Karlsruhe Institute of Technology

Dalitz analysis of $B^- \to D^0\pi^-\pi^0$ and $B^0 \to D^-\pi^+\pi^0$ at the Belle experiment

Zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

bei der Fakultät für Physik des
Karlsruher Institut für Technologie (KIT)

genehmigte

Dissertation

von

Dipl.-Phys. Manuel Heider

aus Bad Friedrichshall

Tag der mündlichen Prüfung:	19. Juni 2015
Referent:	Prof. Dr. Michael Feindt
Korreferent:	Prof. Dr. Günter Quast

DOi : 10.5445/IR/1000049451

Contents

1. Introduction

1.1. Motivation

Curiosity - that is the name of a Mars rover landed in Gale Crater on Mars on August 6, 2012, by the Mars Science Laboratory. Its objectives include the study of the climate and geology of Mars along with collecting data for a possible future manned mission to Mars. To cite the twelve year old girl Clara Ma who named the rover:

> Curiosity is an everlasting flame that burns in everyone's mind. [...] Curiosity is such a powerful force. Without it, we wouldn't be who we are today.[...]

And science would not be either. The reason for investing so much to get a small rover to travel 570 million kilometers to examine the dust of a vast desert is nothing but curiosity. This driving force is not only the heart of science, it is also the reason why science even exists. And the scientific method of generating knowledge is probably the most important invention ever made. It is not a product but a process, an everlasting dialogue between observation and theory, the most effective way of deciding if explanations of phenomena are right or wrong. The current status of this dialogue in the case of particle physics manifests itself in the Standard Model of particle physics, which has been developed over the last half century and is capable of explaining a wide variety of experimental results. The recent discovery of the Higgs boson has given further credence to it, but there are still phenomena unexplained in the scope of the Standard Model like dark matter, dark energy or neutrino oscillations. This led to many particle physics analyses in the last years being focused on probing the Standard Model in the right places, where the sensitivity to a possible influence of *new physics* is sufficient, i.e., where the experimental and theoretical precision allow for a significant discovery. Up to now, none of these analyses have found a discrepancy.

In contrast to these analyses, this analysis aims at strengthening the fundament of the Standard Model by making a precision measurement of branching fractions in the decays $B^- \to D^0\pi^-\pi^0$ and $B^0 \to D^-\pi^+\pi^0$. They both account for about 1% of the decays of B mesons and thereby present a significant contribution not only as background in numerous analyses but more importantly in all the simulations used by many particle physics experiments at particle colliders. But surprisingly, they are poorly measured. The most recent and by far best measurement to date was performed in 1994 on CLEO data [1], using 212 (79) events in total for the branching ratio of $B^- \to D^0\rho^-$ ($B^0 \to D^-\rho^+$). The expected number of signal events in the Belle data sample, based on the CLEO measurement, is approximately 80000 for each channel after reconstruction and selection of B candidates. With the results of the branching

fractions the soon-to-be generated Monte Carlo samples for the Belle II experiment can be significantly improved. Especially the successor of the EKP full reconstruction tool [2], the Full Event Interpretation (FEI) currently in development at Belle II, will profit from the more precise measurements. The FEI utilizes an important property of lepton colliders: The precise knowledge of the initial state due to the well-known energy of the electron and positrons. This allows for measurements of B meson decays including neutrinos, of which many are sensitive to small contributions from new physics. Examples are the analysis of $B^+ \to \tau^+\nu_\tau$, $B^+ \to K^+\nu\nu$ or $B^+ \to D^{(*)}\tau^+\nu_\tau$, where the properties of the D^{**} are also important to estimate the background originating in decays via D^{**} [3, 4, 5].

In this analysis the whole decay structure is examined utilizing the Dalitz plot analysis technique. All quantities that can be measured are discussed in Section 1.2 along with previous measurements of them. The analysis is performed on 772 million $B\bar{B}$ pairs recorded with the Belle detector at the KEKB accelerator in Tsukuba, Japan, described in Chapter 2. The Dalitz plot analysis technique is described in Chapter 3 followed by the reconstruction and selection of B meson candidates in Chapter 4. The neutral pions in the final state present the main challenge for this analysis, as they are highly prone to misreconstruction. To counteract the effects of misreconstruction without losing these candidates as signal, a migration matrix is used to backtrack the effects of the misreconstruction, which is described in Chapter 7. This analysis was performed blindly, the simulated data samples used in the evaluation and verification of the analysis are described in Chapter 5. In the Chapters 6 to 9 the different aspects of the fit model are described along with the extraction of the signal yield. Finally Chapters 10 and 11 contain the Dalitz fit to recorded data, including the results of the measurements, and the estimation of systematic errors, respectively.

1.2. Dalitz analysis of the decays $B^- \to D^0\pi^-\pi^0$ and $B^0 \to D^-\pi^+\pi^0$

Up to this date there were no previous Dalitz analysis performed in these decay modes. In general, dalitz analyses of such decays containing a neutral pion are rare. However, there are various analyses performed in the case of two charged Pions, where the D meson and the B meson are either both charged or both neutral: $B^+ \to D^-\pi^+\pi^+$ or $B^0 \to \bar{D}^0\pi^+\pi^-$. For the latter there is a Belle analysis performing a measurement of the total branching fraction and examining the structure of the decay, observing doubly excited D mesons [6]. So far all the knowledge about excited D mesons extracted from B decays comes from channels with one D meson and two charged Pions [7]. While charged Pions are far easier to reconstruct, the decays with one neutral Pion have one advantage: the branching fraction is about a factor of ten higher compared to the modes with two charged Pions, overcompensating for the lower efficiency and worse resolution of the reconstruction of neutral Pions.

In this analysis we will use the invariant mass pairs $m^2_{13} = m^2_{\pi^0 D}$ and $m^2_{m23} = m^2_{\pi^\pm D}$ where D denotes either $D^\pm$ or D^0. The Dalitz model is constructed with several intermediate states. In both cases, there is a non-resonant contribution, which populates the Dalitz plane uniformly, as well as a contribution of the decay via a charged ρ meson, which is expected to dominate the decay. Depending on the charge of the D meson, there is a $D^*(2007)^0$ or $D^*(2010)^+$ contributing. An interesting part of this analysis is to measure the contribution of doubly excited D mesons, namely $D_0^*(2400)$ and $D_2^*(2462)$ (see Section 1.3). The properties of the neutral doubly excited D meson are difficult to measure, the current world average of its mass is $2.318 \pm 0.029\,\mathrm{GeV\,c^{-2}}$ and of its width $0.267 \pm 0.040\,\mathrm{GeV\,c^{-2}}$, provided by the Particle Data Group (PDG) [7]. Thus, an important goal of this analysis is to measure these parameters of the broad D^{**}.

1.3. Excited D mesons

In this analysis we have the chance to study the production of D meson excited states, referred to as D^{**}. These D^{**} are P-wave excitations of a quark system containing one charmed and one light quark (u, d). The results can provide tests of the heavy quark effective theory (HQET). In the heavy quark limit, the spin of the heavy quark decouples from the light degrees of freedom, meaning they become insensitive to the mass of the heavy quark. Thus, the total angular momentum of the light quark $\vec{j}_q = \vec{L} + \vec{s}_q$ is a good quantum number.

As shown in Fig. 1.1, there are four P-wave states, two with $j_q = 1/2$ and two with $j_q = 3/2$ and a D^* with zero angular momentum between the quarks. The four D^{**} mesons are usually labeled as D_0^*, D_1', D_1, D_2^*. While the two $j_q = 3/2$ states are narrow and well known, the D_0^* and the D_1' have large widths, illustrated in the manner of error bars in Fig. 1.1. In the case of the D_1' and D_1, a decay directly to the D meson with $J = 0$ is not possible because of conservation of the angular momentum. However, they can decay to a D^* meson that has spin 1 and a pion.In this analysis, only final states with a D meson and two pions are reconstructed, thus only decays via D_0^* and D_2^* can be observed. They proceed via an S-wave or D-wave. In both cases, the conservation of parity forbids the decay in a P-wave.

The contributions of the charged $D_0^*(2400)^-$ and $D_2^*(2462)^-$ are only expected to be visible in the second channel $B^0 \to D^-\pi^+\pi^0$ (see Fig. 1.5), as they are doubly Cabibbo suppressed in the first case (see Fig. 1.3, where the $D_0^*(2400)^-$ is displayed exemplary). The neutral D^{**} mesons are expected to be visible in both decay modes. There are two tree-level diagrams contributing in Fig. 1.2, which are similar to Fig. 1.4 and Fig. 1.5, where only the spectator quark is different.

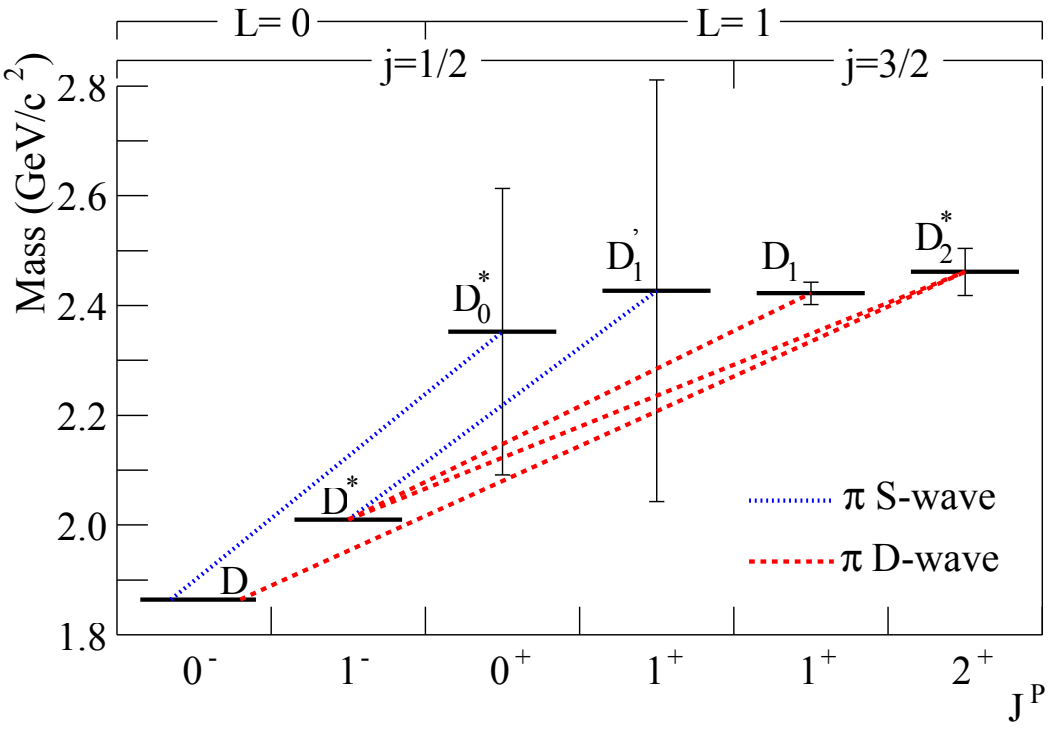


Figure 1.1.: Properties of the D^{**} mesons.

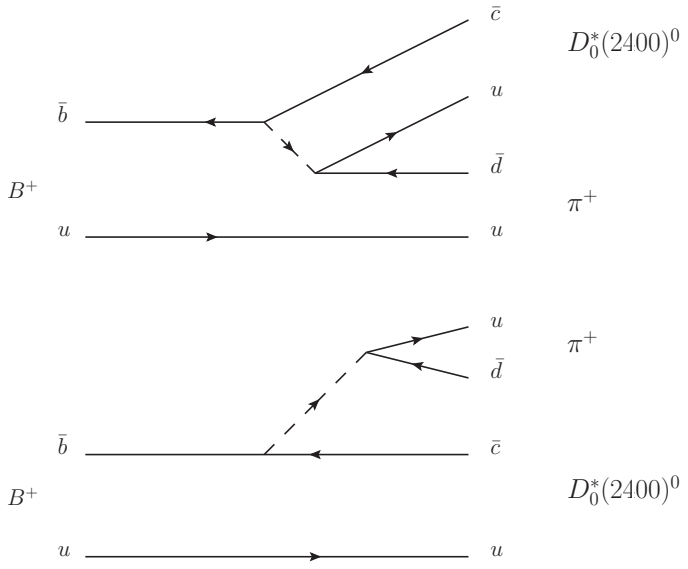


Figure 1.2.: There are two possible tree-level diagrams of the decay to a neutral $D_0^*(2400)$ and a charged π. The first is similar to Fig. 1.4 and the second to Fig. 1.5.

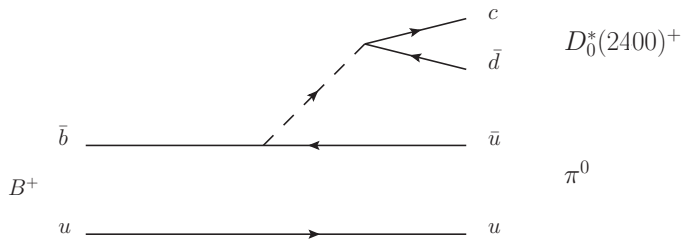


Figure 1.3.: Decay to a charged $D_0^*(2400)$ and a neutral π. The decay is doubly Cabibbo suppressed and will not be visible.

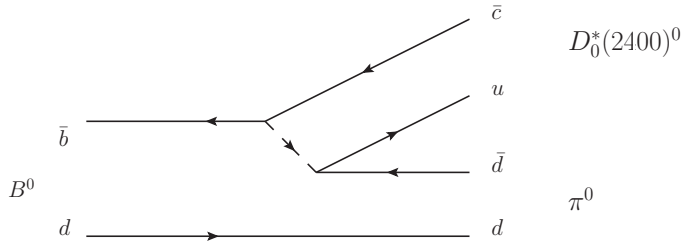


Figure 1.4.: Decay to a neutral $D_0^*(2400)$ and a neutral π.

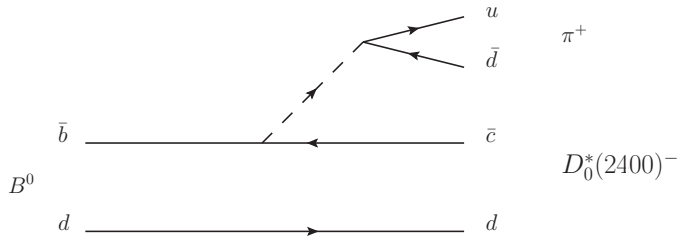


Figure 1.5.: Decay to a charged $D_0^*(2400)$ and a charged π. This decay is color suppressed.

1.4. Expected number of Events

The number of expected events can be calculated by multiplying the number of recorded $B\bar{B}$ pairs by the currently known branching fractions given by the PDG. This also takes the branching fractions of the decay of the $D/D^*/D^{**}$ mesons into account. For the decays via $D_0^*(2400)^0$, $D_2^*(2462)^0$, $D_0^*(2400)^-$ and $D_2^*(2462)^-$, there are no measurements available, but the branching fractions can be obtained from isospin considerations. Although the overall branching fraction of the mode $B^0 \to D^-\pi^+\pi^0$ is smaller and the overall efficiency is 14% compared to 18% for $B^- \to D^0\pi^-\pi^0$, the branching ratio of the charged D decaying in $K\pi\pi$ is 2.8 times higher leading to a comparable amount of signal events in both channels. The branching fractions of daughter decays used are listed in Table 1.3

Resonance	Branching fraction ($\times 10^{-3}$)	Error ($\times 10^{-3}$)	Expected number of events
$\rho(770)$	13.4	1.8	66331
$D_0^*(2400)^0$	0.32	0.07	1584
$D_2^*(2462)^0$	0.175	0.02	866
Non-resonant	-	-	-
$D^*(2007)$	3.21	0.26	15872
Sum	19.075	2.15	84653

Table 1.1.: Branching fraction of decays via different resonances in $B^- \to D^0\pi^-\pi^0$. There is no previous measurement available for the non-resonant decay.

Resonance	Branching fraction ($\times 10^{-3}$)	Error ($\times 10^{-3}$)	Expected number of events
$\rho(770)$	7.8	1.3	70113
$D_0^*(2400)^0$	-	-	-
$D_2^*(2462)^0$	-	-	-
$D_0^*(2400)^-$	0.3	0.15	2697
$D_2^*(2462)^-$	0.11	0.017	966
Non-resonant	-	-	-
$D^*(2010)$	0.85	0.13	7616
Sum	11.11	1.56	82673

Table 1.2.: Branching fraction of decays via different resonances in $B^0 \to D^-\pi^+\pi^0$. There are no measurements available for the non-resonant decay and the decays via neutral D^{**}.

Resonance	$\mathcal{B}$ ($\times 10^{-2}$)	Error ($\times 10^{-2}$)
$D^0 \to K^-\pi^+$	3.88	0.05
$D^+ \to K^-\pi^+\pi^+$	9.13	0.19
$D^*(2007)^0 \to D^0\pi^0$	61.9	2.9
$D^*(2010)^+ \to D^+\pi 0$	30.7	0.5
$\pi^0 \to \gamma\gamma$	98.823	0.034

Table 1.3.: Branching fractions of daughter particles to the reconstructed final state particles.

2. Experimental setup

2.1. Experimental Particle Physics

The knowledge of most particles and their properties was obtained in particle collision experiments. While there are various kinds of particle accelerators, they all have some basic principles in common. Stable charged particles are accelerated using electromagnetic fields and then collided with either a fixed target or other accelerated particles moving in the opposite direction. In such collisions new particles are created. These new particles or their decay products then leave tracks in a detector. Another possibility to utilize accelerated particles would be inelastic scattering, but that topic will not be covered here.

2.1.1. Characteristics of Colliders

The most important characteristic property of an accelerator for the creation of particles is the center of mass energy

$$\sqrt{s} = \sqrt{(\mathbf{p_1} + \mathbf{p_2})^2}, \tag{2.1}$$

where $\mathbf{p_i}$ are the four momenta of the colliding particles. The center of mass energy determines what production mechanisms are possible. In order to produce any number of particles, more energy than the sum of their rest masses has to be available. For a fixed target experiment the center of mass energy is

$$\sqrt{s} = \sqrt{2E_1 m_2 c^2 + (m_1^2 + m_2^2)c^4}, \tag{2.2}$$

so it rises with $\sqrt{E_1}$. Using $E_i^2 - p_i^2 c^2 = m_i^2 c^4$, $m_i c^2 << E_i$ and $E^2 = E_1 E_2$, the same property for a head-on collision of two relativistic particles is

$$\sqrt{s} = \sqrt{2\left(E_1 E_2 - \vec{p_1}\vec{p_2}c^2\right) + m_1^2 c^4 + m_2^2 c^4}, \tag{2.3}$$

$$\sqrt{s} \approx \sqrt{4E_1 E_2}, \tag{2.4}$$

$$\sqrt{s} \sim E. \tag{2.5}$$

Due to the linear rise of the center of mass energy with the energy of the accelerated particles, in modern accelerators only head-on collisions of two particle beams are used. A so-called particle beam consists of thousands of particle bunches with equal spacing in between and

each bunch holding up to millions of particles. This topology is produced by the mechanism of acceleration which typically uses high frequency electromagnetic fields. The center of mass energy limits what processes are possible in a collision, but provides no information how often a collision takes place. This is described by the *Luminosity*.

Luminosity The *luminosity* of a collider depends on the properties of the beam. If two bunches cross, the number of particles per bunch N_i and the spatial dimensions $\sigma_{x,y}$ of the bunch determine the number of particle collisions. Multiplied with the frequency f of the bunch crossing this gives the luminosity:

$$\mathcal{L} = \frac{N_1 N_2 f}{4\pi\sigma_x\sigma_y}. \tag{2.6}$$

For a certain process the interaction rate can be calculated by multiplying the luminosity with the cross section σ, which corresponds to the probability of a physics process to happen at a given energy. Integrated over time this yields the expected number of events of a given process in the accumulated data of an experiment:

$$\frac{dN}{dt} = \mathcal{L} \cdot \sigma \tag{2.7}$$

$$N = \int \mathcal{L} \cdot \sigma dt \tag{2.8}$$

There are two main types of colliders, one colliding leptons and one colliding hadrons.

Electron-Positron Collider The advantage of colliding electrons and positrons is that they do not have a substructure. If a collision takes place, the energy in the center of mass is well known. This allows for high precision measurements as well as for special analysis techniques like the full reconstruction of one "tag" particle in the case of the production of one $B\bar{B}$ pair. Unfortunately the maximal energy achievable on circular colliders is limited due to synchrotron radiation emitted by charged particles when they are accelerated radially. The energy loss per turn can be expressed as

$$\Delta E = \frac{1}{3}\left(\frac{e^2\beta^3\gamma^4}{\rho}\right), \tag{2.9}$$

where e is the electric charge, β the velocity, γ the Lorentz factor and ρ the bending radius. Since $\gamma \sim E/m_0$, the energy loss is high for light particles:

$$\Delta E \sim \frac{1}{m^4}, \tag{2.10}$$

limiting the achievable energies in circular accelerators.

Hadron Collider In hadron colliders, usually proton-proton or proton-anti-proton collisions are examined. They achieve the highest possible energies in accelerators and they are much heavier and are therefore less prone to energy loss via synchrotron radiation. The synchrotron radiation for electrons is about 10^{13} times higher compared to that of protons. In opposition to electrons, protons do have a substructure and thus usually only one constituent of each hadron is colliding, carrying far less energy than the whole proton. Therefore the energy in a particular collision is unknown.

2.2. The KEK B Accelerator

After the electron-proton collider projects TRISTAN I and II at the KEK had ended, in 1994 the construction of the electron-positron accelerator KEK B began Fig. 2.1. The aim was to

build a so-called B-factory, where a large amount of B meson pairs could be produced. For this task the center of mass energy was chosen to be approximately at the energy of the $\Upsilon(4S)$ resonance $\sqrt{s} = 10.58\,\mathrm{GeV}$ which decays in $\sim 96\%$ of the time into a B meson pair. One of the main goals of the KEK B B-factory was to measure CP violation in the decay of B mesons where it is necessary to measure the decay time. However, the B meson lifetime is very short ($\sim 10^{-12}\,\mathrm{s}$) and therefore it is necessary to have the $B\bar{B}$ pair boosted to travel measurable distances before they decay. To account for this, the KEK B accelerator consists of two rings with different beam energies:

- LER: a low energy ring containing positrons at an energy of $E_+ = 3.5\,\mathrm{GeV}$
- HER: a high energy ring containing electrons at an energy of $E_- = 8.0\,\mathrm{GeV}$

They carry a current of $1600\,\mathrm{mA}$ and $1200\,\mathrm{mA}$, respectively. The asymmetry can be expressed by the resulting Lorentz boost:

$$\beta\gamma = \frac{E_- - E_+}{\sqrt{s}} = 0.43 \tag{2.11}$$

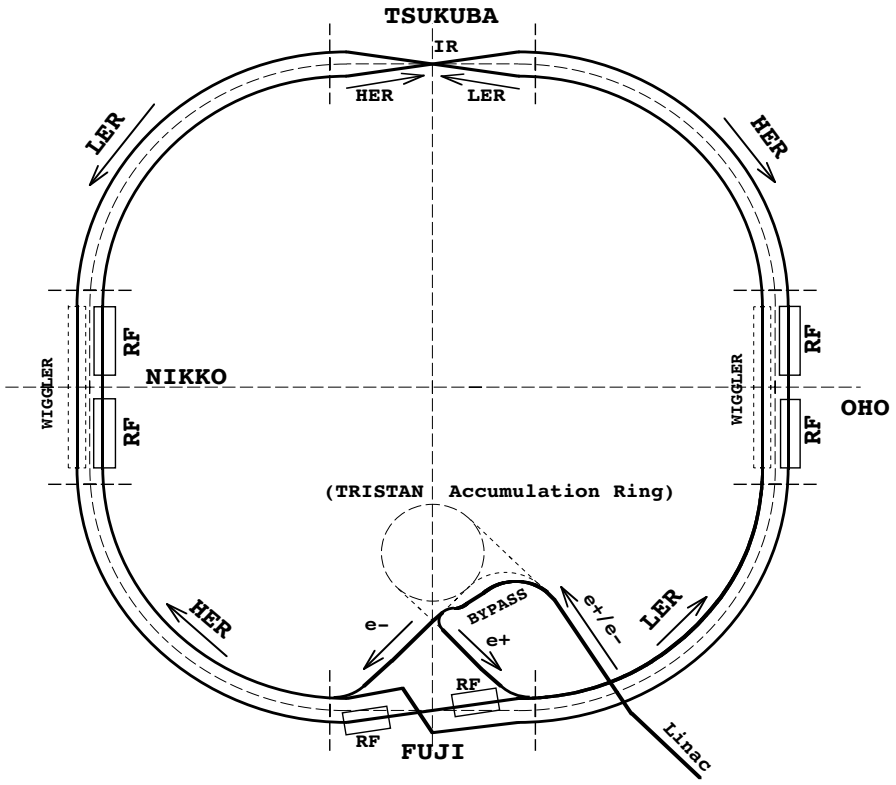


Figure 2.1.: The KEK B accelerator

The beams cross at an angle of $22\,\mathrm{mrad}$ in the Tsukuba interaction area, allowing to separate the beams effectively after collision without a high detector background level. The KEK B B-factory achieved a world record luminosity of $2.11 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, which exceeds the original design luminosity by more than a factor of 2.

2.3. The Belle Detector

The *Belle detector* [8] was designed for event detection in the KEK-B-factory, an asymmetric e^+e^- collider operating at the $\Upsilon(4S)$ resonance. The aim was to study rare B meson decays

and the Kobayashi–Maskawa mechanism for CP-violation with data samples containing $\sim 10^9$ B meson decays. The detector encloses the Tsukuba interaction region and is constructed around a $1.5\,\mathrm{T}$ super conducting solenoid that uses the iron structure as a yoke. Outside the cylindrical beryllium beam pipe the first component of the detection is the silicon vertex detector (SVD) which measures B decay vertices. It is followed by the central drift chamber (CDC), that tracks charged particles and gives more information for particle identification by measuring the energy deposition dE/dx. Outside of the CDC, an aerogel threshold Cerenkov counter (ACC) and a time-of-flight counter (TOF) provide additional information for particle identification (PID). To detect electromagnetic showers, an array of CsI(Tl) crystals serves as the electromagnetic calorimeter. The most distant component of the detector is for K_L^0 and $\mu^\pm$ detection (KLM). It is build of resistive plate counters and is situated outside the iron yoke. To improve the polar angle coverage an extreme forward calorimeter (EFC) is used at the front and end cap of the detector. The overall layout can be seen in figure Fig. 2.2. The above mentioned parts of the detector will be described briefly in the following sections, following and using parts of the extensive description of the detector in the technical design report [8]. The forward direction along the positive z-axis is defined to point in the direction of the high energetic ring (HER) beam. The x-axis lies in the horizontal plane and the y-axis is perpendicular to the x- and z-axis. The angle θ is measured with respect to the z-axis and ϕ is measured with respect to the x-axis.

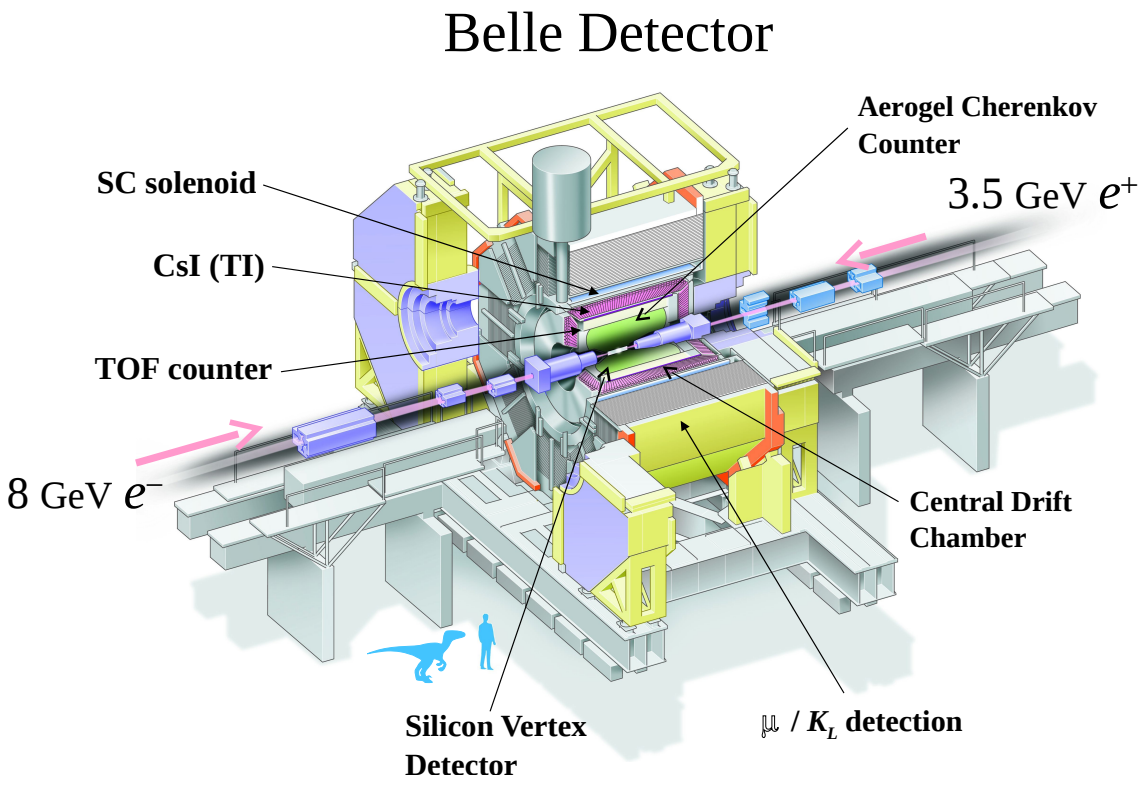


Figure 2.2.: The Belle detector

SVD One of the main goals of Belle was the measurement of time dependent CP violation in B meson decays. Therefore a precise knowledge of the decay vertex position of the B mesons on the z axis is crucial. A resolution of $\approx 100\,\mu\mathrm{m}$ also allows to use the SVD for D meson and τ identification and helps to improve track reconstruction. The first generation (SVD1) was a three layer silicon vertex detector with such a resolution. Its polar angle coverage was $23° < \theta < 139°$ and the radii of the layers were 30, 45.5 and $60.5\,\mathrm{mm}$. Due to massive damage from radiation, it was replaced by the second generation (SVD2) in 2003 [9]. It was decided to improve the SVD by adding two more

layers to cover a larger range in the direction of the radius:

$$r_1 = 15.0\,\mathrm{mm},\ r_2 = 21.1\,\mathrm{mm},\ r_3 = 44.0\,\mathrm{mm},\ r_4 = 70.0\,\mathrm{mm},\ r_5 = 90.0\,\mathrm{mm} \quad (2.12)$$

The coverage on the polar angle of SVD2 is $17^\circ < \theta < 150^\circ$ and the precision of the z-difference measurement improved to $\sim 80\,\mu\mathrm{m}$ because of the new layers closer to the interaction point.

CDC The efficient reconstruction of charged particle tracks and the precise determination of their momenta are the main tasks of the *Central Drift Chamber*. The chamber is filled with a mixture of 50% helium and 50% ethane gas. This ratio has been chosen to minimize multiple coulomb scattering (low-Z gas) and allow also for a good resolution for a dE/dx measurement. The CDC inner radius is $103.5\,\mathrm{mm}$, the outer radius $874\,\mathrm{mm}$. The chamber has 50 cylindrical layers, each containing between three and six either axial or small-angle-stereo layers, and three cathode strip layers. The individual drift cells are nearly square with a maximum drift distance of $8 - 10\,\mathrm{mm}$ and add up to a total of 8400. The performance of the CDC can be seen in the resolution parameters:

$$\sigma_{r\phi} = 130\,\mu\mathrm{m} \quad (2.13)$$

$$\sigma_z = 200 - 1400\,\mu\mathrm{m} \quad (2.14)$$

$$\frac{\sigma_{p_t}}{p_t} = \frac{0.3\%}{\sqrt{p_t^2+1}} \quad (2.15)$$

$$\sigma_{dE/dx} = 6\% \quad (2.16)$$

ACC Particle identification plays a key role in the performance of a detector. Especially light mesons like Kaon and Pion are hard to distinguish. In the momentum range $1.2\,\mathrm{GeV\,c^{-1}} < p < 3.5\,\mathrm{GeV\,c^{-1}}$ the mass of a Kaon or Pion is small enough to let them travel faster than the speed of light in the scintilator material. Thus they emit Cerenkov radiation, the opening angle proportional to the velocity. In combination with the momentum measurement from CDC this allows to estimate the mass of the particle. The ACC consists of 960 counter modules in the ϕ direction and 228 modules for the forward direction. The refractive indices of aerogels are selected to be between 1.01 and 1.03, depending on their polar angle region (figure Fig. 2.3). To get uniform response for light velocity particles, the counters are read out by fine mesh-type photo multiplier tubes with diameters different for different refractive indices.

TOF For momenta below $1.2\,\mathrm{GeV\,c^{-1}}$ even the light mesons are too slow to discriminate them using the ACC. But they are slow enough for the resolution of the trigger ($\sigma_t \sim 100$ps) to suffice to measure the time they need to travel from the interaction point to the Time Of Flight system which is $\sim 1.2\,\mathrm{m}$ away. Because nearly 90% of the particles from the $\Upsilon(4S)$ resonance have a momentum below $1.2\,\mathrm{GeV\,c^{-1}}$, the TOF system can provide a clean and efficient b-flavor tagging. The TOF counters are augmented with *Thin Scintillation Counters* (TSC) forming 64 modules that cover a polar angle of $34^\circ < \Theta < 120^\circ$ and are read out by fine-mesh-dynode photo multiplier tubes.

ECL The main purpose of the Electromagnetic Calorimeter is to detect the energy and position of photons. On the one hand, photons often are the end products of cascade decays and thus a good performance at energies below 500 MeV is required. On the other hand, photons coming from two-body decays like $B \to K^*\gamma$ or $B^0 \to \pi^0\pi^0$ can have energies up to $4\,\mathrm{GeV}$. For π^0 meson decays, a good angular resolution is necessary to separate nearby photons. Since the electron identification relies primarily on the comparison of the charged track momentum and the energy deposition in the ECL, good energy resolution leads to good hadron rejection. To satisfy these requirements, a highly segmented array

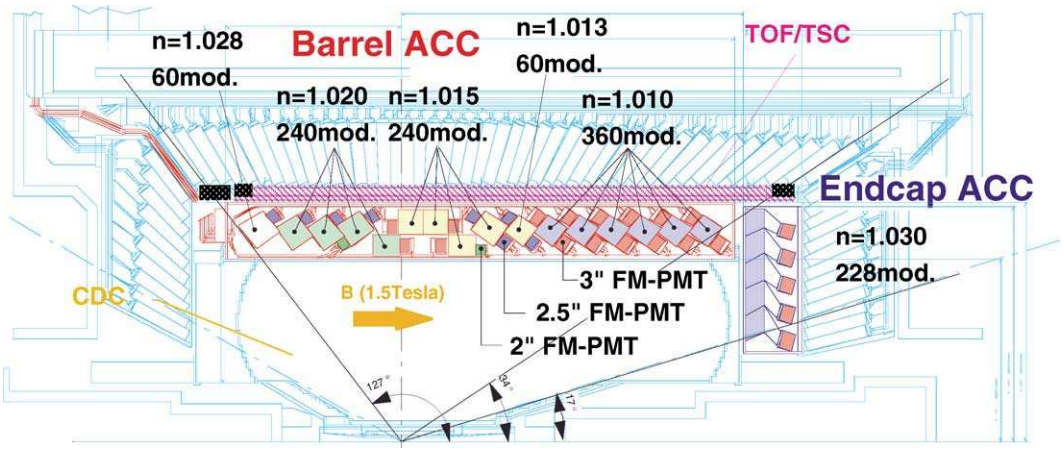


Figure 2.3.: Scheme of the detector components used for PID

of CsI(Tl) crystals with silicon photo diode readout is used. The barrel section is $\sim 3\,\mathrm{m}$ in length and has an inner radius of $1.25\,\mathrm{m}$. The angular coverage and the number of crystals are listed in table Table 2.1.

Item	θ coverage	No. of crystals
Front cap	$12.4 - 31.4°$	1152
Barrel	$32.2 - 128.7°$	6624
Endcap	$130.7 - 155.1°$	960

Table 2.1.: Geometrical parameters of the ECL

EFC In order to improve the experimental sensitivity, the polar angle coverage of the ECL is extended by an *Extreme Forward Calorimeter*. The EFC covers the region from $6.4°$ to $11.5°$ in the forward direction and $163.3°$ to $171.2°$ in the backward direction. It is placed in the very high radiation level area around the beam pipe and therefore is built from Bismuth-Germanate crystals.

Magnet The magnetic field used to force charged particles to a circular path, in order to measure their momenta, is provided by a super conducting NbTi/Cu solenoid. Because it is surrounded by a multi layer structure of iron plates and calorimeters (see KLM) which works as a magnetic return circuit, the magnetic field is located inside a cylindrical volume of $3.4\,\mathrm{m}$ in diameter and $4.4\,\mathrm{m}$ in length. The coil itself has a length of $3.92\,\mathrm{m}$, an effective radius of $1.8\,\mathrm{m}$ and carries a current of $4400\,\mathrm{A}$ corresponding to a stored energy of $35\,\mathrm{MJ}$. While the coil is comparatively lightweighted with $\sim 23\,\mathrm{t}$, the iron yoke weights 1132 tons.

KLM The KLM detection system is the outermost part of the detector. It is built from alternating layers of 15 glass-resistive plate counters and 14 $4.7\,\mathrm{cm}$ thick iron plates. This provides a total of 3.9 interaction lengths in the angular range of $20°$ to $155°$ to convert K_L's. Another 0.8 interaction lengths are added by the ECL. The showers produced by an arriving K_L meson indicate the direction of the K_L and allow to separate it from a muon, which travels much further and with smaller deflections.

3. Analysis techniques

3.1. Dalitz analysis technique

3.1.1. Introduction

In a two-body decay, the energies of the two daughter particles are fully defined because of energy and momentum conservation. In contrast, in a three-body particle decay there are nine degrees of freedom. After requiring momentum and energy conservation, there are five degrees of freedom left. Three of them can be represented by the three Euler angles and thus define arbitrary rotations. In the case of spin zero initial and final state particles, the orientation of the so-called Dalitz plane is irrelevant leaving two degrees of freedom to describe the properties of the decay. The Dalitz plot of the decay is the scatter plot of this pair of parameters. While there are many possibilities for such a pair of parameters, there are some common choices. While in 1953 R. H. Dalitz [10] proposed to use the kinetic energies of two daughter particles in his study of the decay of Kaons into three Pions (which is convenient for non-relativistic decays), it is generally more suitable to use the squared invariant masses of two pairs of final state particles. The differential decay width is given by

$$d\Gamma = \frac{1}{(2\pi)^3}\frac{1}{32M^3}|\mathcal{A}|^2 dm_{12}^2 dm_{23}^2, \tag{3.1}$$

where $\mathcal{A}$ represents the decay amplitude element, M the mass of the mother particle and m_{12}, m_{23} are the invariant masses of the daughters 1 and 2 or 2 and 3, respectively. From Eq. (3.1) it is apparent that in the parametrization of squared invariant masses the phase-space term is constant within the kinematic limits, because for non-resonant decays $\mathcal{A}$ does not depend on m_{12}^2 or m_{23}^2. Any observed structure in the Dalitz plot hence belongs to a resonant decay. In some cases, a different parametrization called Squared Dalitz plot is useful (see Section 3.1.4). The description of the details of the Dalitz analysis technique in the following sections follows the *Physics of the B factories* book [11], written by the Belle and BaBar collaborations. The Dalitz plot technique can provide advantages in the following types of measurements:

- Measurements of properties of resonances like mass, width, and quantum numbers
- Searches for new states
- CP violation searches
- Studies of flavor mixing

3.1.2. Kinematic boundaries

The invariant masses of pairs of the daughter particles are constrained to the masses of the involved particles via

$$m_{12}^2 + m_{23}^2 + m_{13}^2 = M^2 + m_1^2 + m_2^2 + m_3^2. \tag{3.2}$$

Since the involved particle masses are known, all the information of the resonant substructure is contained in two of the invariant masses, yielding three different possibilities to define the Dalitz plot, all containing the same information.

The maximum value of a given invariant mass of a two-body combination occurs, when the third daughter particle is produced at rest in the restframe of the decaying particle. For a given value of one invariant mass, the range of the other invariant masses can be written in terms of the given pair, e.g. m_{12}^2:

$$\begin{aligned}(m_{23}^2)_{\text{max}} &= (E_2^* + E_3^*)^2 - (p_2^* - p_3^*)^2, \\ (m_{23}^2)_{\text{min}} &= (E_2^* + E_3^*)^2 - (p_2^* + p_3^*)^2,\end{aligned} \tag{3.3}$$

where

$$\begin{aligned}p_2^* &= \sqrt{E_2^* - m_2^2}, \\ p_3^* &= \sqrt{E_3^* - m_3^2}\end{aligned} \tag{3.4}$$

are the momenta in the rest frame of the combination of particles 1 and 2 and

$$\begin{aligned}E_2^* &= \frac{m_{12}^2 - m_1^2 + m_2^2}{2m_{12}}, \\ E_3^* &= \frac{M^2 - m_{12}^2 + m_3^2}{2m_{12}}\end{aligned} \tag{3.5}$$

are their corresponding energies. Fig. 3.1 illustrates the positions of different kinematic configurations in the Dalitz plane.

