi(2S)K^-$
fitted value m_{fit}	3685.5 ± 0.2	3685.5 ± 0.1
PDG value m_{pdg}	3686.09 ± 0.03	3686.09 ± 0.03
$m_{fit} - m_{pdg}$	-0.59 ± 0.20	-0.59 ± 0.10

Table 4.8: Fitted mass of the $\psi(2S)$ on data and the difference to the world average. We will use this as a reference for the fitted X-mass.

expected $\psi(2S)$ signal events

$$n_{S,PDG} = \epsilon\, N_{B\overline{B}}\, \mathcal{B}(B \to \psi(2S)K)\, \mathcal{B}(\psi(2S) \to J/\psi \pi^+\pi^-)\, \mathcal{B}(J/\psi \to l^+l^-) \quad (4.12)$$

$$\times \left[\mathcal{B}(K^0_S \to \pi^+\pi^-)\, \mathcal{B}(K^0 \to K^0_S)\right] \quad (4.13)$$

where the secondary branching ratios are taken from the PDG [8] and the factor [...] only applies to the B^0 mode. Comparing the measured and expected yields for n_S we conclude, that the yields agree within one standard deviation. Also note that the given errors only include statistical and errors due to secondary branching ratios.

4.4.5 Fit result

Due to some unexpected fluctuations of the data around 4 GeV/c^2 compared to the MC background prediction (see for example the right plot in Fig. 4.18), it was decided to do the final data fit in the range $3.8 < m_X < 3.95$ GeV/c^2 where the MC gives a good description of the background. All other cuts are the same as in the previous validations and MC experiments.

From the comparison of the m_{ES} parametrization between $\psi(2S)$-data and our MC samples, we can assume that our chosen parametrization matches the data in the $X(3872)$ region within errors. Therefore, we keep the mean and width of the signal and peaking background m_{ES} distributions fixed. The remaining three parameters (m_X parameters c in peaking/ combinatorial background and ARGUS shape) are floated.

For the $X(3872)$ line shape, we use a Lorentzian with floating central value. Due to limited statistics in the K^0_S mode we have to fix the width of the Lorentzian to the value obtained from signal MC in order to get a converging fit. The width in the K^- mode is kept floating. Tab. 4.9 shows the fit result for both modes and Figs. 4.13 and 4.14 show the m_{ES} (a) and m_X (b) projections of the fit. The lower two plots show the same projection but with an additional cut on m_X around the signal region when plotting m_{ES} and vice versa.

Figure 4.15 shows the projection of the negative log-likelihood $-\log(L/L_{max})$ on the number of signal events n_S. We obtain the statistical significance of the signal(not including systematics)x using $\sqrt{-2\log(L(n_S=0)/L_{max})}$. With a NLL-value for zero signal events of 3.38 and 28.1 we obtain a statistical significance (not including any systematic errors) of 2.6σ and 7.5σ for the K^0_S and K^- mode, respectively.

	$B^0 \to X(3872)K^0_S$	$B^- \to X(3872)K^-$
Parameters		
m_X mean (GeV/c^2)	3.8680 ± 0.0012	3.8707 ± 0.0006
m_X FWHM (MeV/c^2)	5.41	5.04 ± 2.23
m_X parameter c_{comb}	-0.14 ± 0.13	-0.31 ± 0.01
m_X parameter c_{peak}	-0.27 ± 0.01	-0.27 ± 0.01
m_{ES} ARGUS κ	-41.3 ± 5.1	-35.3 ± 2.4
Event yields		
signal n_S	8.3 ± 4.5	61.2 ± 15.3
peaking n_P	35.0 ± 11.5	244.8 ± 27.9
comb. n_C	619.6 ± 26.6	2788.0 ± 56.4

Table 4.9: Fit results on $210.6\,\mathrm{fb}^{-1}$ of on-peak data. Parameters not listed or with no errors are kept fixed to the values obtained from MC.

4.4.6 Invariant $J/\psi\pi^+\pi^-$ mass spectrum

Figures 4.16 and 4.17 show the $J/\psi\pi^+\pi^-$ invariant mass spectrum for both modes with the optimized cuts applied for data and generic $B\overline{B}$ Monte Carlo. A clear signal of the $X(3872)$ in the K^- mode can be seen. There might be a 2σ enhancement at slightly lower mass in the K^0_S mode as well. As a further check, Fig. 4.18 shows the $J/\psi\pi^+\pi^-$

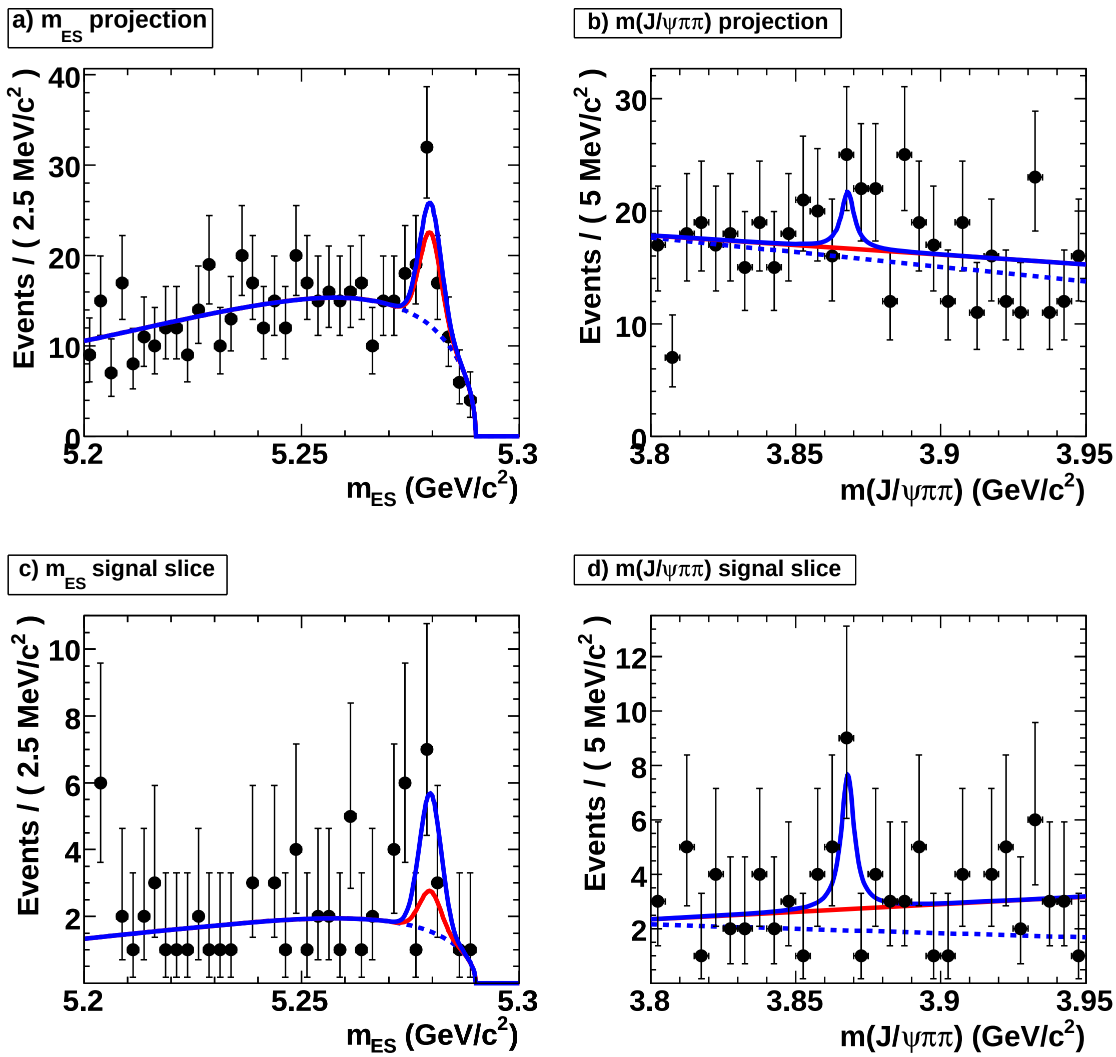


Figure 4.13: The fit result on 210.6 fb^{-1} of data in $B^0 \to X(3872)K_S^0$. The lines represent (from bottom to top) combinatorial background (dotted, blue), peaking background (solid, red) and signal events (solid, blue). a) and b) are projections on m_{ES} and m_X, respectively. Plots c) and d) show the signal band projection with an additional cut on the signal region of the variable not shown.

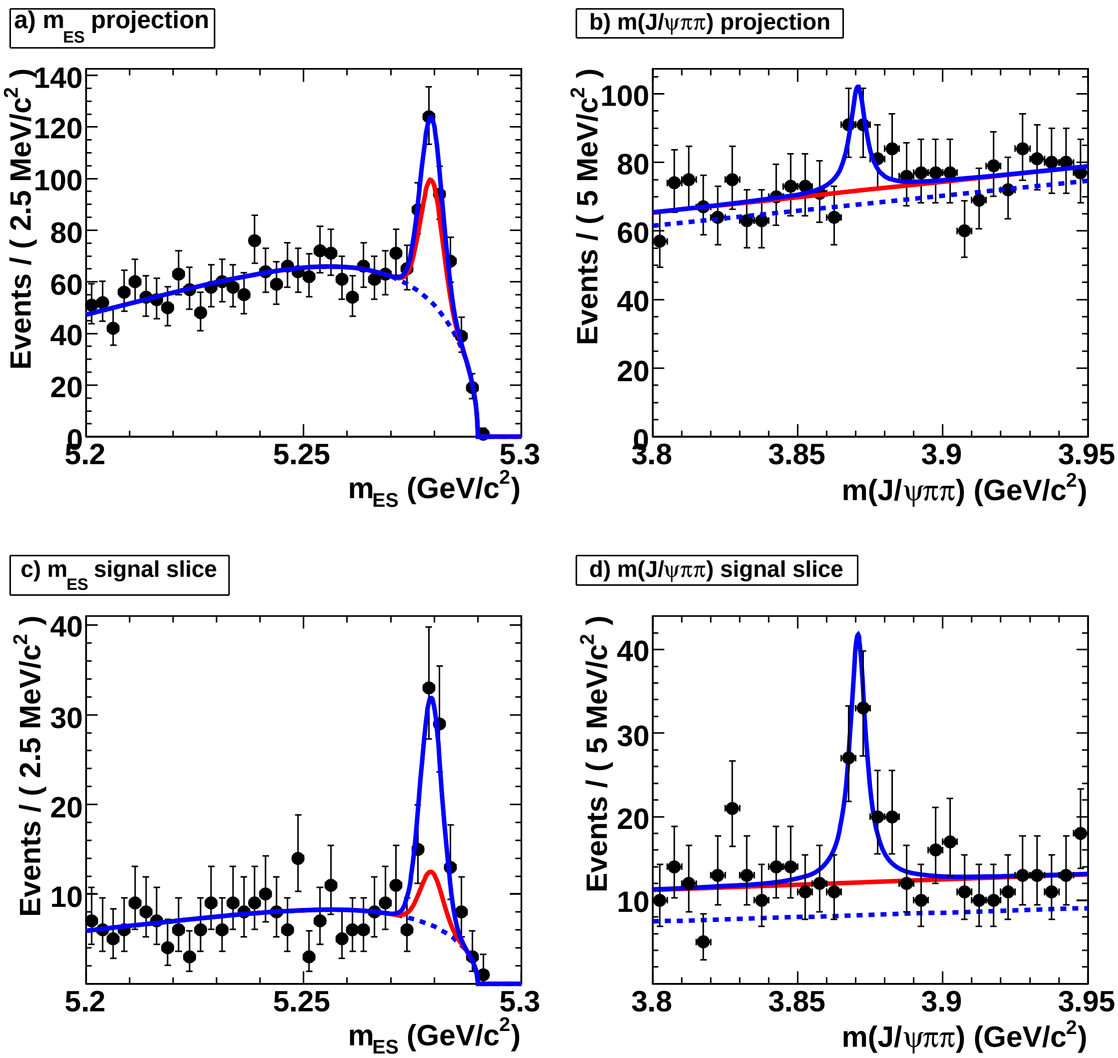


Figure 4.14: The fit result on $210.6\,\mathrm{fb}^{-1}$ of data in $B^- \to X(3872)K^-$. The lines represent (from bottom to top) combinatorial background (dotted, blue), peaking background (solid, red) and signal events (solid, blue). a) and b) are projections on m_{ES} and m_X, respectively. Plots c) and d) show the signal band projection with an additional cut on the signal region of the variable not shown. Note the different bin size in m_X compared to Fig. 4.13.