3.1.3. Amplitude description

Experimental data suggests that hadronic three-body B decays proceed predominantly through resonant two-body decays. Therefore, the decay amplitude $\mathcal{A}(m_{13}^2, m_{23}^2)$, introduced in Eq. (3.1), consists of a coherent sum of two-body amplitudes (subscript r) and a non-resonant contribution (NR):

$$\mathcal{A}(m) = \sum_r a_r e^{i\phi_r} \mathcal{A}_r(m) + a_{NR} e^{i\phi_{NR}} \mathcal{A}_{NR}(m). \tag{3.6}$$

The parameters a and ϕ are the magnitude and the phase of the corresponding component and m is a point in the Dalitz plane: $m \equiv (m_{13}^2, m_{23}^2)$. The functions $\mathcal{A}(m)$ contain the dynamic properties of the decay into the three final state particles via the given resonance r.

3.1.3.1. Isobar formalism

In the so-called isobar formalism, each function $\mathcal{A}_r$ describes a single intermediate resonance. They take the form

$$\mathcal{A}_r = F_P \times F_r \times T_r \times W_r, \tag{3.7}$$

where F_P and F_r are the transition form factors of the parent particle and resonance, respectively. The resonance propagator $T_r \times W_r$ describes the line shape (T_r) and angular dependence (W_r) of the resonance.

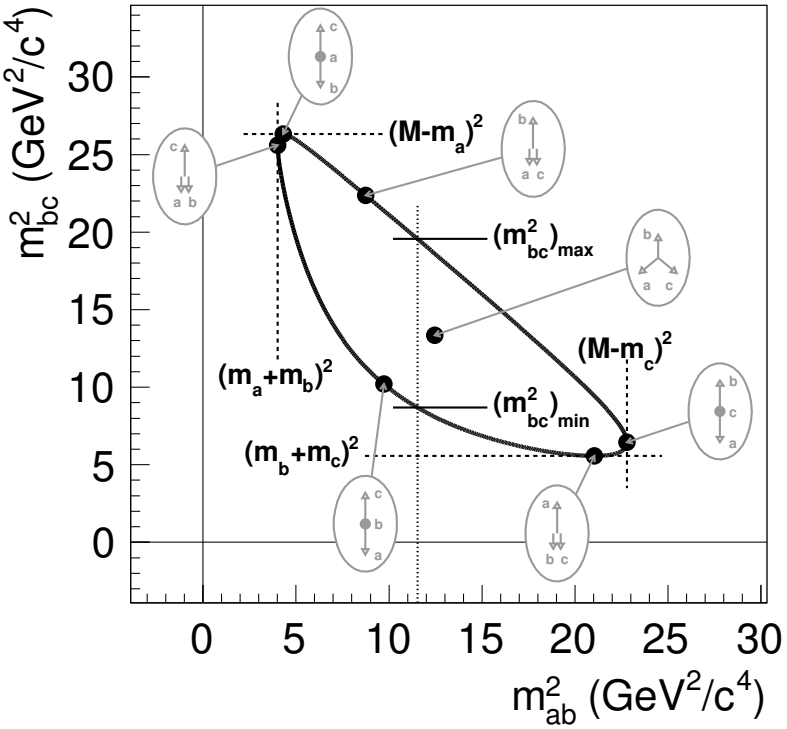


Figure 3.1.: Kinematic boundaries of the three-body decay phase space and illustration of various kinematic configurations [11].

3.1.3.2. Blatt Weisskopf penetration factors

To describe the form factors F_P and F_r, the Blatt Weisskopf parametrization is commonly used [12]. They depend on the spin J of the intermediate resonance and take the form

$$\begin{aligned} J = 0: \quad F &= 1, \\ J = 1: \quad F &= \sqrt{\frac{1 + R^2 q_r^2}{1 + R^2 q_{13}^2}}, \\ J = 2: \quad F &= \sqrt{\frac{9 + 3R^2 q_r^2 + R^4 q_r^4}{9 + 3R^2 q_{13}^2 + R^4 q_{13}^4}}. \end{aligned} \tag{3.8}$$

In this equation q_{13} is the momentum of the daughter particles in the center-of-mass frame of particle 1 and 3, while q_r denotes the momentum these decay products would have in the rest frame of the resonance with mass m_r. R is the so-called meson radius and usually takes values between 1 and $5\,\mathrm{GeV}^{-1}$. As the B meson is a pseudo-scalar, the form factor used is $F_P = 1\,\mathrm{GeV}^{-1}$. For the intermediate states with nonzero spin the effective meson radius used is $R = 4.0\,\mathrm{GeV}^{-1}$ for charmed mesons and $R = 5.3\,\mathrm{GeV}^{-1}$ for the light meson ρ throughout this analysis.

3.1.3.3. Line shapes

The line shapes used for the dynamical function T_r to parametrize resonance shapes are listed below.

Relativistic Breit-Wigner: Most common line shape with a mass-dependent width Γ_{13}:

$$T_r = \frac{1}{m_r^2 - m_{13}^2 - i m_r \Gamma_{13}}, \tag{3.9}$$

with

$$\Gamma_{13} = \Gamma_r \left(\frac{q_{13}}{q_r}\right)^{2J+1} \left(\frac{m_r}{m_{13}}\right) F_r^2, \tag{3.10}$$

where Γ_{13} and J are the width and spin of the resonance and q and m are defined in Eq. (3.8). This shape works well only in the case of narrow states, because the use of the mass-dependent width results in a non-analytic T_r function.

Gounaris Sakurai: An alternative parametrization providing a better description for broad vector resonances such as $\rho(770)$. It is a model of the $\pi\pi$ scattering amplitude [13]:

$$GS(m_{\pi\pi}^2) = \frac{m_\rho^2 + d \cdot m_\rho \Gamma_\rho}{m_\rho^2 - m_{\pi\pi}^2 - i \cdot m_\rho \Gamma_{\pi\pi} + \Gamma_\rho \frac{m_\rho^2}{p_\rho^3} \cdot \left[p_{\pi\pi}^2 \cdot (h_{\pi\pi} - h_\rho) + p_\rho^2 \cdot \frac{dh_{\pi\pi}}{dm_{\pi\pi}^2}|_{m_\rho^2} \cdot (m_\rho^2 - m_{\pi\pi}^2)\right]},$$

where

$$\begin{aligned} d &= \frac{3}{\pi} \cdot \frac{m_\pi^2}{p_\rho^2} \cdot \ln\left(\frac{m_\rho + 2 \cdot p_\rho}{2m_\pi}\right) + \frac{m_\rho}{2\pi p_\rho} - \frac{m_\pi^2 m_\rho}{\pi p_\rho^3}, \\ h_{\pi\pi} &= \frac{2}{\pi} \cdot \frac{p_{\pi\pi}}{m_{\pi\pi}} \cdot \ln\left(\frac{m_{\pi\pi} + 2 \cdot p_{\pi\pi}}{2m_\pi}\right). \end{aligned} \tag{3.11}$$

Non resonant: Every decay not proceeding via an intermediate resonance is assumed to populate the Dalitz plot with equal probability for every position. Therefore the non-resonant contribution can be described with a flat distribution. To account for possible very broad states or final-state interactions, that can make the non-resonant part non-uniform, there were two possible empirical models proposed by Belle [14] and BaBar, where s and t are the Mandelstam variables:

$$\begin{aligned} A_{NR_{Belle}} &\propto e^{-\alpha s} + e^{-\alpha t} \\ A_{NR_{BaBar}} &\propto 1 + \alpha \frac{s+t}{m_B^2} \end{aligned} \tag{3.12}$$

In this analysis, the contribution of the non resonant decay is expected to be very small and the flat distribution is used.

3.1.3.4. Angular distribution

The angular dependence W_r is either described using Zemach tensors, where transversality is enforced, or the helicity formalism, that allows for a longitudinal component. In this analysis the initial and final state particles have all total spin 0 and therefore the Zemach formalism is used. The Zemach expressions are

$$\begin{aligned} L = 0 : T &= 1 \\ L = 1 : T &= -2 \cdot \vec{p} \cdot \vec{q} \\ L = 2 : T &= \frac{4}{3}[3(\vec{p} \cdot \vec{q})^2 - (|\vec{p}| \cdot |\vec{q}|)^2] \end{aligned} \tag{3.13}$$

The vectors $\vec{p}$ and $\vec{q}$ are measured in the rest frame of the decaying resonance, $\vec{p} \cdot \vec{q}$ is proportional to the cosine of the helicity angle $\cos\theta_H$. This explains why the resonance bands in a Dalitz plot show information about the spin of the resonance, a spin 0 resonance is flat, a spin 1 resonance follows $\cos^2\theta_{ij}$ and a tensor resonance is distributed according to $|3\cos^2\theta_{ij} - 1|^2$.

3.1.4. Squared Dalitz plot

The events where the decay proceeds via an intermediate state are of great interest, but because of the low masses of the daughter particles compared to the B meson mass in this Dalitz analysis, they populate the region near the kinematic boundaries of the Dalitz plot. Also the interferences between light meson resonances appear in this area. Thus, the region of importance where one has the greatest sensitivity to relative phases occupies only a small part of the phase space and large variations could occur over small areas. Additionally, if some contribution are represented by two dimensional histograms, it is unavoidable that some bins cross the kinematic boundaries of the Dalitz plot, leaving them underpopulated and prone to large statistical fluctuations. A possible solution is to map the kinematic variables to a rectangle Dalitz plot, the so-called *squared* Dalitz plot, in such a way, that the regions of interest get moved away from the boundaries and populate larger areas. A common definition was proposed by BaBar [15]:

$$dm_{ab}^2 dm_{bc}^2 \longrightarrow |detJ|\, dm' d\theta', \tag{3.14}$$

defining the new variables as

$$\begin{aligned} m' \equiv & \frac{1}{\pi} arccos\left(2\frac{m_{ac} - m_{ac}^{min}}{m_{ac}^{max} - m_{ac}^{min}} - 1\right), \\ \theta' \equiv & \frac{1}{\pi}\theta_{ac}, \end{aligned} \tag{3.15}$$

where $m_{ac}^{max} = M - m_b$ and $m_{ac}^{min} = m_a + m_c$, θ_{ac} is the helicity angle of the ac combination, and J is the Jacobian of the transformation. The transformation is made visible in Fig. 3.2, where the result of the transformation of a flat distribution in the normal Dalitz plot is shown. Throughout this thesis, the terms *mPrime* and *thPrime* will be used for m' and θ', respectively.

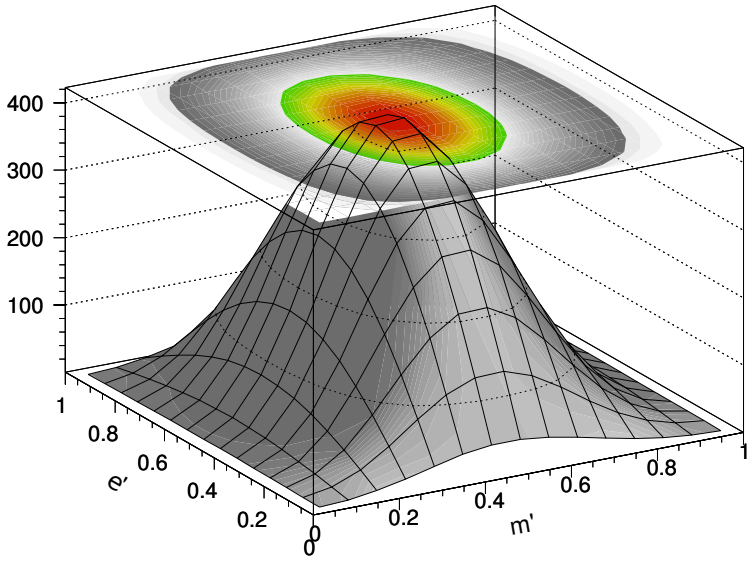


Figure 3.2.: Distribution of uniformly distributed events over the normal Dalitz plot in the squared Dalitz representation [11]

3.2. Kernel density estimation

Kernel density estimation (KDE) estimates the probability density function of a random variable in a non-parametric way. The construction of a kernel density estimation is quite similar to histograms: For histograms, the X-axis is divided in bins, and for each data point that lies within the bin boundaries, the height of the bin is incremented by $1/N_{points}$. To construct a KDE, to each data point, a kernel is assigned and all kernels are then added up to make the KDE (see Fig. 3.3). This procedure ensures smoothness compared to the discrete nature of histograms. It also preserves possible correlations in the data.

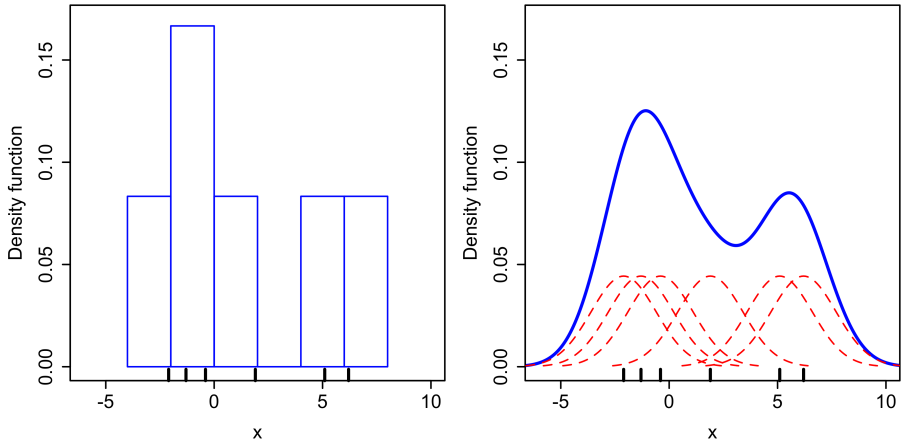


Figure 3.3.: Comparison of the construction of a histogram and a kernel density estimate [16]

The KDE of a function f is

$$\hat{f}_h(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - x_i) = \frac{1}{nh}\sum_{i=1}^{n} K(\frac{x - x_i}{h}), \tag{3.16}$$

where K is the kernel and h a smoothing parameter called *bandwidth*. While there are different possible kernel functions, often a Gaussian distribution is used due to its convenient mathematical properties. In this analysis a KDE with a Gaussian kernel is used to smooth the two dimensional distributions of the continuum and combinatoric background. A wrong choice of the bandwidth could under-smooth (create data artifacts) or over-smooth (obscure underlying structure) the distribution. Therefore an adaptive approach is used, where the bandwidth is varied depending on the local density. For this, the implementation from RooFit [17] is used.

3.3. Maximum Likelihood

When dealing with a measurement of a parameter or a set of parameters, the usual approach is to define an estimator $\hat{a}$ of the true value a_0, that satisfies the following criteria:

1. Consistency: $\lim_{n\to\infty} \hat{a} = a_0$
2. Unbiased: The expectation value $E[\hat{a}]$ should be equal to the true value a_0
3. Efficiency: The variance of $\hat{a}$ should be small
4. Robustness: The method should be robust in the case of wrong data or wrong assumptions

Usually not all of this criteria can be matched, e.g., the maximum likelihood method is very effective but can have large discrepancies if the probability density function is wrong. The maximum likelihood method and the least squares method are the most widely used methods for parameter estimation. The fits in this analysis use the extended unbinned maximum likelihood method, therefore it shall be introduced in this section, following closely [18]. If there are n measurements of a random variable x, that follows a probability density function (PDF) $f(x|a)$ with parameters a, the likelihood function is defined as

$$L(a) = f(X_1|a) \cdot f(X_2|a) \cdots f(X_n|a) = \prod_{i=1}^{n} f(X_i|a). \tag{3.17}$$

It gives the probability to observe the values of x given a set of Parameters a. Maximizing $L(a)$ yields the best estimation of a. The PDF is normalized to one for all values of a, causing the main numeric effort of this method. The maximum is found by differentiating $L(a)$, but for practical reasons, e.g., machine precision among others, the negative logarithm is taken:

$$F(a) = -\sum_{i=1}^{n} \ln f(x_i|a). \tag{3.18}$$

In this analysis this quantity is minimized making use of the likelihood fitting package LAURA++ [19], that has an interface to the Minuit [20] minimization package.

3.3.1. Extended Maximum Likelihood fits

If the yields of different components are to be extracted, the probability density $f(x|a)$ is no longer normalized. $g(x|a)$ is defined in such a way, that the integral of g corresponds to the number of expected events N:

$$N = \int_{\Omega} g(x|a)dx, \tag{3.19}$$

where N depends on the parameter set a. The negative log likelihood function is thus

$$F(a) = -\sum_{i=1}^{n} ln(g(x_i|a) + \int_{\Omega} g(x|a)dx. \tag{3.20}$$

3.3.2. Error calculation in the Maximum Likelihood method

In the case of $n \to \infty$ the likelihood function becomes Gaussian and the variance approaches zero. Then it is possible to expand the log likelihood function at the minimum, where $dF(a)/da = 0$ is true. This yields

$$F(a) = F(\hat{a}) + \frac{1}{2} \cdot \frac{D^2F}{da^2} \cdot (a - \hat{a})^2 + ... \tag{3.21}$$

or

$$L(a) \cong c \cdot \exp\left(-\frac{1}{2} \cdot \frac{d^2F}{da^2} \cdot (a - \hat{a})^2\right) = c \cdot \exp\left(-\frac{(a - \hat{a})^2}{2\sigma^2}\right). \tag{3.22}$$

The negative log likelihood function has the shape of a parabola and its second derivative is constant

$$\sigma(\hat{a}) = \left(\frac{d^2F}{da^2}\bigg|_{\hat{a}}\right)^{-1/2}. \tag{3.23}$$

This can be used to extract errors in the case of parabolic behavior of the log likelihood function, but otherwise a nonlinear transformation of a in a parameter $z = z(a)$ is necessary in such a way that $F(z)$ is parabolic. Because of the invariance of the parameter estimation of the maximum likelihood method, the best value $\hat{z}$ is $z(\hat{a})$ and one can get the left and right standard deviations from

$$F(\hat{z} \pm \sigma_z) = F(\hat{z}) \pm \frac{1}{2} = F(\hat{a}) \pm \frac{1}{2}. \tag{3.24}$$

This can be generalized to multiple parameters, and the covariance matrix $\mathbf{V}$ of the parameter vector $\mathbf{a}$ is then given by

$$\mathbf{V} = \mathbf{G}^{-1} \qquad \text{with} \qquad G_{ik} = \frac{\partial^2 F}{\partial a_i \partial a_k} \tag{3.25}$$

at the minimum $\hat{a}$. $\mathbf{G}$ has the form of the Hesse matrix, and in the non-asymptotic case the inverted Hesse matrix in the minimum is an approximation for the covariance matrix.

The calculation of the asymmetric errors is implemented in the package MINOS, that is contained in Minuit.

4. Reconstruction and selection of B meson candidates

4.1. Reconstruction of B mesons

The full decay chain of a B meson is reconstructed using tracks and clusters left in the electromagnetic calorimeter of the final state particles. The reconstruction is structured hierarchically with selection criteria to be met in each intermediate step.

4.1.1. Reconstructed decay channels

Two decay channels are reconstructed for this analysis, both very similar: $B^- \to D^0\pi^-\pi^0$ and $B^0 \to D^-\pi^+\pi^0$. The branching fractions for the decays via different intermediate states (Tables 1.1 and 1.2) are taken from the Particle Data Group, which provides current world averages [7]. The inclusive branching fractions for the doubly excited D mesons are obtained from measurements of $B \to D\pi^\pm\pi^\pm$ using Clebsch-Gordan coefficients. There are no measurements seeing a non-resonant contribution in both cases. Only the most clean decay mode for the decay of the charged or neutral D meson is reconstructed, to have a high purity in the reconstructed sample (Table 1.3). Throughout this document the properties listed for any decay channel apply as well to the charge conjugate state of this channel. In addition to that, for charged pions/kaons the symbol π/K is used while for neutral pions the lack of charge is indicated (π^0).

4.1.2. Final state particles

The first step of reconstruction is to read out the lists of the final state particles that occur in the B meson decay chain such as Pions, Kaons or neutral Pions. Some preselection requirements have to be met, see Table 4.1. To reduce the misidentification rate (e.g. take a Pion as a Kaon), combined information of ACC, TOF, CDC (dE/dx) is used in order to separate e, μ, K, π, p. The impact parameter cuts reject tracks that do not belong to the current event, e.g. particles originating from other processes in the beam pipe or from cosmic rays.

4.1.3. D mesons

In this step, the final state particles are used to reconstruct D mesons. The particles matching the particle types of the D children are combined with the requirement, that the reconstructed mass lies within a specified interval. The distribution of the D meson mass at this stage can be seen in Figs. 4.6a and 4.7a.

$\pi^{\pm}$	$PID_{\pi:K}$	> 0.4
	impact parameter $\lvert dz \rvert$	$< 4\,\mathrm{cm}$
	impact parameter $\lvert dr \rvert$	$< 2\,\mathrm{cm}$
$K^{\pm}$	$PID_{K:\pi}$	$>= 0.6$
	impact parameter $\lvert dz \rvert$	$< 4\,\mathrm{cm}$
	impact parameter $\lvert dr \rvert$	$< 2\,\mathrm{cm}$
π^0	E_γ barrel region	$> 50\,\mathrm{MeV}$
	E_γ front-/endcap region	$> 100\,\mathrm{MeV}$
	$\lvert M_{\pi^0} - M_{PDG} \rvert$	$< 25\,\mathrm{MeV}$

Table 4.1.: Basic requirements on PID, impact parameters and photon energy

- D^0: $|M_{D^0} - M_{PDG}| < 20\,\mathrm{MeV}$
- D^+: $|M_{D^+} - M_{PDG}| < 30\,\mathrm{MeV}$

4.1.3.1. Vertex reconstruction

The combination of particles to a D meson must only fulfill the requirements from above (Section 4.1.3), while the origin of the tracks is not considered at first. For this task a kinematic fitter is used which returns the most probable decay vertex for a given set of daughter tracks. The fit itself is a minimization of a χ^2 term constructed from the track parameters and the corresponding error matrices. This term also allows for additional constraints, like a mass constraint (see Section 7.2.2). If the fit fails or the probability for the resulting decay vertex does not reach a given level, the candidate is considered to be wrong. After a successful fit, the result is used to recalculate the mass of the mother particle.

For the D mesons the kinematic fitter is used to perform a vertex and a mass constraint fit, while for neutral Pions only a mass constrained fit is performed. For a Dalitz analysis the migration of reconstructed events away from the point in the Dalitz plot, where they have been produced, should be small. Therefore a B mass constrained fit is performed after the reconstructed properties of the B meson candidate have been saved, and the updated invariant masses of the daughters of the B meson are then saved into different variables. The updated invariant masses are then only used in forming the dalitz plot, but not in the selection procedure.

4.1.4. B reconstruction

The last step of the reconstruction is to combine the D meson and final state particles to a B meson. In this combination, the deviation of the reconstructed mass from the nominal value is required to be within 200 MeV.
One of the advantages of B factories is that in a B-$\bar{B}$ event, one of the B mesons has to carry half of the beam energy in the center of mass frame. Therefore it is useful to introduce two variables to take advantage of this knowledge: A cut to the beam constraint mass M_{bc} and the deviation of the energy of the B candidate from the beam energy ΔE is applied:

$$M_{\mathrm{bc}} = \frac{1}{c^2}\sqrt{E^{*2}_{\mathsf{Beam}} - c^2 p_B^2} = \frac{1}{c^2}\sqrt{E^{*2}_{\mathsf{Beam}} - c^2\left(\sum_i^N p_i\right)^2}, \tag{4.1}$$

$$\Delta E = E_B - E^*_{\mathsf{Beam}} = \sum_i \left(\sqrt{c^2 p_i^2 + c^4 m_i^2}\right) - E^*_{\mathsf{Beam}}, \tag{4.2}$$

where p_i are the momenta of the daughter particles used to reconstruct the B meson and E^*_{Beam} is the nominal beam energy corresponding to half of the center of mass energy. Before the B candidates and their properties are saved, a precut is made on these two variables, M_{bc} has to be above $5.24\,\mathrm{GeV\,c^{-2}}$ and the value range of ΔE is $-0.250\,\mathrm{GeV} < \Delta E < 0.250\,\mathrm{GeV}$.

4.2. Selection of B candidates

So far, only the reconstruction and online selection of the B mesons has been described. To further decrease the background and make the signal more prominent, an offline selection is applied. It consists of the following steps, which are described in detail below:

- continuum suppression - reject non-$b\bar{b}$ events
- cut-based selection
- best candidate selection - select the best candidate per event per channel

4.2.1. Continuum suppression

Although the center of mass energy is chosen to be at the $\Upsilon(4S)$ resonance, the e^+e^- collision will produce non $b\bar{b}$ quark pairs, called *continuum* events, with a three times larger cross-section than $e^+e^- \to \Upsilon(4S) \to B\bar{B}$. These quark-anti-quark pairs can hadronize leading to two or more hadronic jets in the event while b-$\bar{b}$ events have a spherical event topology. Particles like D mesons that originate in *continuum* events can also be used in the reconstruction, but surely don't belong to a signal event. It is tried to utilize the difference in the event shapes to reject such events.
In 1987 G. C. Fox and S. Wolfram tried to discriminate between two- and three-jet events by observables which characterize the shape of each event instead of finding a jet-axis by minimization [21]. A set of such variables is given by

$$H_l = \sum_{i,j} |\vec{p_i}|\,|\vec{p_i}|\,P_l(cos\phi_{i,j}) \tag{4.3}$$

where p_i is the momentum of particle i, $\phi_{i,j}$ is the angle between particle i and j in an arbitrary set of axes and P_l are the usual Legendre polynomials. These observables can also be used to distinguish non-$b\bar{b}$ and $b\bar{b}$ events. They are usually normalized by H_0 and are then called *Fox-Wolfram-moments* R_l:

$$R_l = \frac{H_l}{H_0}. \tag{4.4}$$

An enhancement of the discrimination power was achieved by calculating the double sum only for signal events and only for non signal but b-$\bar{b}$ events (called "other B") separately and using these in linear combination (*Super Fox-Wolfram-moments* (SFW)). Another improvement was made by H. Kakuno by adding information about the charge, treating neutral particles and charged tracks separately and by including a missing momentum vector to account for neutrinos, abbreviated as KSFW. Finally, there are 17 variables that can be used (i ranges from 0 to 4):

- E_{T}: scalar sum of transverse energy of all particles
- $H_i^{so,c}$: using only charged tracks of other B
- $H_i^{so,n}$: using only photons of other B
- $H_i^{so,v}$: using only missing momentum of other B
- H_i^{oo}: double sum only over other B tracks

More information about the event shape is given by the so-called *CLEO cones* [22]. They are defined as cones around the thrust axis in 10° intervals and the momenta of the particles inside each cone are added up. This yields a distribution of the momentum carried away by the decay products and thus allows to separate jet-like and spherical event shapes. Furthermore, the directions of the thrust axes of the two B mesons are independent while for two jets they are the same. So the distribution of the cosine of the angle between both thrust axes is flat for b-$\bar{b}$ events while for *continuum* it peaks at one.

Besides the shape variables there is even more information. First, the direction of the B candidate momentum relative to the beam axis. This variable is called $\cos\Theta_B$. For background events it is flat while for $B\bar{B}$ it follows $1-\cos^2\Theta$. This distribution is due to the interaction of the electron and positron being mainly electromagnetic at energies much larger than the electron/positron mass and thus the *vector meson* $\Upsilon(4S)$ can only be produced with the spin-projection $M=\pm1$ along the beam axis. The distance of the decay vertices (called Δz) of the two B mesons provides additional information as well. For signal events it is the shape of an exponential decay folded with a Gaussian compared to only a Gaussian distribution for background. Although it is correlated to the shape variables, one can use the flavor tagger output to reject *continuum*. The task of the flavor tag algorithm is to determine the sign of the B and how likely it is that the candidate really is a B meson. The larger the absolute value of flavor tag output r is, the more B-like the candidate is.

All the information mentioned above was combined using NeuroBayes [23] in a hierarchical system. One neural network was trained with the *KSFW* variables, one with the *CLEO-cones* and one for combining these two networks with $\cos\Theta_B$ and R_2, which is the most powerful variable of the basic Fox-Wolfram-moments. The chosen network was trained using the same variables as the ones mentioned before plus Δz and the flavor tagger output. The performance of these networks is visualized and compared in a purity-efficiency plot (Fig. 4.1), where the *RooKSFW* represents the root version of *KSFW*.

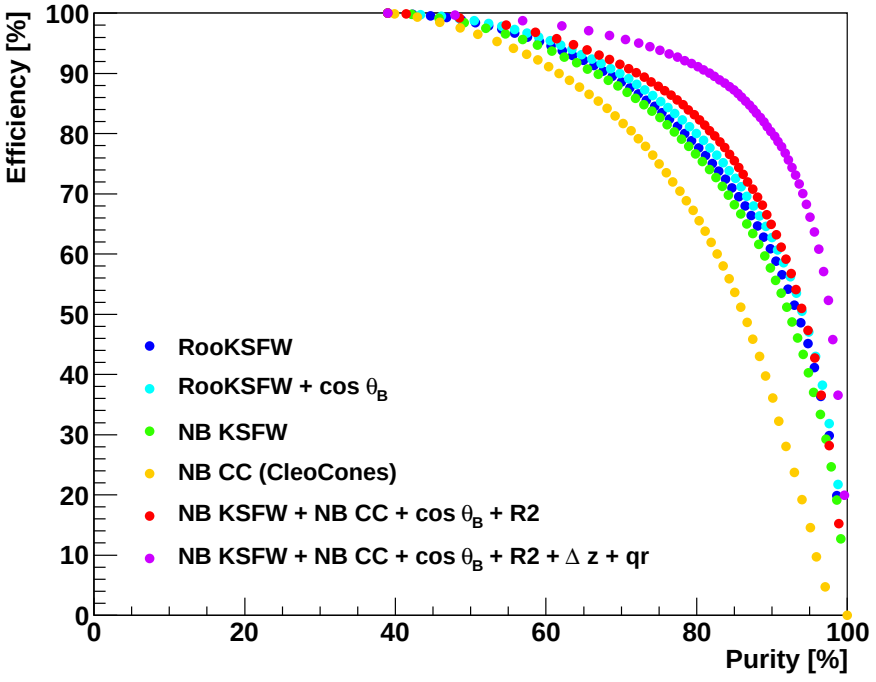


Figure 4.1.: Performance of different continuum suppression networks. The combination network (purple) yields the best performance and is chosen for this analysis.

The output of the neural network in the case of $B^- \to D^0\pi^-\pi^0$ can be seen in Fig. 4.2 for continuum and signal candidates. In Fig. 4.3 a comparison of the output for simulated off-resonance Monte Carlo events and for off-resonance recorded data can be seen. It shows that the network describes both simulated and recorded data in the same way above the cut chosen. While the continuum suppression rejects 61% (66%) of continuum events, 96% (95%) of signal pass the criterion in $B^- \to D^0\pi^-\pi^0$ ($B^0 \to D^-\pi^+\pi^0$). This can be seen in Figs. 4.4 and 4.5.

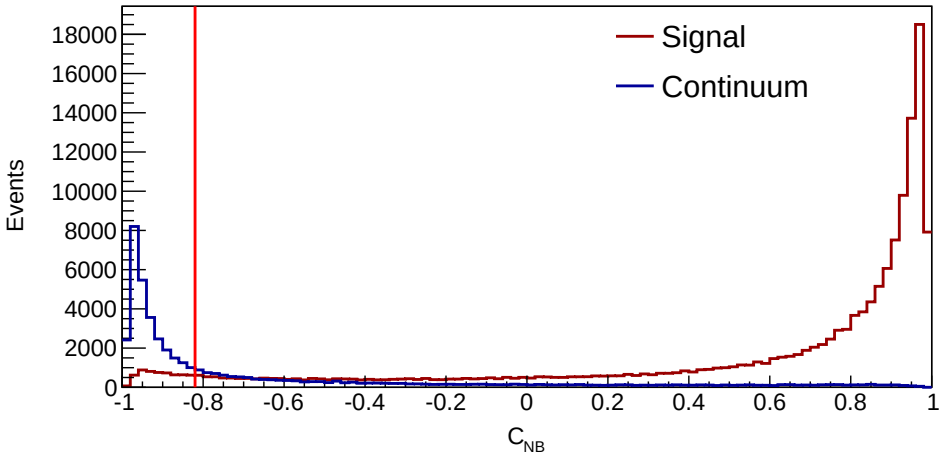


Figure 4.2.: Output of the neural network of the continuum suppression for continuum and signal candidates ($B^- \to D^0\pi^-\pi^0$). The red line indicates where the cut is applied and shows, that continuum background and signal decays can be well separated.

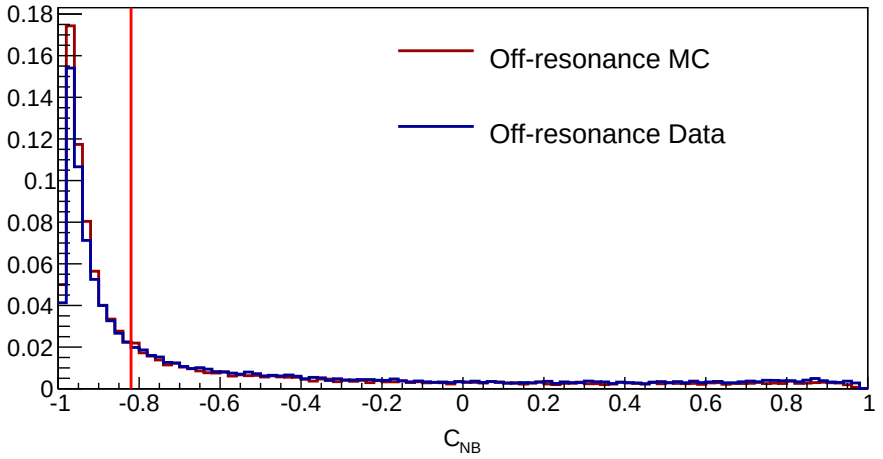
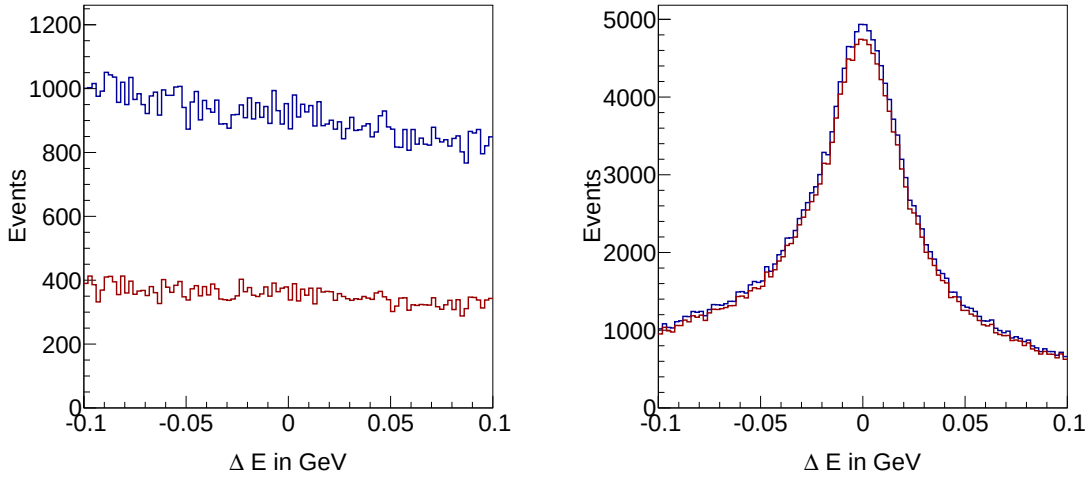


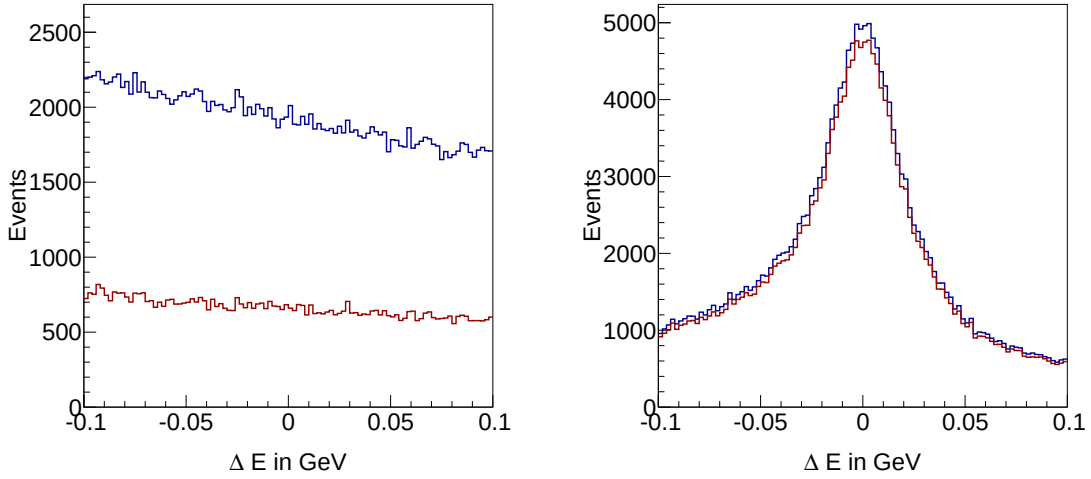
Figure 4.3.: Output of the neural network of the continuum suppression for simulated continuum and recorded off-resonance data ($B^- \to D^0\pi^-\pi^0$). The red line indicates where the cut is applied. Above the cut the output is similar for simulation and recorded data, indicating that the network performs in the same manner in both.



(a) Continuum suppression cut applied to continuum candidates from stream 0 Monte Carlo

(b) Continuum suppression cut applied to signal candidates from signal Monte Carlo

Figure 4.4.: Effect of the cut on the continuum suppression output on the ΔE distribution of $B^- \to D^0\pi^-\pi^0$ of continuum or signal candidates (before and after).



(a) Continuum suppression cut applied to continuum candidates from stream 0 Monte Carlo

(b) Continuum suppression cut applied to signal candidates from signal Monte Carlo

Figure 4.5.: Effect of the cut on the continuum suppression output on the ΔE distribution of $B^0 \to D^-\pi^+\pi^0$ of continuum or signal candidates (before and after).

4.2.2. Cut-based selection

Besides the cut on the continuum suppression network output, there are additional cuts applied. Since the aim of the selection is to reject as much background as possible while keeping as much signal as possible, the cuts were optimized by maximizing the following figure of merit:

$$\alpha = \frac{N_{sig}}{\sqrt{N_{sig} + N_{bkg}}} \tag{4.5}$$

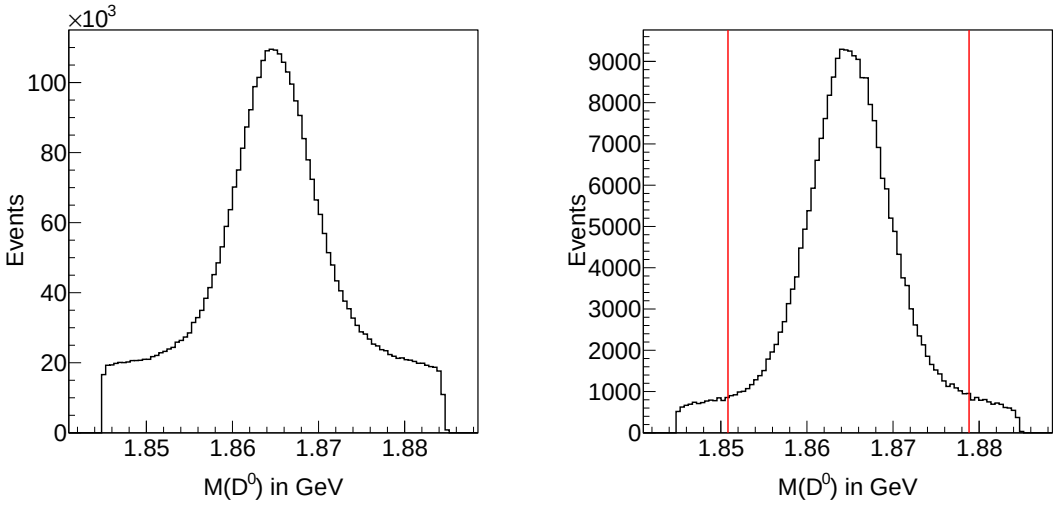
This optimization was done for each decay channel separately. The sample on which the optimization was performed is experiment 41, run $1-1000$ of the first stream of Monte Carlo. The cuts were determined sequentially. The cut values were varied in reasonable steps that are small regarding the changes in the figure of merit. An example for the optimization of the cut on the D mass can be seen in Fig. 4.8. The variables used and the result of the optimization are listed in Table 4.2. To have a sample with better signal to background ratio, an additional cut was made on M_{bc} on both channels:

$$M_{\mathrm{bc}} > 5.27\,\mathrm{GeV\,c^{-2}} \tag{4.6}$$

The effect of the selection and the optimal cut on the D meson mass can be seen in Figs. 4.6b and 4.6b.

decay channel	$B^- \to D^0\pi^-\pi^0$	$B^0 \to D^-\pi^+\pi^0$
$\Delta M(D)$ [GeV]	± 0.014	± 0.011
continuum	> -0.82	> -0.80
$M(\pi^0)$ (GeV)	$0.12298 < M(\pi^0) < 0.14698$	$0.12298 < M(\pi^0) < 0.14798$
$PID_{\pi:k}$	> 0.3	> 0.3
$\lvert\Delta E\rvert$	$< 0.05\,\mathrm{GeV}$	$< 0.05\,\mathrm{GeV}$

Table 4.2.: Optimized cuts that get applied during the cut-based selection.



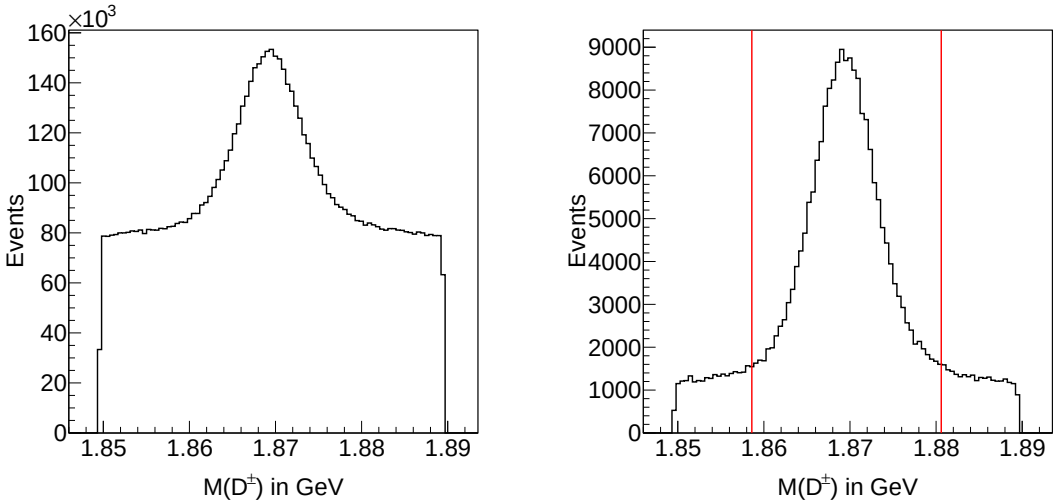
(a) D^0 Meson mass at the reconstruction level.

(b) D^0 Meson mass after selection, cuts on D^0 mass indicated in red.

Figure 4.6.: Distribution of the D^0 mass on reconstruction level and after selection (No best candidate selection applied).

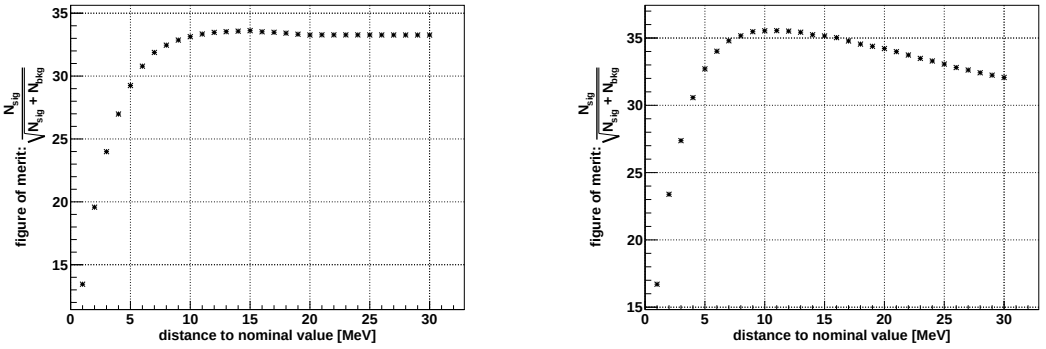
4.2.2.1. Helicity angle

A useful quantity in many analyses is the helicity angle θ. In the reaction $Y \to X \to A + B$, the helicity angle of particle X is the angle measured in the rest frame of the decaying particle, X, between the direction of the daughter A and the direction of the parent particle Y. In two body decays, the helicity angle has a characteristic shape depending on the spin J of a particle. E.g. the decays $B \to D\pi\pi^0$ and $B \to D^{(*)}\rho$, where the ρ decays into $\pi\pi^0$, can be distinguished by only looking at the helicity angle. In the three body decay, the pions can have arbitrary directions, resulting in a flat distribution of $\cos(\theta)$. However, in the decay of



(a) $D^{\pm}$ Meson mass at the reconstruction level. (b) $D^{\pm}$ Meson mass after selection, cuts on $D^{\pm}$ mass indicated in red.

Figure 4.7.: Distribution of the $D^{\pm}$ mass on reconstruction level and after selection (No best candidate selection applied).



(a) $B^- \to D^0\pi^-\pi^0$ (b) $B^0 \to D^-\pi^+\pi^0$

Figure 4.8.: Optimization of the cut to the D meson mass. The results are listed in Table 4.2.

the ρ, the forward/backward direction is favored. This behavior is described by the so-called *d* functions which are the elements of the Wigner D-Matrix and take the angular momentum into account. Because both B and D meson have zero angular momentum, the *d* function $d^1_{0,0}$ describing the helicity amplitude of the decay of ρ is proportional to $\cos(\theta)$.

4.2.3. Best candidate selection

Until now there is no limitation to how many B candidates are allowed to be in one event. It is obvious, that only one candidate can be a real B, neglecting the very small chance that both B mesons decay in the same (charge conjugate) manner. So one best candidate per channel is selected. As stated before, M_{bc} and ΔE have excellent discrimination power for B mesons in B-factories. It would be easy to determine the best candidate by taking e.g. the candidate closest to the nominal ΔE , but since this observable will be later used to extract signal and background yields, this could introduce a bias by changing the shape of the according distribution. It was decided to use the χ^2 of the π^0 mass constraint fit in addition with a χ^2 formed out of M_{bc}. The properties of the best candidate selection were evaluated

using a signal Monte Carlo sample produced with the EvtGenericDalitz model using current world averages for the different contributions. Only truth matched signal candidates were used to evaluate the performance. The multiplicity in the decay mode $B^- \to D^0\pi^-\pi^0$ ($B^0 \to D^-\pi^+\pi^0$) is 2.29 (2.23). In events with more than one candidate, the best candidate selection selects the correct candidate in 75% (77%) of the cases. The probability to randomly select the correct candidate is 38% (39%). This is obtained by weighting two-candidate events with $1/2$, three-candidate events with $1/3$ and so forth. In summary, the use of $M_{\rm bc}$ yields a powerful best candidate selection, that helps to select well reconstructed over misreconstructed signal and thus provides a sample with less migration.

5. Data samples

The Belle experiment ran at the KEKB e^+e^- accelerator during the first decade of the millennium, recording over $1000\,\mathrm{fb}^{-1}$ of data, breaking all integrated and instantaneous luminosity records for high energy accelerators. While most of the data ($711\,\mathrm{fb}^{-1}$) were taken at the $\Upsilon(4S)$ resonance, substantial samples were recorded approximately $60\,\mathrm{MeV}$ below the $\Upsilon(4S)$ resonance, called *off-resonance* data ($89\,\mathrm{fb}^{-1}$) and at the $\Upsilon(5S)$ resonance ($121\,\mathrm{fb}^{-1}$). This analysis will make use of the $\Upsilon(4S)$ sample containing $772 \times 10^6 B\bar{B}$ pairs and the off-resonance sample to extract the shape of continuum events. In general, the data consists of decays of $B\bar{B}$ pairs, continuum events, where e^+e^- annihilates either to light quark (uds) or $c\bar{c}$ pairs and beam background, caused by intrabeam scattering or interactions with the beam gas.

5.1. Simulated data

A technique commonly used in the Belle collaboration are blind analyses. In a blind analysis information (here the data) is concealed from the analyst to avoid preexisting expectations to bias the analysis. Therefore the reconstruction, selection and fitting procedure are developed, optimized and tested on simulated data, called *Monte Carlo* and often abbreviated as *MC*. The name *Monte Carlo* refers to the administrative area of the Principality of Monaco called Monte Carlo, which is famous for the Monte Carlo casino, and was chosen because of the heavy use of random numbers in the simulation. The simulation process consists of two steps: the event generation and the simulation of the detector response. In the event generation the software EvtGen [24] is utilized to simulate the decay of particles, including their momenta and decay vertices, and subsequently simulating the full decay chains for short-lived particles. The interaction of those particles and the response of the detector are then simulated using [25]. An additional component is the beam background, which cannot be easily simulated and thus is recorded with the detector at random points in time between e^+e^- collisions and added to the simulation. There are different kinds of simulated data at hand that satisfy criteria for different tasks occuring in an analysis. Several *streams* of simulated data have been generated, imitating the composition of the recorded data and corresponding to the total integrated luminosity. The individual components are described below.

5.1.1. Generic Monte Carlo

The term generic Monte Carlo describes simulated data samples that try to mimic data, by simulating all known $b \to c$ transitions according to the current knowledge about their

branching fractions. There are ten generic Monte Carlo datasets available, each corresponds to the total integrated luminosity of Belle. They can be divided into a charged (B^+B^-) and mixed ($B^0\bar{B}^0$) sub-sample. This data set provides the possibility to check for B decays that, e.g., mimic the desired decay mode and produce peaking background or to estimate the expected background from similar B decays.

5.1.2. Continuum Monte Carlo

One background is the contribution of $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$), where the quarks hadronize. It is simulated by PYTHIA and JETSET [26]. The level of the continuum contribution depends on the beam energy, while the shape does not vary with it. This allows for determining the contribution of the continuum events to the decay on off-resonance data, or on off-resonance Monte Carlo.

5.1.3. Rare Monte Carlo

In the rare MC samples processes with small branching fraction are enhanced by a factor of 50 to have sufficient statistics to study the properties of these decays. This sample is only used to check if those processes can influence the shape of the generic background in this analysis.

5.1.4. Signal Monte Carlo

The generic MC may be not sufficient in terms of available statistics for all tasks, e.g., for extracting the detector efficiency, a large sample containing only the signal decay mode is necessary. In addition, the Belle MC has several shortcomings: It lacks the simulation of interference and due to the use of the accept/reject method, the tails of resonances can be cut off. Since it has been produced at the beginning of the Belle experiment, branching fractions can also be outdated in the generic MC. Signal Monte Carlo samples can be produced according to the expected branching fractions and with sufficient statistics. Furthermore it allows for using different decay models that allow to simulate interference effects.

To extract the detector efficiency and to model the migration of candidates on the Dalitz plane (see Section 7.2.2), exceptionally large signal MC samples were needed for this analysis. For these tasks, so-called phasespace signal MC is used, where the events are generated with a uniform distribution on the Dalitz plot (referred to as "flat"). In total 610 million events have been produced, 270 million for $B^- \to D^0\pi^-\pi^0$ and due to a smaller overall efficiency 340 million for $B^0 \to D^-\pi^+\pi^0$. Out of this sample, only 140 million events have been used for the efficiency, while all events are used building the migration matrix. For testing the fit procedure on samples containing interference effects, a specialized model for Dalitz (three-body) decays is used.

5.1.5. Dalitz model: EvtGenericDalitz

In contrast to the model used to generate the flat phasespace signal MC, this model takes interference effects into account. The different amplitudes are coherently added up according to Eq. (3.6), and events are generated using an accept/reject method. This model has been used to generate signal MC samples containing the contribution of only one intermediate state in turn as well as samples containing all expected components. For each sample $500,000$ events have been generated, yielding a number of reconstructed and selected events similar to what is expected in recorded data (see Section 1.4). Unfortunately this model has some severe shortcomings, the most prominent being the lack of proper normalization of the individual components. The details of these limitations are discussed in the scope of the description of the fit to the generated samples. This discussion can be found in Section 9.3.1.

6. Description of the background components

The background in the signal box can be divided into two categories, $q\bar{q}$ *continuum* and *combinatoric* $B\bar{B}$ background. The source of continuum background events is described in Section 4.2.1. Combinatoric background is caused by decays of B mesons to different final states, where the combination of three random particles fulfill the requirements of the reconstruction process and thus fake the signal decay. These random particles can originate from different sources, e.g., a B meson decay to a four-body final state, where one daughter particle is missed during reconstruction, a Kaon/Pion misidentification, or using a final state particle produced in the decay of the other B meson.
Another important effect of the reconstruction process is the misreconstruction of true signal, sometimes called *self cross feed* in Dalitz analyses. This can be caused by exchanging one low-energetic particle with another one, especially in the decay of neutral Pions to two photons. These candidates of misreconstructed signal still satisfy all selection criteria, but cause the migration of the event from its true coordinate in the Dalitz plot to its reconstructed position. This is described in Section 7.2.
All backgrounds and the effects of misreconstructed signal decays are studied using the generic Monte Carlo samples, that are produced using a full detector simulation of Belle. The distributions of the background components are used in the fit as two dimensional histograms, that are smoothed using kernel density estimation.
All plots shown in this chapter are for the decay mode $B^- \to D^0\pi^-\pi^0$. The corresponding plots for $B^0 \to D^-\pi^+\pi^0$ can be found in Appendix A. The expected number of combinatoric $B\bar{B}$ background events is about 22000 (24000) and for $q\bar{q}$ continuum events about 13000 (19500) for $B^- \to D^0\pi^-\pi^0$ ($B^0 \to D^-\pi^+\pi^0$). A comparison of the background shape (combinatoric plus continuum) between stream 0 of the MC and the recorded data sample in the M_{bc} sideband between $M_{\mathrm{bc}} > 5.0$ and $M_{\mathrm{bc}} < 5.25$ can be seen in Figs. 6.1 and 6.2.

6.1. Continuum background

Because of the jet-like event shapes of continuum background, the final state particles are often (anti-)collinear and therefor populate the edges and corners of the Dalitz plot. This can be seen in Fig. 6.3a. Here the squared Dalitz plot representation can increase the resolution in these regions (see Fig. 6.3b). This background contribution is described using a two dimensional histogram, which is smoothed using a kernel density estimation, and thus also benefits from the fact, that in the squared Dalitz plot the binning can be chosen in such a way, that

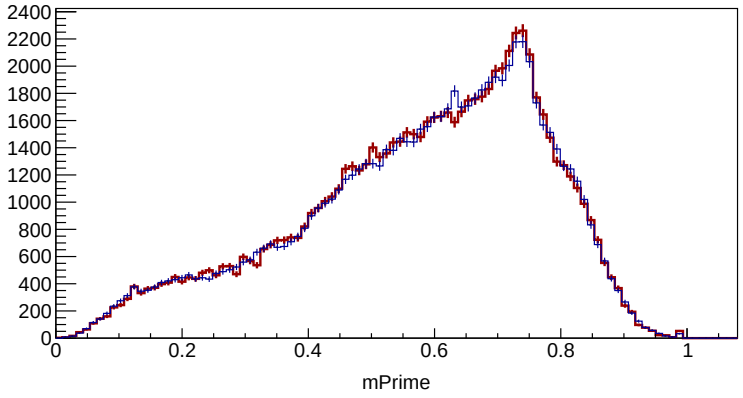


Figure 6.1.: Comparison of the combinatorical background in mPrime in the M_{bc} sideband on recorded data and stream 0 of MC. They show a good agreement between simulation and recorded data.

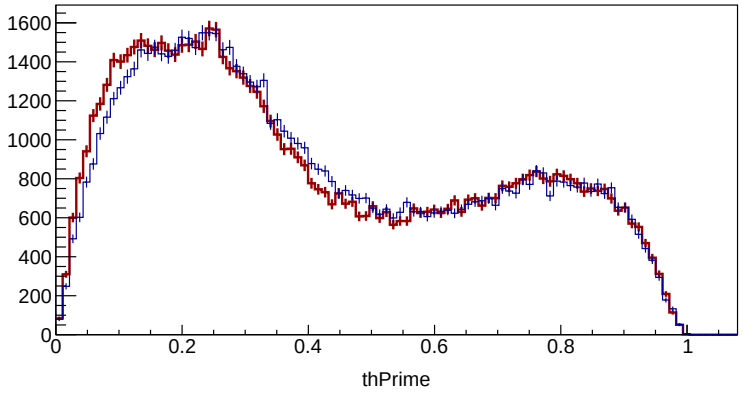
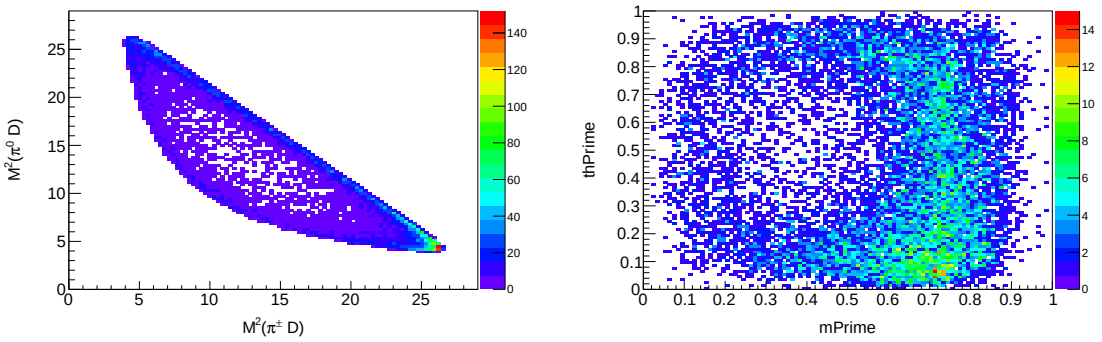


Figure 6.2.: Comparison of the combinatorical background in thPrime in the M_{bc} sideband on recorded data and stream 0 of MC. A small discrepancy can be observed at low values of thPrime, what might be due to the underestimation of background from slow neutral pions.

no bins cross the kinematic limits.

The distribution of the continuum background is extracted from off-resonance data. Off-resonance data is taken at an beam energy $60\,\mathrm{MeV}$ below the $\Upsilon(4S)$ threshold, so that no $B\bar{B}$ pairs can be produced. Since the off-resonance sample is small, it is statistically not sufficient to use only off-resonance data in most analysis. However, after the selection including the continuum suppression, in this analysis a sample with 18045 (35570) events for $B^- \to D^0\pi^-\pi^0$ ($B^0 \to D^-\pi^+\pi^0$) remains. This is enough to take the continuum shape only from off-resonance data instead of using combinatoric background subtracted samples from the M_{bc} side band, which is desirable to keep the systematic errors low.

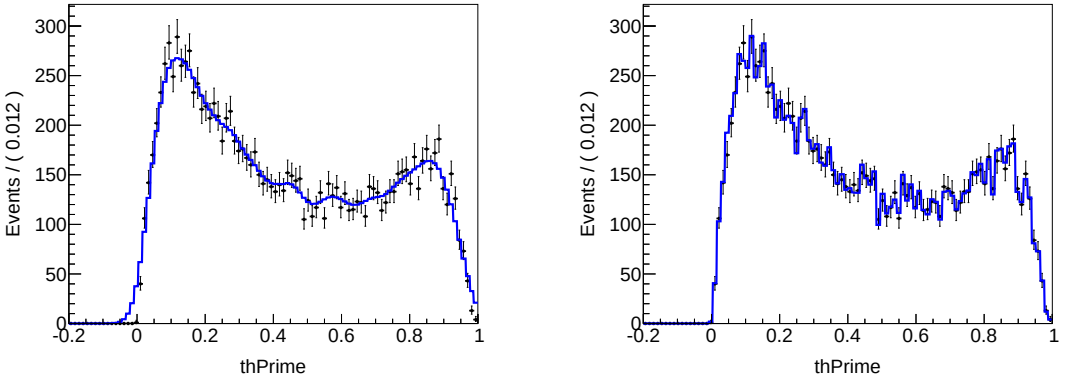
As described above, for the smoothing, the kernel density estimation was performed using an adaptive approach. However, the continuum distribution has a steep slope towards low and high values of θ', and regions without entries adjoining. This causes the KDE to overestimate the entries close to the edges of the θ' range, despite the adaptive approach. Choosing a lower overall bandwidth would solve this problem, but would also make the smoothing prone to not



(a) Normal Dalitz plot (b) Squared Dalitz plot

Figure 6.3.: Comparison of the continuum background from M_{bc} side band (selected with MC truth variables) of stream 0 of generic Monte Carlo in the normal and squared Dalitz plot representation. It clearly shows the advantage of the squared Dalitz plot, the region of the corners in the normal Dalitz plot are distributed over a much larger area.

smooth statistical fluctuation in other parts of the squared Dalitz plot. Finally a combination of both was made. The merging was done using two KDE results, one with the bandwidths 1.2 in the direction of m' and a bandwidth of 0.85 in the direction of θ' (K_{high}), and the second (K_{low}) with the same bandwidth in m', but a highly reduced bandwidth of 0.1 for θ' (see Fig. 6.4).



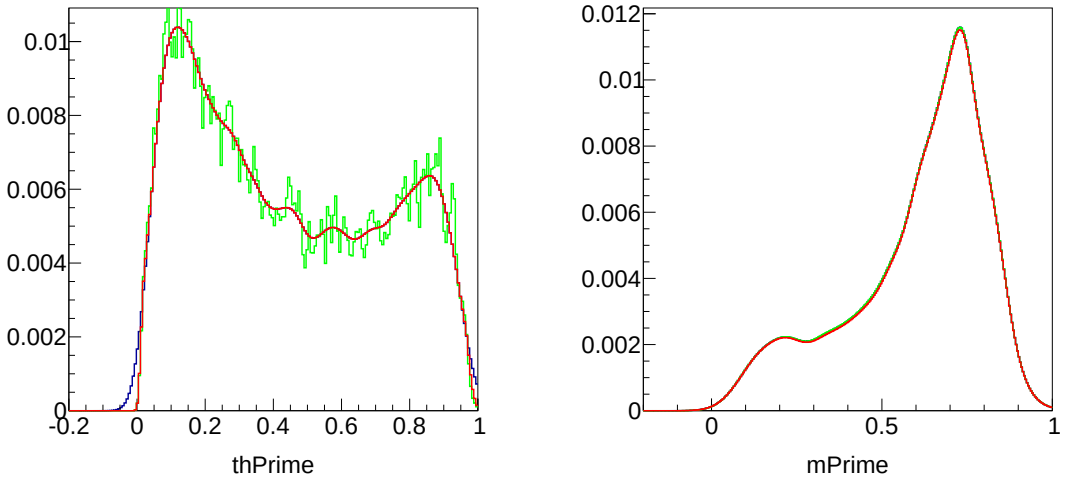
(a) High-bandwidth KDE (b) Low-bandwidth KDE

Figure 6.4.: Comparison of the high- and low-bandwidth KDE in the projection to θ' in the off-resonance Monte Carlo sample. The low-bandwidth KDE describes the steep slope at the edges, but does not smooth the distribution in between.

The KDE probability density functions were both saved as two dimensional histograms with a binning of 300×300. Out of these two histograms, a new one was formed, where for small or high values of θ' the low-bandwidth KDE, and for everything else the high-bandwidth KDE is used. The limits were set to be $L_{\mathrm{low}} = 0.025$ and $L_{\mathrm{high}} = 0.975$, with an overlapping zone of ± 0.02 and ± 0.025, respectively. In this overlapping area the entries of the histograms were added up with a weight, that is linearly depending on the distance to the limit, for example:

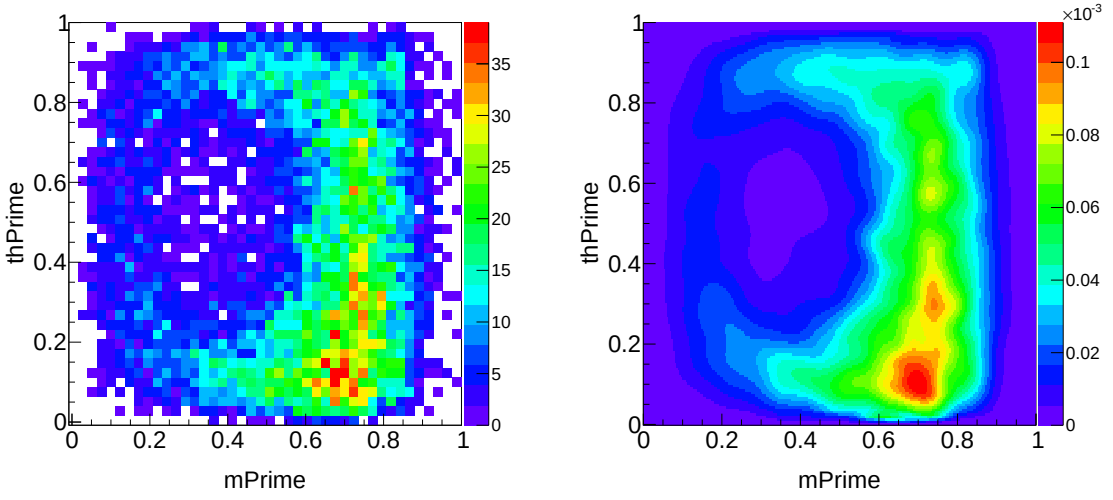
- At L_{low}: $K_{\mathrm{low}} \cdot 0.5 + K_{\mathrm{high}} \cdot 0.5$
- At $L_{\mathrm{low}} - 0.02$: $K_{\mathrm{low}} \cdot 1 + K_{\mathrm{high}} \cdot 0$
- At $L_{\mathrm{low}} + 0.02$: $K_{\mathrm{low}} \cdot 0 + K_{\mathrm{high}} \cdot 1$

As can be seen in Fig. 6.5a, this approach takes K_{low} for low and high values of θ', while for the majority the high-bandwidth KDE K_{high} is used. The projection to m' shows no difference in the distributions of K_{high}, K_{low} and the merged result (Fig. 6.5b), supporting the method. The corresponding two dimensional plots can be seen in Fig. 6.6. They show that the approach describes the given two dimensional distribution well in both the edges and central regions of the squared Dalitz plot. The results using simulated off-resonance data is used in the fits to the simulated data, while for the fit to recorded data the histogram is produced using off-resonance data. The results for off-resonance data can be seen in Fig. 6.7.



(a) (b)

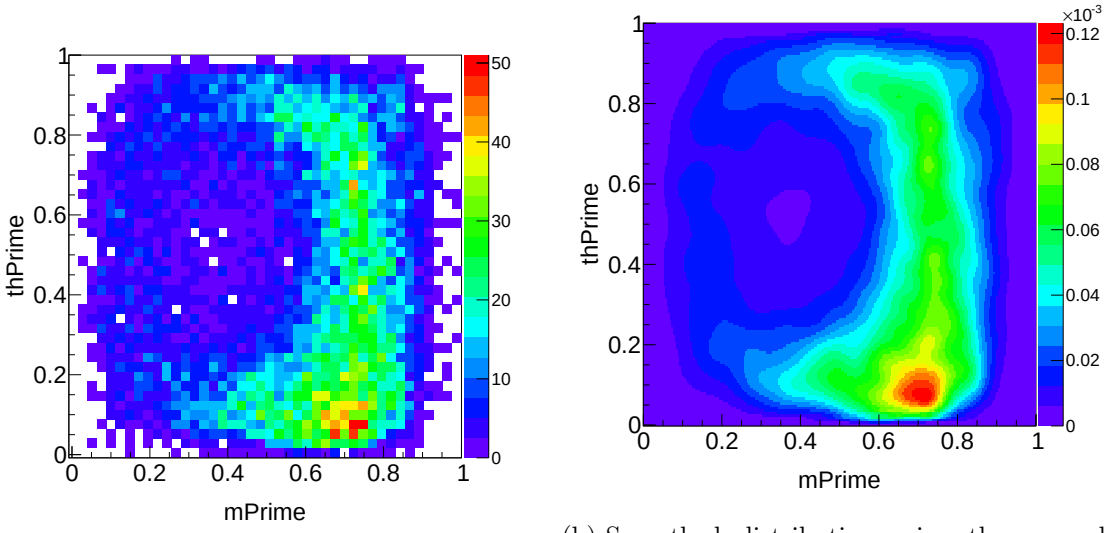
Figure 6.5.: Result of the merging (red) of the low-bandwidth KDE (green) and high-bandwidth KDE (blue) (Continuum background taken from off-resonance Monte Carlo). No difference is observed in mPrime while the overshooting of the high-bandwidth smoothing is corrected at the edges.



(a) Continuum in M_{bc} side band, raw distribution

(b) Smoothed distribution using the merged KDE

Figure 6.6.: Comparison of simulated off-resonance continuum background and its smoothed distribution.



(a) Off-resonance data, raw distribution

(b) Smoothed distribution using the merged KDE

Figure 6.7.: Comparison of off-resonance data and its smoothed distribution. The distribution looks similar to the simulated data, see Fig. 6.6.

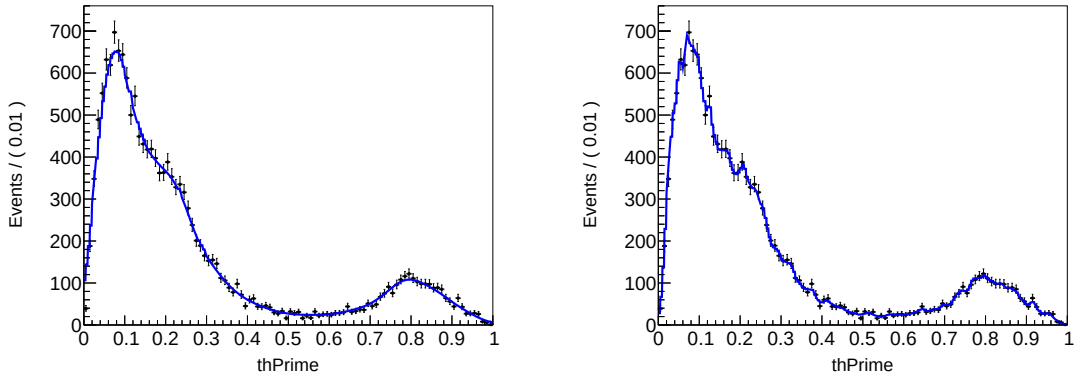
6.2. Combinatoric background

In contrast to the continuum background, the combinatoric background is caused by making wrong combinations of particles originating from B mesons in the reconstruction process. Therefore it is not possible to get the shape of the distribution from off-resonance data. For combinatoric background one has to rely on the generic Monte Carlo, where it is tried to simulate all possible B meson decays and the full detector. The combinatoric background from generic Monte Carlo consists of decays of neutral or charged B mesons, referred to as *mixed* and *charged*, respectively. The relative contribution of mixed and charged can be different on recorded data, therefore two background models are used in the Dalitz fit to describe the combinatoric background. Both background distributions are obtained by using generator information stored along with all the other event information. The generator information, often called *MC Truth*, allows to deselect all signal in the signal box, leaving only the desired distributions.
Again the squared Dalitz plot is utilized, for the same reasons as stated in the description of the continuum background, namely an increased resolution of the corners of the normal Dalitz plot and for the elimination of binning effects at the boundaries.

The two dimensional histograms for mixed and charged background are again smoothed using KDE with different bandwidths and then merged afterwards. Compared with the continuum background, in this case it is sufficient to use an increased bandwidth in the direction of θ' of 0.2. The merging is done in the same way as described in Section 6.1. However, the size of the available samples is larger. As the time consumption of a kernel density estimation rises with the number of entries, it is not possible to smooth a combined histogram of multiple streams of generic Monte Carlo. Instead, such a smoothing is done separately for four streams, along with the merging of the different KDE PDFs. Afterwards, the results for all four streams are averaged to get the two-dimensional histogram as input to the Dalitz fit.

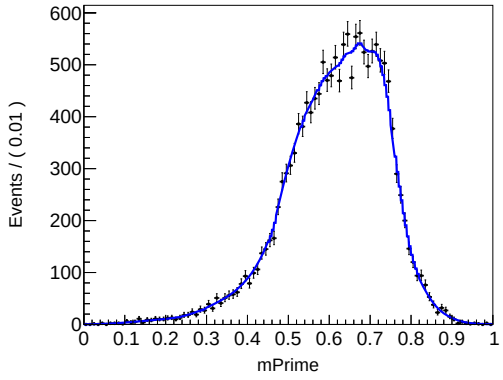
In Fig. 6.8a it is apparent that the high-bandwidth KDE is not able to describe the data points at the very edge. Here again the low-bandwidth KDE helps to correctly describe the behavior at the low edge. The different bandwidth for the θ' direction does not affect the other dimension, as can be seen in Figs. 6.8c and 6.8d. Here only the results of the charged combinatoric

background for $B^- \to D^0\pi^-\pi^0$ are shown, the plots for the mixed background along with all plots for the second decay channel, $B^0 \to D^-\pi^+\pi^0$, can be found in Appendix A.2.

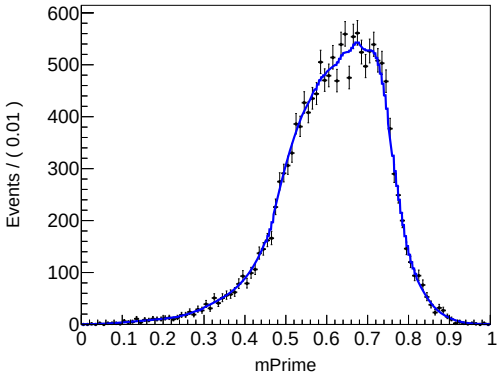


(a) High-bandwidth KDE - θ'

(b) Low-bandwidth KDE - θ'



(c) High-bandwidth KDE - m'

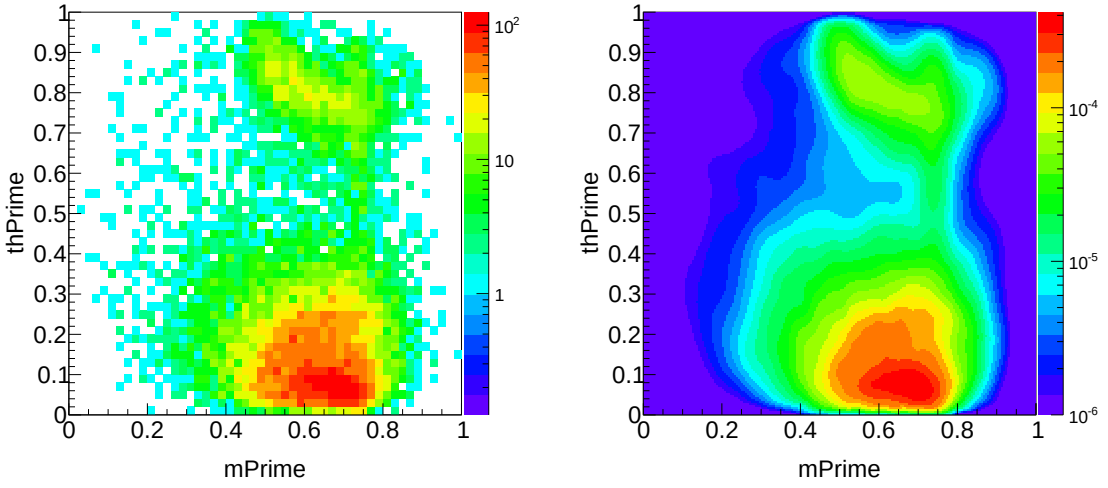


(d) Low-bandwidth KDE - m'

Figure 6.8.: Comparison of the High- and Low-bandwidth KDE in the projection to θ' of combinatoric background in stream 0 of generic Monte Carlo. The lower plots show that in the projection to m' the θ' bandwidth change has no effect.

6.3. Rare $B\bar{B}$ background

To rule out the possibility for decay channels ending up in the signal region, that are not covered by the generic Monte Carlo, the Rare Monte Carlo sample was processed in the same way as generic Monte Carlo. This sample consists of rare mixed and charged decays, where the branching fraction was enhanced by a factor of 50 and thus corresponds to 50 times what is expected in data. The amount of each different decay mode ending up in $D\pi\pi^0$ was compared with the amount in the first stream of generic Monte Carlo. If a decay mode in the rare Monte Carlo made no contribution to generic Monte Carlo or has a larger contribution in rare than in the generic Monte Carlo, it was added up. The total contribution of rare decays to the combinatoric background is 0.018% (0.015%) in the case of $B^- \to D^0\pi^-\pi^0$ ($B^0 \to D^-\pi^+\pi^0$). This contribution is dominated by wrong reconstructed D mesons, were the Pion used in the D meson originates from a $a_1(1260)$ or a $h_1(1170)$. The 14 (10) events are distributed over a large area in the Dalitz plot and therefore can be neglected.



(a) Combinatoric background in stream 0 of generic Monte Carlo, raw distribution

(b) Smoothed distribution using the merged KDE

Figure 6.9.: Comparison of combinatoric background in stream 0 of generic Monte Carlo (selected with MC truth variables) and its smoothed distribution (Log scale).

7. Detector response

To extract the correct results from the measurement it is important that the effects of the detector are taken into account. For instance the detector behaves different for the same particle type at different momenta. Especially low energetic (e.g. slow pions) particles can be hard to detect. While the major effect is to completely miss such a particle in the reconstruction, reducing the number of detected events, there is also the possibility to misreconstruct it. As the kinematic properties of the daughter particles differ for different positions on the Dalitz plot, the description of the detector acceptance and misreconstruction behavior depend on the Dalitz variables.

7.1. Efficiency

The detector acceptance and efficiency are determined using a large sample of signal Monte Carlo, generated in a way, that it is uniform in the Dalitz plane. This was done using *EvtGen* [24] to generate a sample with 140 million events for each decay mode. After these events were processed using *gsim*, a software package utilizing *Geant3* [27] to simulate the detector effects, they were subjected to the same reconstruction and selection criteria as the data.
The overall reconstruction efficiency is defined as

$$\epsilon = \frac{N_{\text{reco}}}{N_{\text{gen}}}, \tag{7.1}$$

where N_{reco} is the number of reconstructed events, while N_{gen} represents the number of generated events. There are two important issues that are crucial in properly describing the effect of the detector. First, all events that pass the reconstruction and selection are used to calculate the efficiency, even if they are misreconstructed. The second issue is to use the true position of the event; during reconstruction, the position of the event in the Dalitz plane can be shifted. The efficiency for a given point should describe the probability that an event generated at this point in the Dalitz plane is reconstructed and selected at all, no matter where it ends up after reconstruction. The countermeasure to the migration of an event is treated separately and is described in Section 7.2.

To account for the varying efficiency over the dalitz plot, this efficiency will be evaluated in bins i of a two dimensional histogram:

$$\epsilon_i(m_{13}^2, m_{23}^2) = \frac{N_{\text{reco}}^i}{N_{\text{gen}}^i}, \tag{7.2}$$

where N^i_{reco} gives the number of reconstructed events that were generated at a point within bin i, independent of their reconstructed position. For this, two approaches were tested and compared. To avoid binning effects and the kinematic limits of the dalitz plot, it was tried to evaluate the efficiency in the squared dalitz plot representation. Since the transformation leaves the allowed kinematic configurations to be within $[0, 1]$ in both dimensions (see Section 3.1.4), it is possible to chose the binning in such a way, that no bin crosses the boundaries. However, this approach introduces the problem, that while the distribution of the generated events is flat in the normal dalitz plot representation, they are not uniformly distributed in the squared Dalitz plot. In addition, after the transformation there are only few events left near the boundaries. Any reasonable binning thus leaves empty efficiency bins at the edge. Using an adaptive binning results in huge bins in this region with large errors on the efficiency.
The second approach is to use the efficiency histogram in the normal dalitz representation, where the generated events are equally distributed. As mentioned before, the kinematic boundaries can be crossed by some bins, leaving them underpopulated compared to bins in the center. This is still correct, since the binning for the 2D histogram of the generated and reconstructed events is the same for both. However, the efficiency in these bins would be prone to high statistical fluctuations, which could even lead to nonphysical values. To counteract this, an adaptive binning technique was used. In the case of an underpopulated bin i, the neighboring eight bins were examined. Since the underpopulated bins occur only at the dalitz boundaries, some of the neighbor bins can be also underpopulated or empty. These bins were neglected and the bin in question was then merged with the remaining neighboring bins j. The efficiency for the merged bin then is given by

$$\epsilon_{\text{merged}} = \frac{N^i_{\text{reco}} + \sum_j N^j_{\text{reco}}}{N^i_{\text{gen}} + \sum_j N^j_{\text{gen}}} \tag{7.3}$$

The initial binning was chosen to be 150×150 bins in the range $0 - 29$ GeV^2, where the dalitz plane covers approximately 17500 bins with an average bin content of 7750 events. A bin is taken as underpopulated, when the number of entries is below 6000, what corresponds to being 20 sigma away from the average ($N - 20 \cdot \sigma \cdot \sqrt{N}$) and only selects underpopulated bins. The efficiency usually drops towards the kinematic limits of a Dalitz plot, but the definition of an underpopulated bin is chosen in such a way, that bins where the whole area is within the kinematic limits are not affected by the adaptive binning method. The result of the correction can be seen in Figs. 7.1 and 7.2, where the left hand plot shows the efficiency without correction. It clearly shows some bins (in dark blue or red) at the boundaries where the statistic fluctuations caused highly over-/underestimated efficiency. In the corrected plot on the right hand side, this fluctuations are corrected.