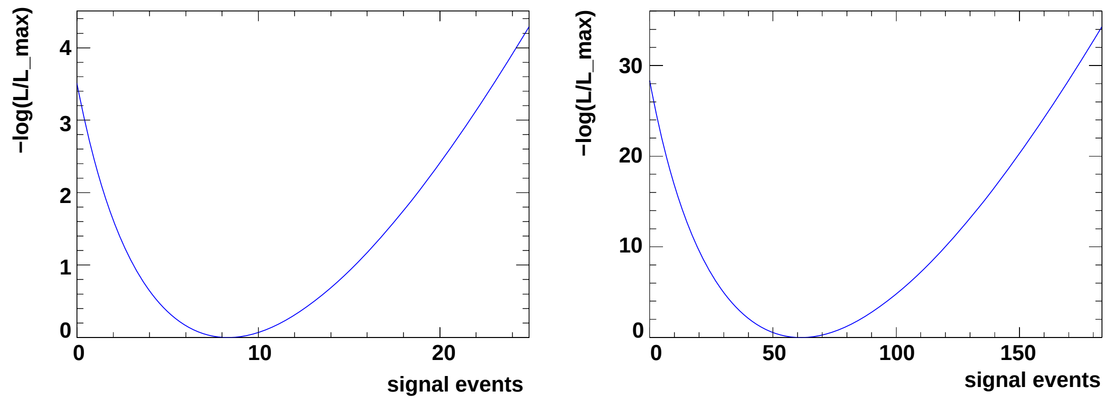


Figure 4.15: Projection of the negative log-likelihood of the data fit on the number of signal events n_S. The left plots shows $B^0 \to X(3872)K^0_S$ and the right plot shows the same for $B^- \to X(3872)K^-$.

invariant mass together with a m_{ES} sideband ($|m_{ES} - 5.26| < 0.006$). We can see that the background behaves "nicely" in the X-region and that there is an excess of events in the around $3.87\,\mathrm{GeV}/c^2$ in both modes.

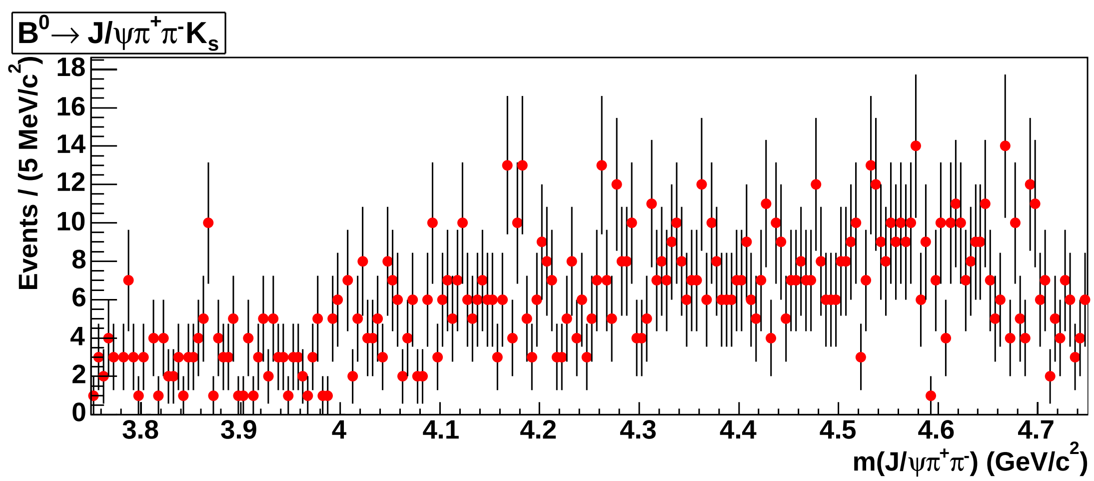


Figure 4.16: $J/\psi\,\pi^+\pi^-$ invariant mass spectrum between 3.75 and $4.75\,\mathrm{GeV}/c^2$ in $5\,\mathrm{MeV}/c^2$ bins with the optimized cuts applied for $B^0 \to J/\psi\,\pi^+\pi^- K^0_S$. The dots represent the data and the solid line is the prediction from generic $B\overline{B}$ Monte Carlo.

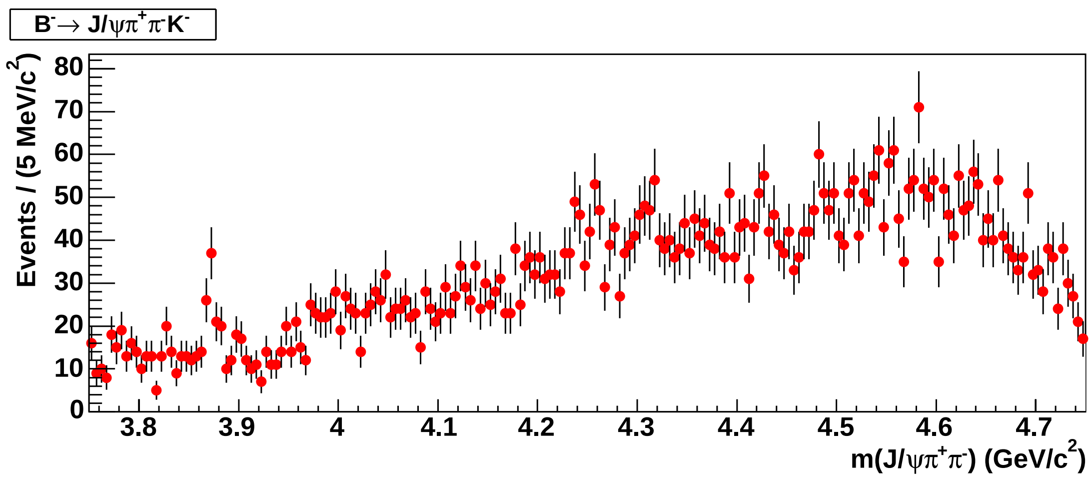


Figure 4.17: $J/\psi\,\pi^+\pi^-$ invariant mass spectrum between 3.75 and 4.75 GeV/c^2 in 5 MeV/c^2 bins with the optimized cuts applied for $B^- \to J/\psi\,\pi^+\pi^- K^-$. The dots represent the data and the solid line is the prediction from generic $B\overline{B}$ Monte Carlo.

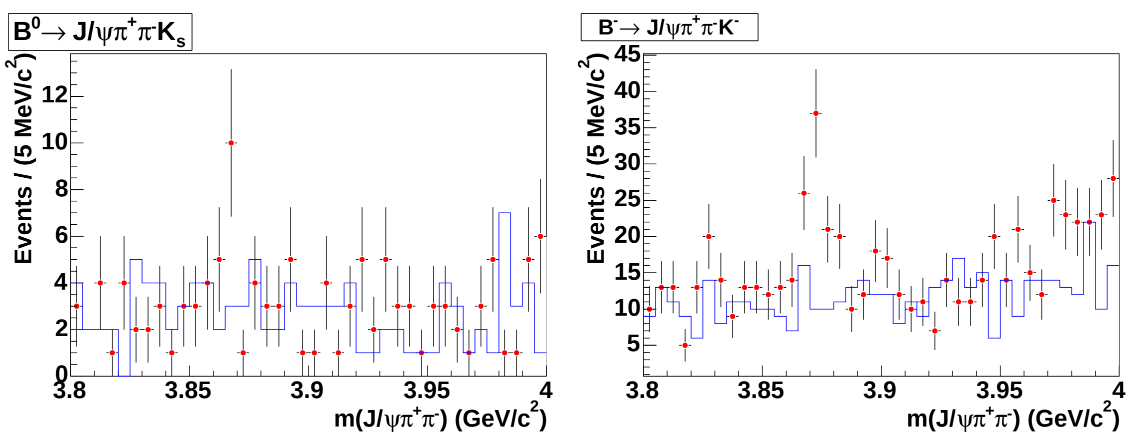


Figure 4.18: $J/\psi\,\pi^+\pi^-$ invariant mass spectrum between 3.8 and 4.0 GeV/c^2 in 5 MeV/c^2 bins. The dots represent the data and the solid line shows an m_{ES} sideband. The left plot shows $B^0 \to J/\psi\,\pi^+\pi^- K^0_S$ and the right plot shows $B^- \to J/\psi\,\pi^+\pi^- K^-$.

4.5 Efficiency corrections and systematic errors

Since particle ID, tracking and neutrals reconstruction have a slightly different performance on Monte Carlo and data, corrections are applied to the efficiency obtained from MC to match the efficiency expected on data. With each of those corrections comes an error that we include in our overall systematic error. Efficiency correction factor given here are meant to be multiplied with the MC efficiency. We also discuss other sources of systematic errors. See tables 4.10 and 4.11 for a summary of all numbers.

4.5.1 Number of $B\overline{B}$ events

The number of $B\overline{B}$ events obtained from B-counting is assigned the standard systematic (fractional) error of 1.10%.

4.5.2 Secondary branching ratios

The secondary branching ratios and their errors (fractional errors in brackets) we use in our calculations are [8]

$$\mathcal{B}(J/\psi \to l^+l^-) = \mathcal{B}(J/\psi \to e^+e^-) + \mathcal{B}(J/\psi \to \mu^+\mu^-) \tag{4.14}$$

$$= (11.81 \pm 0.20) \times 10^{-2} \quad (1.69\%) \tag{4.15}$$

$$\mathcal{B}(K^0_S \to \pi^+\pi^-) = (68.95 \pm 0.14) \times 10^{-2} \quad (0.39\%) \tag{4.16}$$

We assume fully correlated errors in the leptonic branching ratios of the J/ψ. Furthermore, we assign an error for the uncertainty in the production rate of B^0 and B^+ from the $\Upsilon(4S)$ meson. The result of a recent *BABAR* measurement [76] for the production ratio is $R^{+/0} = 1.006 \pm 0.036 \pm 0.031$ which corresponds to a fractional error of 4.72%.

4.5.3 MC statistics

The fixed parameters in the final fit were obtained from limited statistics MC samples. To account for this, we vary each parameter i by $\pm 1\sigma$ and repeat the fit on data. From the new results $n^{\pm}_{S,i}$ we calculate the fractional systematic error as

$$\sigma^2_{syst} = \frac{1}{2}\sum_i \left[\left(\frac{n^{+}_{S,i}}{n_{S,0}} - 1\right)^2 + \left(\frac{n^{-}_{S,i}}{n_{S,0}} - 1\right)^2\right] \tag{4.17}$$

where $n_{S,0}$ is the number of signal events returned by the fit with all parameters fixed to the final parameters in the data fit. Table 4.11 lists the results.