7.2. Misreconstructed signal

The reconstruction of events can cause a shift of the coordinate in the Dalitz plot. For well reconstructed signal events this corresponds to the resolution and is negligible relative to the widths of the involved resonances. This scenario does not hold anymore for misreconstructed signal events. The treatment of misreconstructed signal events follows the procedure suggested in [11].

7.2.1. Properties of misreconstructed signal

The majority of misreconstructed events are caused by using a particle originating from the other B meson in the event to form a candidate. This candidate still satisfies all selection criteria, but can cause a significant migration over the dalitz plot. As expected, the largest contribution originates from low energetic neutral pions, where one photon is exchanged with

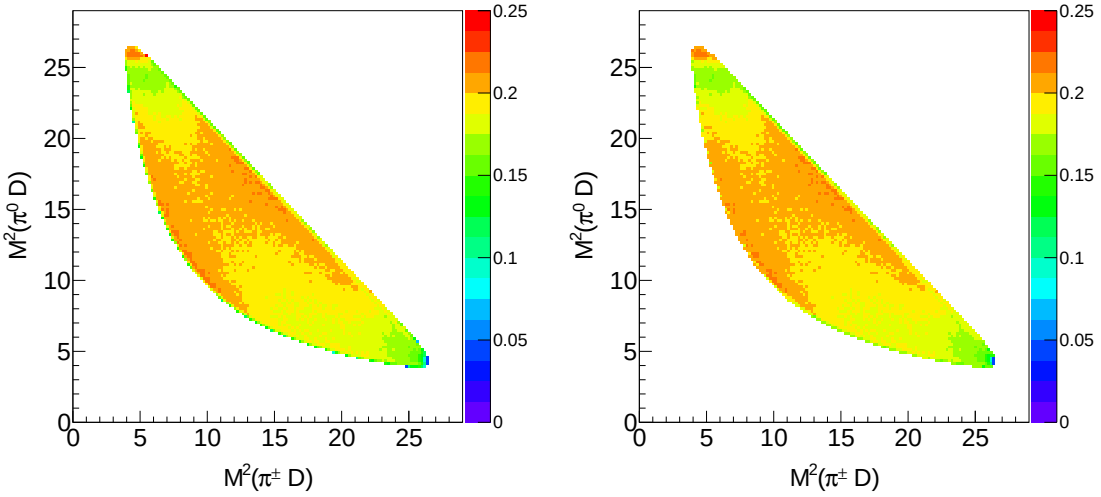


Figure 7.1.: Efficiency over the Dalitz plot. The right hand side shows the plot with corrections for low statistics bins. ($B^- \to D^0\pi^-\pi^0$)

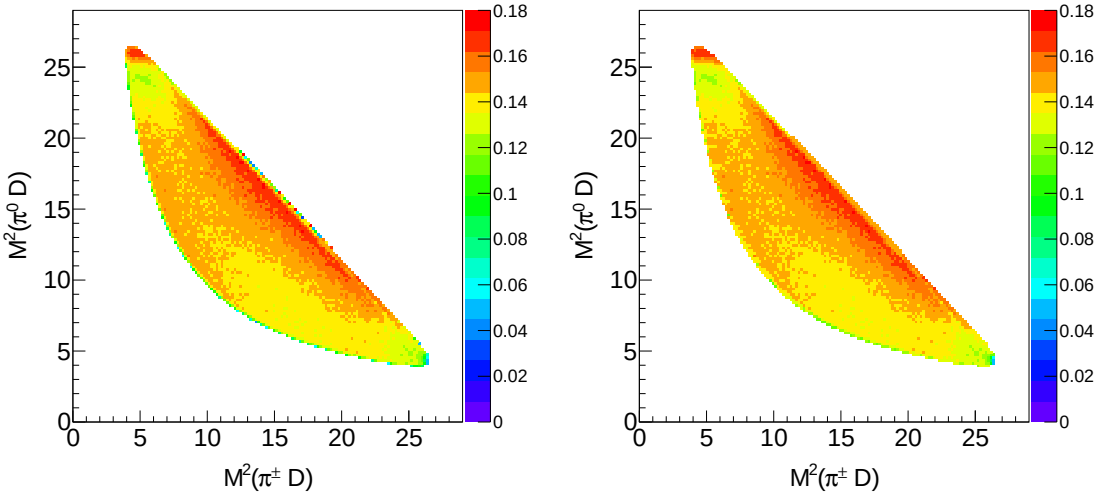


Figure 7.2.: Efficiency over the Dalitz plot. The right hand side shows the plot with corrections for low statistics bins. ($B^0 \to D^-\pi^+\pi^0$)

a random photon in the event. The fraction of misreconstructed signal is highly dependent on the position in the Dalitz plot and peaks at the region where neutral pions are produced almost at rest (so-called slow pions). This effect even counteracts the effect of decreasing efficiency towards the corner of the Dalitz plot by increasing the number of available combinations. This can be seen in the efficiency plots in Figs. 7.1 and 7.2, where the efficiency increases again towards the upper left corner of the Dalitz plot.

7.2.2. Migration

The properties of the migration are studied using a Monte Carlo sample that has both the generated (*true*) Dalitz plot coordinates and the reconstructed (*reco*) coordinates available. The distance of the shift caused by migration is defined as:

$$d_{\text{migration}} = \sqrt{(M_{\text{true}}(D\pi^{\pm}) - M_{\text{reco}}(D\pi^{\pm}))^2 + (M_{\text{true}}(D\pi^0) - M_{\text{reco}}(D\pi^0))^2} \qquad (7.4)$$

The migration can cause events to move several hundred MeV during reconstruction, making them leave resonance bands and consequently regions of interference. This will diminish the ability of the fit to get the phases right and distinguish between different resonances. As a first countermeasure, the feature to constrain the mass in the kinematic fitter is used to refit the B meson candidates after the reconstructed properties other than the invariant masses have been saved. The constraint takes care that the recalculated invariant masses of the combination of daughter particles add up to the nominal B mass and lie within the kinematic limits of the Dalitz plot. This greatly reduces the overall migration of well reconstructed events in the Dalitz plane. However, there is the chance to misreconstruct signal events which produce candidates that migrate significantly. A comparison of $d_{\text{migration}}$ of truth matched and misreconstructed signal can bee seen in Fig. 7.3.

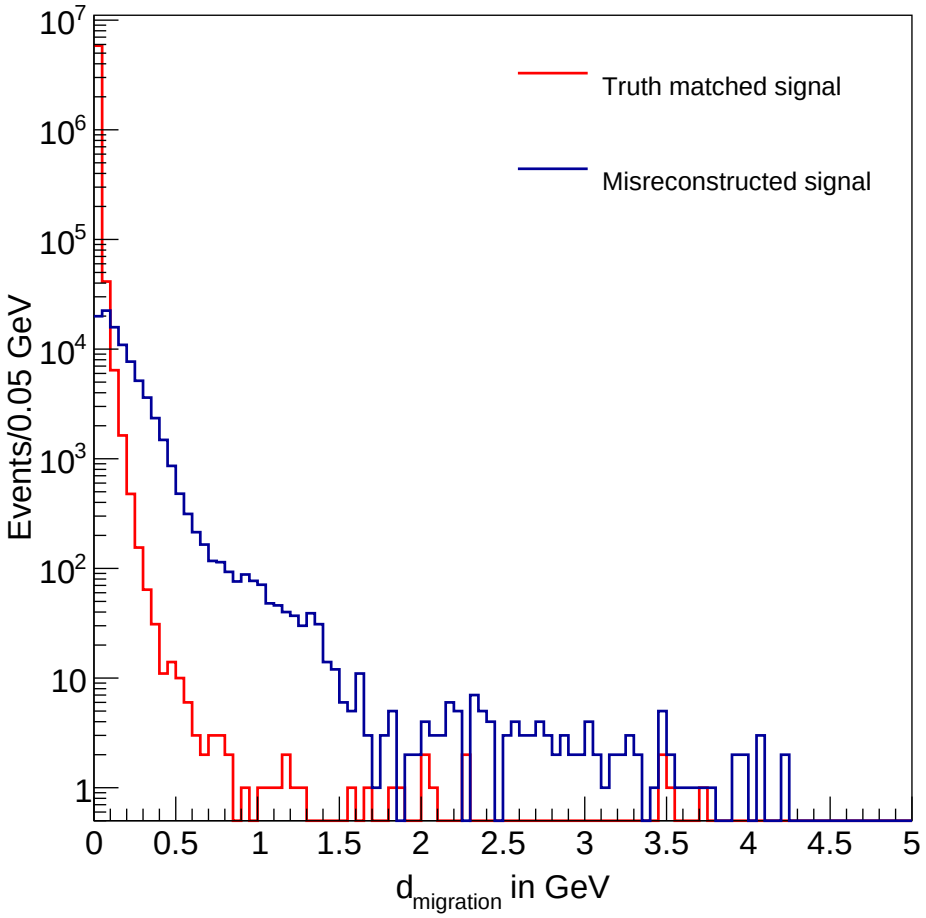


Figure 7.3.: Distribution of migration distance of truth matched and misreconstructed signal candidates, N.B. the log scale on the Y axis.

In order to correctly model this behavior, a convolution with a resolution function would be necessary. This resolution function should not only represent the detector resolution, but also the effects of misreconstructing signal events. Here, an approximation is made, where only events that migrate significantly are treated, while the well reconstructed events are assumed to have been reconstructed at their true position. Therefore the signal is divided in two categories, depending on the fact if a significant migration occurs or not. Instead of using Monte Carlo truth information to discriminate between misreconstructed and well reconstructed signal, the migration distance is utilized for this task. This makes sure that events where the misreconstruction has a small effect are taken as signal, and the properties of all events that

cause a large migration are described. As mentioned before, for signal events the migration should be small compared to the smallest width of an involved resonance. Since the width of a D^* meson is far smaller than the detector resolution, the detector resolution is the smallest width present in the contributions of the signal model. The defining cut for misreconstructed signal is thus chosen to be $d_{\text{migration}} > 5\text{MeV}$, which is approximately the detector resolution in the region of slow neutral pions. This cut defines an overall fraction of 63% in a generated phase space signal Monte Carlo sample to be "misreconstructed" or rather "migrating".

7.2.3. Treatment of misreconstructed signal

The overall detector response can be written as a combination of the efficiency and the migration:

$$\mathcal{R}_{\text{detector}} = \epsilon(m_{13}^2, m_{23}^2) \cdot \mathcal{R}_{\text{migration}}. \tag{7.5}$$

The efficiency $\epsilon(m_{13}^2, m_{23}^2)$ has been described in detail in Section 7.1 while $\mathcal{R}_{\text{migration}}$ will be defined and explained below. As motivated above, the signal events are divided into two categories, depending on whether they significantly migrate or not. Since the true position for an event is not known in data, the probability for a candidate to be misreconstructed or well reconstructed signal depending on the Dalitz plot coordinate is determined using a phase space signal Monte Carlo sample, similar to the one used for the efficiency, but with 270 million events for $B^- \to D^0\pi^-\pi^0$ and 340 million events for $B^0 \to D^-\pi^+\pi^0$. The fraction is defined as

$$f_{MR} = \frac{N_{MR}}{N_{MR} + N_{CR}}, \tag{7.6}$$

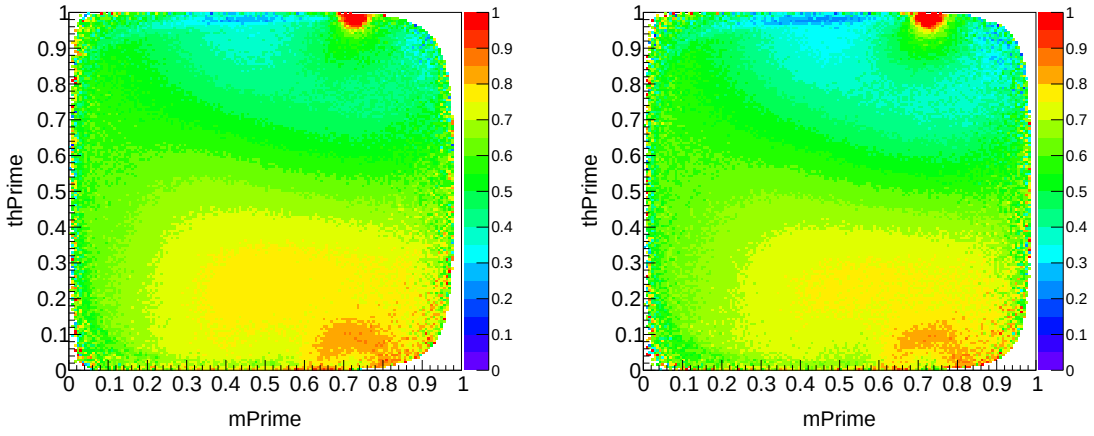
where N_{MR} is the number of misreconstructed signal events and N_{CR} the number of correctly reconstructed signal events. It is evaluated in each bin and supplied to the fit in a two dimensional histogram (see Fig. 11.2). Again it is important to calculate the fraction for a given bin i using the true position of generated B mesons. The fraction then denotes the probability for an event with a true position within the bin boundaries to get misreconstructed. Using the abbreviation $x \equiv m_{13}^2$ and $y \equiv m_{23}^2$, $\mathcal{R}_{\text{migration}}$ can then be written as

$$\begin{aligned}\mathcal{R}_{\text{migration}}(x_{\text{reco}}, y_{\text{reco}}; x_{\text{true}}, y_{\text{true}}) = (1 - f_{MR}(x_{\text{true}}, y_{\text{true}}))\, \delta(x_{\text{reco}} - x_{\text{true}})\delta(y_{\text{reco}} - y_{\text{true}}) \\ + f_{MR}(x_{\text{true}}, y_{\text{true}})\mathcal{R}_{\text{MR}}\left(x_{\text{reco}}, y_{\text{reco}}; x_{\text{true}}, y_{\text{true}}\right),\end{aligned} \tag{7.7}$$

where $\mathcal{R}_{\text{MR}}$ denotes the migration matrix that represents the probability for an event originating at $(x_{\text{true}}, y_{\text{true}})$ to migrate to $(x_{\text{reco}}, y_{\text{reco}})$.

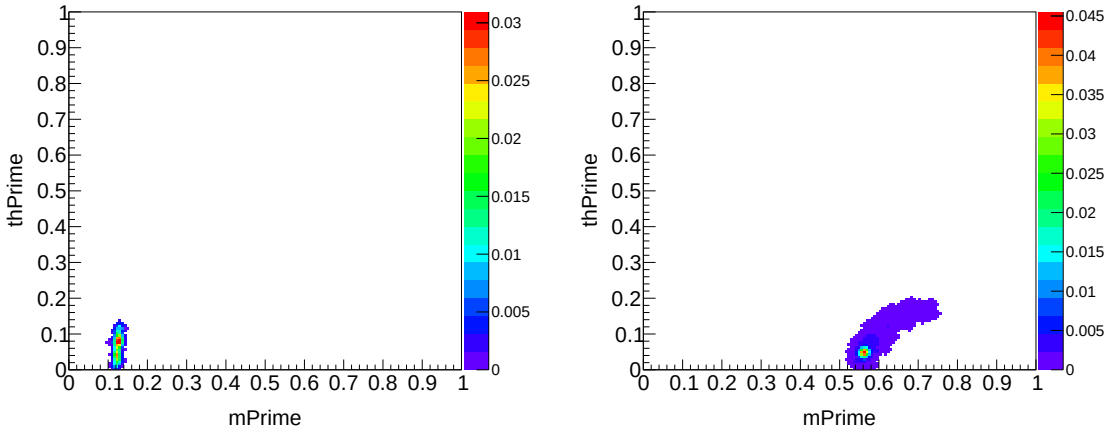
7.2.4. Migration matrix

This migration matrix is built in a binned way, using a 270 (340) million events phase space signal Monte Carlo sample for $B^- \to D^0\pi^-\pi^0$ ($B^0 \to D^-\pi^+\pi^0$) which was also used for the determination of the misreconstructed signal fraction histogram. For every bin in the squared Dalitz plot representation, a migration probability histogram is generated in the following way: The reconstructed position of all events that were generated within the bin boundaries of a given bin are plotted into a histogram, which is then normalized. An example can be seen in Fig. 7.5, where the bin with the highest probability is where all events originate. It is visible that the migration distance depends on the position in the dalitz Plot and that it is large at the region with slow neutral pions (Fig. 7.5b) compared to other regions (Fig. 7.5a). To



(a) $B^- \to D^0\pi^-\pi^0$ (b) $B^0 \to D^-\pi^+\pi^0$

Figure 7.4.: Fraction of migrating events (in the squared Dalitz plot). The limit of $5\,\mathrm{MeV\,c^{-2}}$ selects 30% up to 100% of the signal events as migrating signal. The fraction depends on the kinematic configuration which is similar in both decay channels.



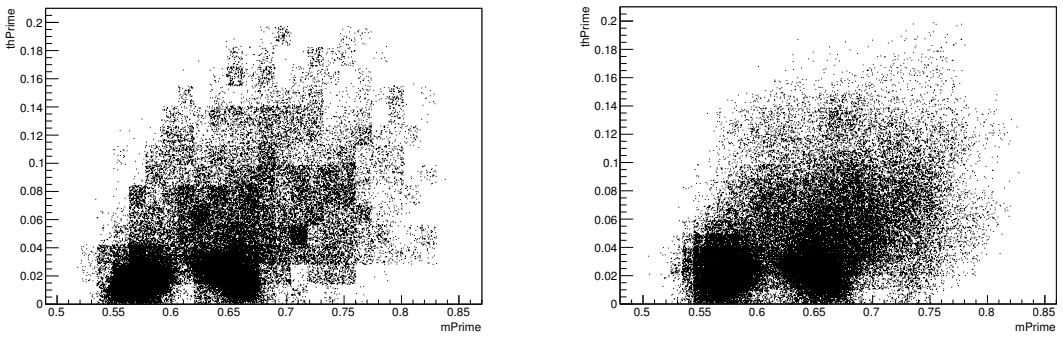
(a) Bin 1235 (b) Bin 1301

Figure 7.5.: Examples of migration probability histograms (in the squared Dalitz plot). The bin 1301 lies in the region of slow neutral pions, resulting in a distribution of the migration probability over a larger area.

reduce statistical fluctuations, a smoothing algorithm was applied to each individual migration histogram. A comparison can be seen in Fig. 7.6, where events of a model containing only the D^* are generated using the unsmoothed (Fig. 7.6a) or the smoothed (Fig. 7.6b) individual histograms.

To properly describe the effect of the misreconstruction, the binning of the two dimensional histograms used for constructing the migration matrix has to be fine enough. Different binnings have been used and tested in ensemble tests with the result that a binning below 100×100 produced biases in the amplitudes and again an exchange of events between the two background components. With increasing number of bins these effects vanished and it was decided to use 150×150 bins.

In addition, it was found that in some of the migration histograms one or several single events



(a) No smoothing applied (71 bins).

(b) Smoothing applied (100 bins).

Figure 7.6.: Comparison of a generated D^* distribution in the squared Dalitz plot with unsmoothed and smoothed migration histograms.

have a large migration distance. This results in a region, where the migration probability distribution is smooth around the true bin plus the bin where the outlier (an entry in a bin disconnected from the distribution of the other entries) was reconstructed. Although the probability for an event to migrate there is very small, it turned out to produce a significant bias in the ensemble studies for the amplitude of the D^*. In addition, it caused the background components to exchange events, leading to a bias in the yield of both. Therefore these outliers are not used in building the migration matrix. At first, a cut was utilized to remove these outliers, but since the migration can be large in only one direction of the Dalitz variables while small in the other, there was a risk of choosing the cut too strict and cut away events that have a large but still reasonable migration distance. With this setting, the fit is not able to properly describe the migration in the region of slow pions. This was solved by introducing a different technique to reject events with extreme migration distance: Events were rejected if they are disconnected from the main distribution of migrating events. This was checked by looking at the neighborhood of bins, that had only one or two entries. For one entry bins, the adjacent two bins in every direction were checked for entries and the event is rejected if there are none or only one (what happened in a few cases only). It also occurred rarely that two events ended up in a bin far from the main distribution. They were rejected if in the surrounding four bins in every direction no other event was found.

8. Extraction of the signal yield

8.1. Fit model of the fit to the ΔE distribution

The signal yields are determined via an unbinned extended maximum likelihood fit to the ΔE (see Eq. (4.2)) distribution. The range of the fit is limited to $-0.1 < \Delta E < 0.1$ GeV due to the possibility of peaking structures at the distance of one Pion mass below or above zero (Fig. 8.1).

Although the ΔE cuts for the signal box for the Dalitz fit are set to $-0.05\,\mathrm{GeV} < \Delta E < 0.05\,\mathrm{GeV}$, the wider range in the ΔE fit is necessary to determine the level of background. To extract the yield for the region used in the Dalitz fit, the resulting function of the signal component is integrated from $-0.05\,\mathrm{GeV}$ to $0.05\,\mathrm{GeV}$, but the error on the yield is conservatively taken to be the error on the yield of the fit to the whole region. The signal yield of the ΔE fit is used in the Dalitz fit to constrain the total signal using a Gaussian constraint with the signal yield as central value and the error of the yield as width.

The typical shape of the ΔE distribution after reconstruction and the selection is relatively broad and asymmetric. This is due to the fact that neutral Pions leave no track in the central drift chamber resulting in a worse momentum resolution. The probability density function used for the signal shape is a sum of one Gaussian and a Student's t distribution. The signal shape is determined on simulated signal events and the shape parameters are fixed except for the width. Usually the width of simulated distributions can differ from the real resolution in data, therefore the width of the signal shape is left free in the fit to data. Here the Gaussian and the Student's t distribution share the same width parameter, but an additional parameter is introduced to constrain the width of the Student-t distribution to the Gaussian width. This constraint is determined on the fit to the signal MC sample and fixed in the fit to the combined sample or reconstructed data. The overall width is left free in those fits. Note that the signal shape is determined on a signal Monte Carlo sample also containing misreconstructed signal. The distribution of the misreconstructed signal events is relatively broad (see Fig. 8.2), giving the possibility to exchange events with the background. As a countermeasure the shape of the combined continuum and combinatoric background component is also determined from simulated data. The combined background is parametrized by a Chebychev polynomial of second order. The contribution of the continuum background is linear in ΔE (see Fig. 8.3) and allows to combine both backgrounds in one component. The combinatoric background has a curvature, but no peaking structure, which can be seen in Fig. 8.4.

All fits in this analysis are performed using RooFit [17].

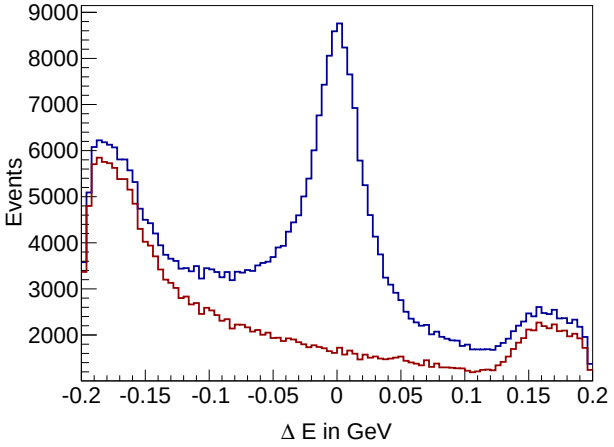


Figure 8.1.: ΔE distribution in a wider range. Blue is signal plus background, red is background only from stream 0 of MC ($B^- \to D^0\pi^-\pi^0$). The peaks at the edges are mainly caused by either missing a pion of a B meson decay to $D\pi\pi\pi$ or adding a random pion to a B meson decay to $D\pi$ in the reconstruction.

Channel	Yield	MC Truth	Deviation in σ
$B^- \to D^0\pi^-\pi^0$	90613 ± 886	91373	-0.86
$B^0 \to D^-\pi^+\pi^0$	71084 ± 941	71629	-0.58

Table 8.1.: Results of ΔE fits on a combined sample of signal Monte Carlo and backgrounds from stream 3 of MC.

8.2. Results on simulated events

The number of fitted signal events for both decay channels is shown in Table 8.1, where the yield is compared to the true number from Monte Carlo information. The fits are in good agreement with the values from the signal Monte Carlo sample. In Fig. 8.5 the fitting procedure is illustrated: First the signal shape is determined by a fit to the signal component from signal Monte Carlo (top left). The signal Monte Carlo is generated according to the current world average of the branching fractions. The shape of the background is obtained in the same way. Then a fit to a combined sample of signal Monte Carlo and background from generic Monte Carlo is performed. In this fit the yields and the width are left floating. The background was taken from stream three of generic Monte Carlo, while deselecting signal decays with Monte Carlo truth flags.

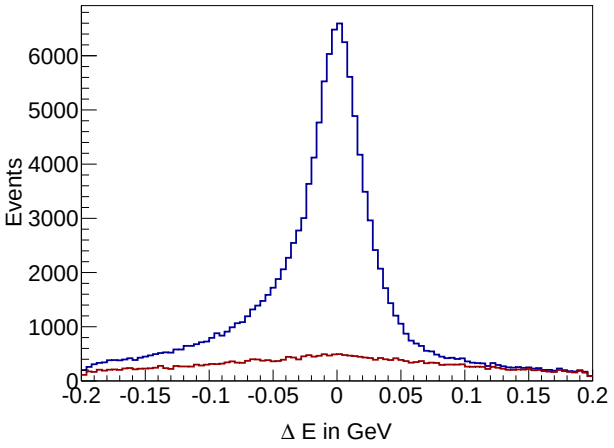
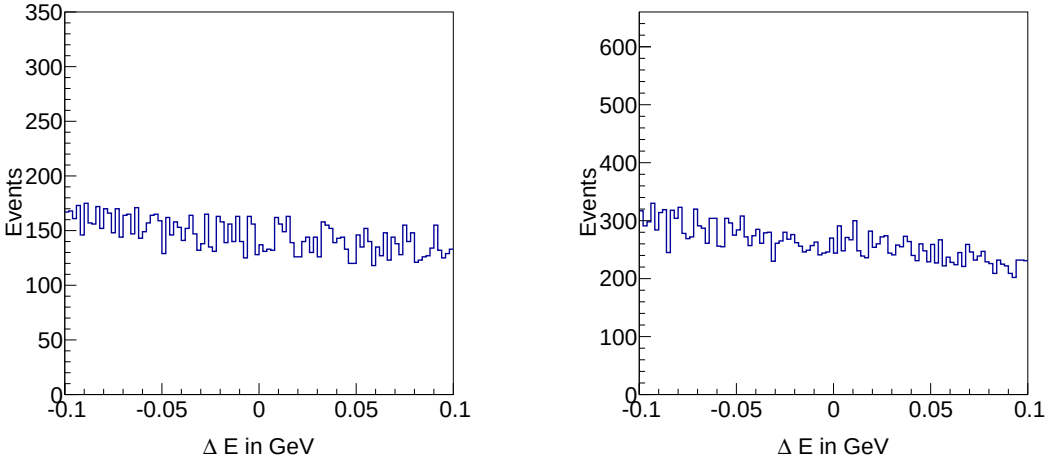
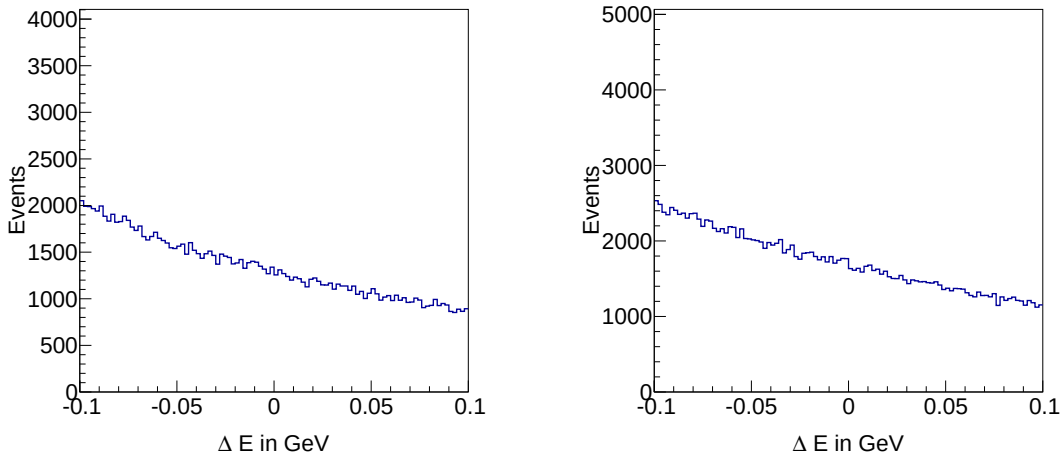


Figure 8.2.: All signal and misreconstructed signal only (red) ($B^- \to D^0\pi^-\pi^0$). The misreconstructed signal is very broad making it difficult to separate background and misreconstructed signal.



(a) $B^- \to D^0\pi^-\pi^0$ (b) $B^0 \to D^-\pi^+\pi^0$

Figure 8.3.: Continuum background in the ΔE distribution, taken from off-resonance data. It is linear and can therefore be combined with the combinatoric background.



(a) $B^- \to D^0\pi^-\pi^0$ (b) $B^0 \to D^-\pi^+\pi^0$

Figure 8.4.: Combinatoric background in the ΔE distribution taken from stream 0 of generic MC. It shows a slight curvature, the rise towards the edges is caused by the missing/additional pions (see Fig. 8.1).

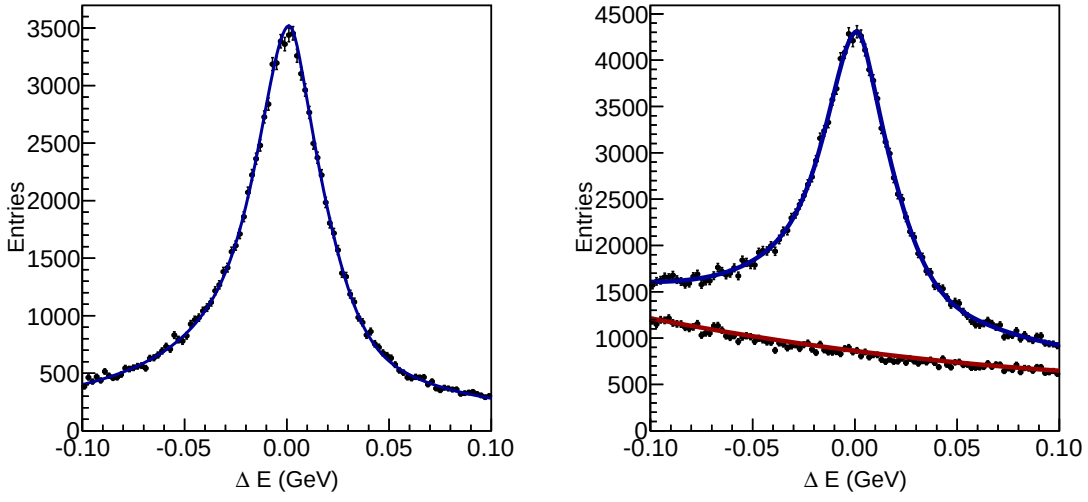


Figure 8.5.: Fit to signal Monte Carlo (left) and fit to a combined sample of stream 0 Monte Carlo background and signal Monte Carlo (right) in $B^- \to D^0\pi^-\pi^0$.

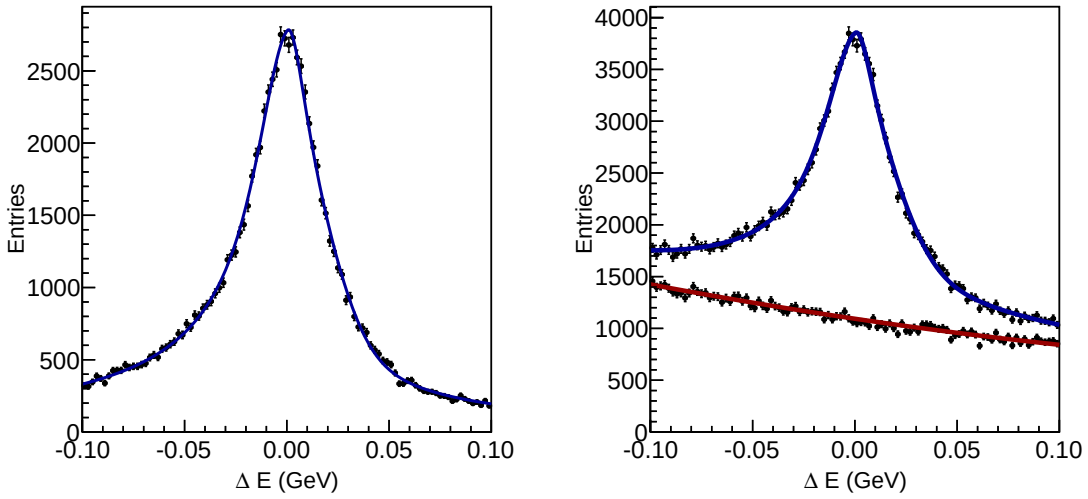


Figure 8.6.: Fit to signal Monte Carlo (left) and fit to a combined sample of stream 0 Monte Carlo background and signal Monte Carlo (right) in $B^0 \to D^- \pi^+ \pi^0$.

8.3. Results on recorded Data

The number of fitted signal events on recorded data for both decay channels is shown in Table 8.2. The fits are in good agreement with the data points, the error on the signal yield is slightly higher than on the fit to simulated events. The width was left free in the fit and increased compared to the results on Monte Carlo by 20% (25%) for $B^- \to D^0\pi^-\pi^0$ ($B^0 \to D^-\pi^+\pi^0$). The expected number of events for $B^- \to D^0\pi^-\pi^0$ is calculated using the result of a CLEO measurement performed in 1994 using 80 events. While validating the EKP full reconstruction tool [2], a similar discrepancy was found between the Belle Monte Carlo (that contains ≈ 75000 signal events) and recorded data.

The yield together with the combined statistical and systematic (see Section 8.4) error is then used to constrain the total signal yield in the Dalitz fit.

Channel	Yield	Expectation
$B^- \to D^0\pi^-\pi^0$	66195 ± 984	84653
$B^0 \to D^-\pi^+\pi^0$	74246 ± 1060	82673

Table 8.2.: Results of ΔE fits on recorded Data. The result in comparison with the expectation shows, that the previous measurements overestimated the branching fractions of the dominant decay via the ρ meson.

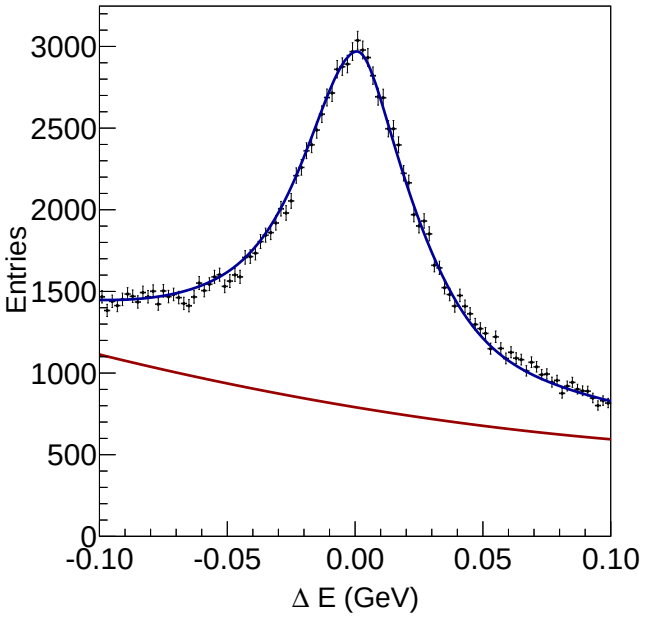


Figure 8.7.: Fit to recorded data ($B^- \to D^0\pi^-\pi^0$).

8.4. Systematic error on the signal yield

Each parameter that is determined on signal Monte Carlo and fixed before the fit to recorded data has an uncertainty on its value. For each parameter the fit is repeated twice, once with the parameter value increased by the value of the error and once with the parameter value decreased by the same amount. The yield of these fits is compared to the actual fit result. The relative deviation is then added quadratically to those from the variation of other parameters. The resulting systematic errors caused by the fixing of signal shape parameters are 0.34% for $B^- \to D^0\pi^-\pi^0$ and 0.32% for $B^0 \to D^-\pi^+\pi^0$. For the fixed parameters in the background description, the systematic error is 0.29% in both channels.

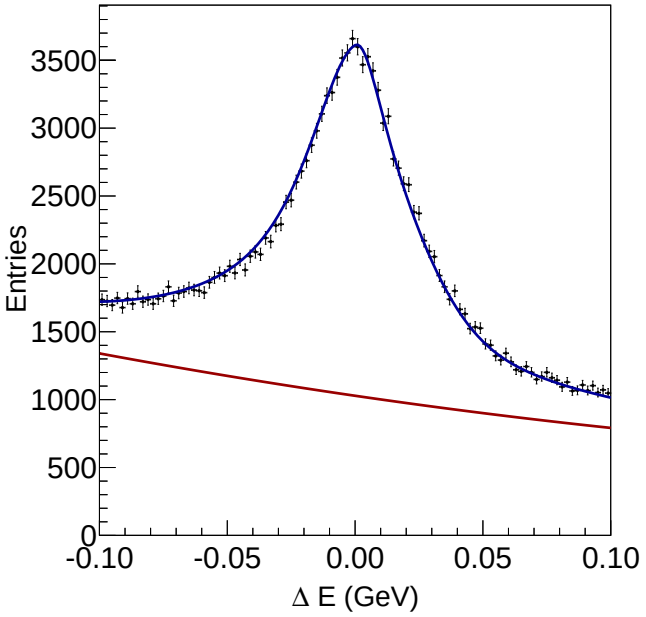


Figure 8.8.: Fit to recorded data ($B^0 \to D^-\pi^+\pi^0$).

The misreconstructed signal in the ΔE distribution is prone to exchange events with the background component. To estimate a systematic error for this, the ΔE fit is repeated changing the fraction of misreconstructed signal by $\pm 10\%$ in the simulated sample and the signal yield is compared with the result for the central value. The largest deviation is taken as systematic error and is listed in Table 8.3.

Decay Mode	Signal PDF (%)	BG PDF (%)	Signal fraction (%)	Sum (%)
$B^- \to D^0\pi^-\pi^0$	0.34	0.29	3.13	3.16
$B^0 \to D^-\pi^+\pi^0$	0.32	0.29	2.76	2.79

Table 8.3.: Systematic errors on the signal yield from the ΔE fit, BG is the abbreviation for background.

9. Dalitz Fit model and validation on simulated data

9.1. Fit framework Laura++

The dalitz fit has been performed utilizing the *Laura++* [19] (Likelihood Analysis Unbinned Reconstruction of Amplitudes) software package. It is designed in particular for Dalitz plot analysis and was used in several analysis published by BaBar, e.g. [28] or [29]. It is implemented in C++ and incorporates classes from *Root* [30], e.g. the interface TMinuit to the *Minuit* [20] fitting routine. Its main features are:

- Laura++ is an extended likelihood fitter.
- It can form an isobar model for a spin 0 parent particle decay to three spin 0 children.
- Various resonance line shapes are implemented.
- Laura++ supports amplitude/phase or real/imaginary coefficients.
- Laura++ supports the squared Dalitz plot.
- Tools for performing ensemble tests are included.
- Efficiency variation can be modeled using a two-dimensional histogram.
- Migration of misreconstructed events can be handled using a migration matrix.
- Different background models can be added via two-dimensional histograms.

To speed up the fit, the integrals for the normalization of the amplitude can be cached, similar to the method described in [31]. One property worth mentioning is, that Laura++ interpolates between the bin centers in the efficiency and misreconstructed signal fraction histograms.

9.2. Fit model

The fit model consists of the following components:

- The signal model containing the different resonances and a non-resonant component (see Table 9.1).
- Three background models for continuum and mixed/charged combinatoric background, supplied as smoothed 2D histograms.

- The detector response: A 2D efficiency histogram, a 2D histogram for the fraction of migrating signal and a migration matrix, consisting of 22500 migration probability histograms.

All signal components are added coherently to account for possible interference terms except for the contribution of the D^*. The width of both D^* mesons is very small compared to the detector resolution and thus the region for possible interference is very small. The contribution is therefore incoherently added as two 2D histograms to properly account for the detector resolution. The histogram is formed from a sample of three million generated events containing only the decay via an excited D meson.

The use of magnitude and phase to describe the amplitudes can exhibit biases when the magnitude of the contribution is small. Therefore the real and imaginary part (X and Y) are used instead to describe the amplitudes, which also increases the fit stability.

Resonance	Mass (GeV)	Width (MeV)	line shape	form factor	Coherent
$\rho(770)$	0.775	149	Gounaris Sakurai	5.3	yes
$D_0^*(2400)^0$	2.318	267	Rel. B-W	1.0	yes
$D_2^*(2462)^0$	2.460	49	Rel. B-w	4.0	yes
$D_0^*(2400)^-$	2.403	283	Rel. B-W	1.0	yes
$D_2^*(2462)^-$	2.464	37	Rel. B-w	4.0	yes
Non-resonant	-	-	Flat	-	yes
$D^*(2007)$	2.007	< 2.1	2 Histograms	-	no
$D^*(2010)$	2.010	0.08	2 Histograms	-	no

Table 9.1.: Components in the signal model and their properties.