4.5.4 Differences between Monte Carlo and data

For the parameters which we have fixed in the fit (m_{ES} mean and width for signal and peaking background), we correct for possible differences between the data and Monte Carlo. Therefore, we obtain an alternative set of parameters from some suitable sample (see below) and repeat the data fit with all other parameters fixed to their values from the initial data fit. Each of those fits, gives a new value $n_{S,i}$ for the number of signal events. We take

$$\sigma^2_{syst} = \sum_i \left(\frac{n_{S,i}}{n_S} - 1\right)^2 \tag{4.18}$$

as the fractional systematic error, where n_S is the result from our initial fit. For the signal m_{ES} shape, we use the parameters (m_{ES} mean and width) obtained from the $\psi(2S)$ fit (Tab. 4.7). For the peaking background shape, we fit the m_{ES} distribution in the mass region $3.7 < m_X < 4.5\,\mathrm{GeV}/c^2$ with a $10\,\mathrm{MeV}/c^2$ veto-cut around the $X(3872)$ and use those m_{ES} parameters as an alternative set. Table 4.11 lists the final systematic error. We did the same check for the central value of the fitted X-mass, but there was no notable change within $0.1\,\mathrm{MeV}/c^2$ due to different m_{ES} shapes. The same is true for the $\psi(2S)$-fit which we use as a reference.

4.5.5 PID corrections

We are using the efficiency correction tables provided by the PID group to assign each signal-MC event a weight w_{PID} that is the product of the individual PID selector weights. We repeat the procedure described in section 4.4.3 with those weighted events and obtain the PID-corrected efficiency and the resulting correction factor which can be found in Tab. 4.10. We estimate a very conservative systematic error of 5% for the efficiency corrections due to PID.

4.5.6 Tracking corrections

The tracking corrections to the efficiency including the systematic errors are taken from the recipes provided by the Track Efficiency Task Force [77].

- **Corrections for $\pi^{\pm}, K^{\pm}, l^{\pm}$**

  All of our charged tracks (pions, kaons, leptons) are taken from lists that are derived from `ChargedTracks`. In this case an efficiency correction of 0.25% with a systematic error of 1.2% per track is recommended. The total efficiency correction factor for a mode with n tracks (not counting tracks resulting from $K_S^0 \to \pi^+\pi^-$) is therefore 0.9975^n (multiplied) with a systematic error of $1.2\% \times n$ (fully correlated).

- **K_S^0 correction**

  The K_S^0 correction is treated separately from the other tracking corrections and is obtained from the appropriate efficiency correction tables. We are using the `3DSign3_noAlpha` tables and apply them to our signal MC sample. Using the provided root macros for run 1-3 and run 4 data we obtain the luminosity weighted average for the efficiency correction. The result for $B^0 \to J/\psi\,\pi^+\pi^- K_S^0$ is 0.977 ± 0.016 and we include the fractional error of 1.64% in our systematics.

4.5.7 Monte Carlo model

The model for the $X(3872)$ used in our signal Monte Carlo generates the decays according to two-body phase space with flat angular distributions. To investigate the effect of different decay models we weight each event with different angular distributions and recalculate the efficiencies. We consider two possible assignments for the $X(3872)$:

- If the X is a DD^* molecule the most likely quantum numbers are $J^{PC} = 1^{++}$ [47].
- For charmonium, we consider the 2^{--} 1^3D_2 as a possible candidate.

Pakvasa and Suzuki [42] calculate the angular distributions of $\rho \to \pi\pi$ for those two different quantum numbers:

$$d\Gamma/d\cos\theta_\pi \sim \cos^2\theta_\pi \sin^2\theta_\pi \qquad (J^{PC} = 2^{--}) \tag{4.19}$$

$$d\Gamma/d\cos\theta_\pi \sim \sin^2\theta_\pi \qquad (J^{PC} = 1^{++}) \tag{4.20}$$

where θ_π is the angle between the pion and the recoiling X in the restframe of the ρ. We assign the difference between the maximum and minimum efficiency as a systematic error. From this method we obtain a fractional systematic error of 1.56% and 1.00% for the B^0 and B^- mode, respectively.

	$J/\psi\,\pi^+\pi^-K^0_S$	$J/\psi\,\pi^+\pi^-K^-$
Number of tracks	4	5
Particle ID	0.9591	0.9603
Tracking	0.9900	0.9876
K^0_S correction	0.9777	-
Total (multiplied)	0.9284	0.9484

Table 4.10: Efficiency correction factors applied to the raw efficiency obtained from signal MC samples.

	$J/\psi\,\pi^+\pi^- K^0_S$	$J/\psi\,\pi^+\pi^- K^-$
No. of $B\overline{B}$ events	1.10	1.10
Branching ratios	5.02	5.02
MC statistics	1.92	0.68
MC decay model	1.56	1.00
MC-data difference	8.94	1.77
Particle ID	5.00	5.00
Tracking	4.80	6.00
K^0_S correction	1.64	-
Total (quadrature)	12.77	9.59

Table 4.11: Summary of systematic errors in %.

4.6 Study of $J/\psi\,\pi^+\pi^-$ invariant mass above $4\,\mathrm{GeV}/c^2$

Recent observations by *BABAR* [78] in initial state radiation (ISR) events provide evidence for at least one broad resonance in the invariant mass spectrum of $J/\psi\,\pi^+\pi^-$ at $4.259\,\mathrm{GeV}/c^2$ that can be characterized by a single resonance with a full width of $88\,\mathrm{MeV}/c^2$. This structure is referred to as $Y(4260)$. Alternatively, the data might support two narrow resonances at $4.26\,\mathrm{GeV}/c^2$ and $4.33\,\mathrm{GeV}/c^2$. This section describes studies of the $J/\psi\,\pi^+\pi^-$ invariant mass above $4\,\mathrm{GeV}/c^2$ and the search for the $Y(4260)$ in B-decays. Furthermore, we investigate the possibility of more than one resonance.

4.6.1 Monte Carlo studies

To study the region above $4\,\mathrm{GeV}/c^2$ in $m(J/\psi\,\pi^+\pi^-)$ we perform another MC experiment. From the $X(3872)$ signal MC, we create a MC sample with a signal at $4.264\,\mathrm{GeV}/c^2$ ($Y(4264)$) and a natural width of $90\,\mathrm{MeV}/c^2$. Furthermore we create another (narrow) state at $4.315\,\mathrm{GeV}/c^2$ ($Z(4315)$). No changes in the selection criteria compared to the $X(3872)$ were found after optimization.

The fit behavior is evaluated by adding a fixed amount from each of the newly created signal MC samples to the $B\overline{B}$ generic MC sample. We perform a fit in the mass region $4.15 < m(J/\psi\,\pi^+\pi^-) < 4.4\,\mathrm{GeV}/c^2$ with two Breit-Wigner components for the m_X signal PDF. Figure 4.19 shows the fit result on a MC cocktail consisting of $B\overline{B}$ generic MC, $Y(4264)$ and $Z(4315)$ and Fig. 4.20 shows the same but without the additional narrow resonance at $4.3\,\mathrm{GeV}/c^2$. For the first experiment with both resonances, we obtain a fraction $f_Y = (86.7 \pm 5.4)\%$ and the fit to the second sample (only one resonance) yields $f_Y = (94.1 \pm 4.2)\%$, where f_Y is the fraction of $Y(4264)$ events. The initial MC cocktails consisted of $f_Y = 96.8\%$ and $f_Y = 100\%$ Y-events, respectively. We generated a dedicated signal MC sample for a state at $4246\,\mathrm{MeV}/c^2$ similar to our signal MC samples used for the $X(3872)$. From this sample we obtain the detector resolution (again parameterized as a Lorentzian) $\Gamma_{res} = (5.4 \pm 0.1)\,\mathrm{MeV}/c^2$ that is exactly the same as for the $X(3872)$.

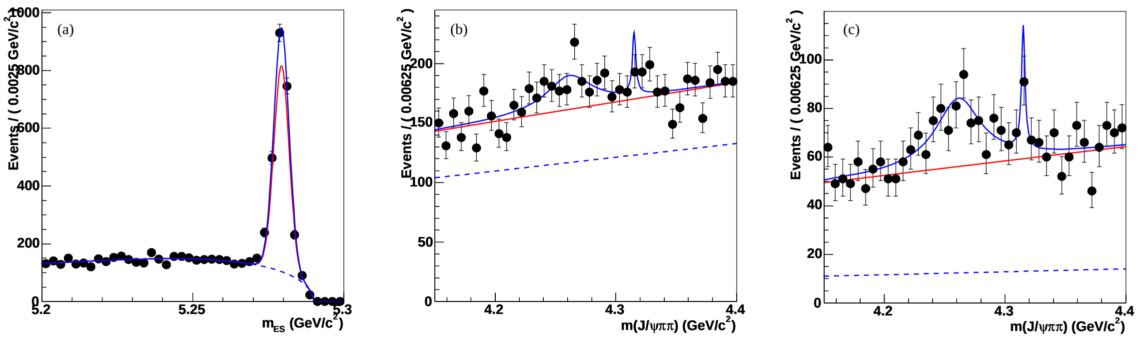


Figure 4.19: Fit to a MC cocktail of $B\overline{B}$ generic MC, $Y(4264)$ and $Z(4315)$. Plots a) and b) show the fit projections for m_{ES} and m_X whereas plot c) shows the m_X projection with a cut on the m_{ES} signal region applied.

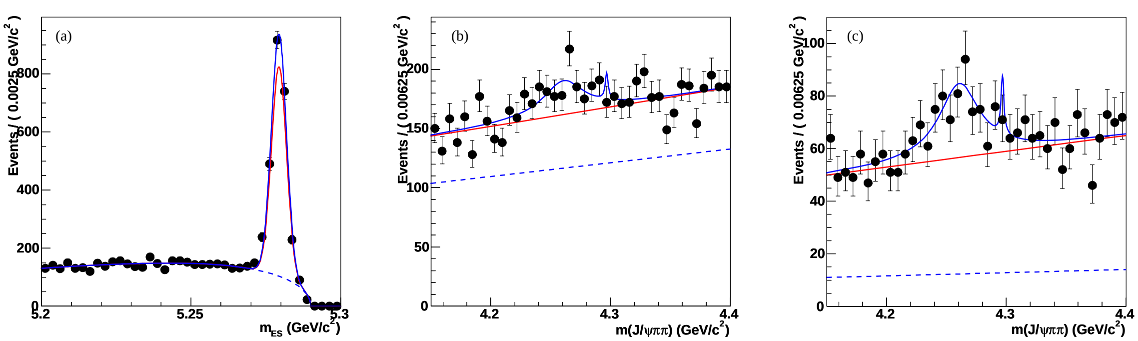


Figure 4.20: Fit to a MC cocktail of $B\overline{B}$ generic MC and $Y(4264)$ without the additional narrow resonance. Plots a) and b) show the fit projections for m_{ES} and m_X whereas plot c) shows the m_X projection with a cut on the m_{ES} signal region applied.

4.6.2 Fit on data (I)

With the same fit configuration as established in our MC studies we fit the data in the B^- mode. Our MC studies were carried out in the region above $4.15\,\mathrm{GeV}/c^2$. To meet up with the upper end of the mass region in our $X(3872)$ studies, we perform the fit on data in the region $4.0 - 4.45\,\mathrm{GeV}/c^2$ and show the fit projection in $15\,\mathrm{MeV}/c^2$ bins. Since the m_{ES} projection does not show any valuable information (except that there is a large amount of peaking background), we only show the m_X projections of the fit result. The mean and width of both m_X signal-PDFs are kept floating in the fit, as well as the slopes of the combinatorial and peaking background shapes in m_X. We perform the fit in four different configurations:

- I) two signal peaks in m_X
- II) no signal peak
- III) one peak (Y) below $4.3\,\mathrm{GeV}/c^2$
- IV) one peak (Z) above $4.3\,\mathrm{GeV}/c^2$

Figs. 4.21 and 4.22 show the fit projections and Tab. 4.12 lists the final values of the fit parameters. The numbers of signal events n_S, the central value of the peak m and the total width Γ are listed. Since the detector resolution is given by a Lorentzian (=Breit-Wigner) with full width $\Gamma_R = (5.4 \pm 0.1)\,\mathrm{MeV}/c^2$, we can obtain the natural width from $\Gamma = \Gamma_{\mathrm{fitted}} - \Gamma_R$. From the likelihood value $\mathcal{L}_0$ of the null hypothesis from configuration (II) we derive the significance of each of the other fit configurations by calculating $\sqrt{-2\ln\mathcal{L}_{\mathrm{max}}/\mathcal{L}_0}$. We also give the corresponding values from the ISR analysis [79].

The data seems to favor two separate states around $4.25\,\mathrm{GeV}/c^2$ and $4.31\,\mathrm{GeV}/c^2$. Both states have the same full width within errors and a significance of 2.9σ and 2.8σ, respectively. The lower state is at the same mass as the lower peak in the "two-peak interpretation" of the ISR result. However, the central values of the higher peaks differ by more than four standard deviations.

configuration	(I)	(II)	(III)	(IV)	(ISR)
$n_{S,Y}$	79 ± 32	-	77 ± 36	-	
$n_{S,Z}$	63 ± 25	-	-	58 ± 27	
$m_Y(\mathrm{MeV}/c^2)$	4246 ± 7	-	4248 ± 8	-	4254 ± 6
$m_Z(\mathrm{MeV}/c^2)$	4313 ± 4	-	-	4312 ± 4	4334 ± 1
$\Gamma_Y(\mathrm{MeV}/c^2)$	20 ± 9	-	22 ± 12	-	55 ± 18
$\Gamma_Z(\mathrm{MeV}/c^2)$	10 ± 5	-	-	10 ± 5	0 ± 4
$\sqrt{-2\ln\mathcal{L}_{\max}/\mathcal{L}_0}$	4.2	0.0	2.9	2.8	

Table 4.12: Fit results for the four different fit configurations: (I) two peaks, (II) no peak, (III) one peak below $4.3\,\mathrm{GeV}/c^2$ and (IV) one peak above $4.3\,\mathrm{GeV}/c^2$. The last column lists the results from the ISR analysis [79] in the two-peak fit configuration.

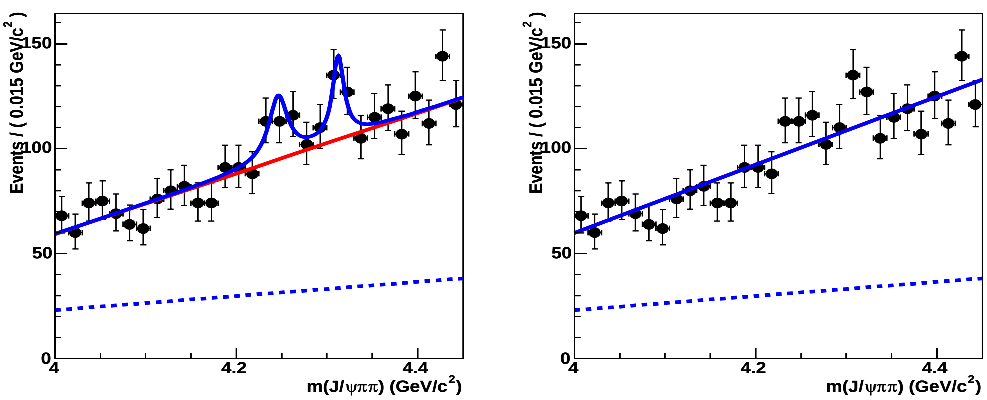


Figure 4.21: m_X projections of the fit result in configuration (I) (left) and the null hypothesis (II, right). See Tab. 4.12 for the fit parameters.

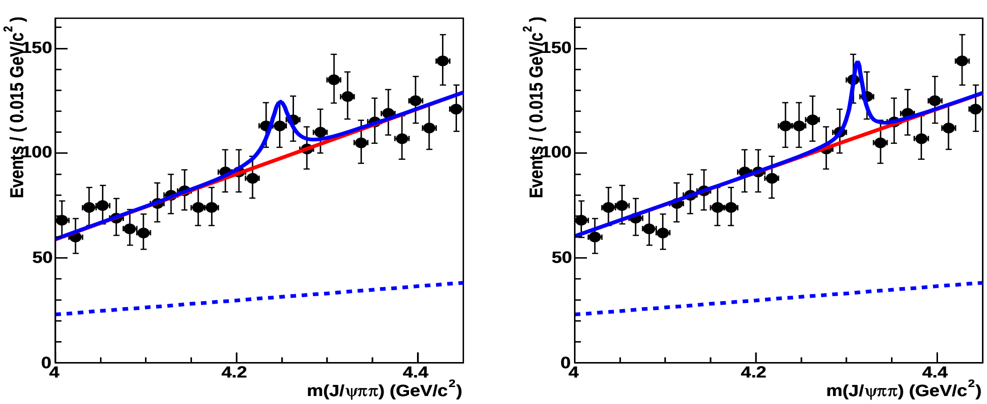


Figure 4.22: m_X projections of the fit result in configuration (III) (left) and (IV) (right). See Tab. 4.12 for the fit parameters.

4.6.3 Additional studies

After unblinding and with the help of appropriately generated signal MC for the $Y(4264)$, we are now able to perform some further studies in this mass region. First we investigate the angle cos2Pi, which is defined in the following way:

- cos2Pi is the cosine of the angle between the momentum direction of the $(\pi\pi)$-system and the B momentum direction in the $(J/\psi\,\pi\pi)$ restframe.

Fig. 4.23 shows the cos2Pi distribution of the $\psi(2S)$ as a reference. This angle is possibly

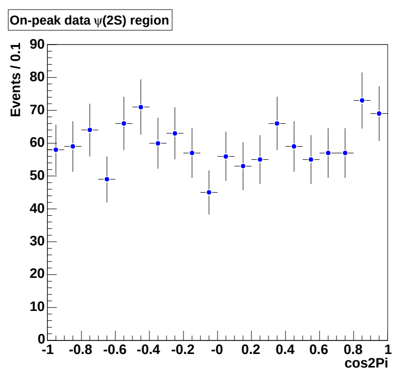


Figure 4.23: cos2Pi distribution for $\psi(2S)$ events in on-peak data.

helpful in suppressing backgrounds from $K_1(1270) \to K\rho$. decays. To investigate this, Fig. 4.24 shows plots of cos2Pi versus the $K\pi\pi$ invariant mass for signal MC (left), $B\overline{B}$ generic MC (middle) and data (right). All plots show events within the region $3.8 < m(J/\psi\,\pi\pi) < 4.5\,\mathrm{GeV}/c^2$. One can see a nice separation of the signal and K_1 background in this plane. Furthermore, Fig. 4.25 shows the same angle but versus the $J/\psi\,\pi^+\pi^-$ invariant mass.

We re-optimized our signal selection criteria including this angle. It turns out that the optimum is reached for the selection cos2Pi< 0.5. Fig. 4.26 shows the $J/\psi\,\pi^+\pi^-$ invariant with different cuts on this angle. It seems that the two-peak structure above $4.2\,\mathrm{GeV}/c^2$ is more pronounced above smaller backgrounds for cos2Pi> 0.5. Notice that the $X(3872)$ signal disappears with this cut applied. Fig. 4.27 shows the effect of the cos2Pi-cut on

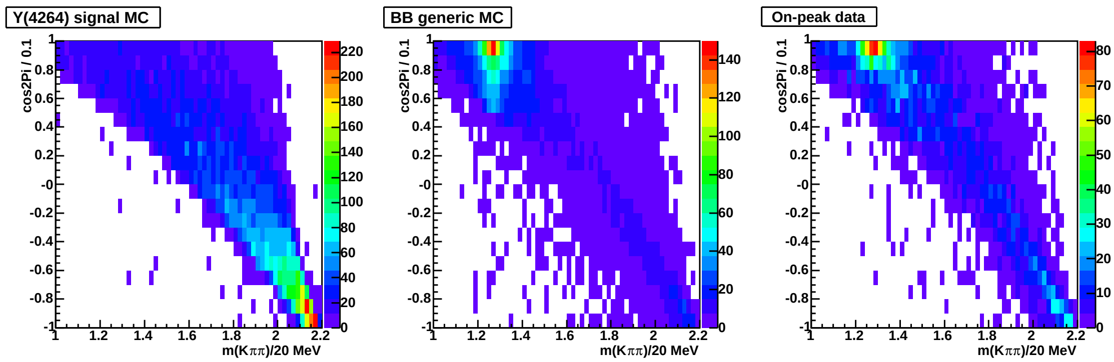


Figure 4.24: The angle cos2Pi versus the $K\pi^+\pi^-$ invariant mass for signal MC (left), $B\overline{B}$ generic MC (middle) and data (right) for events with $3.8 < m(J/\psi\pi^+\pi^-) < 4.5\,\mathrm{GeV}/c^2$.

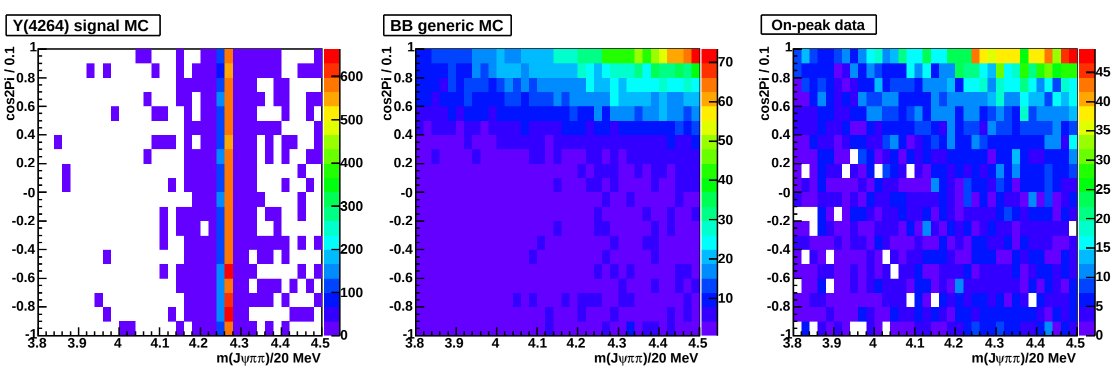


Figure 4.25: The angle cos2Pi versus the $J/\psi\pi^+\pi^-$ invariant mass for signal MC (left), $B\overline{B}$ generic MC (middle) and data (right)

the K_1 invariant mass. As expected, the cut cos2Pi$>$ 0.5 (right) selects a clean sample of K_1-events. To further investigate the influence of K_1-decays, Fig.4.28 shows the $J/\psi\pi^+\pi^-$ invariant mass with different cuts on the $K\pi^+\pi^-$ invariant mass. The left plot shows the spectrum with the initial selection criteria, the middle plot has a K_1 veto applied and the right plot shows the spectrum with a $150\,\mathrm{MeV}/c^2$ wide K_1 mass cut.