9.3. Fit to fully simulated data

In order to check the fit procedure and that the different components describe the effects of the detector properly, fits to signal Monte Carlo samples with full detector simulation are performed. This is not possible on the generic Monte Carlo since no interference effects are taken into account there.
Although there is a dedicated EvtGen model for Dalitz decays (called EvtGenericDalitz), this task still possesses various difficulties, mainly due to imperfections in the EvtGenericDalitz model.

9.3.1. The EvtGenericDalitz model

While this model performs the basic task of calculating and coherently adding up amplitudes quite well, it has severe limitations. First, the individual resonances described by a relativistic Breit Wigner (Gounaris-Sakurai not available) are not normalized to unity over the whole Dalitz area. This makes it impossible to compare the input values of amplitudes with the fit results from Laura++, where each resonance is normalized to one. For example, setting the amplitudes of the ρ and the $D_0^*(2400)^0$ to the same value in EvtGen, does not yield an equal contribution in the generated sample, but rather a ratio of $\approx 50/1$. Since both EvtGen and Laura++ use the accept/reject method to generate samples according to the calculated PDF, one can get an approximation of this ratio by generating samples containing only one resonance and comparing the highest value of the PDF, which is used to define the ceiling for the accept/reject method. With this approximation it is possible to simulate EvtGen samples, that roughly resemble the expected distributions.
However, for single resonances it is possible to use *EvtGenericDalitz* to test that the fit is able to describe the distribution and the detector effects, especially the migration of events. The

test was performed using Monte Carlo truth matched samples, where no events with migration besides the detector resolution are present, and samples with no Monte Carlo truth cuts, thus containing all migrating events in addition to good signal. The results of these fits, where the amplitude of the resonance in question was fixed to 1.0, while the amplitudes of the other contributions were allowed to float, can be seen in Figures 9.1 to 9.12. In all cases the fitted contribution from the resonances not present in the data sample were small and compatible with zero.
The results for the second channel $B^0 \to D^-\pi^+\pi^0$ can be seen in Figures 9.13 to 9.32 for the fits to the single resonances with and without migrating signal. All plots follow the same scheme: The resonance and the decay mode that is fitted is given in the headline. On the first page the fits to a truth matched signal sample is shown, without migrating events, followed by another page with the fit to the full signal sample including migrating signal events and the treatment of the migration. In the first row, the projections of the two-dimensional Dalitz scatter plot to the squared invariant masses of $D\pi^0$ (left) and $D\pi^\pm$ (right) are shown for the simulated data and the fit results. In the second row, the projections to the squared Dalitz plot variables are shown, mPrime on the left and thPrime on the right. Note that all fit results are visualized by generating toy events with the accept/reject method according to the total fitted two-dimensional probability density function. Due to limited statistics and the binning, it occurs that for narrow structures the data points seem to be not described properly by the fit. This has been checked and can be attributed to the visualization of the fit result rather than to the actual shape of the fit result.

In the last row the scatter plots in the squared Dalitz plot are displayed, where the left hand side shows the (simulated) data distribution, and the right hand side the results of the fit. The scatter plots of the normal Dalitz plots are in Appendix B.

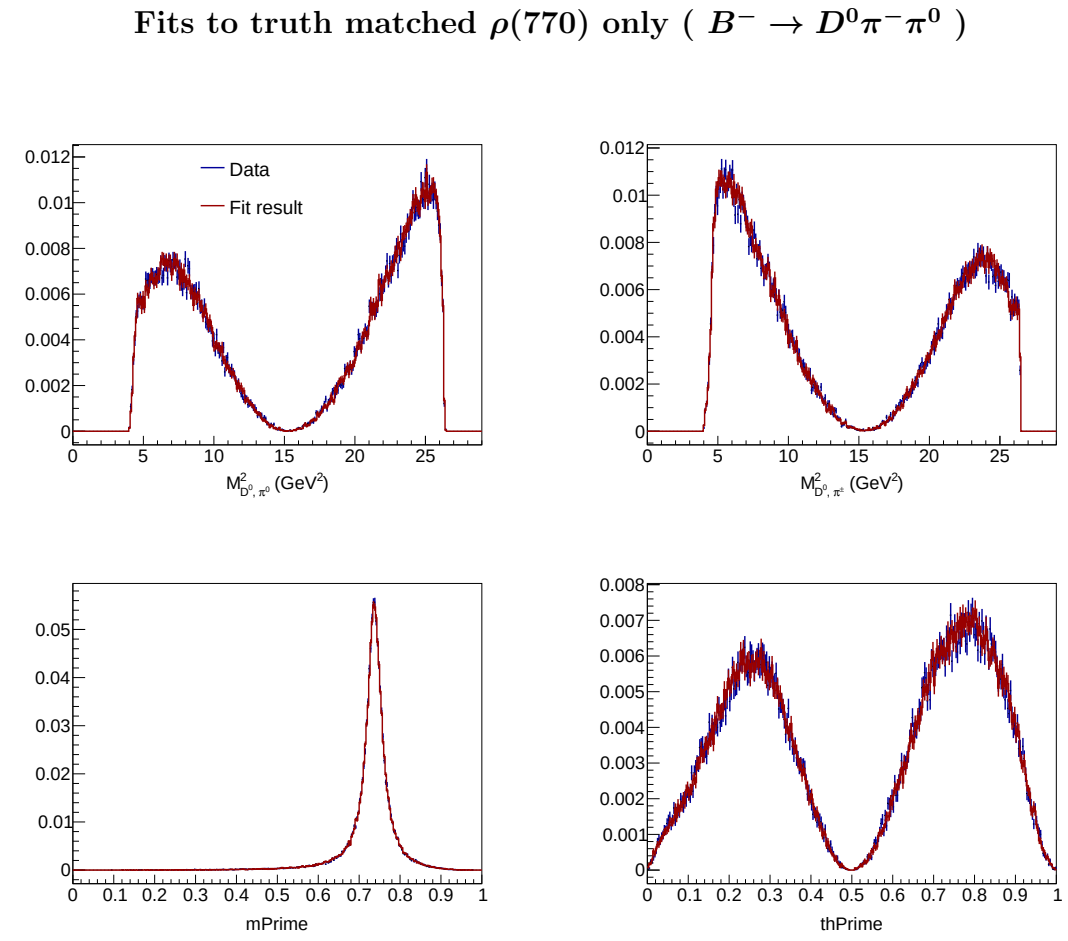


Figure 9.1.: Projections to the normal and squared Dalitz plot variables of the fit to truth matched signal sample.

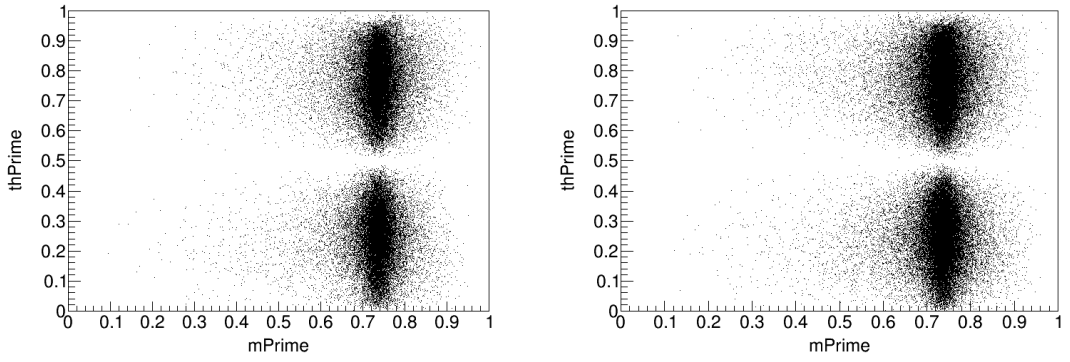


Figure 9.2.: Scatter plots in the squared Dalitz plot showing the input sample (left) and the fit result (right).

Fits to full $\rho(770)$ sample ($B^- \to D^0\pi^-\pi^0$)

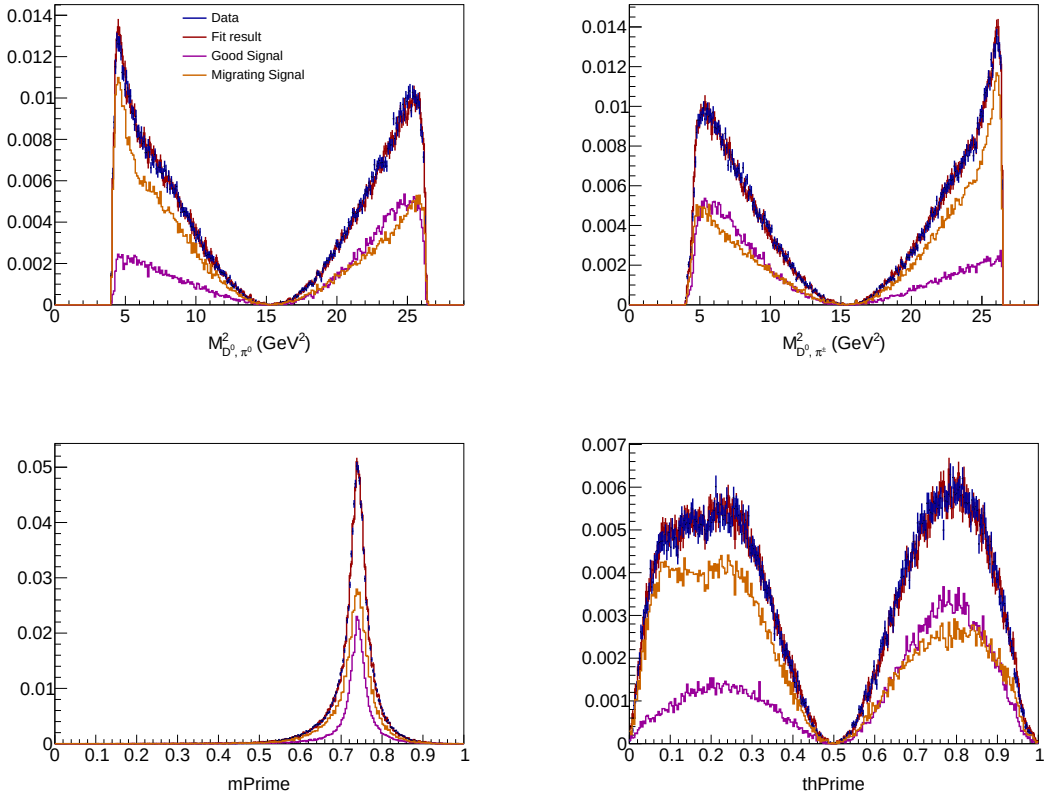


Figure 9.3.: Projections to the normal and squared Dalitz plot variables of the fit to signal Monte Carlo including misreconstructed events.

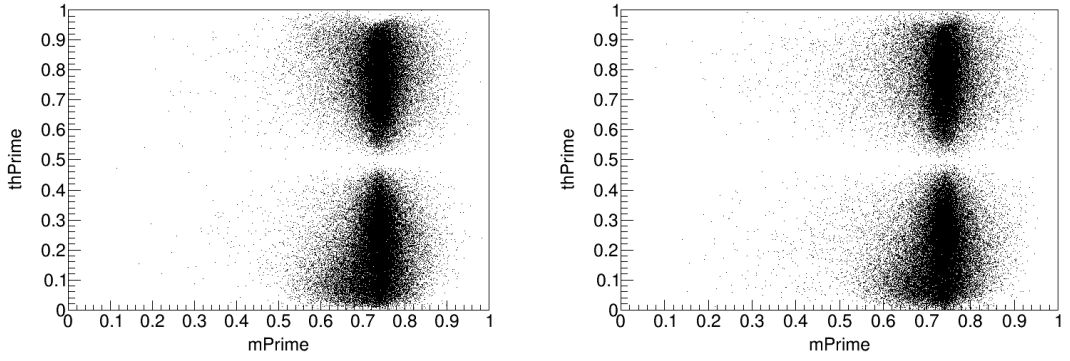


Figure 9.4.: Scatter plots in the squared Dalitz plot showing the input sample (left) and the fit result (right).

Fits to truth matched $D_0^*(2400)^0$ only ($B^- \to D^0\pi^-\pi^0$)

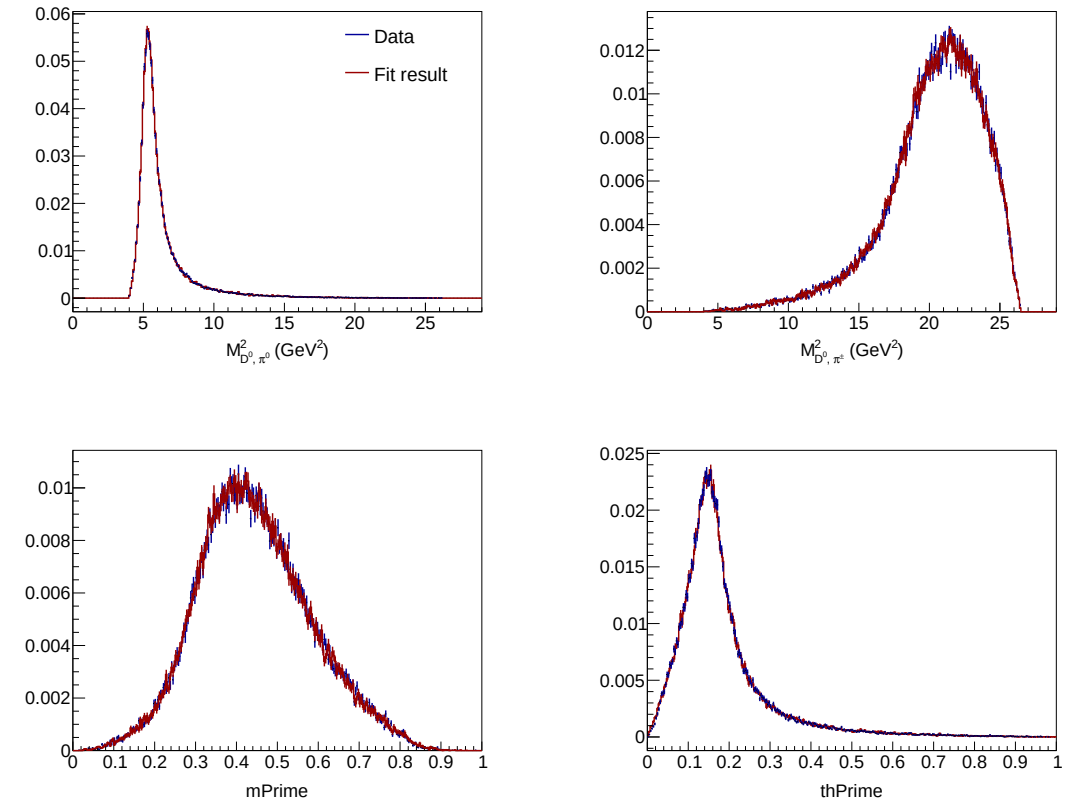


Figure 9.5.: Projections to the normal and squared Dalitz plot variables of the fit to truth matched signal sample.

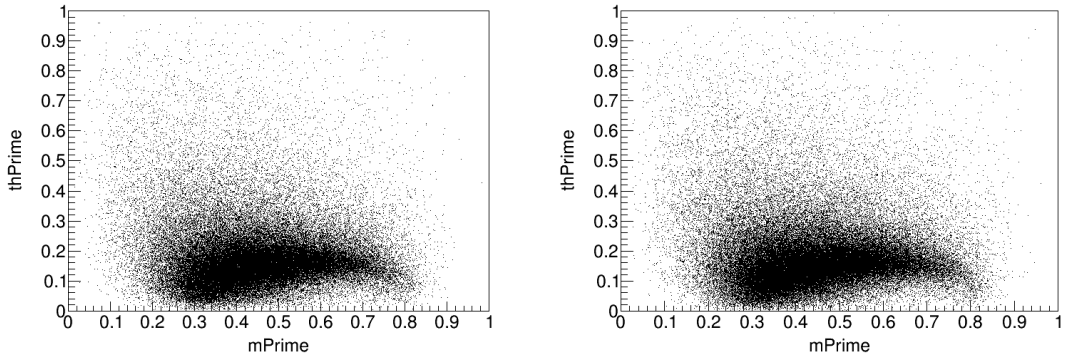


Figure 9.6.: Scatter plots in the squared Dalitz plot showing the input sample (left) and the fit result (right).

Fits to full $D_0^*(2400)^0$ sample ($B^- \to D^0\pi^-\pi^0$)

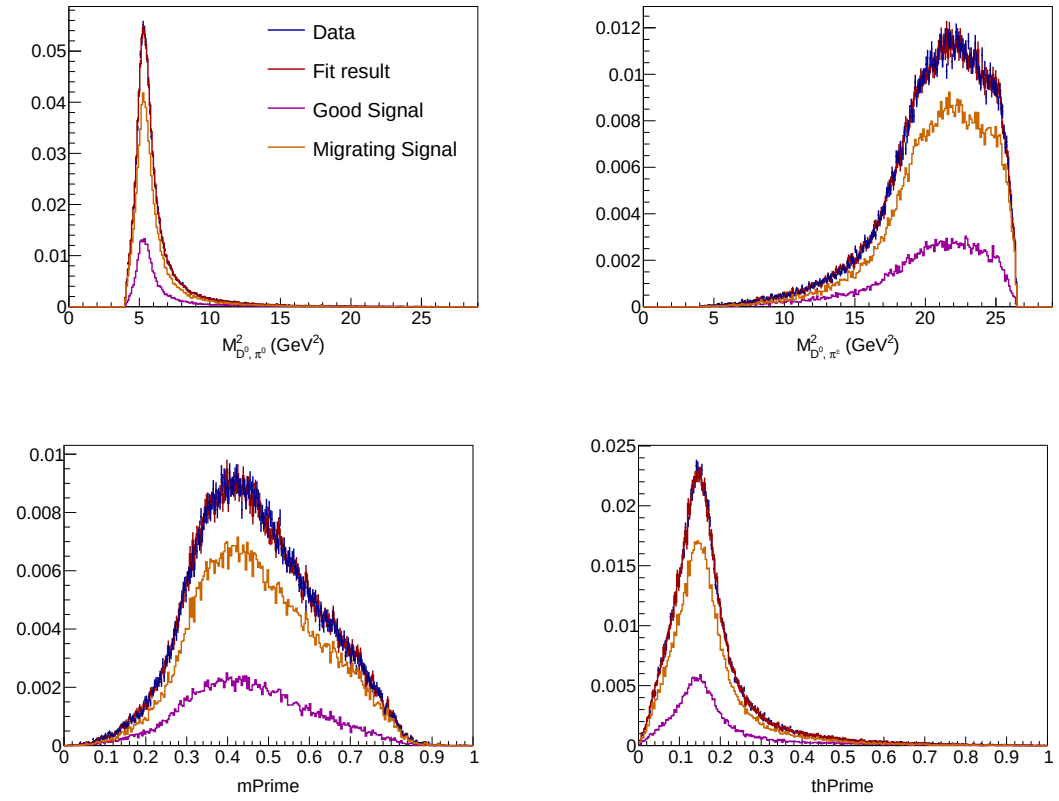


Figure 9.7.: Projections to the normal and squared Dalitz plot variables of the fit to signal Monte Carlo including misreconstructed events.

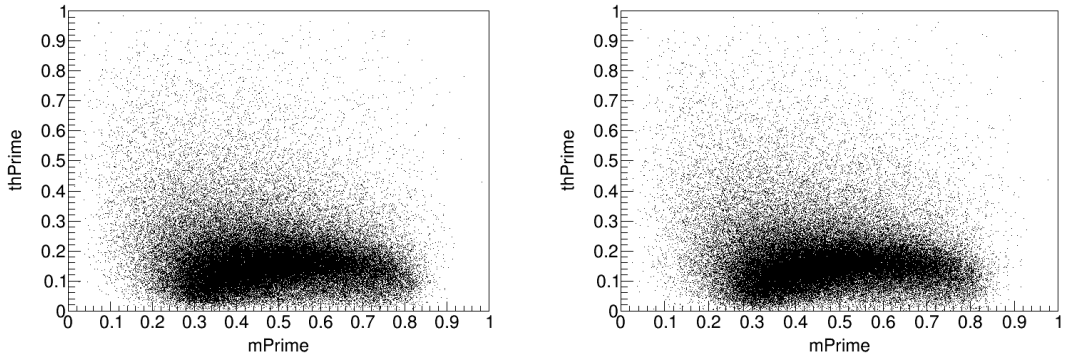


Figure 9.8.: Scatter plots in the squared Dalitz plot showing the input sample (left) and the fit result (right).

Fits to truth matched $D_2^*(2462)^0$ only ($B^- \to D^0\pi^-\pi^0$)

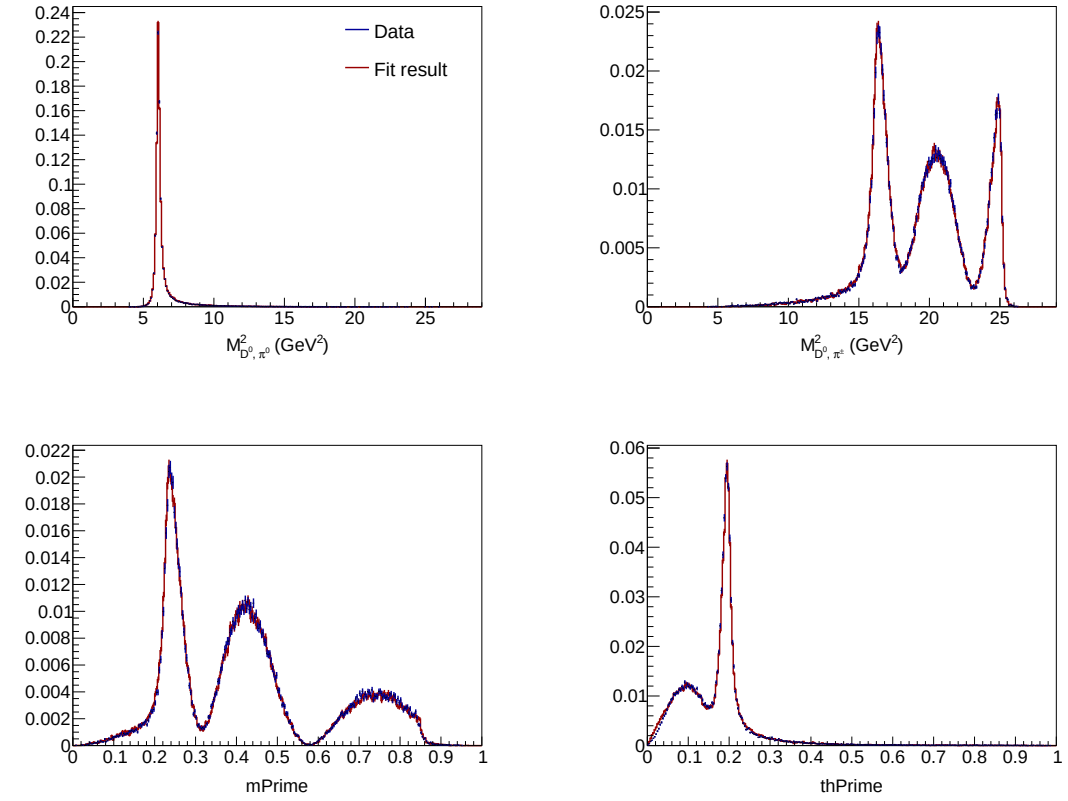


Figure 9.9.: Projections to the normal and squared Dalitz plot variables of the fit to truth matched signal sample.

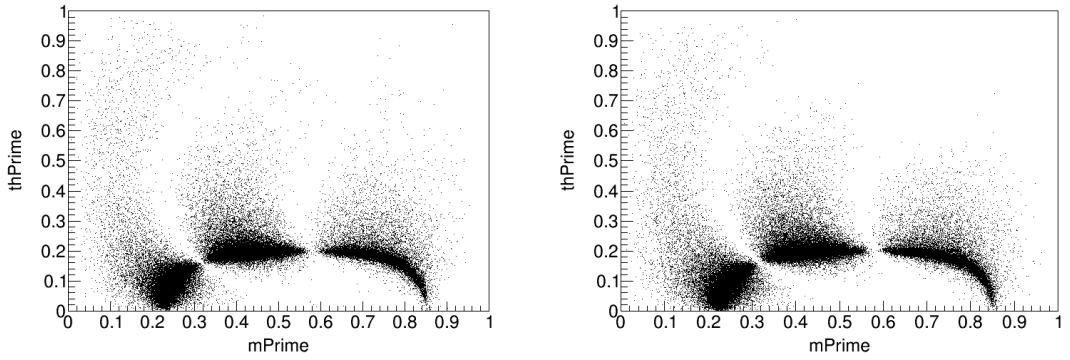


Figure 9.10.: Scatter plots in the squared Dalitz plot showing the input sample (left) and the fit result (right).

Fits to full $D_2^*(2462)^0$ sample ($B^- \to D^0\pi^-\pi^0$)

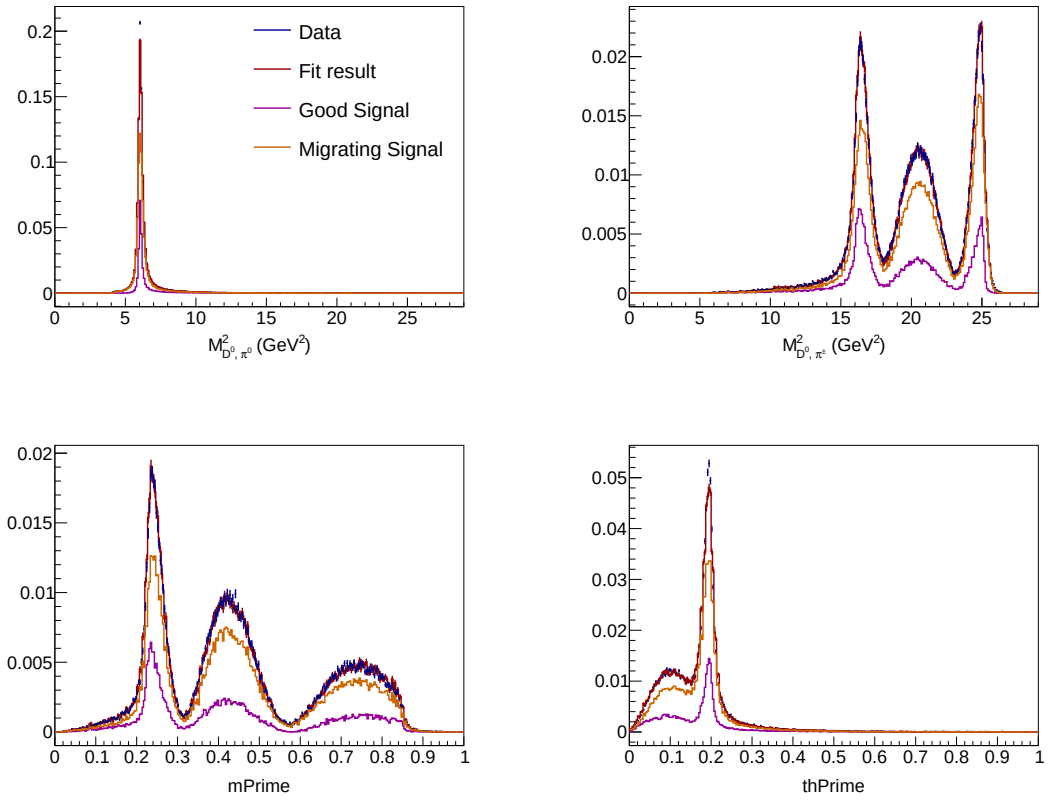


Figure 9.11.: Projections to the normal and squared Dalitz plot variables of the fit to signal Monte Carlo including misreconstructed events.

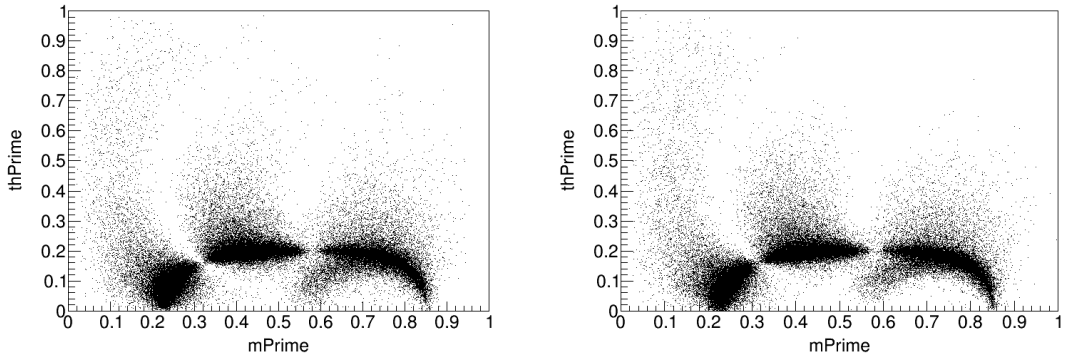


Figure 9.12.: Scatter plots in the squared Dalitz plot showing the input sample (left) and the fit result (right).

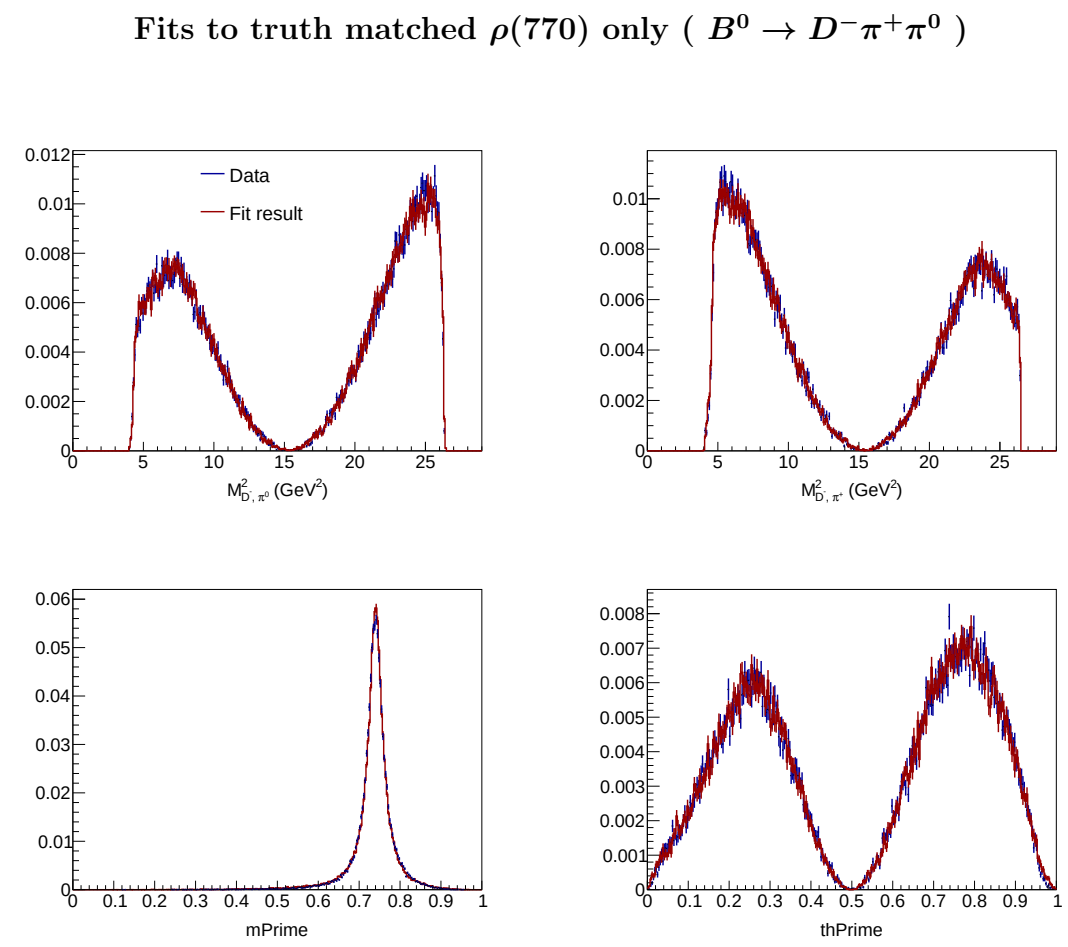


Figure 9.13.: Projections to the normal in the squared Dalitz plot variables of the fit to truth matched signal sample.

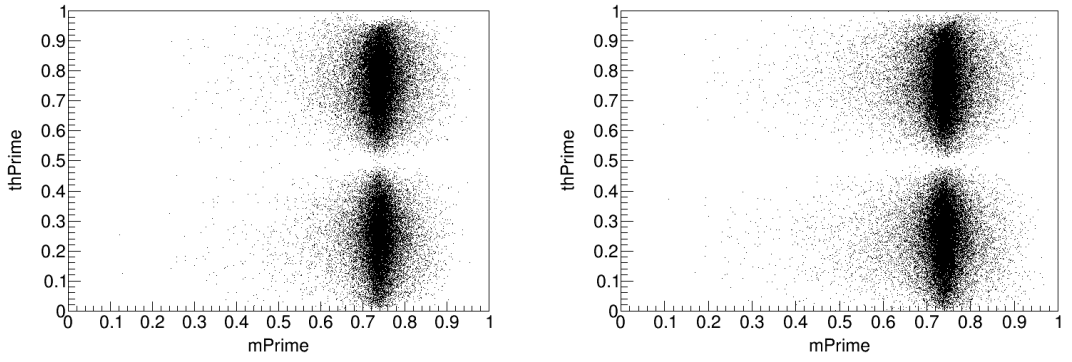


Figure 9.14.: Scatter plots in the squared Dalitz plot showing the input sample (left) and the fit result (right).

Fits to full $\rho(770)$ sample ($B^0 \to D^- \pi^+ \pi^0$)

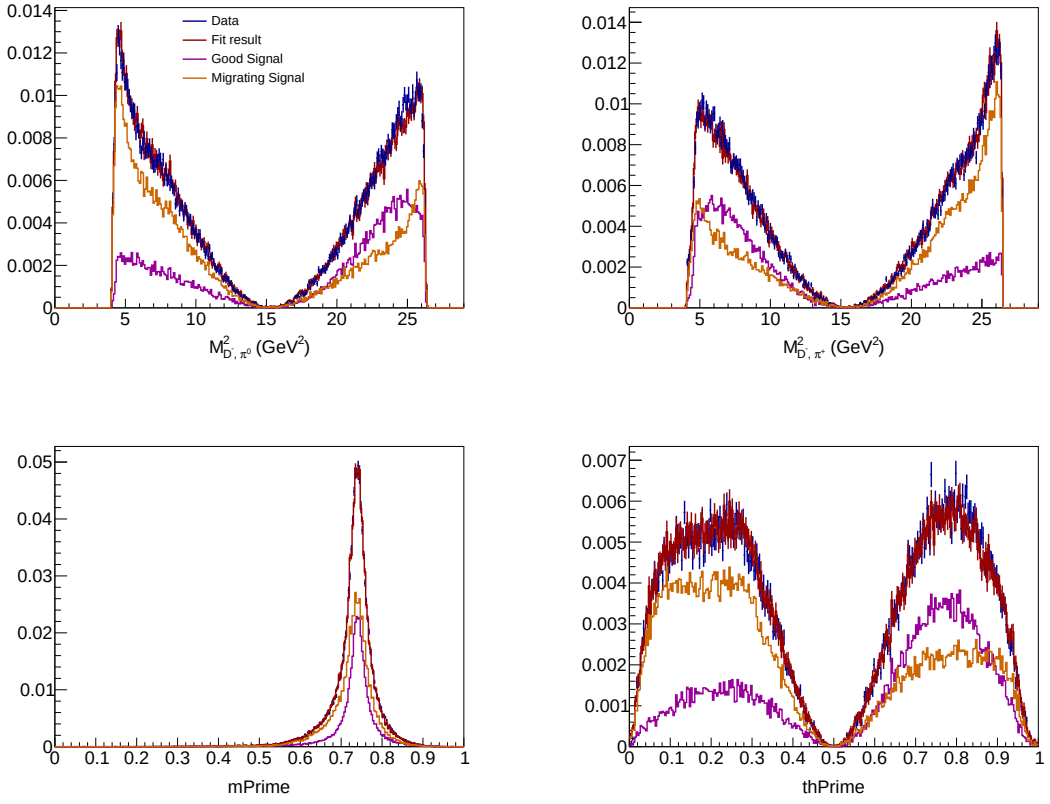


Figure 9.15.: Projections to the normal and squared Dalitz plot variables of the fit to signal Monte Carlo including misreconstructed events.

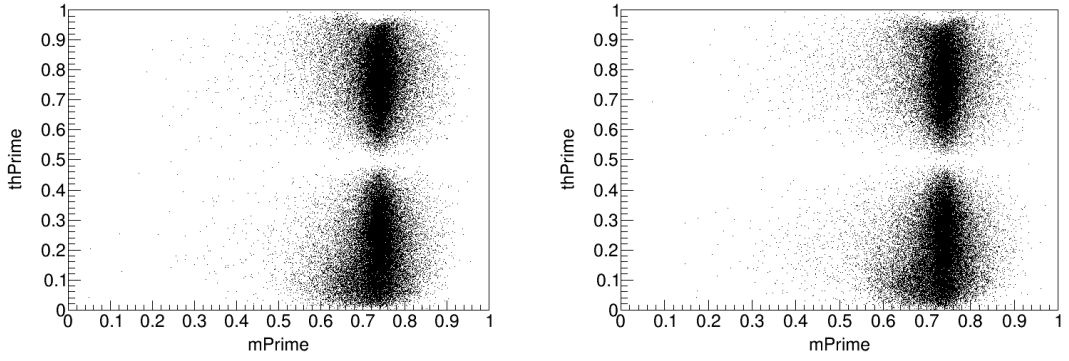


Figure 9.16.: Scatter plots in the squared Dalitz plot showing the input sample (left) and the fit result (right).

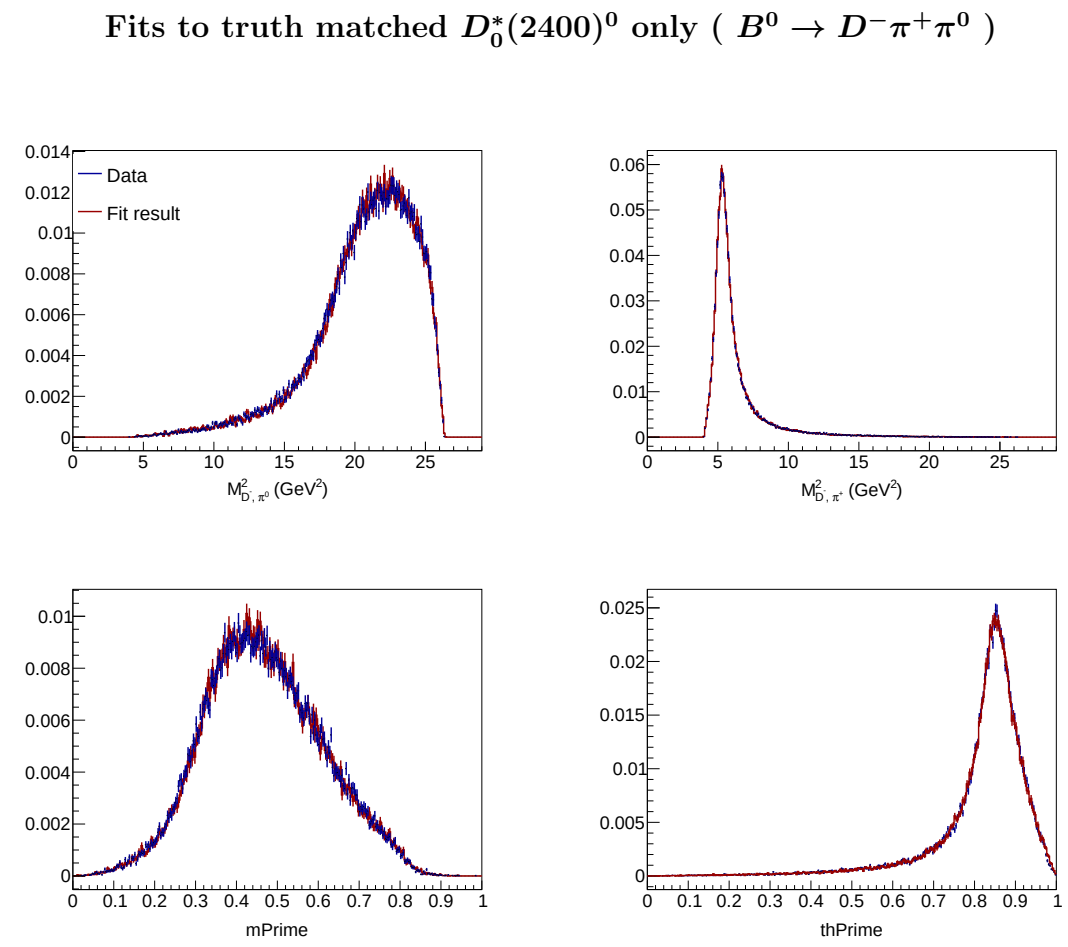


Figure 9.17.: Projections to the normal and squared Dalitz plot variables of the fit to truth matched signal sample.

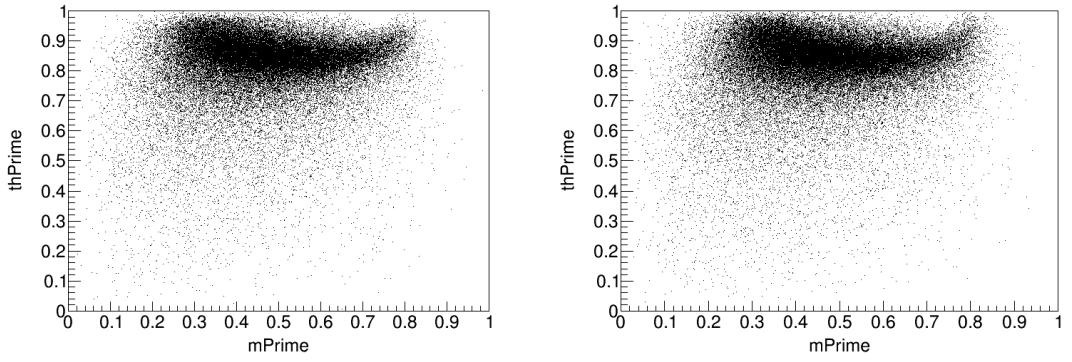


Figure 9.18.: Scatter plots in the squared Dalitz plot showing the input sample (left) and the fit result (right).

Fits to full $D_0^*(2400)^0$ sample ($B^0 \to D^-\pi^+\pi^0$)

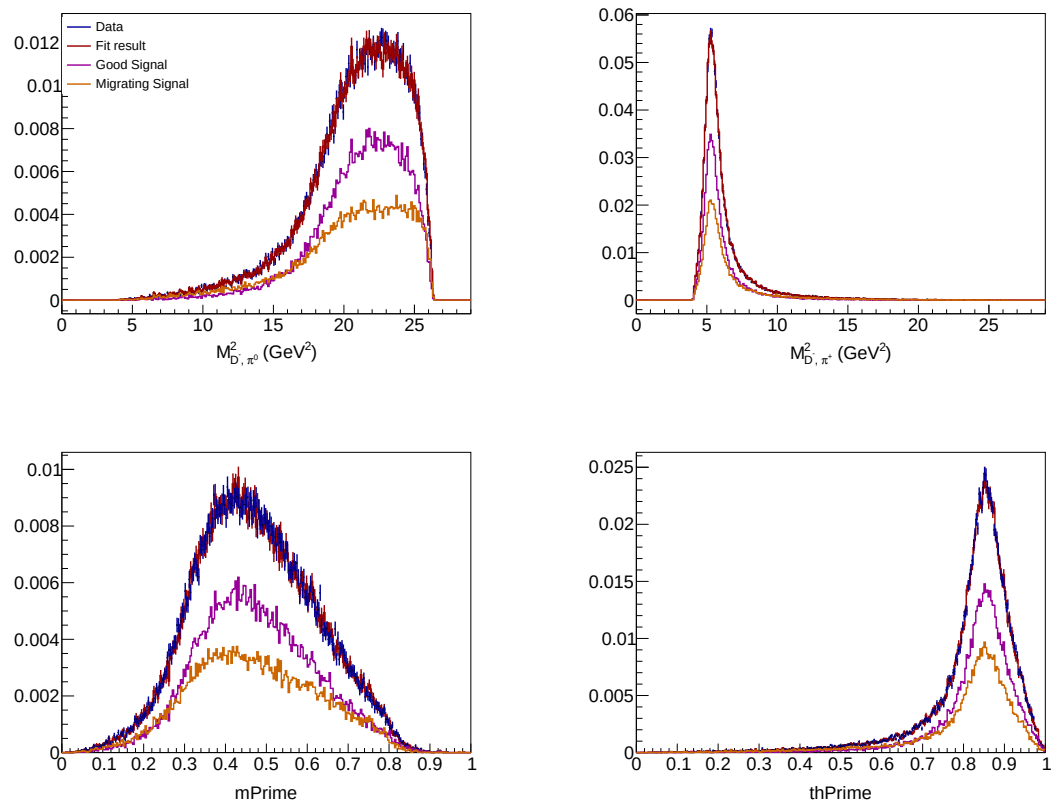


Figure 9.19.: Projections to the normal and squared Dalitz plot variables of the fit to signal Monte Carlo including misreconstructed events.

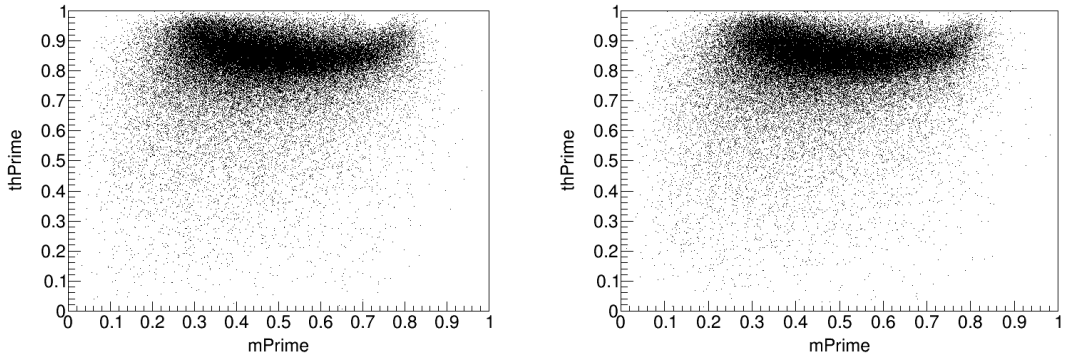


Figure 9.20.: Scatter plots in the squared Dalitz plot showing the input sample (left) and the fit result (right).

Fits to truth matched $D_2^*(2462)^0$ only ($B^0 \to D^-\pi^+\pi^0$)

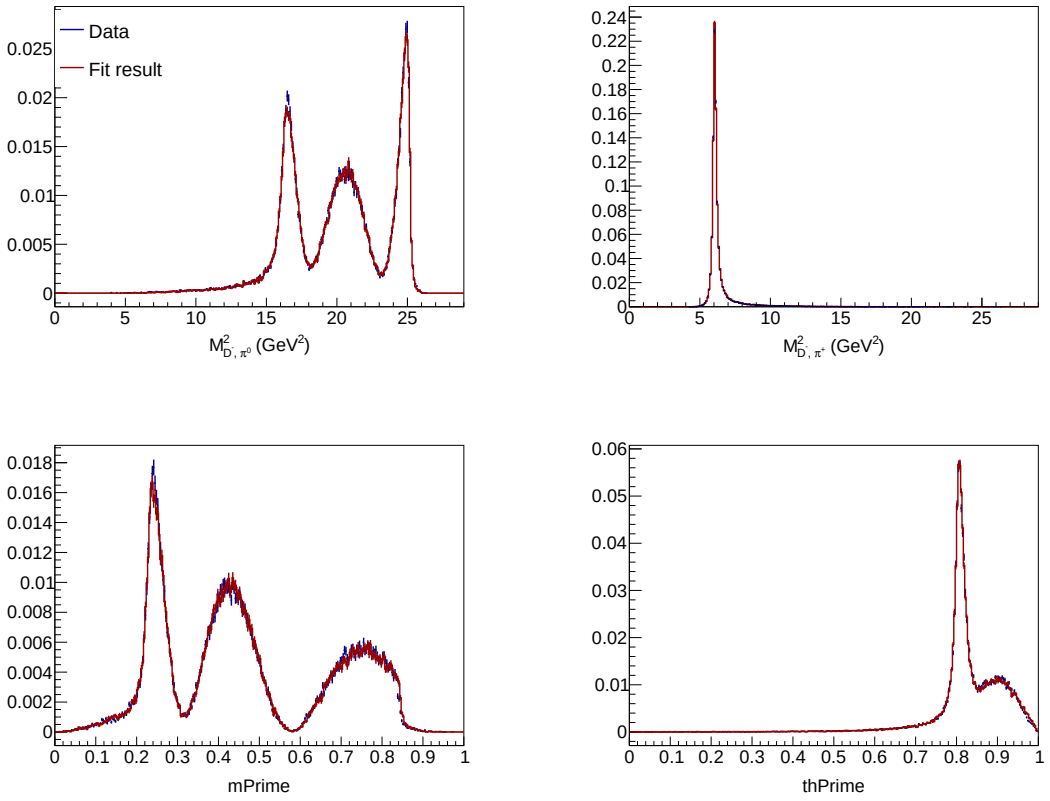


Figure 9.21.: Projections to the normal and squared Dalitz plot variables of the fit to truth matched signal sample.

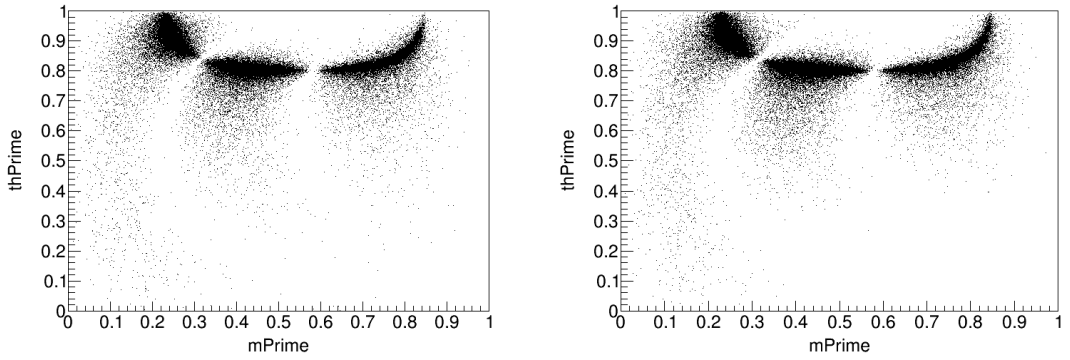


Figure 9.22.: Scatter plots in the squared Dalitz plot showing the input sample (left) and the fit result (right).

Fits to full $D_2^*(2462)^0$ sample ($B^0 \to D^-\pi^+\pi^0$)

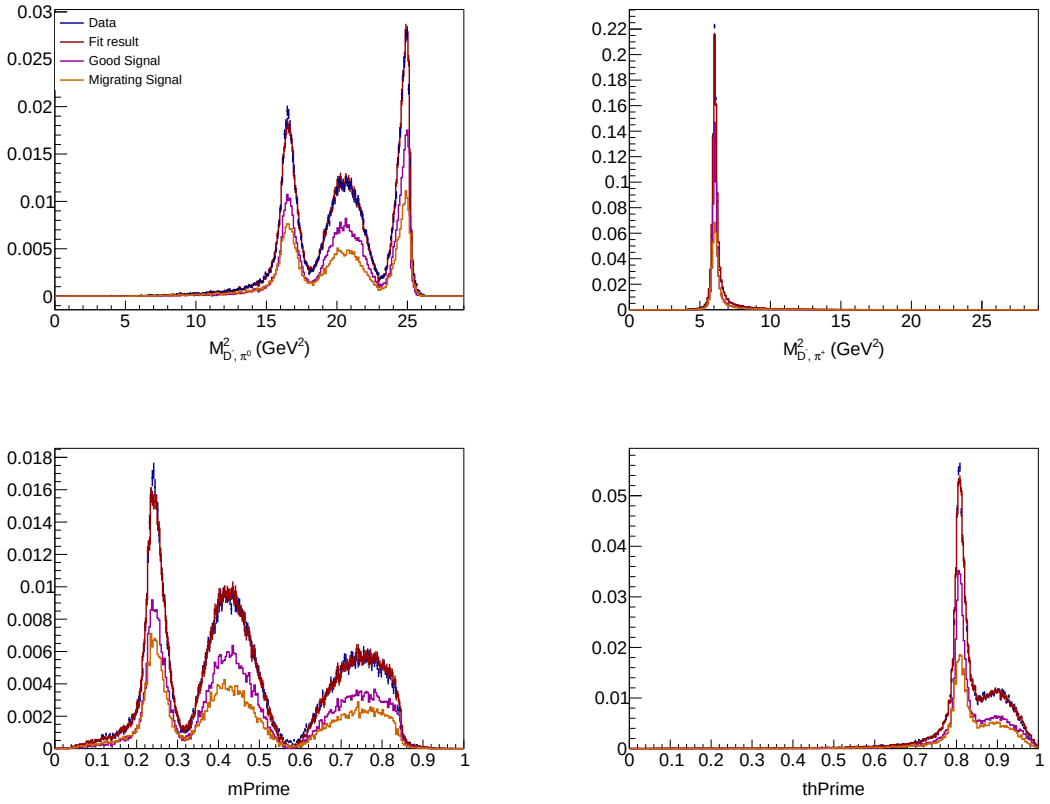


Figure 9.23.: Projections to the normal and squared Dalitz plot variables of the fit to signal Monte Carlo including misreconstructed events.

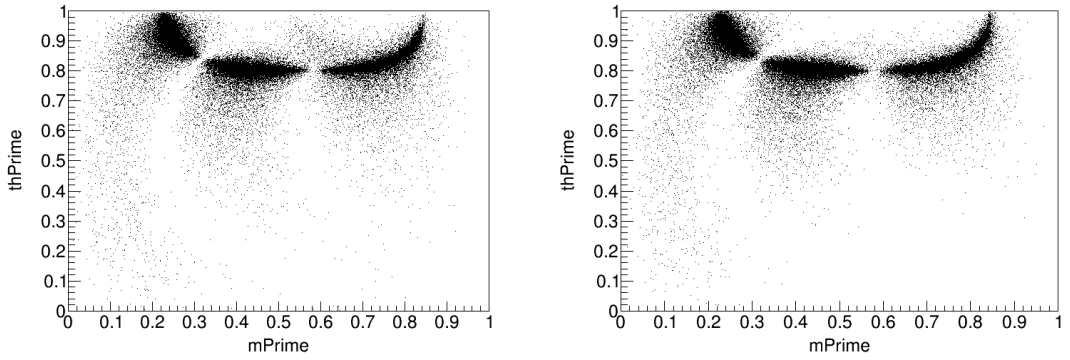


Figure 9.24.: Scatter plots in the squared Dalitz plot showing the input sample (left) and the fit result (right).

Fits to truth matched $D_0^*(2400)^-$ only ($B^0 \to D^-\pi^+\pi^0$)

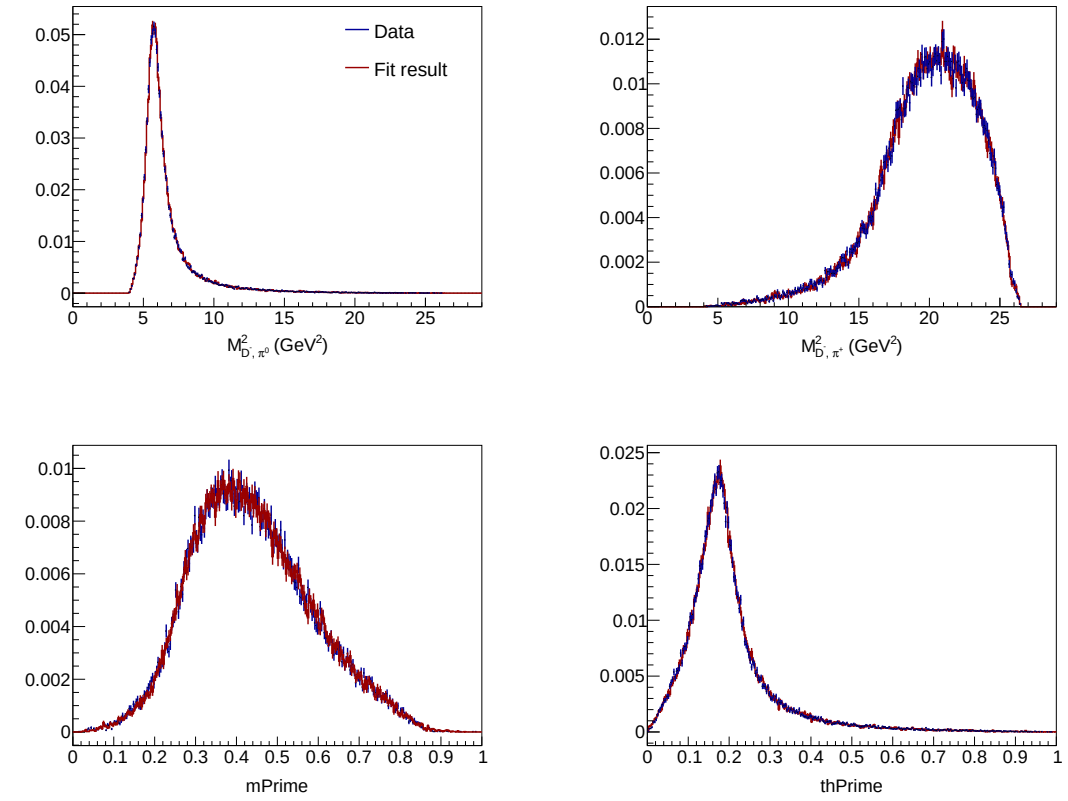


Figure 9.25.: Projections to the normal and squared Dalitz plot variables of the fit to truth matched signal sample.

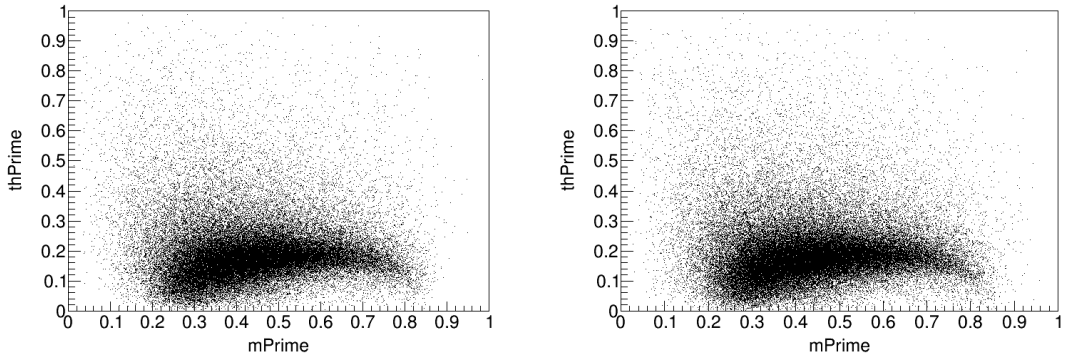


Figure 9.26.: Scatter plots in the squared Dalitz plot showing the input sample (left) and the fit result (right).

Fits to full $D_0^*(2400)^-$ sample ($B^0 \to D^-\pi^+\pi^0$)

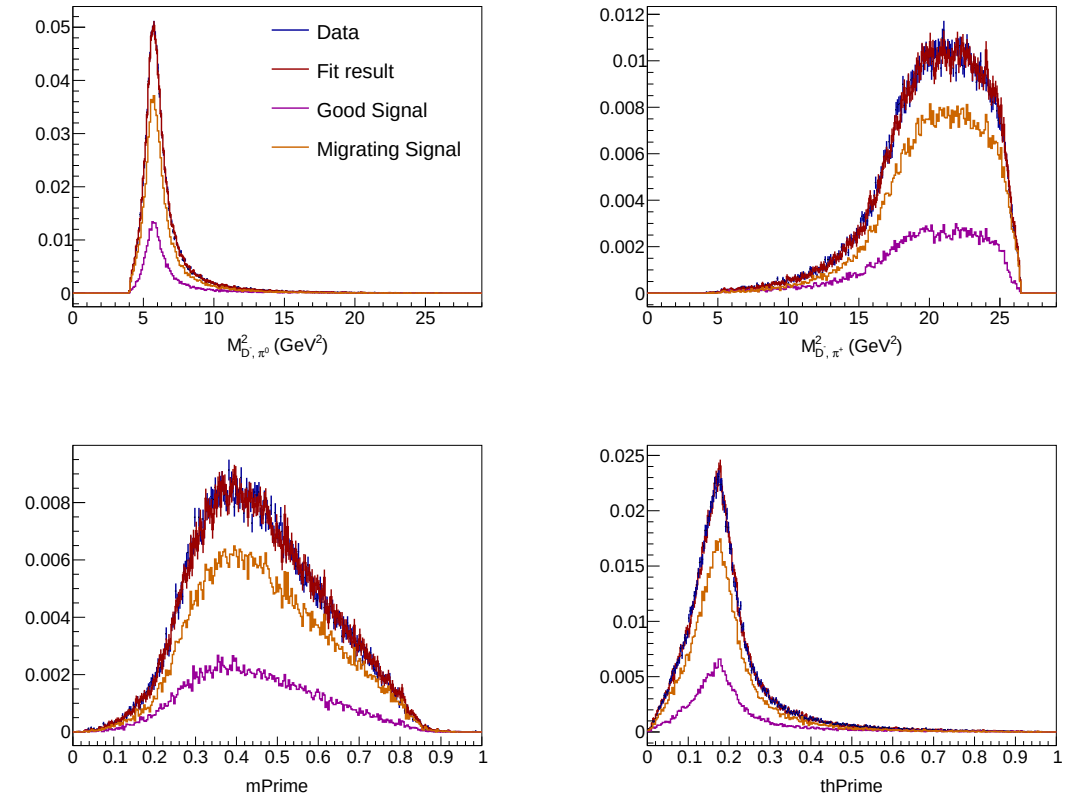


Figure 9.27.: Projections to the normal and squared Dalitz plot variables of the fit to signal Monte Carlo including misreconstructed events.

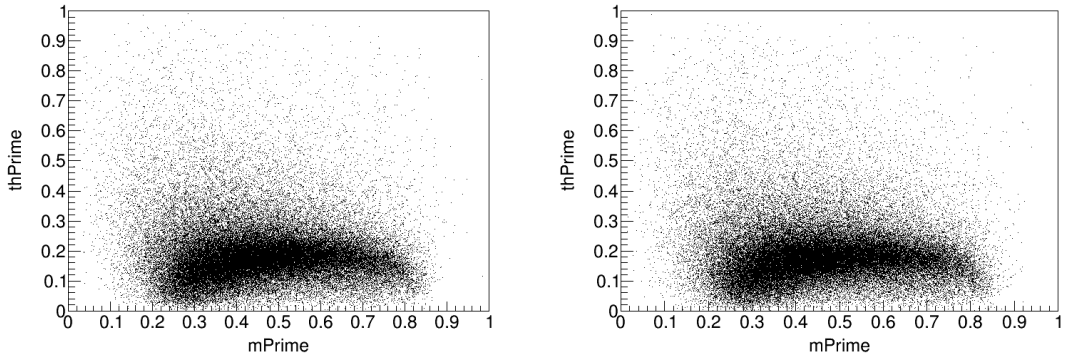


Figure 9.28.: Scatter plots in the squared Dalitz plot showing the input sample (left) and the fit result (right).

Fits to truth matched $D_2^*(2462)^-$ only ($B^0 \to D^-\pi^+\pi^0$)

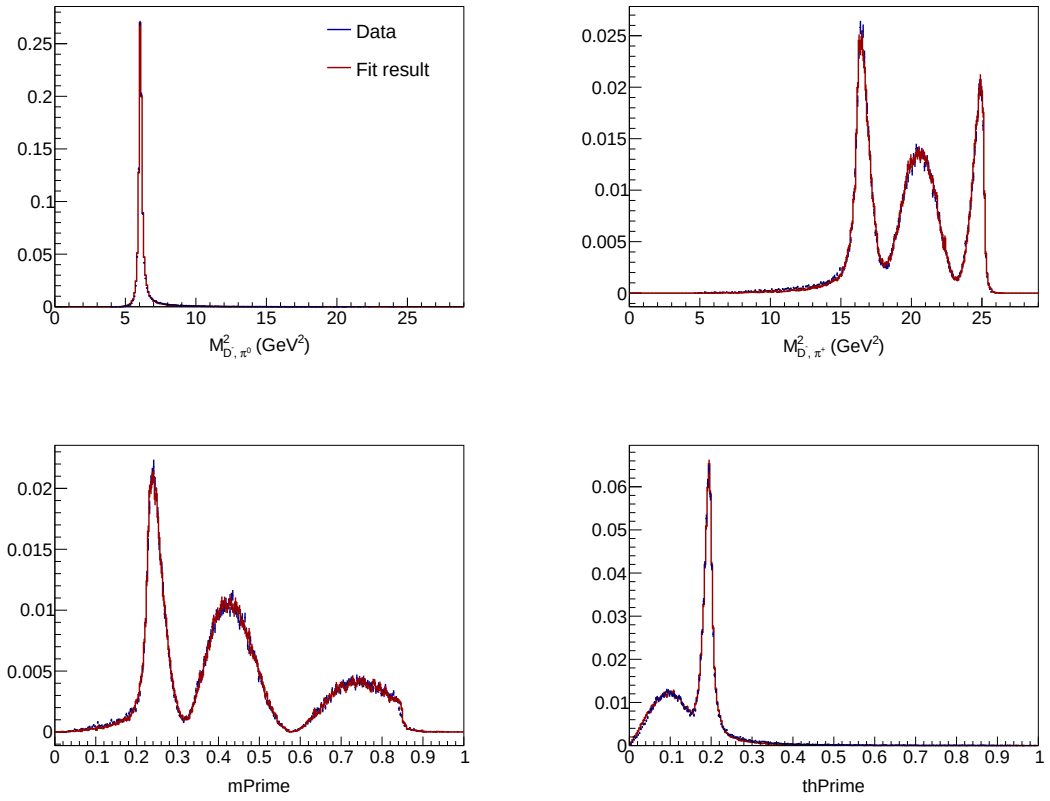


Figure 9.29.: Projections to the normal and squared Dalitz plot variables of the fit to truth matched signal sample.

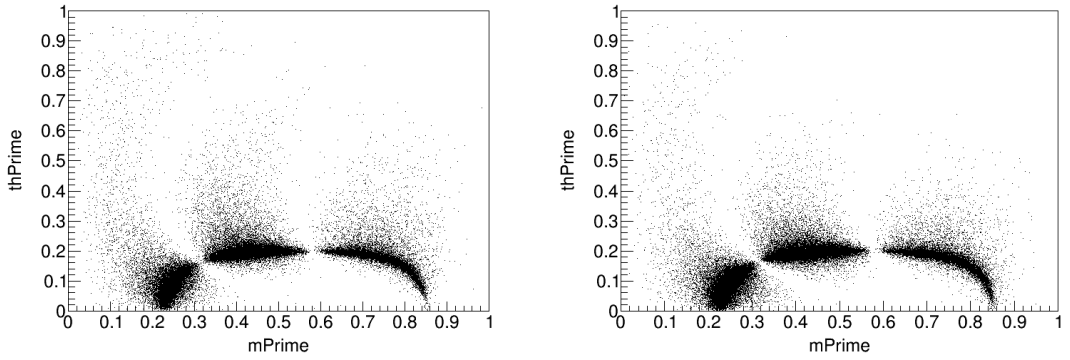


Figure 9.30.: Scatter plots in the squared Dalitz plot showing the input sample (left) and the fit result (right).

Fits to full $D_2^*(2462)^-$ sample ($B^0 \to D^-\pi^+\pi^0$)

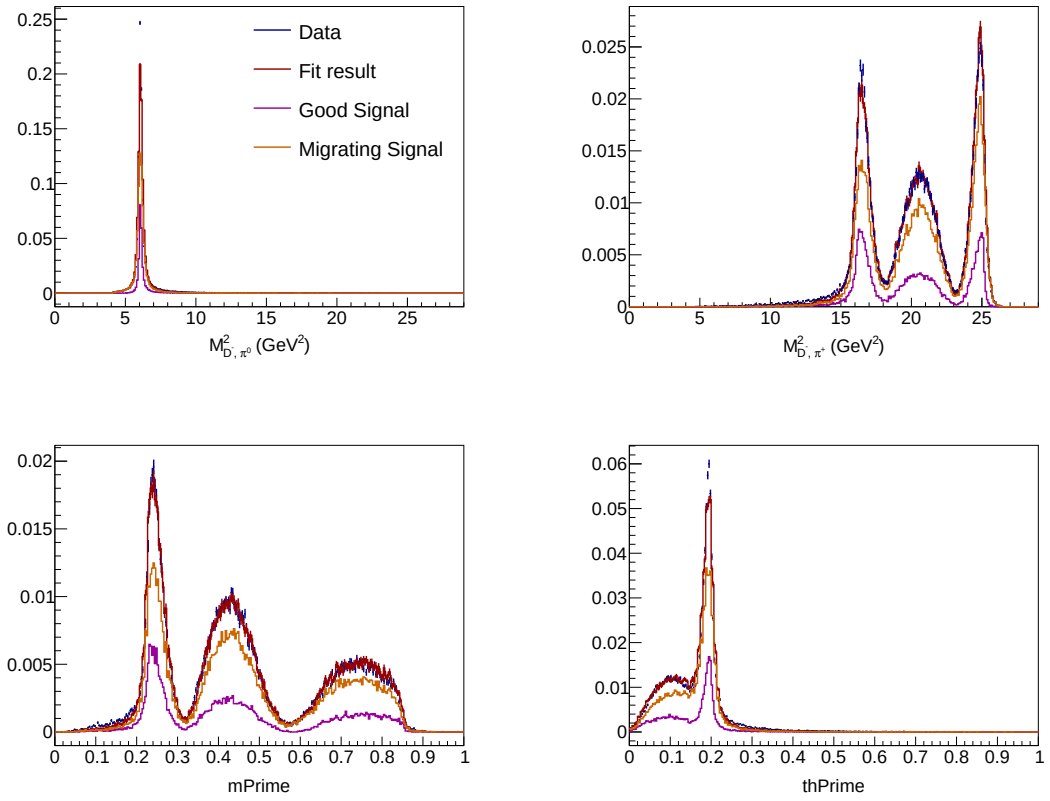


Figure 9.31.: Projections to the normal and squared Dalitz plot variables of the fit to signal Monte Carlo including misreconstructed events.

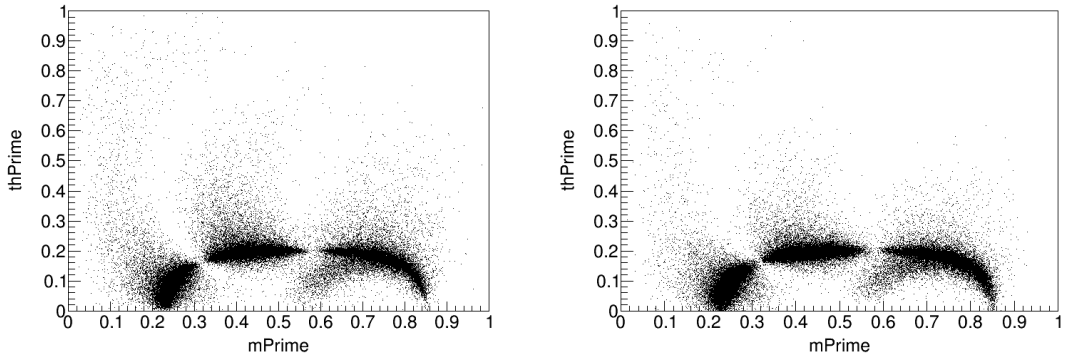


Figure 9.32.: Scatter plots in the squared Dalitz plot showing the input sample (left) and the fit result (right).

9.3.2. Conclusion of the fits to EvtGen samples

The fits to the single resonances, where the normalization is irrelevant, prove that the migration matrix, combined with the efficiency histogram and the migrating signal fraction histogram are able to describe the detector effects. The line shapes of the fit result and the simulated data for truth matched signal or for all signal are all in good agreement.

Also a fit to a EvtGen sample generated according to the current known branching fractions has been performed, and the fit is able to describe the data points, but given that the EvtGen input values for the amplitudes and phases are not comparable with the fit results, this is not a reliable test and therefore not included here. However, the comparison of the fit result with the data points can be found in Appendix C. Note: In the second channel $B^0 \to D^-\pi^+\pi^0$ a large contribution of $D_2^*(2462)^0$ was falsely assumed in these tests instead of a large $D_2^*(2462)^-$.

9.4. Ensemble tests before unblinding

To check if the fit procedure introduces any biases, so-called toy studies or ensemble tests were performed prior to the unblinding of recorded data. This was done to proof that the fitting procedure is able to reproduce the input values in a range around the expected values. To test these different realistic scenarios, e.g., different combinations of amplitudes and phases, the currently best known branching fractions from the Particle Data Group are used (see Tables 1.1 and 1.2) to build *random models* where the relative magnitudes and phases between the resonances are varied: The amplitudes take a value randomly taken from a normal distribution around the nominal mass value, with their width corresponding to the error on the nominal value. The phase is randomly taken to be within $-\pi$ and π. Since there is no branching fraction available for the non-resonant case, it is assumed that it is small compared to the $\rho(770)$ contribution and therefore set to 0.1×10^{-3} with an error of 100%.
For each model 300 experiments are generated and fitted, and the fit results are compared with the input values. The distribution of these differences should have a Gaussian form with mean zero and width unity. The results are discussed in Section 9.4.2. The ensemble tests on these random models were also used to evaluate different settings of the migration matrix, see Section 9.4.1.

9.4.1. Different parameters to build the migration matrix

Ensemble tests were also used to verify that the migration matrix has a sufficiently fine binning and that the cut on the migration distance at 5 MeV is reasonable. One crucial result of this test was that extreme outliers in the individual migration histograms can cause a large bias in the amplitudes and in the yields of the background components. This was independent of the binning of the migration histograms or the cut on the migration distance. In addition to not using the outlier in building the migration matrix, various settings have been tested. Binnings below 100 bins for each dimension were fastly ruled out, the individual bins were not narrow enough to properly describe the structures in the Dalitz plot. For 125 bins and 150 bins different cuts on the migration distance were evaluated: 20, 10, 5 and 2 MeV. While for ensemble tests using the 125×125 bins migration matrix significant biases up to $0.4 \pm 0.06\sigma$ were found, the results using 150 bins were superior, see Table 9.2 in comparison with Table 9.3 for the difference between 150 and 125 bins. For both binnings it was found that the cuts at 5 or 2 MeV produced equally good results, with the 5 MeV slightly ahead (see Tables 9.4 and 9.5), while the other cuts introduced minor biases. Due to the higher computation time needed for the 2 MeV cut, as it selects almost all events as migrating, it was decided to stick with the 5 MeV cut. That also resembles the transition from the migration caused by the detector resolution to the migration caused by the misreconstruction of signal events.

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	500	0.000 25	0.012 ± 0.048	0.96 ± 0.041
$D_0^*(2400)^0$ Y	500	0.022	-0.13 ± 0.047	0.94 ± 0.041
$D_2^*(2462)^0$ X	500	0.0087	-0.17 ± 0.045	0.92 ± 0.038
$D_2^*(2462)^0$ Y	500	−0.0052	0.080 ± 0.052	1.1 ± 0.044
NR	500	0.0094	-0.043 ± 0.043	0.86 ± 0.041
D^*	500	0.28	0.20 ± 0.047	0.96 ± 0.040
Continuum BG	500	13 000	0.094 ± 0.040	0.84 ± 0.033
Combinatoric BG	500	22 000	-0.13 ± 0.035	0.74 ± 0.031
ρ X	-	0.078	-	-
ρ Y	-	0.41	-	-

Table 9.2.: Results of ensemble tests for the comparison of different binnings in the migration matrix: 150 bins, $d_{\text{migration}} > 5$ MeV, ($B^- \to D^0\pi^-\pi^0$). To be compared with Table 9.3.

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	300	0.000 25	-0.094 ± 0.053	0.85 ± 0.045
$D_0^*(2400)^0$ Y	300	0.022	-0.11 ± 0.066	1.0 ± 0.057
$D_2^*(2462)^0$ X	300	0.0087	-0.21 ± 0.059	0.94 ± 0.048
$D_2^*(2462)^0$ Y	300	−0.0052	-0.087 ± 0.070	1.1 ± 0.061
NR	300	0.0094	-0.24 ± 0.060	0.94 ± 0.056
D^*	300	0.28	0.27 ± 0.067	1.0 ± 0.059
Continuum BG	300	13 000	0.062 ± 0.054	0.87 ± 0.039
Combinatoric BG	300	22 000	-0.17 ± 0.048	0.78 ± 0.040
ρ X	-	0.078	-	-
ρ Y	-	0.41	-	-

Table 9.3.: Results of ensemble tests for the comparison of different binnings in the migration matrix: 125 bins, $d_{\text{migration}} > 5$ MeV, ($B^- \to D^0\pi^-\pi^0$). To be compared with Table 9.2.

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	600	0.012	-0.074 ± 0.039	0.85 ± 0.036
$D_0^*(2400)^0$ Y	600	−0.0099	0.079 ± 0.045	1.0 ± 0.038
$D_2^*(2462)^0$ X	600	−0.0019	-0.021 ± 0.044	0.96 ± 0.036
$D_2^*(2462)^0$ Y	600	0.010	-0.12 ± 0.042	0.93 ± 0.037
NR	600	0.0063	$-0.000\,37 \pm 0.045$	1.00 ± 0.037
D^*	600	0.28	-0.067 ± 0.043	0.94 ± 0.036
Continuum BG	600	13 000	0.038 ± 0.035	0.80 ± 0.033
Combinatoric BG	600	38 000	0.040 ± 0.027	0.64 ± 0.021
ρ X	-	0.078	-	-
ρ Y	-	0.41	-	-

Table 9.4.: Results of ensemble tests for the comparison between the limit on the migration distance: $d_{\text{migration}} > 5$ MeV, ($B^- \to D^0\pi^-\pi^0$). To be compared with Table 9.5.

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	600	0.012	-0.032 ± 0.041	0.89 ± 0.035
$D_0^*(2400)^0$ Y	600	-0.0099	0.23 ± 0.047	1.0 ± 0.039
$D_2^*(2462)^0$ X	600	-0.0019	-0.0078 ± 0.044	0.99 ± 0.039
$D_2^*(2462)^0$ Y	600	0.010	-0.061 ± 0.041	0.91 ± 0.037
NR	600	0.0063	-0.066 ± 0.041	0.92 ± 0.039
D^*	600	0.28	-0.061 ± 0.044	0.99 ± 0.038
Continuum BG	600	13 000	0.028 ± 0.038	0.86 ± 0.034
Combinatoric BG	600	38 000	0.12 ± 0.030	0.67 ± 0.024
ρ X	-	0.078	-	-
ρ Y	-	0.41	-	-

Table 9.5.: Results of ensemble tests for the comparison between the limit on the migration distance: $d_{\text{migration}} > 2$ MeV, ($B^- \to D^0\pi^-\pi^0$). To be compared with Table 9.4.

9.4.2. Results of the ensemble tests

To describe the amplitudes and phases of the different contributions, real and imaginary part $(X + iY)$ are used.

All ensemble test have been performed using 300 samples. The number of entries given in the tables give the number of successfully finished fits. Note that the amplitude and phase of the $\rho(770)$ were fixed to their input values, and therefore have no results in the tables of the pulls. In the fit to data they are fixed to amplitude 1 and phase 0 and all other parameters are determined relative to the ρ contribution. For some parameters, especially the yield of the combinatoric background, the width of the pull distribution is significantly lower than 1, indicating an overestimation of the error on the yield. As the calculation of asymmetric errors is very time consuming, and the impact on the results of the ensemble tests is small (compare Table 9.8 with Table 9.7), it will only be performed on the toy model, where the parameters are set to the results of the fit to data, see Section 11.7.

The results of ensemble tests, where the input parameters have been set to the expectation from previous measurements (see Section 1.4), are given below (Tables 9.6 and 9.9). Since no information about the relative phases is available, they are set to zero, but left floating in the fit, to make it comparable by having the same number of free parameters. From now on this sample will be labeled *PDG sample*. The results of this ensemble study are compared with the results of the study using the same generated toy samples, but with the $D_0^*(2400)^0$ mass and width fixed to their input values (Tables 9.7 and 9.10). It is visible, that adding two free parameters to the fit slightly increases the small deviations of the pull values, but the fit exhibits no significant bias in the measurement of the mass or width of the $D_0^*(2400)^0$.

The tables with results of additional different random models can be seen in the Appendices D.1 and D.2. As the floating of the mass and width of the $D_0^*(2400)^0$ renders it impossible to cache integrals in the fit, the cpu time increases significantly, and therefore the random models are fitted with the mass and width fixed to their input value. To determine possible biases and the related systematic errors, once the data is fitted, an ensemble test will be performed using the results of the fit to data. In this fit the mass and width of the $D_0^*(2400)^0$ is left floating.

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	283	0.16	0.14 ± 0.060	0.93 ± 0.049
$D_0^*(2400)^0$ Y	283	0	-0.056 ± 0.069	1.0 ± 0.065
$D_2^*(2462)^0$ X	283	0.11	-0.46 ± 0.060	0.90 ± 0.051
$D_2^*(2462)^0$ Y	283	0	0.31 ± 0.061	0.94 ± 0.061
NR	283	0.087	0.15 ± 0.068	1.0 ± 0.063
A3_Y_Pull	283	0	0.14 ± 0.071	1.0 ± 0.076
D^*	283	0.49	0.23 ± 0.061	0.93 ± 0.052
$M(D_0^*(2400)^0)$	283	2.3	0.077 ± 0.073	1.1 ± 0.063
$\Gamma(D_0^*(2400)^0)$	283	0.27	0.19 ± 0.060	0.90 ± 0.050
Continuum BG	283	14 000	0.011 ± 0.058	0.90 ± 0.055
Combinatoric BG	283	25 000	-0.15 ± 0.048	0.77 ± 0.041
ρ X	-	1	-	-
ρ Y	-	0	-	-

Table 9.6.: Results of ensemble tests for the PDG model of $B^- \to D^0\pi^-\pi^0$. The mass and width of the $D_0^*(2400)^0$ are floated and given in $\mathrm{GeV\,c^{-2}}$.