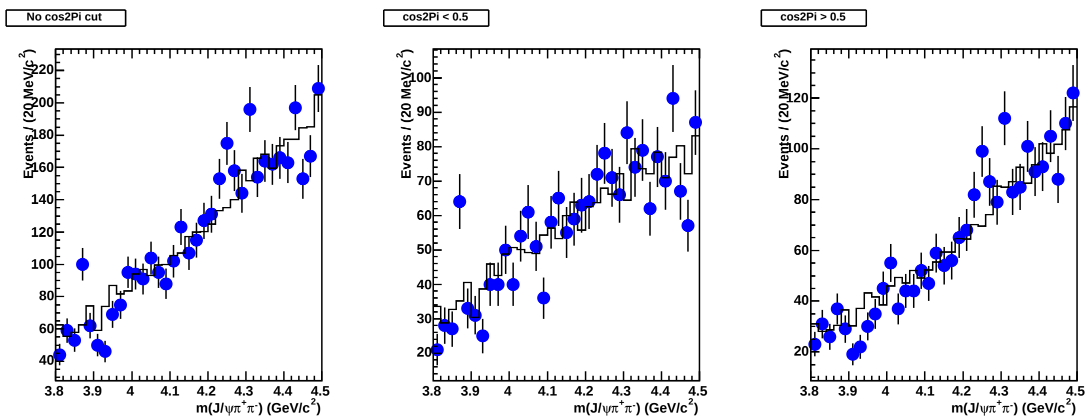


Figure 4.26: The $J/\psi\pi^+\pi^-$ invariant mass without cos2Pi cut (left), cos2Pi$<$ 0.5 (middle) and cos2Pi$>$ 0.5 (right). The dots are on-peak data and the solid line shows $B\overline{B}$ generic MC scaled to the same number of events.

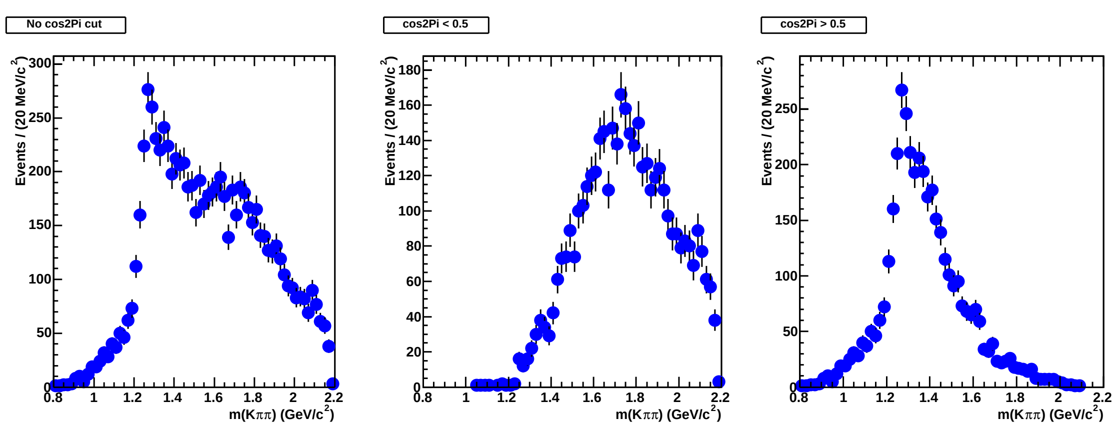


Figure 4.27: The $K\pi^+\pi^-$ invariant mass without cos2Pi cut (left), cos2Pi$<$ 0.5 (middle) and cos2Pi$>$ 0.5 (right) for on-peak data.

Next, we study the lepton decay angle $\cos\theta_l$ that is defined as follows:

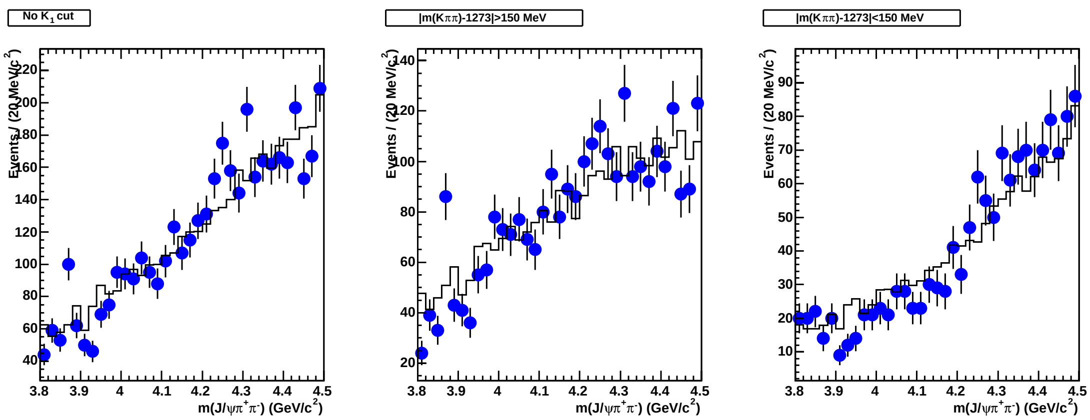


Figure 4.28: The $J/\psi\,\pi^+\pi^-$ invariant mass without K_1 cut (left), K_1 veto $|m(K\pi^+\pi^-) - 1273| > 150\,\mathrm{MeV}/c^2$ (middle) and $|m(K\pi^+\pi^-) - 1273| < 150\,\mathrm{MeV}/c^2$ (right). The dots are on-peak data and the solid line shows $B\overline{B}$ generic MC scaled to the same number of events.

- $\cos\theta_l$ is the cosine of the angle between the momentum direction of one of the leptons and the momentum direction of the recoiling kaon in the J/ψ rest frame

The $\cos\theta_l$ distribution for the $\psi(2S)$ ($J^{PC} = 1^{--}$) is shown in Fig. 4.29 (left) and the right plot shows $\cos\theta_l$ versus the $J/\psi\,\pi^+\pi^-$ invariant mass. Fig. 4.30 shows once more the $J/\psi\,\pi^+\pi^-$ invariant mass with different cuts on $\cos\theta_l$ and Fig. 4.31 shows the same but with a cos2Pi<0.5 cut applied.

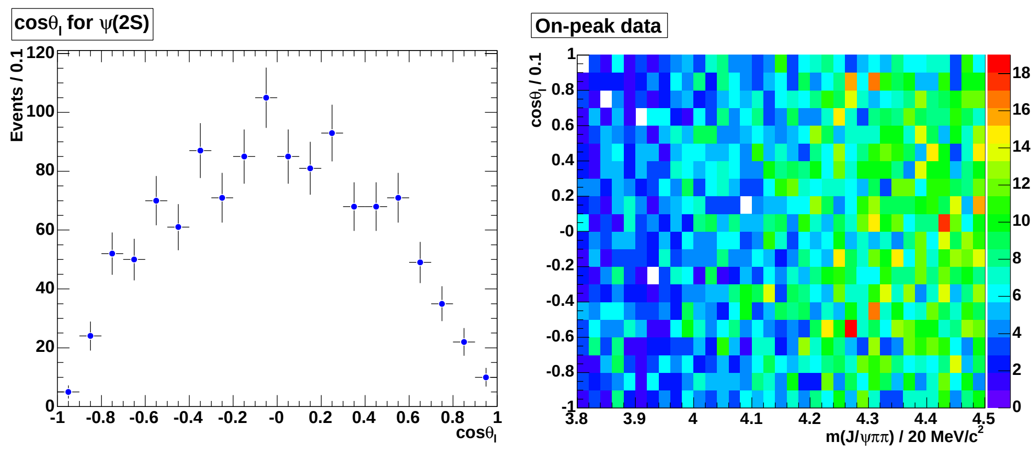


Figure 4.29: (left) $\cos\theta_l$ distribution for the $\psi(2S)$ as reference. (right) $\cos\theta_l$ versus the $J/\psi\,\pi^+\pi^-$ invariant mass.

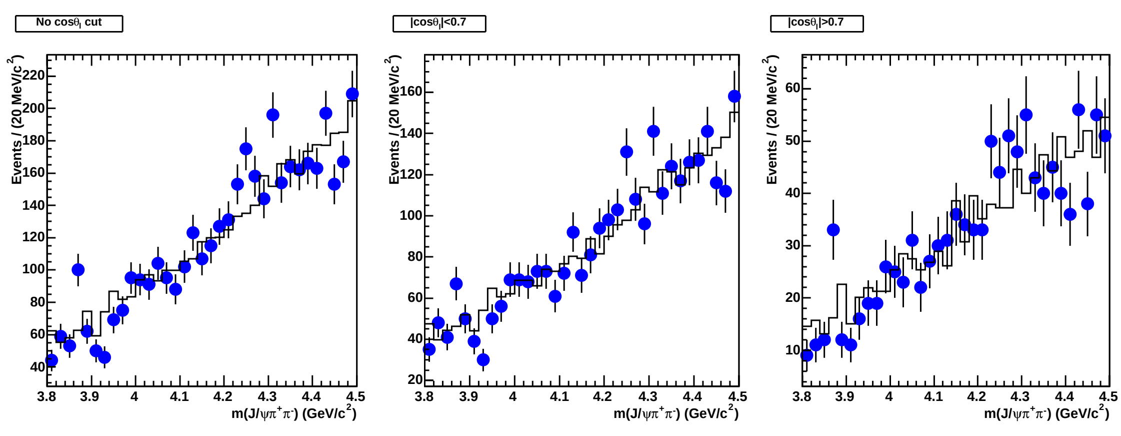


Figure 4.30: The $J/\psi\,\pi^+\pi^-$ invariant mass without $\cos\theta_l$ cut (left), $|\cos\theta_l| < 0.7$ (middle) and $|\cos\theta_l| > 0.7$ (right). The dots are on-peak data and the solid line shows $B\overline{B}$ generic MC scaled to the same number of events.

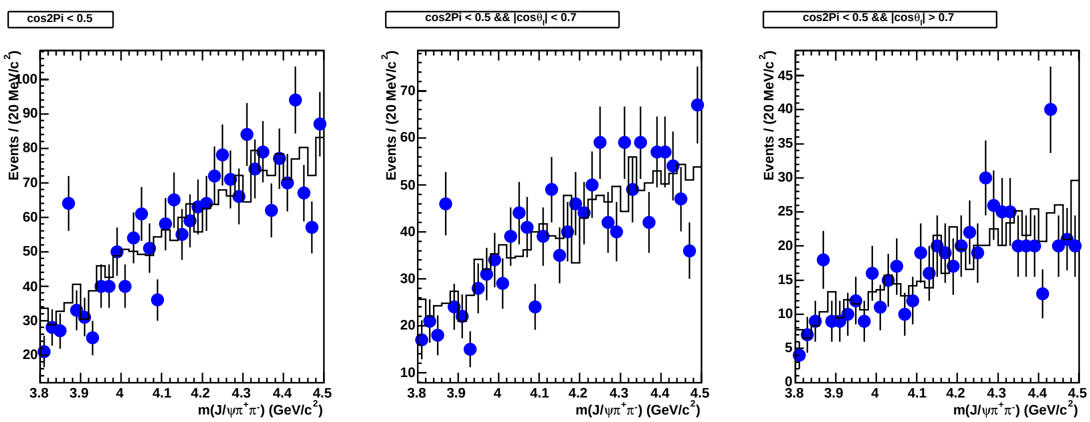


Figure 4.31: The $J/\psi\,\pi^+\pi^-$ invariant mass without $\cos\theta_l$ cut (left), $|\cos\theta_l| < 0.7$ (middle) and $|\cos\theta_l| > 0.7$ (right) for events with cos2Pi$<$ 0.5. The dots are on-peak data and the solid line shows $B\overline{B}$ generic MC scaled to the same number of events.

4.6.4 Fit on data (II)

After the additional studies described in the previous section, we decided to

1. reject $K_1 \rightarrow K\pi^+\pi^-$ backgrounds with a K_1 mass cut

2. present results with an additional cut on the lepton angle $\cos\theta_l$ optimized for a $J^{PC} = 1^{--}$ state.

The optimization for the K_1 veto was carried out on MC and results in $m(K\pi^+\pi^- - 1273) > 250\,\mathrm{MeV}/c^2$. Since we do not have appropriate MC for a 1^{--} state at $4.26\,\mathrm{GeV}/c^2$ we use the (shifted) $\psi(2S)$ signal MC for the optimization of the lepton angle. The optimized cut is $|\cos\theta_l| < 0.8$. Both optimizations included all the cuts, except PID. No changes in the other selection criteria were found.

Fig. 4.32 (left) shows the fit result with the K_1 veto applied but without the lepton angle cut. The final values of the fit parameters can be found in Tab. 4.13. The statistical significance for both peaks over the background hypothesis shown in Fig. 4.32 (right) is 4.0σ. The statistical significance for the lower peak is 3.4σ and 1.9σ for the higher peak, which was obtained by repeating the fit for one peak with all parameters but the yields fixed to the values obtained from the two-peak fit. Fig. 4.33 shows the same with an additional cut on the lepton angle $|\cos\theta_l| < 0.8$. The signal PDF for this fit only includes one signal peak. A fit to two signal peaks results in a fitted width of zero. The statistical significance is 3.5σ.

	No $\cos\theta_l$ cut	$\lvert\cos\theta_l\rvert < 0.8$
$n_{S,Y}$	73 ± 29	59 ± 24
$n_{S,Z}$	31 ± 13	-
$m_Y(\mathrm{MeV}/c^2)$	4249 ± 7	4243 ± 6
$m_Z(\mathrm{MeV}/c^2)$	4305 ± 6	-
$\Gamma_Y(\mathrm{MeV}/c^2)$	26 ± 13	21 ± 10
$\Gamma_Z(\mathrm{MeV}/c^2)$	11 ± 7	-

Table 4.13: Fit results with and without lepton angle cut. Γ denotes the natural width after subtraction of the mass resolution.

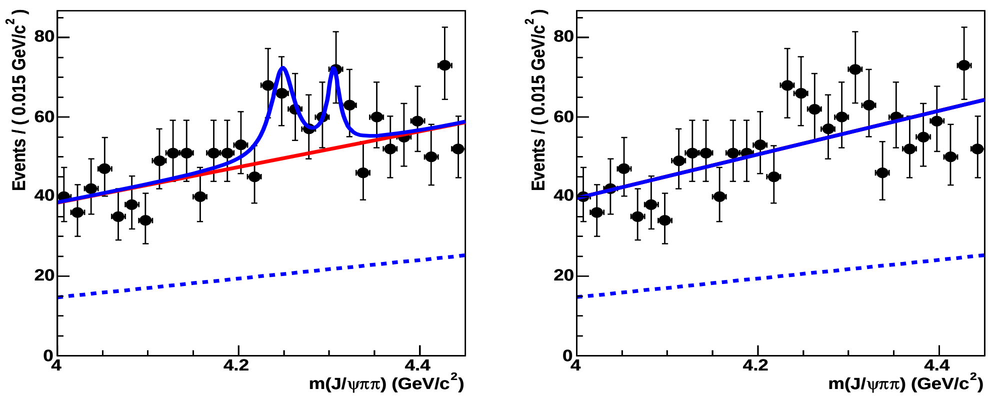


Figure 4.32: m_X projections of the fit result with K_1 veto but no lepton angle cut for the signal (left) and background (right) hypothesis. See Tab. 4.13 for the fit parameters.

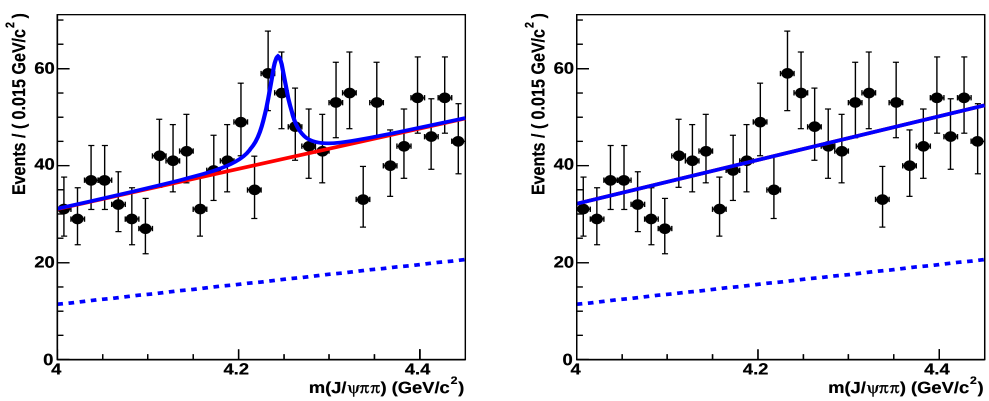


Figure 4.33: m_X projections of the fit result with K_1 veto and $|\cos\theta_l| < 0.8$ for the signal (left) and background (right) hypothesis. See Tab. 4.13 for the fit parameters.

4.6.5 Fit on data (III)

Since there is no compelling evidence for a signal based on these studies a more conservative approach is necessary. In the "single-peak" interpretation, the excess of events is centered around $(4259 \pm 10)\,\mathrm{MeV}/c^2$ with a width of $(88 \pm 24)\,\mathrm{MeV}/c^2$. Figure 4.34 shows the fit result to a single signal peak with mean at $4259\,\mathrm{MeV}/c^2$ and width of $93.4\,\mathrm{MeV}/c^2$ that includes the detector resolution of $5.4\,\mathrm{MeV}/c^2$. The fit result is 128 ± 42 signal events with a significance of 3.1σ compared to the background hypothesis fit. Only the K_1 mass-veto from the previous section has been applied.

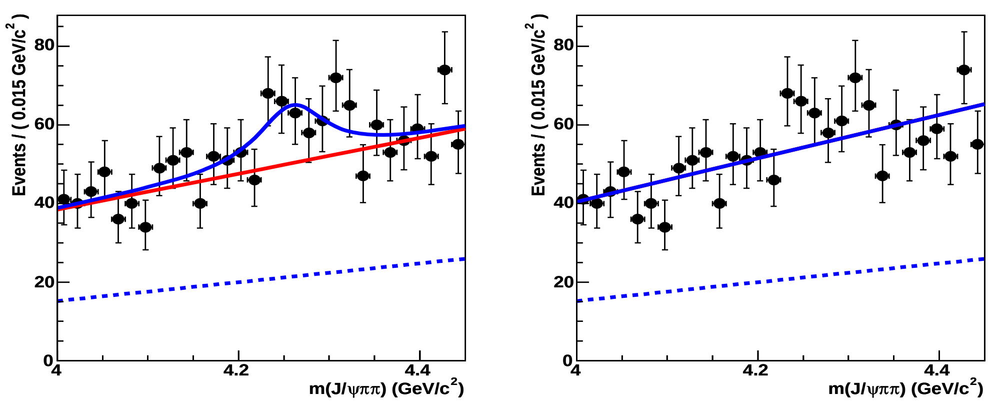


Figure 4.34: m_X projections of the fit result with K_1 veto and signal parameters fixed to the results in the ISR analysis for the signal (left) and background (right) hypothesis.

Using the same systematic uncertainties and efficiency corrections as for the $X(3872)$ we obtain a branching fraction of

$$\mathcal{B}(B^- \to Y(4260)K^-, Y(4260) \to J/\psi\pi^+\pi^-) = (2.0 \pm 0.7 \pm 0.2) \times 10^{-5} \tag{4.21}$$

where the first error is statistical and the second systematics. The efficiency was determined on a zero-width, phase-space MC. We calculate a 90% CL interval on the branching ratio with the same method as described in section 4.7.2 (Eq. 4.31):

$$1.2 < \mathcal{B}(B^- \to Y(4260)K^-, Y(4260) \to J/\psi\pi^+\pi^-) < 2.9 \times 10^{-5} \qquad (90\%\ \mathrm{CL}) \tag{4.22}$$

4.7 Physics results and conclusions

Here we summarize all of our result and possible physics implications.

4.7.1 Previous results

There are no previous results for the K^0_S mode, but both Belle [22] and *BABAR* [27] have measurements for $X(3872) \to J/\psi\,\pi^+\pi^-$ in the K^- mode. The Belle result is

$$\text{Belle:} \quad \frac{\mathcal{B}(B^- \to X(3872)K^-, X(3872) \to J/\psi\,\pi^+\pi^-)}{\mathcal{B}(B^- \to \psi(2S)K^-)\mathcal{B}(\psi(2S) \to J/\psi\,\pi^+\pi^-)} = 0.063 \pm 0.012 \pm 0.007 \tag{4.23}$$

using 152 million $B\overline{B}$ events. This translates into a central value of

$$\text{Belle:} \quad \mathcal{B}(B^- \to X(3872)K^-, X(3872) \to J/\psi\,\pi^+\pi^-) = (13.6 \pm 3.1) \times 10^{-6} \tag{4.24}$$

if we use the uncertainty from the above ratio and on the secondary branching ratios only. Furthermore they report a 90% CL upper limit on the width of $\Gamma < 2.3\,\text{MeV}$. The *BABAR* result for the branching ratio is

$$\textit{BABAR}: \quad \mathcal{B}(B^- \to X(3872)K^-, X(3872) \to J/\psi\,\pi^+\pi^-) = (12.8 \pm 4.1) \times 10^{-6} \tag{4.25}$$

using 117 million $B\overline{B}$ events from run 1-3. A check of the event yields can be found in [80].

4.7.2 Branching ratios

From the fit result $n_S = 8.3 \pm 4.5$ and $n_S = 61.2 \pm 15.3$ (Tab. 4.9) we calculate the product branching ratio

$$\mathcal{B}(B \to X(3872)K, X(3872) \to J/\psi\,\pi^+\pi^-) \quad = \quad \frac{n_S / \left[0.5\,\mathcal{B}(K^0_S \to \pi^+\pi^-)\right]}{\epsilon\, N_{B\overline{B}}\, \mathcal{B}(J/\psi \to l^+l^-)} \tag{4.26}$$

where $N_{B\overline{B}}$ the number of $B\overline{B}$ events (Tab. 4.1), ϵ the corrected efficiency (Tab. 4.5) and we use the secondary branching ratios from the PDG (see [8] and section 4.5.2). The additional factor in square brackets [...] only applies to the K_S^0 mode and we assume that K^0 decays into K_S^0 in 50% of the time.

Our result for the two charged pion modes are

$$\mathcal{B}(B^0 \to X(3872)K^0, X(3872) \to J/\psi\,\pi^+\pi^-) = (5.1 \pm 2.8 \pm 0.7) \times 10^{-6} \quad (4.27)$$

$$\mathcal{B}(B^- \to X(3872)K^-, X(3872) \to J/\psi\,\pi^+\pi^-) = (10.1 \pm 2.5 \pm 1.0) \times 10^{-6} \quad (4.28)$$

where the first error is the statistical error (on n_S from the fit and on ϵ) and the second error includes all the systematic and secondary branching ratio errors. All individual errors have been added in quadrature if not otherwise noted in the systematics section. From this we obtain the interesting ratio

$$R = \frac{\mathcal{B}^0}{\mathcal{B}^-} = \frac{\mathcal{B}(B^0 \to X(3872)K^0, X(3872) \to J/\psi\,\pi^+\pi^-)}{\mathcal{B}(B^- \to X(3872)K^-, X(3872) \to J/\psi\,\pi^+\pi^-)} = 0.50 \pm 0.30 \pm 0.05. \quad (4.29)$$

In this ratio most of the systematic uncertainties cancel each other. Therefore, only the uncertainties due to the K_S^0 (branching ratio and correction), B^-/B^0 production as well as MC statistics and MC-data difference are included. This gives a fractional systematic uncertainty of 9.5%.