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	300	0.16	-0.0016 ± 0.069	1.00 ± 0.060
$D_0^*(2400)^0$ Y	300	0	0.021 ± 0.073	1.0 ± 0.056
$D_2^*(2462)^0$ X	300	0.11	-0.24 ± 0.060	0.95 ± 0.047
$D_2^*(2462)^0$ Y	300	0	0.083 ± 0.062	0.96 ± 0.062
NR	300	0.087	-0.060 ± 0.059	0.94 ± 0.053
A3_Y_Pull	300	0	0.14 ± 0.067	1.0 ± 0.066
D^*	300	0.49	0.23 ± 0.061	0.97 ± 0.050
Continuum BG	300	14 000	0.094 ± 0.052	0.84 ± 0.045
Combinatoric BG	300	25 000	-0.12 ± 0.050	0.78 ± 0.051
ρ X	-	1	-	-
ρ Y	-	0	-	-

Table 9.7.: Results of ensemble tests for the PDG model of $B^- \to D^0\pi^-\pi^0$. Here the mass and width of the $D_0^*(2400)^0$ are fixed, no significant deviation is observed in the comparison with Table 9.6, where they are floated.

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	300	0.16	-0.0075 ± 0.066	0.98 ± 0.053
$D_0^*(2400)^0$ Y	300	0	0.038 ± 0.075	1.0 ± 0.058
$D_2^*(2462)^0$ X	300	0.11	-0.23 ± 0.060	0.95 ± 0.045
$D_2^*(2462)^0$ Y	300	0	0.10 ± 0.063	0.97 ± 0.060
NR	300	0.087	-0.063 ± 0.058	0.93 ± 0.052
A3_Y_Pull	300	0	0.14 ± 0.063	0.97 ± 0.060
D^*	300	0.49	0.24 ± 0.061	0.97 ± 0.049
Continuum BG	300	14 000	0.098 ± 0.052	0.84 ± 0.046
Combinatoric BG	300	25 000	-0.14 ± 0.050	0.75 ± 0.051
ρ X	-	1	-	-
ρ Y	-	0	-	-

Table 9.8.: Results of ensemble tests for the PDG model of $B^- \to D^0\pi^-\pi^0$, where the asymmetric error calculation is used. The results are similar to Table 9.7, where the errors are only approximated assuming a parabolic behavior of the likelihood.

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	256	0.12	-0.12 ± 0.076	1.0 ± 0.079
$D_0^*(2400)^0$ Y	256	0	0.037 ± 0.075	1.0 ± 0.059
$D_2^*(2462)^0$ X	256	0.24	-0.22 ± 0.079	1.1 ± 0.065
$D_2^*(2462)^0$ Y	256	0	-0.16 ± 0.078	1.1 ± 0.074
NR	256	0.11	0.17 ± 0.065	0.89 ± 0.052
D^*	256	0.33	0.086 ± 0.068	1.00 ± 0.054
$M(D_0^*(2400)^0)$	256	2.3	0.13 ± 0.063	0.91 ± 0.053
$\Gamma(D_0^*(2400)^0)$	253	0.27	0.23 ± 0.048	0.71 ± 0.042
Continuum BG	256	14 000	0.062 ± 0.057	0.82 ± 0.045
Combinatoric BG	256	25 000	-0.16 ± 0.054	0.80 ± 0.044
ρ X	-	1	-	-
ρ Y	-	0	-	-

Table 9.9.: Results of ensemble tests for the PDG model of $B^0 \to D^-\pi^+\pi^0$. The mass and width of the $D_0^*(2400)^0$ are floated and given in $\mathrm{GeV\,c^{-2}}$.

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	300	0.12	-0.061 ± 0.064	0.99 ± 0.053
$D_0^*(2400)^0$ Y	300	0	-0.18 ± 0.075	1.1 ± 0.068
$D_2^*(2462)^0$ X	300	0.24	-0.21 ± 0.066	1.1 ± 0.055
$D_2^*(2462)^0$ Y	300	0	-0.11 ± 0.072	1.1 ± 0.064
NR	300	0.11	0.017 ± 0.057	0.93 ± 0.044
D^*	300	0.33	0.045 ± 0.061	0.97 ± 0.047
Continuum BG	300	14 000	0.094 ± 0.055	0.89 ± 0.044
Combinatoric BG	300	25 000	-0.24 ± 0.044	0.71 ± 0.037
ρ X	-	1	-	-
ρ Y	-	0	-	-

Table 9.10.: Results of ensemble tests for the PDG model of $B^0 \to D^-\pi^+\pi^0$. Here the mass and width of the $D_0^*(2400)^0$ are fixed, no significant deviation is observed in the comparison with Table 9.9, where they are floated.

10. Dalitz fit on recorded data

This analysis is a *blind* analysis, all detector effects and the fitting procedure have been thoroughly studied and validated using simulated events before performing the Dalitz fit on recorded data. The yield of the coherently added signal components and the yield of the D^* component, that is treated separately, are constrained to the signal yield extracted from the ΔE fit (Section 8.3) using a Gaussian constraint with the width corresponding to the total error on the signal yield. The two dimensional histogram used to describe the continuum background is extracted from off-resonance data (see Section 6.1).

10.1. First fit to recorded data

The first fit to data converged but it was obvious that especially in the area of low θ' values the fit was not able to describe data. This can be seen in Figs. 10.1 and 10.2. The fit relies on the simulation in the description of the combinatoric backgrounds and the fraction of misreconstructed signal. A detailed study of possible peaking backgrounds in the region where the discrepancies are observed is described in the following section, followed by the study of the effect of misreconstructed (migrating) signal focused on the D^* contribution.

10.2. Study of possible peaking backgrounds

A thorough study of the components of the generic background has been performed to rule out the possibility of a peaking background component at the regions of low θ', corresponding to low values of the invariant mass of the D meson and the neutral pion $M(D\pi^0)$. For this, all background contributions have been plotted in their invariant masses. Only dominant backgrounds that peak in the described region are then examined in more detail. Backgrounds that had a contribution in a wide range on the Dalitz plane cannot account for the observed discrepancy, as they would also alter the background shape all over the Dalitz plot. As described below, none of the peaking backgrounds can account for the disagreement between fit and data.

10.2.1. Peaking backgrounds in $B^- \to D^0\pi^-\pi^0$

There are two main backgrounds contributing in the region of low θ'. The first one being the decay $B^- \to D^{0*}(D^0\gamma)\pi^-$, which is very similar to the signal decay and is misreconstructed by adding one random photon to build a π^0. The second major contribution is built using decay

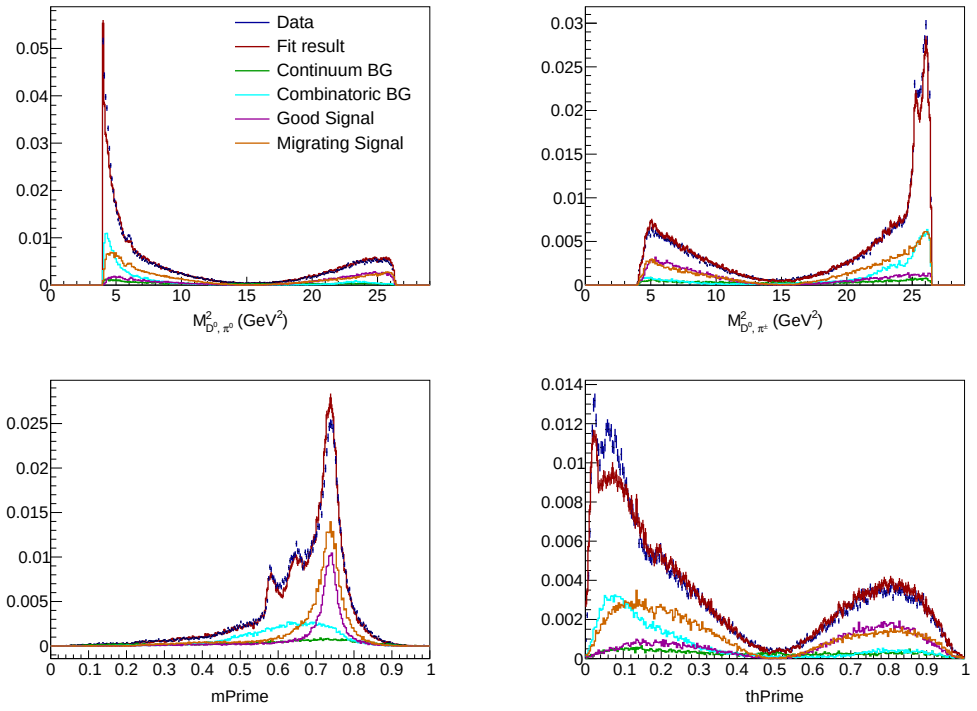


Figure 10.1.: Projections to the normal and squared Dalitz plot variables of the first fit to data, $B^- \to D^0\pi^-\pi^0$. The data is not described well by the fit, especially at low values of thPrime.

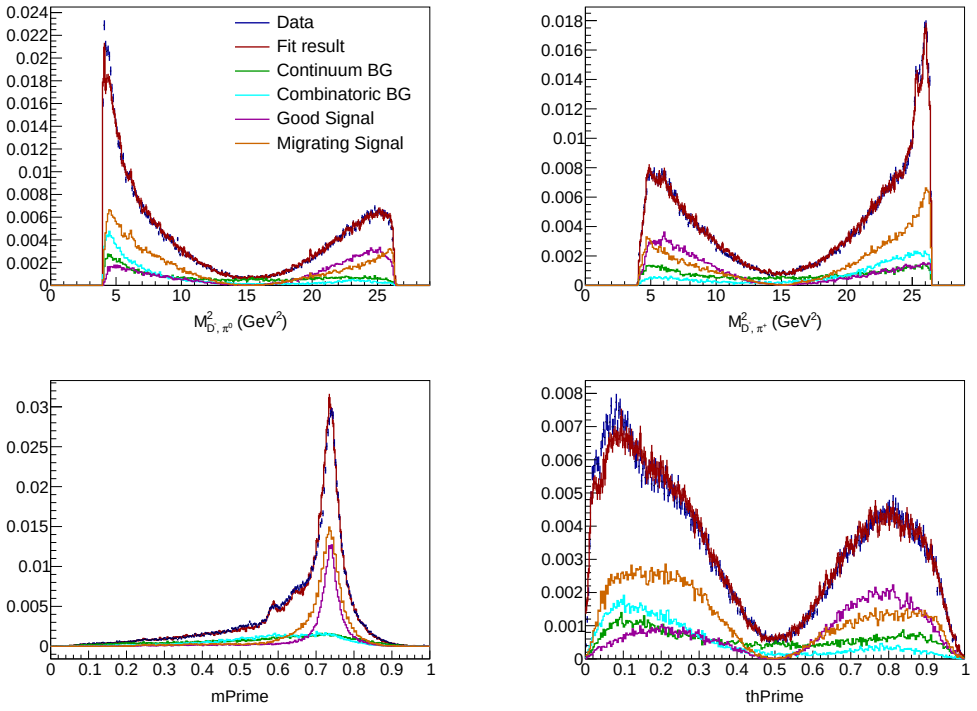


Figure 10.2.: Projections to the normal and squared Dalitz plot variables of the fit to data, $B^0 \to D^-\pi^+\pi^0$. A similar discrepancy is observed between the fit result and the data points at low values of thPrime.

products of the decay $\bar{B}^0 \to D^{*+}(D^0\pi^+)\pi^-$, where a slow charged pion is exchanged with a random π^0. The branching fraction used in the generic Monte Carlo of both background components is identical to the current world average given by the PDG, which has a small error, rendering the possibility of these backgrounds to account for the discrepancy impossible. All other combinatoric backgrounds are not peaking in the region in question, and the branching fractions of the dominant ones are also in agreement with PDG values.

10.2.2. Peaking backgrounds in $B^0 \to D^-\pi^+\pi^0$

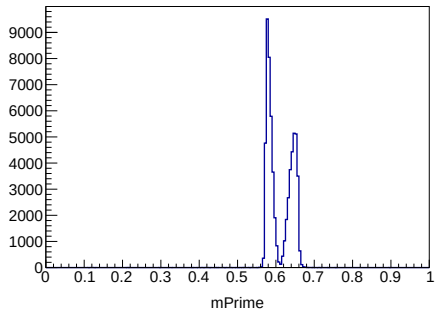
In contrast to the first channel there are four backgrounds in question here:

- $B^- \to D^{0*}(D^0\pi^0)\pi^-$ and $B^- \to D^{0*}(D^0\gamma)\pi^-$
- $\bar{B}^0 \to D^{*+}(D^0\pi^+)\pi^-$
- $B^+ \to \bar{D}^0(K^+\pi^-\pi 0)\pi^+$

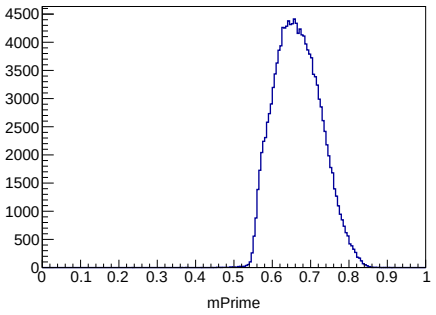
The branching fractions of all of them are in perfect agreement between the generic Monte Carlo and PDG and have sufficiently small errors to be ruled out as source of the discrepancy between fit and data.

10.3. Study of misreconstructed D^* signal

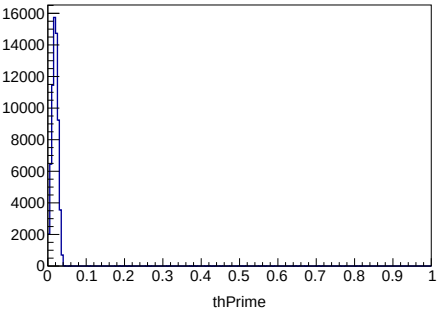
While the well reconstructed signal of the decay $B^- \to D^{*0}\pi^-$ or $B^0 \to D^{*-}\pi^+$ has a narrow width corresponding to the detector resolution, the misreconstruction of the slow neutral pions causes a large migration. The distribution of these misreconstructed events in comparison with the well reconstructed can be seen in Figs. 10.3 and 10.5 as well as in the scatter plots for both in Fig. 10.4 and Fig. 10.6. The misreconstructed D^* signal peaks in the region of interest and therefore is a possible explanation of the discrepancy observed in the first fit to data. To test this hypothesis, the D^* component in the fit is divided in two components representing well reconstructed and misreconstructed signal decays and the yields are left floating in the fit. This is done by utilizing the MC truth information of simulation. The resulting fits are now in good agreement with the data points (see next section). But the ratio of misreconstructed signal to well reconstructed signal is measured to be 3.6 (3.7) in comparison to 2.2 (2.3) in the generic Monte Carlo for $B^- \to D^0\pi^-\pi^0$ ($B^0 \to D^-\pi^+\pi^0$). The fraction of misreconstruction in the region of slow neutral pions is measured to be $\approx 50\%$ larger than the generic Monte Carlo suggests. This is backed up by the fact that with this large misreconstruction, the extracted branching fractions for the decay via the D^* mesons are in agreement with the world average. Especially in the first decay channel it was previously underestimated by $\approx 35\%$. The results presented in the following sections are thus obtained by leaving the ratio of well- and misreconstructed signal of the D^* floating in the fit. In the following this two components are called *good* and *bad* D^*.



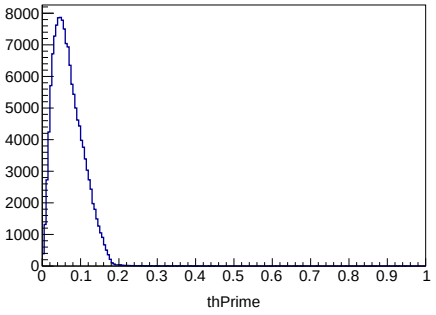
(a) Well reconstructed signal



(b) Misreconstructed signal

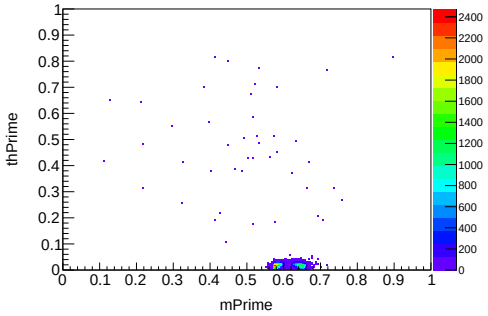


(c) Well reconstructed signal

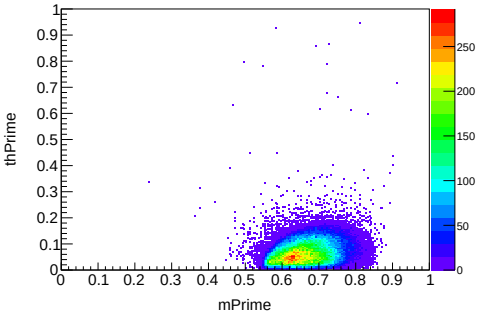


(d) Misreconstructed signal

Figure 10.3.: Comparison of well reconstructed and misreconstructed signal of the decay via an D^* meson in the projections to the squared Dalitz plot variables, $B^- \to D^0\pi^-\pi^0$. The misreconstructed D^* component peaks in the region were a discrepancy is observed in the fit to data.

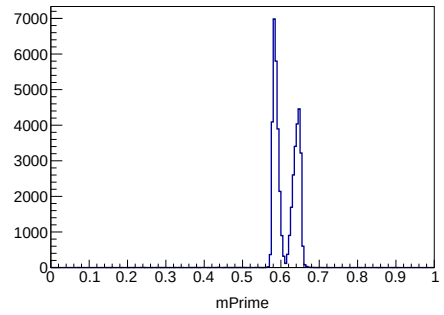


(a) Well reconstructed signal

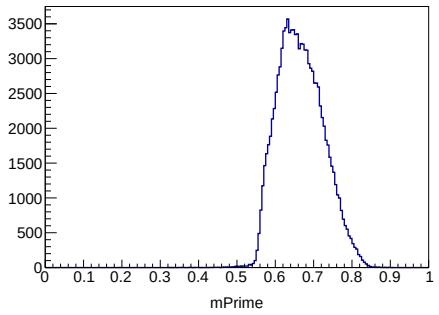


(b) Misreconstructed signal

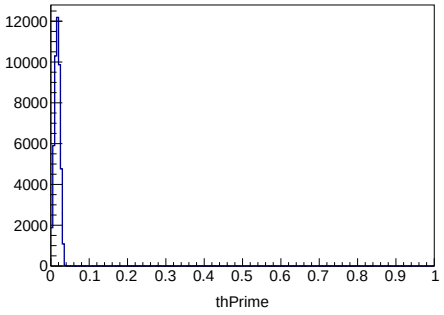
Figure 10.4.: Comparison of well reconstructed and misreconstructed signal of the decay via an D^* meson in scatter plot, $B^- \to D^0\pi^-\pi^0$. These histograms are then used in the fit to recorded data.



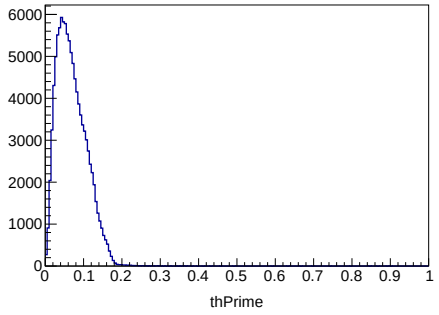
(a) Well reconstructed signal



(b) Misreconstructed signal

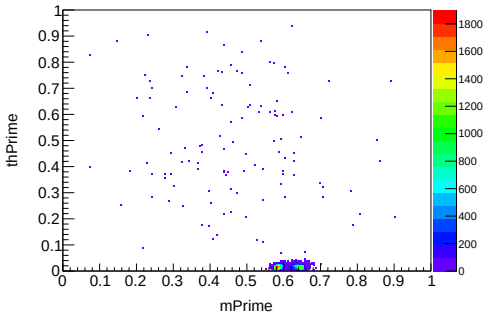


(c) Well reconstructed signal

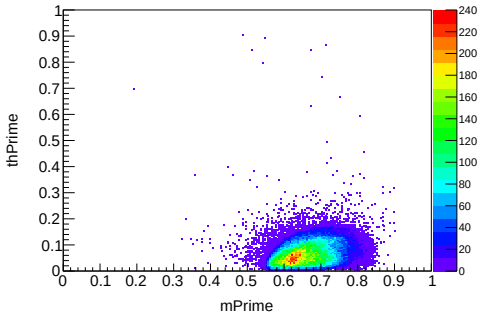


(d) Misreconstructed signal

Figure 10.5.: Comparison of well reconstructed and misreconstructed signal of the decay via an D^* meson in the projections to the squared Dalitz plot variables, $B^0 \to D^-\pi^+\pi^0$. The misreconstructed D^* component peaks in the region were a discrepancy is observed in the fit to data.



(a) Well reconstructed signal



(b) Misreconstructed signal

Figure 10.6.: Comparison of well reconstructed and misreconstructed signal of the decay via an D^* meson in the scatter plot, $B^0 \to D^-\pi^+\pi^0$. These histograms are then used in the fit to recorded data.

10.4. Results

With the good and bad D^* components left floating, the fit is in good agreement with the data points in the projections to the invariant masses or the variables of the squared Dalitz plot. Again, the projections along with the scatter plots are shown in Figs. 10.7 and 10.8 for the first channel $B^- \to D^0\pi^-\pi^0$ and in Figs. 10.9 and 10.10 for $B^0 \to D^-\pi^+\pi^0$. The fit results for all parameters left floating in the fit are listed in Tables 10.1 and 10.2. The correlation of the floating parameters can be seen in Figs. 10.11 and 10.12. As the real and imaginary part for the ρ contribution is fixed to 1 and 0, respectively, they are not listed in the table. In the projections in Fig. 10.9, a small discrepancy is observed between the fit result and data at high θ' values and accordingly at low $M^2(D\pi^+)$ or high $M^2(D\pi^0)$. This can be attributed to the fraction of migrating signal fixed to the MC value. To account for the influence of this fixed fraction, it is varied by $\pm 20\%$ to evaluate a systematic error in Section 11.6. In the comparison of the scatter plots in Figs. 10.8 and 10.10, the precense of the $D_0^*(2400)^0$ can be seen in the second channel at high values of θ', while in the first channel at the same position the $D_0^*(2400)^-$ is not observed.

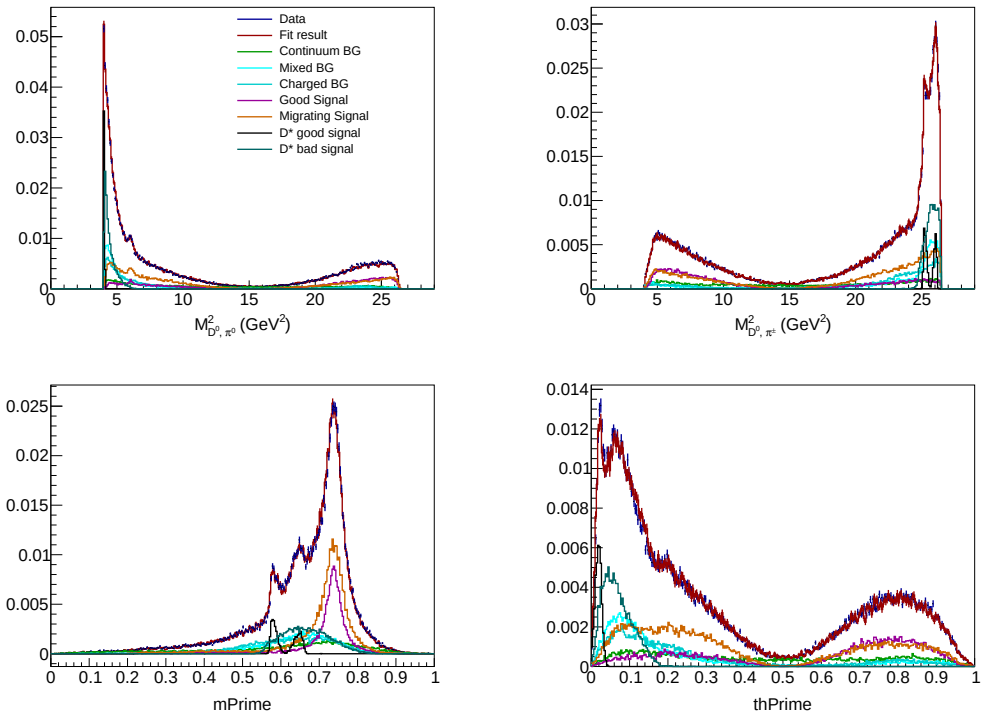


Figure 10.7.: Projections to the normal and squared Dalitz plot variables of the final fit to recorded data ($B^- \to D^0\pi^-\pi^0$). Both the good and bad D^* signal are left floating in the fit, significantly improving it in the region of low thPrime values compared to Fig. 10.1.

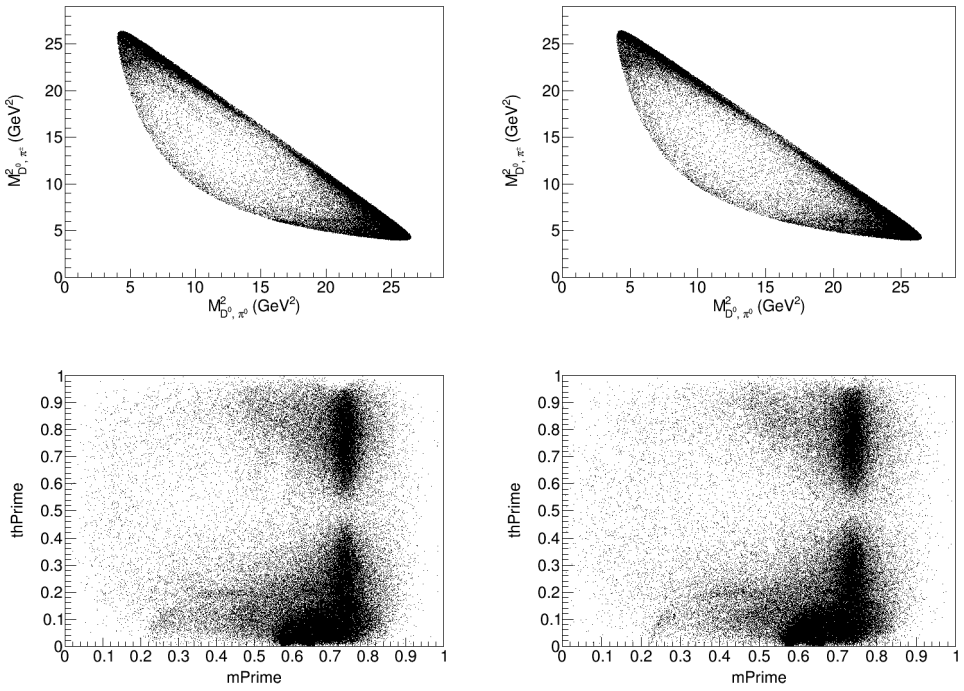


Figure 10.8.: Scatter plots of the final fit to recorded data in the normal (top) and squared (bottom) Dalitz plot showing the input sample (left) and the fit result (right) ($B^- \to D^0\pi^-\pi^0$).

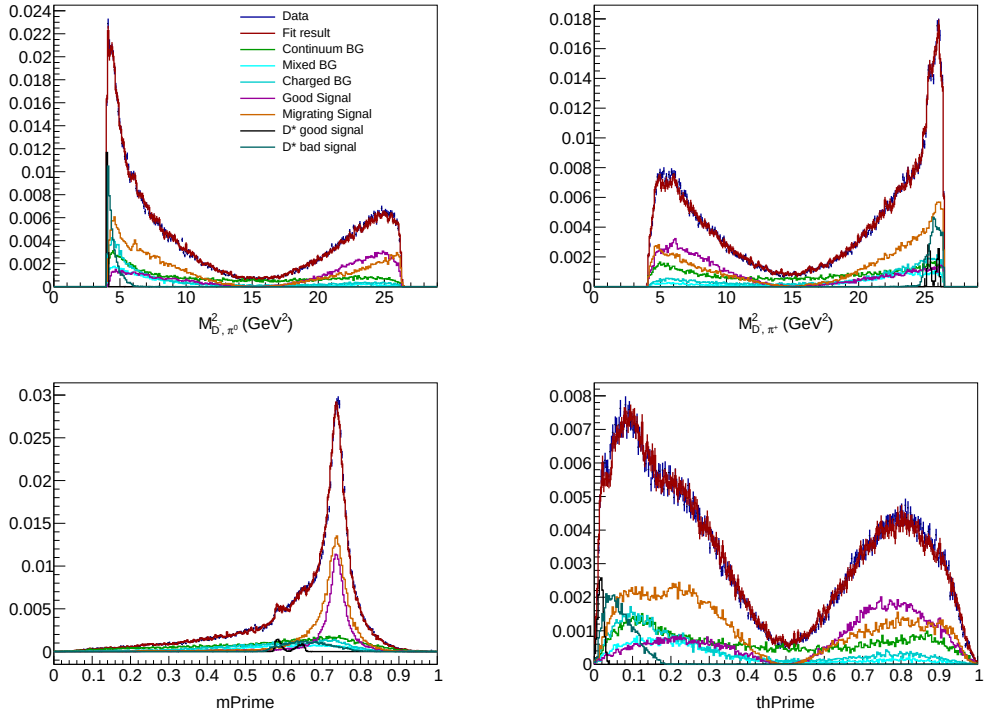


Figure 10.9.: Projections to the normal and squared Dalitz plot variables of the final fit to recorded data ($B^0 \to D^-\pi^+\pi^0$). Both the good and bad D^* signal are left floating in the fit, significantly improving it in the region of low thPrime values compared to Fig. 10.2.

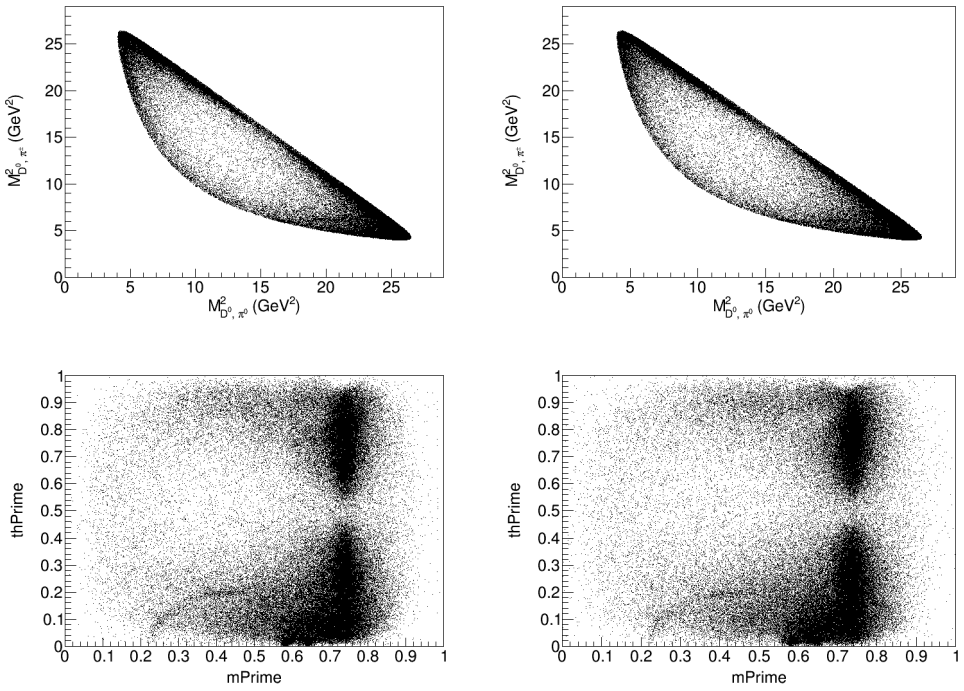


Figure 10.10.: Scatter plots of the final fit to recorded data in the normal (top) and squared (bottom) Dalitz plot showing the input sample (left) and the fit result (right) ($B^0 \to D^-\pi^+\pi^0$).

	Value	Error
$D_0^*(2400)^0$ X	0.228	0.012
$D_0^*(2400)^0$ Y	0.208	0.014
$D_2^*(2462)^0$ X	-0.144	0.004
$D_2^*(2462)^0$ Y	-0.006	0.015
NR X	0.054	0.01
Signal yield	45839	311
Continuum BG	12646	279
Mixed BG	14104	1114
Charged BG	13352	817
Bad D^*	12824	394
Good D^*	4075	78
$M(D_0^*(2400)^0)$	2.256	0.012
$\Gamma(D_0^*(2400)^0)$	0.354	0.03

Table 10.1.: Fit result of all floated parameters of the Dalitz fit on recorded data, $B^- \to D^0\pi^-\pi^0$

	Value	Error
$D_0^*(2400)^0$ X	-0.194	0.007
$D_0^*(2400)^0$ Y	-0.152	0.013
$D_2^*(2462)^0$ X	0.048	0.005
$D_2^*(2462)^0$ Y	-0.004	0.006
$D_0^*(2400)^-$ X	0.042	0.01
$D_0^*(2400)^-$ Y	0.018	0.011
$D_2^*(2462)^-$ X	0.123	0.005
$D_2^*(2462)^-$ Y	0.028	0.012
NR X	-0.045	0.01
Signal yield	64787	380
Continuum BG	23549	450
Mixed BG	8949	791
Charged BG	14764	951
Bad D^*	6442	166
Good D^*	1770	51
$M(D_0^*(2400)^0)$	2.245	0.007
$\Gamma(D_0^*(2400)^0)$	0.305	0.028

Table 10.2.: Fit result of all floated parameters of the the Dalitz fit on recorded data, $B^0 \to D^-\pi^+\pi^0$

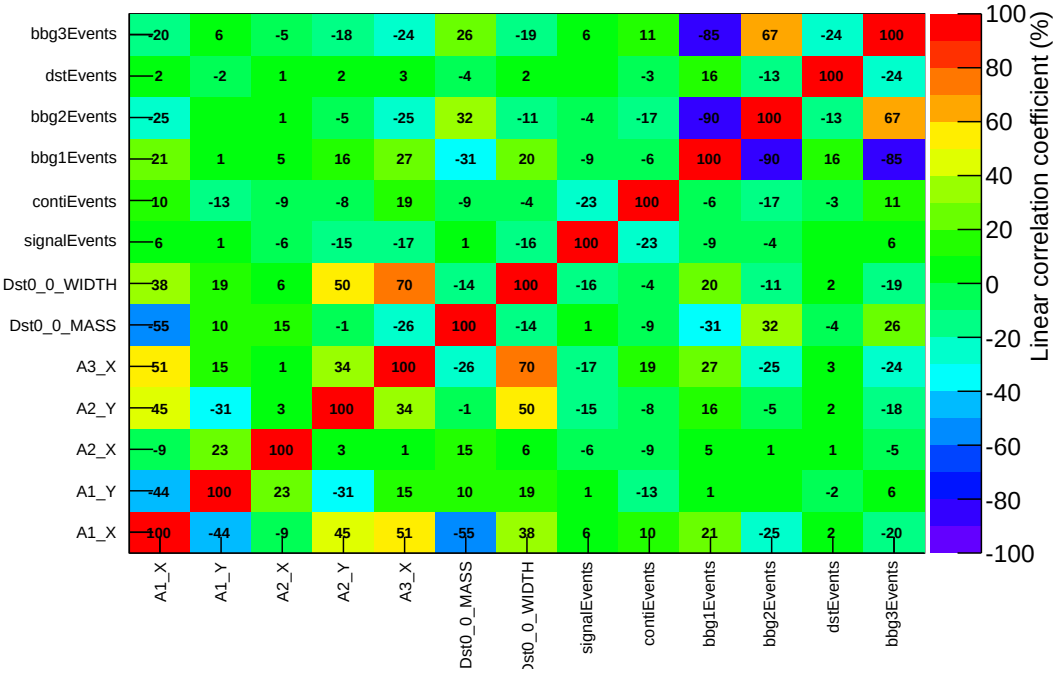


Figure 10.11.: Correlation of the fit parameters in the fit to $B^- \to D^0\pi^-\pi^0$. The labels are: A1: $D_0^*(2400)^0$, A2: $D_2^*(2462)^0$, A3: NR, bbg1 and bbg2: mixed and charged BG, conti: continuum BG, dst: good D^*, bbg3: bad D^*.

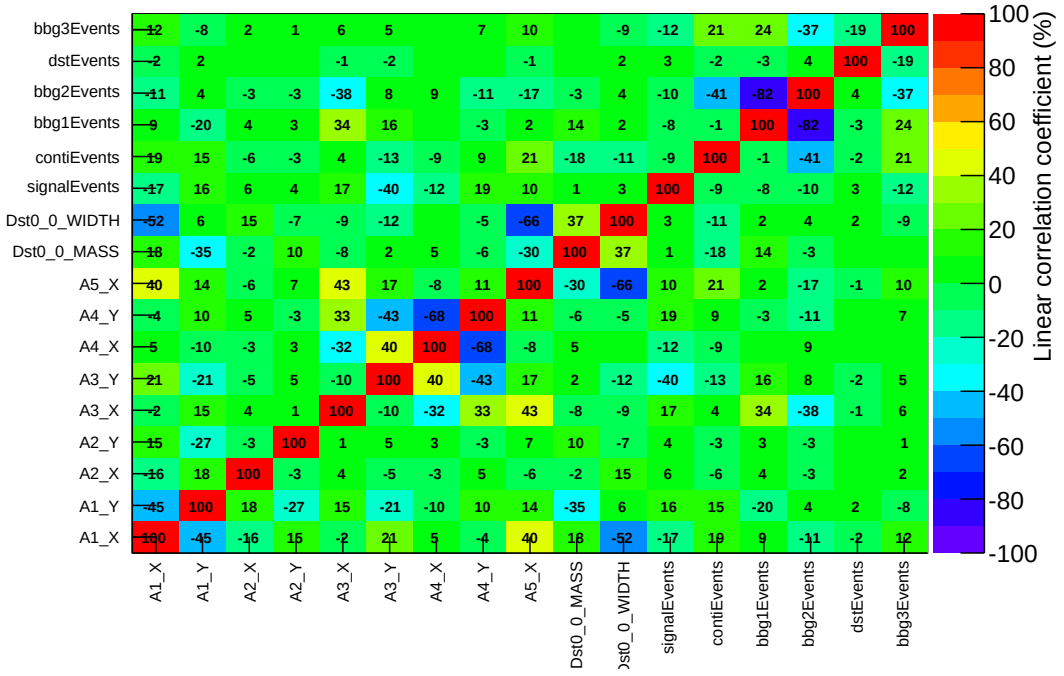


Figure 10.12.: Correlation of the fit parameters in the fit to $B^0 \to D^-\pi^+\pi^0$. The labels are: A1: $D_0^*(2400)^0$, A2: $D_2^*(2462)^0$, A3: $D_0^*(2400)^-$, A4: $D_2^*(2462)^-$, A5: NR, bbg1 and bbg2: mixed and charged BG, conti: continuum BG, dst: good D^*, bbg3: bad D^*.

10.4.1. Branching fractions

The branching fractions are directly proportional to the fit fractions obtained in the Dalitz fit. Due to interference effects, the overall Dalitz rate needs not be unity. In both channels there are no large interference effects visible in the scatter plots and accordingly the deviation from 1 is small: 1.03 for $B^- \to D^0\pi^-\pi^0$ and 0.99 for $B^0 \to D^-\pi^+\pi^0$. The fit fractions *ff* presented are already corrected for the varying efficiency over the Dalitz plot, and thus the branching fraction for a given decay i can be calculated using the average reconstruction efficiency ϵ_{rec} as

$$\mathcal{B}_i = \frac{ff_i \cdot N_{\text{signal}}}{N_{B\bar{B}} \times \epsilon_{rec} \times \epsilon_{PID} \times \mathcal{B}_{\text{daughter}}}. \tag{10.1}$$

The well known branching fractions for the decay of the D, ρ, π^0 and the D^* are already incorporated in the results (see Table 1.3 for the branching fractions used), while for the decay via the D^{**} mesons the branching fraction is given only for the decays to the final states of this analysis (Tables 10.3 and 10.4). In the case of the decay via the ρ and D^* the results are compared with the PDG values. For the well known decays $B^- \to D^{*0}\pi^-$ and $B^0 \to D^{*-}\pi^+$ the values are in perfect agreement. The branching fraction of $B^- \to D^0\rho^-$ is significantly lower than the PDG value, but the observation of a significantly lower number of signal events in the validation of the EKP full reconstruction tool [2] strengthens the result of this analysis. In the second decay mode $B^0 \to D^-\rho^+$ no discrepancy was observed before, and the values are still in agreement here. In the case of the neutral D^{**} in $B^- \to D^0\pi^-\pi^0$ and the charged D^{**} in $B^0 \to D^-\pi^+\pi^0$, there are previous measurements available, where they decay to $D^+\pi^-$ and $D^0\pi^-$, respectively. Using isospin conservation in strong interactions, it is possible to calculate comparable values using Clebsch-Gordan coefficients. The results of the previous measurements are given in parentheses in the table.