To obtain the significance of the signal including systematic errors we use standard error propagation and assume Gaussian systematic errors. We define $L(n) = -\ln(\mathcal{L}(n)/\mathcal{L}_{\max})$ as the negative log-likelihood (NLL) obtained from the fit as in Fig. 4.15. From this we calculate the NLL including systematics as

$$L_{sys}(n) = \left(\frac{1}{L(n)} + \frac{2\sigma_{sys}^2}{(n - n_S)^2} \right)^{-1} \quad (4.30)$$

where σ_{sys} is the systematic error on the number of signal events n_S from the fit. With the fractional systematic errors from Tab. 4.11 we obtain $\sigma_{sys} = 1.1$ and $\sigma_{sys} = 4.9$ as the systematic uncertainties on n_S for the B^0 and B^- mode, respectively. Using

$\sqrt{2L_{sys}(n=0)}$ the significance of the signal including systematic uncertainties becomes 2.5σ for the B^0 mode and 6.1σ for the B^- mode.

We give a two-sided 90% confidence level (CL) interval obtained from the likelihood function including systematics (Eq. 4.30). In the large sample limit, one can show [58] that the confidence interval $[n_1, n_2]$ obtained from

$$L_{sys}(n_{1,2}) = \frac{1}{2}\left(\Phi^{-1}(1-\gamma/2)\right)^2 \tag{4.31}$$

approximates the classical confidence interval. $\Phi^{-1}(1-\gamma/2)$ is the quantile of the standard Gaussian corresponding to the confidence level $1-\gamma$. With $1-\gamma = 0.90$, Eq. 4.31 becomes $L_{sys}(n_{1,2}) = 1.35$. The two solutions in the K_S^0 mode are $n_1 = 2.2$ and $n_2 = 16.9$ events. Using (4.26) we obtain the 90% confidence level interval on the branching ratio (including systematics)

$$1.34 < \mathcal{B}(B^0 \to X(3872)K^0, X(3872) \to J/\psi\,\pi^+\pi^-) < 10.3 \times 10^{-6} \qquad (90\%\ \text{CL}). \tag{4.32}$$

We specifically note, that this confidence interval has to be interpreted under the 2.5σ signal-significance in this mode. With the same strategy, we set limits on the ratio

$$R = \frac{\mathcal{B}^0}{\mathcal{B}^-} = \alpha\frac{n_S^0}{n_S^-} \qquad \text{with} \qquad \alpha = \frac{\epsilon^-}{\epsilon^0}\frac{1}{0.5\ \mathcal{B}(K_S^0 \to \pi^+\pi^-)} = 3.71 \tag{4.33}$$

where n_S^0, n_S^- and ϵ^0, ϵ^- are the signal yields and efficiencies in the B^0 and B^- mode. Using standard error propagation we obtain the NLL for R including systematics

$$L_{sys}(R) = \left(\frac{1}{L^0(Rn_S^-/\alpha)} + \frac{1}{L^-(\alpha n_S^0/R)} + \frac{2\sigma_{sys,R}^2}{(R-\alpha n_S^0/n_S^-)^2}\right)^{-1} \tag{4.34}$$

where $L^{0/-}$ are the NLL in the $B^{0/-}$ mode and $\sigma_{sys,R}$ the systematic uncertainty on R from Eq. 4.29. Using Eq. 4.31 we calculate the 90% CL interval on R including systematics as

$$0.13 < R < 1.10 \qquad (90\% \text{ CL}). \tag{4.35}$$

4.7.3 Mass and natural width

We measure the mass of the $X(3872)$ in reference to the fitted $\psi(2S)$ mass:

$$m_X = m_{X,fit} - m_{\psi(2S),fit} + m_{\psi(2S)} \tag{4.36}$$

where $m_{\psi(2S)}$ is the world average of the $\psi(2S)$ mass. Using the values from Tab. 4.8 for the mass shift and the fitted value from Tab. 4.9 we get

$$m(X(3872) \text{ in } B^0 \to X(3872)K_S^0) = (3868.6 \pm 1.2 \pm 0.2)\,\text{MeV}/c^2 \tag{4.37}$$

$$m(X(3872) \text{ in } B^- \to X(3872)K^-) = (3871.3 \pm 0.6 \pm 0.1)\,\text{MeV}/c^2 \tag{4.38}$$

where the first (statistical) error is due to the m_X fit itself and the second (systematic) error is from the $\psi(2S)$ fit and the error on the $m_{\psi(2S)}$ from the PDG. With this we obtain the mass difference

$$\Delta m = (2.7 \pm 1.3 \pm 0.2)\,\text{MeV}/c^2 \tag{4.39}$$

which is compatible with zero within two standard deviations.

We now determine the natural width of the X-particle in the B^- mode. Note, that the width (FWHM) obtained from the fit on data $\Gamma = (5.04 \pm 2.23)\,\text{MeV}/c^2$ (Tab. 4.9) is compatible with the resolution determined on MC $\Gamma_R = (5.38 \pm 0.06)\,\text{MeV}/c^2$ (Tab. 4.4). Furthermore note, that our resolution function is a Lorentzian (=Breit-Wigner with FWHM Γ_R) and the line-shape of the $X(3872)$ is modelled by a Breit-Wigner function with a natural width Γ_N. The convolution of those two functions will result in a Breit-Wigner with FWHM $\Gamma = \Gamma_R + \Gamma_N$ (this can easily be proven using the convolution theorem).

From this we infer a natural width consistent width zero:

$$\Gamma_N = \Gamma - \Gamma_R = (-0.34 \pm 4.98)\,\mathrm{MeV}/c^2 \tag{4.40}$$

The 90% CL upper limit on the natural width is therefore

$$\Gamma_N < (1.28 \times 4.98 - 0.34)\,\mathrm{MeV}/c^2 = 6.03\,\mathrm{MeV}/c^2 \qquad (90\%\ \mathrm{CL}) \tag{4.41}$$

4.7.4 Limits in the Δm-R plane

Our limits on R depend of course on the mass of the X-particle in the K_S^0 mode. Since the signal significance is only 2.5σ we repeat the fit on data, but fix the central value of m_X to values in the range $3.86 - 3.88\,\mathrm{GeV}/c^2$. We then calculate R and its limits based on the new number of signal events. Figure 4.35 shows the result in the Δm-R-plane, where $\Delta m = 3871.3 - m_X$. The lower limit on R is restricted to positive values since we do not have likelihood projections for negative values of n_S.

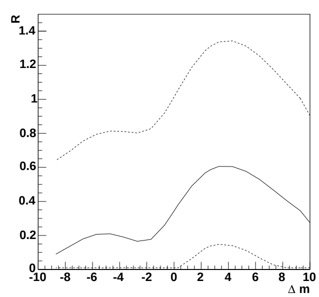


Figure 4.35: The solid line represent the central value of R for different values of Δm. The two dashed lines are the 90% CL interval on this value.

4.7.5 Study of the $J/\psi\,\pi^+\pi^-$ invariant mass above $4\,\mathrm{GeV}/c^2$

In the $J/\psi\,\pi^+\pi^-$ invariant mass region above $4\,\mathrm{GeV}/c^2$ we observe an excess of events above background between 4.2 and $4.4\,\mathrm{GeV}/c^2$. These events are consistent with the broad structure observed in ISR events [78]. However, our result by itself does not warrant the claim of a discovery of the $Y(4260)$ in B decays.

4.7.6 Conclusions

The first separate analysis of $X(3872) \to J/\psi\pi^+\pi^-$ in charged and neutral B decays has been presented and the results were published in [28]. The measured product branching ratio in the charged B mode is consistent with previous measurement from Belle and *BABAR*. Moreover, we observed a 2.5σ excess of events above backgrounds in neutral B decays. Assuming this is the $X(3872)$ we calculated the mass difference and ratio of branching fractions. The difference in masses is consistent both with zero and the prediction from the Diquark-Antidiquark model (Eq. 1.41). The prediction that this ratio should be less than 0.1 in case the $X(3872)$ is a DD^* deuson [40] is challenged by this measurement. However, other decay models for the $X(3872)$ do not predict such a small ratio. With more data to be delivered by *BABAR* and Belle this type of measurement will help to distinguish between different models for the $X(3872)$. Especially, the question about the mass difference between the $X(3872)$ in neutral and charged B decays that is a central prediction of the Diquark-Antidiquark will certainly be answered in the near future.

Bibliography

[1] F. Winklmeier, "Search for a charged partner of the X(3872) in $B \to J/\psi\, \pi^{\pm}\pi^{0} K$ with the Babar detector," Master's thesis, Colorado State University, 2004.

[2] ***BABAR*** Collaboration, B. Aubert *et al.*, "The *BABAR* detector," *Nucl. Instrum. Meth.* **A479** (2002) 1–116, `hep-ex/0105044`.

[3] O. W. Greenberg, "Why is CPT fundamental?," `hep-ph/0309309`.

[4] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, "EVIDENCE FOR THE 2 PI DECAY OF THE K(2)0 MESON," *Phys. Rev. Lett.* **13** (1964) 138–140.

[5] M. Kobayashi, "CP violation and three generations," in *CP violation: In and beyond the standard model: Proceedings of the 27th SLAC Summer Institute on Particle Physics (SSI 99): Stanford, USA, 7-16 Jul 1999.* 1999. SLAC-R-719.

[6] N. Cabibbo, "Unitary symmetry and leptonic decays," *Phys. Rev. Lett.* **10** (1963) 531–532.

[7] M. Kobayashi and T. Maskawa, "CP violation in the renormalizable theory of weak interaction," *Prog. Theor. Phys.* **49** (1973) 652–657.

[8] **Particle Data Group** Collaboration, S. Eidelman *et al.*, "Review of particle physics," *Phys. Lett.* **B592** (2004) 1.

[9] **CKMfitter Group** Collaboration, J. Charles *et al.*, "CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories," *Eur. Phys. J.* **C41** (2005) 1–131, hep-ph/0406184.

[10] **The Heavy Flavor Averaging Group** Collaboration, E. Barberio *et al.*, "Averages of b-hadron properties at the end of 2005," hep-ex/0603003.

[11] ***BABAR*** Collaboration, B. Aubert *et al.*, "Measurement of gamma in B-+ → D(*) K-+ decays with a Dalitz analysis of D → K0(S) pi- pi+," *Phys. Rev. Lett.* **95** (2005) 121802, hep-ex/0504039.

[12] A. Soffer, "Discrete ambiguities in the measurement of the weak phase gamma," *Phys. Rev.* **D60** (1999) 054032, hep-ph/9902313.

[13] M. Gronau and D. Wyler, "On determining a weak phase from CP asymmetries in charged B decays," *Phys. Lett.* **B265** (1991) 172–176.

[14] M. Gronau and D. London, "Isospin analysis of CP asymmetries in B decays," *Phys. Rev. Lett.* **65** (1990) 3381–3384.

[15] D. Atwood, I. Dunietz, and A. Soni, "Enhanced CP violation with B → K D0 (anti-D0) modes and extraction of the CKM angle gamma," *Phys. Rev. Lett.* **78** (1997) 3257–3260, hep-ph/9612433.

[16] D. Atwood, I. Dunietz, and A. Soni, "Improved methods for observing CP violation in B+- → K D and measuring the CKM phase gamma," *Phys. Rev.* **D63** (2001) 036005, hep-ph/0008090.

[17] A. Giri, Y. Grossman, A. Soffer, and J. Zupan, "Determining gamma using B+- → D K+- with multibody D decays," *Phys. Rev.* **D68** (2003) 054018, hep-ph/0303187.

[18] A. Giri, Y. Grossman, A. Soffer, and J. Zupan, "Determination of the angle gamma using multibody D decays in B+- → D K+-," *ECONF* **C0304052** (2003) WG424, hep-ph/0306286.

[19] **CLEO** Collaboration, D. Cronin-Hennessy *et al.*, "Searches for CP violation and pi pi S-wave in the Dalitz- plot of D0 → pi+ pi- pi0," *Phys. Rev.* **D72** (2005) 031102, `hep-ex/0503052`.

[20] **CDF II** Collaboration, D. Acosta *et al.*, "Observation of the narrow state X(3872) → J/psi pi+ pi- in anti-p p collisions at s**(1/2) = 1.96-TeV," *Phys. Rev. Lett.* **93** (2004) 072001, `hep-ex/0312021`.

[21] "SPIRES HEP database." `http://www.slac.stanford.edu/spires/hep/`.

[22] **Belle** Collaboration, S. K. Choi *et al.*, "Observation of a Narrow Charmoniumlike State in Exclusive B+- → K+- pi+ pi- J/psi Decays," *Phys. Rev. Lett.* **91** (2003) 262001, `hep-ex/0309032`.

[23] **Belle** Collaboration, K. Abe *et al.*, "Experimental constraints on the possible J(PC) quantum numbers of the X(3872)," `hep-ex/0505038`.

[24] **Belle** Collaboration, K. Abe *et al.*, "Properties of the X(3872) at Belle," `hep-ex/0408116`.

[25] **Belle** Collaboration, K. Abe *et al.*, "Evidence for X(3872) → gamma J/psi and the sub-threshold decay X(3872) → omega J/psi," `hep-ex/0505037`.

[26] **Belle** Collaboration, T. Lesiak, "Spectroscopy and new particles," *AIP Conf. Proc.* **814** (2006) 493–497, `hep-ex/0511003`.

[27] ***BABAR*** Collaboration, B. Aubert *et al.*, "Study of the B → J/psi K- pi+ pi- decay and measurement of the B → X(3872) K- branching fraction," *Phys. Rev.* **D71** (2005) 071103, `hep-ex/0406022`.

[28] ***BABAR*** Collaboration, B. Aubert *et al.*, "Study of the X(3872) and Y(4260) in B0 → J/psi pi+ pi- K0 and B- → J/psi pi+ pi- K- decays," *Phys. Rev.* **D73** (2006) 011101, `hep-ex/0507090`.

[29] ***BABAR*** Collaboration, B. Aubert *et al.*, "Search for a charged partner of the X(3872) in the B meson decay B → X- K, X- → J/psi pi- pi0," *Phys. Rev.* **D71** (2005) 031501, `hep-ex/0412051`.

[30] ***BABAR*** Collaboration, B. Aubert *et al.*, "Measurements of the absolute branching fractions of B+- → K+- X/(c anti-c)," *Phys. Rev. Lett.* **96** (2006) 052002, `hep-ex/0510070`.

[31] **CDF** Collaboration, A. Abulencia *et al.*, "Measurement of the dipion mass spectrum in X(3872) → J/psi pi+ pi- decays," *Phys. Rev. Lett.* **96** (2006) 102002, `hep-ex/0512074`.

[32] R. L. Jaffe, "MULTI - QUARK HADRONS. 1. THE PHENOMENOLOGY OF (2 quark 2 anti-quark) MESONS," *Phys. Rev.* **D15** (1977) 267.

[33] R. L. Jaffe, "MULTI - QUARK HADRONS. 2. METHODS," *Phys. Rev.* **D15** (1977) 281.

[34] R. L. Jaffe, "Exotica," *Phys. Rept.* **409** (2005) 1–45, `hep-ph/0409065`.

[35] D. J. Griffiths, *Introduction to quantum mechanics.* Prentice Hall, 1995.

[36] L. Maiani, F. Piccinini, A. D. Polosa, and V. Riquer, "Diquark-antidiquarks with hidden or open charm and the nature of X(3872)," *Phys. Rev.* **D71** (2005) 014028, `hep-ph/0412098`.

[37] N. A. Tornqvist, "Possible large deuteron - like meson meson states bound by pions," *Phys. Rev. Lett.* **67** (1991) 556–559.

[38] N. A. Tornqvist, "From the deuteron to deusons, an analysis of deuteron - like meson meson bound states," *Z. Phys.* **C61** (1994) 525–537, `hep-ph/9310247`.

[39] E. Braaten and M. Kusunoki, "Low-energy universality and the new charmonium resonance at 3870-MeV," *Phys. Rev.* **D69** (2004) 074005, `hep-ph/0311147`.

[40] E. Braaten and M. Kusunoki, "Exclusive production of the X(3872) in B meson decay," *Phys. Rev.* **D71** (2005) 074005, hep-ph/0412268.

[41] F. E. Close and P. R. Page, "The D*0 anti-D0 threshold resonance," *Phys. Lett.* **B578** (2004) 119–123, hep-ph/0309253.

[42] S. Pakvasa and M. Suzuki, "On the hidden charm state at 3872-MeV," *Phys. Lett.* **B579** (2004) 67–73, hep-ph/0309294.

[43] E. S. Swanson, "Short range structure in the X(3872)," *Phys. Lett.* **B588** (2004) 189–195, hep-ph/0311229.

[44] E. S. Swanson, "Diagnostic decays of the X(3872)," *Phys. Lett.* **B598** (2004) 197–202, hep-ph/0406080.

[45] E. S. Swanson, "The new heavy mesons: A status report," hep-ph/0601110.

[46] N. A. Tornqvist, "Comment on the narrow charmonium state of Belle at 3871.8-MeV as a deuson," hep-ph/0308277.

[47] N. A. Tornqvist, "Isospin breaking of the narrow charmonium state of Belle at 3872-MeV as a deuson," *Phys. Lett.* **B590** (2004) 209–215, hep-ph/0402237.

[48] M. B. Voloshin, "Interference and binding effects in decays of possible molecular component of X(3872)," *Phys. Lett.* **B579** (2004) 316–320, hep-ph/0309307.

[49] C.-Y. Wong, "Molecular states of heavy quark mesons," *Phys. Rev.* **C69** (2004) 055202, hep-ph/0311088.

[50] E. S. Swanson. private communications.

[51] W. T. Ford, "Choice of Kinematic Variables in B Meson Reconstruction-Take 3," *BAD 53 Version 5* (2000).

[52] B. Dunwoodie.
`http://www.slac.stanford.edu/~wmd/bbkinematics/kinematics.note`.

[53] G. C. Fox and S. Wolfram, "Event shapes in e+ e- annihilation," *Nucl. Phys.* **B149** (1979) 413.

[54] M. Jacob and G. C. Wick, "On the general theory of collisions for particles with spin," *Ann. Phys.* **7** (1959) 404–428.

[55] H. E. Haber, "Spin Formalism and Applications to New Physics Searches," *SLAC-R-44 Proceedings of the 21st SLAC Summer Institute on Particle Physics: Spin Structure in High Energy Processes* (1993).

[56] J. Richman, "An Experimentalist's Guide to the Helicity Formalism," *Cal. Tech Preprint CALT-68-1148* (1984).

[57] R. Kutschke, "An Angular Distribution Cookbook,".

[58] G. Cowan, *Statistical Data Analysis.* Oxford Science Publications, 1998.

[59] "CERN Program Library Long Writeup D506, MINUIT."
`http://wwwasdoc.web.cern.ch/wwwasdoc/minuit/`.

[60] ***BABAR*** Collaboration, B. Aubert *et al.*, "Measurement of the branching fraction and decay rate asymmetry of B- $\rightarrow$ D (pi+ pi- pi0) K-," *Phys. Rev.* **D72** (2005) 071102, `hep-ex/0505084`.

[61] A. Soffer and Q. Zeng, "Study of B$\rightarrow$ D K with singly Cabibbo suppressed D decays to three body final states," *BAD 825* (2005).

[62] **Belle** Collaboration, A. Poluektov *et al.*, "Measurement of phi(3) with Dalitz plot analysis of B+ $\rightarrow$ D(*) K(*)+ decay," `hep-ex/0604054`.

[63] "Numerical Recipes in C."
`http://www.library.cornell.edu/nr/cbookcpdf.html`.

[64] **CLEO** Collaboration, S. Kopp *et al.*, "Dalitz analysis of the decay D0→ K- pi+ pi0," *Phys. Rev.* **D63** (2001) 092001, `hep-ex/0011065`.

[65] **E791** Collaboration, E. M. Aitala *et al.*, "Study of the D/s+ → pi- pi+ pi+ decay and measurement of f0 masses and widths," *Phys. Rev. Lett.* **86** (2001) 765–769, `hep-ex/0007027`.

[66] K. Mishra, A. Soffer, F. Winklmeier, and J. Zhang, "Measurement of the CKM angle gamma in B-→ D0K- decays using a D0 Dalitz analysis of D0→ pi+ pi- pi0," *BAD 1174* (2006).

[67] G. Mancinelli, B. T. Meadows, K. Mishra, and M. D. Sokoloff, "Branching Ratio Measurements of the Decays D0→ pi- pi+ pi0 and D0→ K- K+ pi0," *BAD 1549* (2006).

[68] **Particle Data Group** Collaboration, S. Eidelman *et al.*, "Review of particle physics," *Phys. Lett.* **B592** (2004) 1.

[69] G. Mancinelli, B. T. Meadows, K. Mishra, and M. D. Sokoloff, "Branching Ratio Measurements of the Decays $D^0 \to \pi^-\pi^+\pi^0$and $D^0 \to K^-K^+\pi^0$," *BAD 1469* (2006).

[70] E. J. Eichten, K. Lane, and C. Quigg, "B meson gateways to missing charmonium levels," *Phys. Rev. Lett.* **89** (2002) 162002, `hep-ph/0206018`.

[71] T. Barnes and S. Godfrey, "Charmonium options for the X(3872)," *Phys. Rev.* **D69** (2004) 054008, `hep-ph/0311162`.

[72] E. J. Eichten, K. Lane, and C. Quigg, "Charmonium levels near threshold and the narrow state X(3872) → pi+ pi- J/psi," *Phys. Rev.* **D69** (2004) 094019, `hep-ph/0401210`.

[73] G. Punzi, "Sensitivity of searches for new signals and its optimization," `physics/0308063`.

[74] F. Winklmeier and W. Toki, "Supporting document for the Search for a charged of $X(3872)$ in $B \to J/\psi\,\pi^{\pm}\pi^{0}K$," *BAD 868 Version 10* (2004).

[75] W. Verkerke and D. Kirkby, "The RooFit toolkit for data modeling," *ECONF* **C0303241** (2003) MOLT007, `physics/0306116`.

[76] ***BABAR*** Collaboration, B. Aubert *et al.*, "Measurement of the B+/B0 production ratio from the Upsilon(4S) meson using B+ → J/psi K+ and B0 → J/psi K0(S) decays," *Phys. Rev.* **D69** (2004) 071101, `hep-ex/0401028`.

[77] "Track Efficiency Task Force." `http://www.slac.stanford.edu/BFROOT/www/Physics/TrackEfficTaskForce/Tra%ckingTaskForce-2004.html`.

[78] ***BABAR*** Collaboration, B. Aubert *et al.*, "Observation of a broad structure in the pi+ pi- J/psi mass spectrum around 4.26-GeV/c**2," *Phys. Rev. Lett.* **95** (2005) 142001, `hep-ex/0506081`.

[79] J. Izen, X. Lou, G. Williams, and S. Ye, "Search for heavy particles decaying into $\pi^{+}\pi^{-}J/\psi$ produced in initial state radiation interaction," *BAD 1121 Version 13* (2005).

[80] F. Winklmeier and W. Toki, "Search for $B \to X(3872)K, X(3872) \to J/\psi\,\pi\pi$," *BAD 1133 Version 17* (2005).