Component	Value (10^{-3})	Error (10^{-3})	rel. Error	PDG (10^{-3})
$\mathcal{B}(B^- \to D^0\rho^-)$	8.3	± 0.12	1.46%	13.4 ± 1.8
$\mathcal{B}(B^- \to D_0^*(2400)^0(D^0\pi^0)\pi^-)$	0.79	± 0.067	8.45%	(0.32 ± 0.07)
$\mathcal{B}(B^- \to D_2^*(2400)^0(D^0\pi^0)\pi^-)$	0.17	± 0.01	6.06%	(0.18 ± 0.02)
$\mathcal{B}(B^- \to D^0\pi^-\pi^0(NR))$	0.024	± 0.0086	35.16%	-
$\mathcal{B}(B^- \to D^{*0}\pi^-)$	5.1	± 0.065	1.3%	5.18 ± 0.26
Sum	11.7	± 0.042	0.4%	-

Table 10.3.: Branching fractions for $B^- \to D^0\pi^-\pi^0$. PDG values in parentheses are obtained using isospin considerations.

10.4.2. Mass and width measurement of the $D_0^*(2400)^0$

The Mass and Width of the $D_0^*(2400)^0$ are measured in both Dalitz fits. The results are listed in Table 10.5 along with the combination of both measurements. For the combination, the two results are averaged according to their total error, consisting of the quadratic sum of the statistical error and the total systematic error given in Tables 11.1 and 11.2. The evaluation of systematic errors is described in the next section Chapter 11. The final result including systematic errors is given in the conclusion, Chapter 12.

Component	Value (10^{-3})	Error (10^{-3})	rel. Error	PDG (10^{-3})
$\mathcal{B}(B^0 \to D^-\rho^+)$	6.5	$\pm$0.14	2.16%	7.8 ± 1.3
$\mathcal{B}(B^0 \to D_0^*(2400)^0(D^-\pi^+)\pi^0)$	0.39	$\pm$0.033	8.24%	-
$\mathcal{B}(B^0 \to D_2^*(2400)^0(D^-\pi^+)\pi^0)$	0.015	$\pm$0.003	19.80%	-
$\mathcal{B}(B^0 \to D_0^*(2400)^-(D^-\pi^0)\pi^+)$	0.014	$\pm$0.0061	44.96%	(0.3 ± 0.15)
$\mathcal{B}(B^0 \to D_2^*(2400)^-(D^-\pi^0)\pi^+)$	0.1	$\pm$0.01	9.70%	(0.11 ± 0.02)
$\mathcal{B}(B^0 \to D^-\pi^+\pi^0(NR))$	0.013	$\pm$0.0057	43.18%	-
$\mathcal{B}(B^0 \to D^{*-}\pi^+)$	2.7	$\pm$0.055	2.1%	2.76 ± 0.13
Sum	7.27	$\pm$0.034	0.5%	-

Table 10.4.: Branching fractions for $B^0 \to D^-\pi^+\pi^0$. PDG values in parentheses are obtained using isospin considerations.

	$M(D_0^*(2400)^0)$ in GeV c^{-2}		$\Gamma(D_0^*(2400)^0)$ in GeV c^{-2}	
	Value	Error	Value	Error
$B^- \to D^0\pi^-\pi^0$	2.256	0.012	0.354	0.029
$B^0 \to D^-\pi^+\pi^0$	2.245	0.007	0.305	0.028
Combined	2.248	0.006	0.321	0.011

Table 10.5.: Combination of the two individual measurements of the mass and width of the $D_0^*(2400)^0$.

10.4.3. Isospin analysis of the $B \to D\rho$ system

The amplitudes of the three different decays of a B meson to a D and ρ meson are related with each other because of isospin symmetry. The decays $B^- \to D^0\rho^-$, $B^0 \to D^-\rho^+$ and $B^0 \to \bar{D}^0\rho^0$ can be written as linear combinations of the isospin eigenstates A_I with $I = 1/2$ and $I = 3/2$:

$$\begin{aligned} A(\bar{D}^0\rho+) &= \sqrt{3}A_{3/2}, \\ A(D^-\rho+) &= \sqrt{1/3}A_{3/2} + \sqrt{2/3}A_{1/2}, \\ A(\bar{D}^0\rho0) &= \sqrt{2/3}A_{3/2} - \sqrt{1/3}A_{1/2}, \end{aligned} \tag{10.2}$$

or as a triangle relation

$$A(\bar{D}^0\rho+) = A(D^-\rho+) + \sqrt{2}A(\bar{D}^0\rho0). \tag{10.3}$$

In the heavy-quark limit, the factorization model predicts a difference in the strong phase

$$\delta_{D\rho} = \left|\delta_{1/2} - \delta_{3/2}\right| = \mathcal{O}(\Lambda_{QCD}/m_b) \tag{10.4}$$

and an amplitude ratio of

$$R_{D\rho} = \frac{|A_{1/2}|}{|\sqrt{2}|A_{3/2}|} = 1 + \mathcal{O}(\Lambda_{QCD}/m_b), \tag{10.5}$$

where m_b represents the b quark mass and Λ_{QCD} the QCD scale. Final-state interactions between the states $\bar{D}^0\rho^0$ and $D^-\rho^+$ could lead to a larger value of $\delta_{D\rho}$.

Two of the three related amplitudes have been measured in this analysis, the branching fraction of the third one is taken from a recent analysis of LHCb [32] (using their isobar results):

$$\mathcal{B}(B^0 \to \bar{D}^0\rho^0) = 0.321 \pm 0.023 \times 10^{-4}. \tag{10.6}$$

The amplitude ratio can thus be obtained by

$$R_{D\rho} = \sqrt{\frac{1}{2}} \left(\frac{3(\mathcal{B}(D^-\rho^+) + \mathcal{B}(\bar{D}^0\rho^0))}{\mathcal{B}(\bar{D}^0\rho^+)} \times \frac{\tau_{B^+}}{\tau_{B^0}} - 1 \right)^{1/2} \tag{10.7}$$

and

$$\cos\delta_{D\rho} = \frac{1}{4R_{D\rho}} \left(\frac{3(\mathcal{B}(D^-\rho^+) - 2\mathcal{B}(\bar{D}^0\rho^0))}{\mathcal{B}(\bar{D}^0\rho^+)} \times \frac{\tau_{B^+}}{\tau_{B^0}} + 1 \right). \tag{10.8}$$

A comparison of this values for the branching fractions from the previous measurement and using the branching fraction of this analysis is in Table 10.6. The previous numbers are in contrast to the numbers of the $D\pi$ system [33] and the theoretical expectations [34, 35], while the new results suggest that non-factorisable final-state interaction effects introduce a phase difference between the isopsin amplitudes of about $26°$.

Analysis	$R_{D\rho}$	$\cos\delta_{D\rho}$
Previous	0.69 ± 0.15	$0.984^{+0.113}_{-0.048}$
This	0.91 ± 0.07	0.90 ± 0.02

Table 10.6.: Results for $R_{D\rho}$ and $\cos\delta_{D\rho}$ obtained from previous measurements and the results of this analysis.

11. Estimation of Systematic uncertainties

The measurements performed in this analysis depend on various inputs, such as the description of the backgrounds and the migration of signal obtained from simulation, or the parameters of resonances. In addition, the calibration of the detector could also be a source of a systematic uncertainty. Systematic errors try to estimate the effect that, e.g., a wrong assumption or fixed shape from simulation could cause in the measured parameters. To estimate these effects, various tests have been performed. The overall results can be found in Tables 11.1 and 11.2. The columns contain the errors on the number of $B\bar{B}$ pairs ($N_{B\bar{B}}$), the uncertainty of the PID calibration (PID cal.), the efficiencies of charged tracks and neutral pions, the results of fluctuating the histograms used in the fit (Fluct.), the results of the variation of fixed resonance parameters (Reso.) and the influence of the fixed fraction of migrating signal events obtained on simulation (Migration). All this tests are described in more detail in the subsequent sections.

	$N_{B\bar{B}}$	PID cal.	Tracks/π^0	Fluct.	Reso.	Migration	Sum
ρ	-	-	-	0.03	0.05	0.39	0.39
$D_0^*(2400)^0$	-	-	-	0.50	1.57	0.52	1.73
$D_2^*(2462)^0$	-	-	-	0.38	0.81	4.69	4.77
NR	-	-	-	2.94	9.83	31.70	33.32
Signal yield	1.4	3.0	4.13	0.05	0.16	2.41	5.8
Good D^*	1.4	3.0	7.08	0.27	0.05	0.03	7.8
Bad D^*	1.4	3.0	7.08	0.67	0.19	0.44	7.86
$M(D_0^*(2400)^0)$	-	-	-	0.05	0.06	0.14	0.16
$\Gamma(D_0^*(2400)^0)$	-	-	-	0.76	2.23	8.93	9.24

Table 11.1.: Summary of systematic errors for $B^- \to D^0\pi^-\pi^0$ given in percent.

11.1. Number of $B\bar{B}$ pairs on recorded data

The total number of recorded $B-\bar{B}$ pairs is needed for the calculation of the absolute branching fraction of the measured channels. The number is measured by the Belle collaboration to be 771.581 million $B-\bar{B}$ pairs while the uncertainty of this number is 10.566 million or 1.4 % [36].

	$N_{B\bar{B}}$	PID cal.	Tracks/π^0	Fluct.	Reso.	Migration	Sum
ρ	-	-	-	0.04	0.26	0.22	0.34
$D_0^*(2400)^0$	-	-	-	0.60	2.29	3.98	4.63
$D_2^*(2462)^0$	-	-	-	1.86	3.16	6.44	7.41
$D_0^*(2400)^-$	-	-	-	4.38	51.83	63.02	81.71
$D_2^*(2462)^-$	-	-	-	0.40	6.19	2.04	6.53
NR	-	-	-	4.92	28.73	3.08	29.31
Signal yield	1.4	6.0	4.24	0.04	0.36	2.07	7.77
Good D^*	1.4	6.0	7.14	0.29	0.14	0.99	9.49
Bad D^*	1.4	6.0	7.14	0.40	0.35	6.89	11.69
$M(D_0^*(2400)^0)$	-	-	-	0.03	0.07	0.04	0.09
$\Gamma(D_0^*(2400)^0)$	-	-	-	1.35	2.05	3.14	3.99

Table 11.2.: Summary of systematic errors for $B^0 \to D^-\pi^+\pi^0$ given in percent

11.2. $K^\pm/\pi^\pm$ selection efficiency

The efficiency and the misidentification rate in the particle identification are momentum and polar angle dependent. The efficiencies obtained from MC simulation can have a systematic shift due to differences between the simulation and recorded data. Therefore a calibration of the PID efficiency has to be performed. The PID group of the Belle collaboration [37] offers the information and calculations to make this calibration. The efficiency ratio between recorded data and MC simulation depends mostly on the particle momentum in the lab frame and the polar angle between the direction of the particle and the z-axis. Because of this, the calibration has to be performed independently for each decay channel and charged particle type. For example a pion coming from a D meson has a different momentum distribution than a pion from a B meson. In addition, different intermediate states, or more specifically, their relative contributions, also influence the distributions of the momenta. The signal MC sample used to evaluate the calibration is simulated according to the world averages of the branching ratios from PDG. The calibration also depends on which version of the SVD has been used, thus the events from SVD1 and SVD2 have to be treated separately. The results are combined using the fraction of events measured with SVD1 which is 19.7 %. The calibration factor (Table 11.3) is then applied and the error of the calibration is taken into account in the systematic error calculation.

Decay Mode	Overall efficiency (%)	PID calibration	Error on calibration
$B^- \to D^0\pi^-\pi^0$	0.18	0.91	$\pm$0.03
$B^0 \to D^-\pi^+\pi^0$	0.14	0.88	$\pm$0.06

Table 11.3.: Overall efficiency of reconstruction and selection, and correction for PID cut efficiency.

11.3. Reconstruction efficiency of charged tracks, π^0 and slow π^0

The Belle collaboration also tested the consistency of efficiencies for charged tracks, neutral pions and slow pions. It was shown that the ratio of the efficiencies on recorded data and MC simulation does not differ significantly from one. However, the uncertainty of this ratio is taken as a systematic error for each track. The determination of the ratios and the corresponding errors is described in [38] for high P_t tracks and in [39] for slow charged and neutral pions. In Table 11.4 the systematic errors per track are listed, where the error is multiplied by the number of corresponding tracks for each decay channel to obtain the total systematic errors.

	error per track (%)
charged tracks	0.35
slow $\pi^{\pm}$	1.99
π^0	4.0
slow π^0	2.997

Table 11.4.: Errors on the consistency of efficiencies per track or neutral pion.

11.4. Statistical fluctuations in histograms

A part of the systematic uncertainties of the Dalitz fit originate in statistical fluctuations of the bin contents of the unsmoothed histograms used in the fit. To evaluate these errors, the entries in each bin are randomly fluctuated within their errors according to a Poisson distribution and the fit is repeated many times. The width of the distribution of the fit results of floated parameters is taken as systematic error. The histograms for the efficiency, the good D^* and the bad D^* are probed.

In the case of the efficiency, another systematic error is evaluated, which is due to the limited statistics used to calculate it. This can be estimated using

$$\sigma_\epsilon = \sqrt{\frac{N_{rec}(N_{gen} - N_{rec})}{N_{gen}^3}} \quad . \tag{11.1}$$

Due to the size of the generated samples used, this relative error is negligible: 3.4×10^{-5} for $B^- \to D^0\pi^-\pi^0$ and 3.0×10^{-5} for $B^0 \to D^-\pi^+\pi^0$. All the results are listed in Tables 11.5 and 11.6.

	Efficiency	bad D^*	good D^*	Sum
ρ	0.02	0.02	0.00	0.03
$D_0^*(2400)^0$	0.27	0.42	0.02	0.50
$D_2^*(2462)^0$	0.35	0.13	0.01	0.38
NR	1.93	2.20	0.25	2.94
Signal yield	0.05	0.02	0.00	0.05
Good D^*	0.02	0.23	0.14	0.27
Bad D^*	0.14	0.64	0.15	0.67
$M(D_0^*(2400)^0)$	0.03	0.05	0.00	0.05
$\Gamma(D_0^*(2400)^0)$	0.53	0.54	0.05	0.76

Table 11.5.: Errors on fit fractions, signal yields and $D_0^*(2400)^0$ parameters from fluctuating histograms used in the fit, given in percent ($B^- \to D^0\pi^-\pi^0$).

11.5. Fixed parameters of resonances

The width and masses of contributing resonances are fixed to their PDG value in the fit. To evaluate the systematic error, the masses and widths are varied according to the uncertainties on the PDG values. The variation of parameters of the well known $\rho(770)$ and the $D_2^*(2462)^0$ is small and thus the impact is also small. But especially in the $B^0 \to D^-\pi^+\pi^0$ decay, the large error of the width and mass of the $D_0^*(2400)^-$ cause a dominant systematic error in the fit fractions of the non resonant and $D_0^*(2400)^-$ contribution. However, both of these contributions are very small and distributed over a large area in the Dalitz plot, so making a significant change on one of them can lead to an exchange of events with the other, but

	Efficiency	bad D^*	good D^*	Sum
ρ	0.04	0.01	0.00	0.04
$D_0^*(2400)^0$	0.59	0.06	0.01	0.60
$D_2^*(2462)^0$	1.86	0.06	0.02	1.86
$D_0^*(2400)^-$	4.16	1.36	0.12	4.38
$D_2^*(2462)^-$	0.39	0.08	0.01	0.40
NR	4.78	1.17	0.11	4.92
Signal yield	0.04	0.02	0.00	0.04
Good D^*	0.06	0.24	0.16	0.29
Bad D^*	0.27	0.28	0.08	0.40
$M(D_0^*(2400)^0)$	0.03	0.00	0.00	0.03
$\Gamma(D_0^*(2400)^0)$	1.34	0.20	0.03	1.35

Table 11.6.: Errors on fit fractions, signal yields and $D_0^*(2400)^0$ parameters from fluctuating histograms used in the fit, given in percent ($B^0 \to D^-\pi^+\pi^0$).

having only small effects on the remaining contributions. Although the relative error is large, there are only a few events exchanged so this error does not propagate and influence the other measurements. The results are listed in Tables 11.7 and 11.8.

	ρ	$D_2^*(2462)^0$	Sum
ρ	0.05	0.02	0.05
$D_0^*(2400)^0$	1.55	0.21	1.57
$D_2^*(2462)^0$	0.47	0.66	0.81
NR	9.82	0.28	9.83
$M(D_0^*(2400)^0)$	0.04	0.04	0.06
$\Gamma(D_0^*(2400)^0)$	2.22	0.19	2.23

Table 11.7.: Errors from varying the resonance parameters mass and width, that are fixed in the Dalitz fit to recorded data, given in percent ($B^- \to D^0\pi^-\pi^0$).

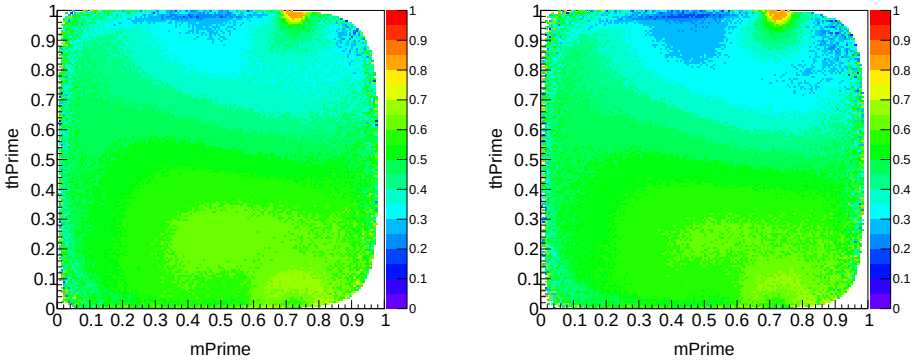
	ρ	$D_2^*(2462)^0$	$D_0^*(2400)^-$	$D_2^*(2462)^-$	Sum
ρ	0.03	0.01	0.25	0.08	0.26
$D_0^*(2400)^0$	1.73	0.14	1.34	0.68	2.29
$D_2^*(2462)^0$	2.11	1.48	1.68	0.75	3.16
$D_0^*(2400)^-$	6.32	0.17	47.56	19.61	51.83
$D_2^*(2462)^-$	0.46	0.03	4.90	3.76	6.19
NR	0.92	0.99	27.02	9.67	28.73
$M(D_0^*(2400)^0)$	0.04	0.02	0.04	0.03	0.07
$\Gamma(D_0^*(2400)^0)$	1.14	0.56	1.50	0.57	2.05

Table 11.8.: Errors from varying the resonance parameters mass and width, that are fixed in the Dalitz fit to recorded data, given in percent ($B^0 \to D^-\pi^+\pi^0$).

11.6. Fraction of migrating signal

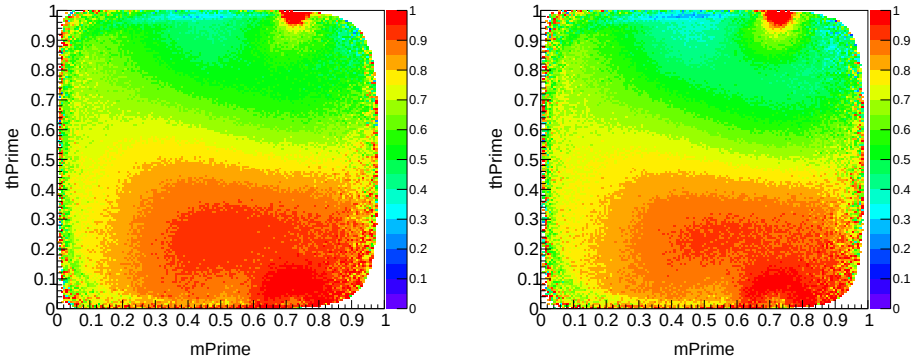
The fraction of migrating signal is determined on a generic MC sample. To account for a possible difference on recorded data, the fraction is varied by $\pm 20\%$ and the fit is repeated. The difference in the fit results is assigned as a systematic error. Note that, as described in

Section 7.2, signal is classified as migrating, if the migration distance is larger than $5\,\mathrm{MeV\,c^{-1}}$. This is fundamentally different compared to the use of the MC information, as done in the case of the D^* contribution in Section 10.3, as also good signal can significantly migrate. The impact of such a large variation in the fraction of migrating signal has consequently a large impact on measured parameters, especially in the case of small contributions, e.g., the non-resonant part in $B^- \to D^0\pi^-\pi^0$ or the $D_0^*(2400)^-$ in $B^0 \to D^-\pi^+\pi^0$. It is the dominant error for the measurement of the width of the $D_0^*(2400)^0$ in $B^- \to D^0\pi^-\pi^0$. The results are listed in Tables 11.9 and 11.10. The projections and scatter plots can be seen in Appendix E.3.



(a) $B^- \to D^0\pi^-\pi^0$ (b) $B^0 \to D^-\pi^+\pi^0$

Figure 11.1.: Fraction of migrating events (in the squared Dalitz plot), reduced by 20%.



(a) $B^- \to D^0\pi^-\pi^0$ (b) $B^0 \to D^-\pi^+\pi^0$

Figure 11.2.: Fraction of migrating events (in the squared Dalitz plot), increased by 20%.

	+20% Mig. Sig.	-20% Mig. Sig.	Sum
ρ	0.27	0.28	0.39
$D_0^*(2400)^0$	0.49	0.19	0.52
$D_2^*(2462)^0$	2.81	3.75	4.69
NR	19.33	25.13	31.70
Signal yield	1.66	1.74	2.41
Good D^*	0.03	0.02	0.03
Bad D^*	0.18	0.41	0.44
$M(D_0^*(2400)^0)$	0.12	0.07	0.14
$\Gamma(D_0^*(2400)^0)$	5.80	6.79	8.93

Table 11.9.: Errors from varying the fraction of migrating signal, given in percent ($B^- \to D^0\pi^-\pi^0$).

	+20% Mig. Sig.	-20% Mig. Sig.	Sum
ρ	0.16	0.16	0.22
$D_0^*(2400)^0$	3.23	2.32	3.98
$D_2^*(2462)^0$	5.51	3.33	6.44
$D_0^*(2400)^-$	34.46	52.77	63.02
$D_2^*(2462)^-$	0.97	1.79	2.04
NR	2.96	0.85	3.08
Signal yield	1.33	1.59	2.07
Good D^*	0.64	0.75	0.99
Bad D^*	4.62	5.11	6.89
$M(D_0^*(2400)^0)$	0.00	0.04	0.04
$\Gamma(D_0^*(2400)^0)$	2.62	1.73	3.14

Table 11.10.: Errors from varying the fraction of migrating signal, given in percent ($B^0 \to D^-\pi^+\pi^0$).

11.7. Fit bias

The ensemble tests described in Section 9.4 were used to check if the fit model exhibits biases prior to unblinding the data. While they show that the fit model works for a range of parameters around the world average, the errors and biases are evaluated here for the model of the results of the fit to recorded data. The results in Tables 11.11 and 11.12 show no significant bias and no underestimation of the error on the floated parameters. However, due to the constraint of the signal yield, good D^* yield and the bad D^* yield to the ΔE fit result, the variation of these parameters is limited. Therefore the width of the pull is significantly lower than one for the signal yield and the good D^* yield. In addition, in the case of $B^- \to D^0\pi^-\pi^0$, the fit result for the combined signal yield is at the lower edge of the constraint window. As for the generation of the samples only the central value is used, and fluctuated according to its error without obeying the constraint, this results in an asymmetric distribution in the pull and a mean different from zero. In the second decay channel $B^0 \to D^-\pi^+\pi^0$, where the result for the signal yield lies close to the central value of the constraint, this behavior is not observed. The plots of the individual pull distribution can be found in Appendices E.1 and E.2.

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	300	0.23	-0.041 ± 0.072	1.1 ± 0.064
$D_0^*(2400)^0$ Y	300	0.21	0.11 ± 0.066	0.98 ± 0.052
$D_2^*(2462)^0$ X	300	-0.14	0.095 ± 0.066	1.0 ± 0.053
$D_2^*(2462)^0$ Y	300	-0.0056	-0.094 ± 0.061	0.97 ± 0.050
NR	300	0.054	-0.088 ± 0.066	1.0 ± 0.055
$M(D_0^*(2400)^0)$	300	2.256	0.053 ± 0.061	0.98 ± 0.053
$\Gamma(D_0^*(2400)^0)$	300	0.354	-0.098 ± 0.062	0.99 ± 0.055
Signal yield	300	45838	0.44 ± 0.039	0.59 ± 0.032
Continuum BG	300	12646	-0.11 ± 0.065	0.97 ± 0.063
Mixed BG	300	14107	-0.17 ± 0.062	0.98 ± 0.048
Charged BG	300	13349	0.027 ± 0.064	1.0 ± 0.048
Good D^*	300	4074	0.058 ± 0.035	0.56 ± 0.027
Bad D^*	300	12823	0.080 ± 0.064	0.98 ± 0.051

Table 11.11.: Results of the ensemble studies at the central values obtained from the fit to recorded data ($B^- \to D^0\pi^-\pi^0$).

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	299	-0.19	0.018 ± 0.067	1.0 ± 0.059
$D_0^*(2400)^0$ Y	299	-0.15	-0.025 ± 0.061	0.96 ± 0.052
$D_2^*(2462)^0$ X	299	0.048	-0.028 ± 0.060	0.98 ± 0.051
$D_2^*(2462)^0$ Y	299	-0.0037	-0.045 ± 0.060	0.95 ± 0.051
$D_0^*(2400)^0$ X	299	0.042	0.16 ± 0.060	0.96 ± 0.055
$D_0^*(2400)^0$ Y	299	0.018	-0.073 ± 0.062	0.98 ± 0.049
$D_2^*(2462)^0$ X	299	0.12	-0.12 ± 0.063	0.97 ± 0.056
$D_2^*(2462)^0$ Y	299	0.028	0.013 ± 0.063	1.0 ± 0.054
NR	299	-0.045	-0.090 ± 0.066	1.0 ± 0.056
$M(D_0^*(2400)^0)$	299	2.245	-0.080 ± 0.071	1.1 ± 0.061
$\Gamma(D_0^*(2400)^0)$	299	0.305	0.0010 ± 0.068	1.0 ± 0.060
Signal yield	299	64787	0.12 ± 0.049	0.76 ± 0.044
Continuum BG	299	23548	-0.016 ± 0.059	0.93 ± 0.048
Mixed BG	299	8948	0.039 ± 0.060	0.92 ± 0.052
Charged BG	299	14764	-0.060 ± 0.058	0.95 ± 0.049
Good D^*	299	1770	0.043 ± 0.029	0.49 ± 0.024
Bad D^*	299	6442	-0.022 ± 0.055	0.87 ± 0.050

Table 11.12.: Results of the ensemble studies at the central values obtained from the fit to recorded data ($B^0 \to D^-\pi^+\pi^0$).

12. Conclusion

In summary, two independent Dalitz plot analysis have been performed of the decay modes $B^- \to D^0\pi^-\pi^0$ and $B^0 \to D^-\pi^+\pi^0$, being the first coherent analysis of both. It is performed on the full dataset recorded by the Belle detector at the $\Upsilon(4S)$ resonance that consists of 772×10^6 $B\bar{B}$ pairs. The presence of a neutral pion was a big challenge in this analysis, as it is produced almost at rest in parts of the Dalitz plot and therefore easy to misreconstruct. A sophisticated approach utilizing a migration matrix built from 22500 2D histograms, each containing 22500 bins, was implemented to track down the shift in the invariant masses in the case of a misreconstructed signal candidate. This allowed for the measurement to be performed on the whole Dalitz plane with a high efficiency of 18% and 14%, respectively. Introducing cuts to reduce the contribution of misreconstructed signal would have rendered a measurement of the properties of the broad $D_0^*(2400)^0$ impossible.

The branching fractions for all involved decays have been measured and are listed in Tables 12.1 and 12.2. The results for $B^- \to D^{*0}\pi^-$ and $B^0 \to D^{*-}\pi^+$ are in perfect agreement with previous measurements and hence confirm the result, that the fraction of misreconstructed slow neutral pions is underestimated by $\approx 40\%$ in the simulation of the Belle detector. The branching fractions measured in the decays via the ρ meson are far superior compared to the CLEO measurement from 1994, where especially in the first decay the central value changed remarkably. The previous measurement claimed that an additional 0.5% of all B^+ decays proceed via the $D\rho$ two-body decay, which is not the case. The errors in both measurements are dominated by the systematic errors originating in the PID calibration and the uncertainty on the efficiency of charged tracks and neutral pions.

In the case of the two non-resonant decays, the branching fractions have been measured for the first time, and are found to be very small compared to the dominant decay via the ρ. The branching fractions of decays via D^{**} mesons have been measured for the first time in the listed D^{**} decay modes. As described in Section 1.3, the decay of the B^- via the $D_0^*(2400)^0$ has similar tree-level decay diagrams as $B^0 \to D_0^*(2400)^0(D^-\pi^+)\pi^0$ and $B^0 \to D_0^*(2400)^-(D^-\pi^0)\pi^+$. Considering the different decays of the $D_0^*(2400)^0$, the results for the branching fractions of $B^- \to D_0^*(2400)^0(D^0\pi^0)\pi^-$ and of $B^0 \to D_0^*(2400)^-(D^-\pi^0)\pi^+$ plus $B^0 \to D_0^*(2400)^0(D^-\pi^+)\pi^0$ are similar.

The measurements of the properties of the broad $D_0^*(2400)^0$ state is difficult, but with the coherent description and the treatment of migrating signal, I was able to make two individual measurements of the mass and width of the $D_0^*(2400)^0$, where the results are in agreement

Mode	Value	Stat. Error	Syst. Error	PDG
$\mathcal{B}(B^- \to D^0\rho^-)$	8.3	± 0.12	± 0.48	13.4 ± 1.8
$\mathcal{B}(B^- \to D_0^*(2400)^0(D^0\pi^0)\pi^-)$	0.79	± 0.067	± 0.048	(0.32 ± 0.07)
$\mathcal{B}(B^- \to D_2^*(2400)^0(D^0\pi^0)\pi^-)$	0.17	± 0.01	± 0.013	(0.18 ± 0.02)
$\mathcal{B}(B^- \to D^0\pi^-\pi^0(NR))$	0.024	± 0.0086	± 0.0081	-
$\mathcal{B}(B^- \to D^{*0}\pi^-)$	5.1	± 0.065	± 0.4	5.18 ± 0.26
Sum ($B^- \to D^0\pi^-\pi^0$)	11.7	± 0.4	± 0.743	

Table 12.1.: Final results for branching fraction measurements in the decay mode $B^- \to D^0\pi^-\pi^0$.

Mode	Value	Stat. Error	Syst. Error	PDG
$\mathcal{B}(B^0 \to D^-\rho^+)$	6.5	± 0.14	± 0.51	7.8 ± 1.3
$\mathcal{B}(B^0 \to D_0^*(2400)^0(D^-\pi^+)\pi^0)$	0.39	± 0.033	± 0.035	-
$\mathcal{B}(B^0 \to D_2^*(2400)^0(D^-\pi^+)\pi^0)$	0.015	± 0.003	± 0.0016	-
$\mathcal{B}(B^0 \to D_0^*(2400)^-(D^-\pi^0)\pi^+)$	0.014	± 0.006	± 0.011	(0.3 ± 0.15)
$\mathcal{B}(B^0 \to D_2^*(2400)^-(D^-\pi^0)\pi^+)$	0.10	± 0.01	± 0.01	(0.11 ± 0.02)
$\mathcal{B}(B^0 \to D^-\pi^+\pi^0(NR))$	0.013	± 0.0057	± 0.0039	-
$\mathcal{B}(B^0 \to D^{*-}\pi^+)$	2.7	± 0.055	± 0.303	2.76 ± 0.13
Sum ($B^0 \to D^-\pi^+\pi^0$)	7.27	± 0.034	± 0.593	-

Table 12.2.: Final results for branching fraction measurements in the decay mode $B^0 \to D^-\pi^+\pi^0$.

with each other. The combined result for the mass is

$$M(D_0^*(2400)^0) = 2.248 \pm 0.006 \pm 0.002 \quad \mathrm{GeV\,c^{-2}}, \tag{12.1}$$

where the first error given is the statistical error followed by the systematic error. Compared with the previous world average of 2.318 ± 0.029, my measurement significantly improved the error, by reducing it by a factor of four. The result of the measurement of the width is

$$\Gamma(D_0^*(2400)^0) = 0.321 \pm 0.020 \pm 0.011 \quad \mathrm{GeV\,c^{-2}}, \tag{12.2}$$

also superseding the previous world average of 0.267 ± 0.040 by almost halving the combined error.

Using the two obtained branching fractions for $B^- \to D^0\rho^-$ and $B^0 \to D^-\rho^+$, the amplitude ratio $R_{D\rho}$ and strong phase difference $\cos\delta_{D\rho}$, defined in Section 10.4.3, can be calculated:

$$\begin{aligned} R_{D\rho} &= 0.91 \pm 0.07 \\ \cos\delta_{D\rho} &= 0.90 \pm 0.02 \end{aligned} \tag{12.3}$$

In contrast to using the results of previous measurements, they are now in agreement with results from the $D\pi$ system and theoretical expectations [34, 35]. The results confirm a phase difference between the isospin amplitudes introduced by non-factorizable final-state interaction effects between the $\bar{D}^0\rho^0$ and $D^-\rho^+$ states.

Looking forward, the results of this analysis will be published in the autumn of 2015 as a peer-reviewed article. It will influence future analysis at particle experiments by helping to provide simulations that are much closer to the recorded data. For the successor of Belle, Belle II, an improved simulation will not only enhance the quality of blind analyses but also strengthen the Full Event Interpretation used in many analyses of decays including neutrinos.

Bibliography

[1] M.S. Alam et al. Exclusive hadronic B decays to charm and charmonium final states. *Phys.Rev.*, D50:43–68, 1994.

[2] M. Feindt, F. Keller, M. Kreps, T. Kuhr, S. Neubauer, D. Zander, and A. Zupanc. A hierarchical neurobayes-based algorithm for full reconstruction of B mesons at B factories. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 654(1):432 – 440, 2011.

[3] B. Kronenbitter et al. Measurement of the branching fraction of $B^+ \to \tau^+\nu_\tau$ decays with the semileptonic tagging method. 2015.

[4] Matthias Huschle. *Measurement of the branching ratio of $B \to D^{(*)}\tau\nu_\tau$ relative to $B \to D^{(*)}\ell\nu_\ell$ decays with hadronic tagging at Belle*. Dr., Karlsruher Institut für Technologie (KIT), 2015. Karlsruher Institut für Technologie (KIT), Diss., 2015.

[5] Oksana Lutz. Search for $B \to h^{(*)}\nu\overline{\nu}$ decays at Belle and development of track finding for Belle II.

[6] A. Kuzmin et al. Study of anti-B0 to D0 pi+ pi- decays. *Phys.Rev.*, D76:012006, 2007.

[7] K. Nakamura et al. . *(Particle Data Group), Journal of Physics*, G37, 075021, 2010 and 2011 partial update for the 2012 edition.

[8] A. Abashian et al. The belle detector. *Nucl. Instrum. Meth.*, A79:117–232, 2002.

[9] Inner tracker upgrade task force. Conceptual design report on the belle inner tracker upgrade. *Internal Note: Belle Note*, 435, June 2001.

[10] R.H. Dalitz. CXII. on the analysis of τ-meson data and the nature of the τ-meson. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 44(357):1068–1080, oct 1953.

[11] A.J. Bevan et al. The Physics of the B Factories. *Eur.Phys.J.*, C74(11):3026, 2014.

[12] G. BREIT. Theoretical nuclear physics. john m. blatt and victor f. weisskopf. new york: Wiley, london: Chapman & hall, 1952. 864 pp. illus. $12.50. *Science*, 117(3042):419–420, apr 1953.

[13] G. J. Gounaris and J. J. Sakurai. Finite-width corrections to the vector-meson-dominance prediction for $\rho \to e^+e^-$. *Phys. Rev. Lett.*, 21:244–247, Jul 1968.

[14] A. Garmash and K. Abe. Dalitz analysis of the three-body charmless decays $B^+ \to K^+\pi^+\pi^-$ and $B^+ \to K^+K^+K^-$. *Phys. Rev. D*, 71:092003, May 2005.

[15] Bernard Aubert et al. An amplitude analysis of the decay $B^\pm \to \pi^\pm\pi^\pm\pi^\mp$. *Phys.Rev.*, D72:052002, 2005.

[16] Wikipedia. Kernel density estimation — Wikipedia, the free encyclopedia, 2015. [Online; accessed 20-January-2015].

[17] Wouter Verkerke. The roofit toolkit for data modeling. June 2003. arXiv:physics/0306116v1.

[18] Volker Blobel and Erich Lohrmann. *Statistische und numerische Methoden der Datenanalyse*. Teubner Verlag, 1998.

[19] Tom Latham J.J. Back. Laura++, a maximum likelihood fitting package for dalitz plot analysis, 2015.

[20] F. James and M. Roos. Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations. *Comput.Phys.Commun.*, 10:343–367, 1975.

[21] G. C. Fox and S. Wolfram. Observables for the analysis of event shapes in e^+e^- annihilation and other processes. *Physical Review Letters*, 41:1581–1585, 1978.

[22] D. M. Asner, M. Athanas, and Bliss. Search for exclusive charmless hadronic *B* decays. *Phys. Rev. D*, 53:1039–1050, Feb 1996.

[23] M. Feindt and U. Kerzel. The neurobayes neural network package. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 559(1):190 – 194, 2006.

[24] S. Agostinelli et al. Geant4: A simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.

[25] D. J. Lange et al. The evtgen particle decay simulation package. *Nucl. Instrum. Meth.*, A462:152–155, 2001.

[26] Torbjorn Sjostrand. PYTHIA 5.7 and JETSET 7.4: Physics and manual. 1995.

[27] R. Brun, F. Bruyant, M. Maire, A.C. McPherson, and P. Zanarini. GEANT3. 1987.

[28] P. del Amo Sanchez et al. Dalitz-plot Analysis of $B^0 -> \bar{D}^0\pi^+\pi^-$. *PoS*, ICHEP2010:250, 2010.

[29] Roel Aaij et al. Amplitude analysis of $B^0 \to \bar{D}^0 K^+\pi^-$ decays. 2015.

[30] Rene Brun and Fons Rademakers. Root - an object oriented data analysis framework. *Proceedings AIHENP'96 Workshop, Lausanne, Nucl. Inst. & Meth.*, A389:81–86, 1997. `http://root.cern.ch`.

[31] M. Prim et al. Angular analysis of $B^0 \to \phi K^*$ decays and search for CP violation at Belle. *Phys.Rev.*, D88(7):072004, 2013.

[32] Roel Aaij et al. Dalitz plot analysis of $B^0 \to \overline{D}^0\pi^+\pi^-$ decays. 2015.

[33] J.P. Lees et al. Branching Fraction Measurements of the Color-Suppressed Decays $\bar{B}^0 \to D^{(*)0}\pi^0$, $D^{(*)0}\eta$, $D^{(*)0}\omega$, and $D^{(*)0}\eta'$ and Measurement of the Polarization in the Decay $\bar{B}^0 \to D^{*0}\omega$. *Phys.Rev.*, D84(3):112007, 2011.

[34] Matthias Neubert and Alexey A. Petrov. Comments on color suppressed hadronic B decays. *Phys.Lett.*, B519:50–56, 2001.

[35] Jonathan L. Rosner. On large final state phases in heavy meson decays. *Phys.Rev.*, D60:074029, 1999.

[36] Number of $B\bar{B}$ events in data. `http://belle.kek.jp/secured/nbb/nbb.html`.

[37] S. Nishida. Study of kaon and pion identification using inclusive D^* sample. *Internal Note: Belle Note*, 779, Jan. 2005.

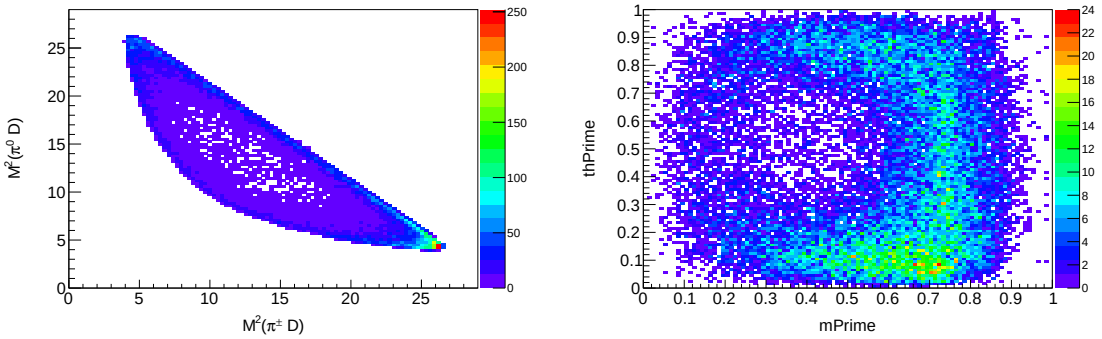
[38] Bipul Bhuyan. High p_t tracking efficiency using partially reconstructed D^* decays. *Internal Note: Belle Note*, 1165, Dec. 2010.

[39] Wolfgang Dungel. Systematic investigation of the reconstruction efficiency of low momentum $\pi^\pm$ and π^0. *Internal Note: Belle Note*, 1176, Jan. 2011.

Appendix

A. Backgrounds

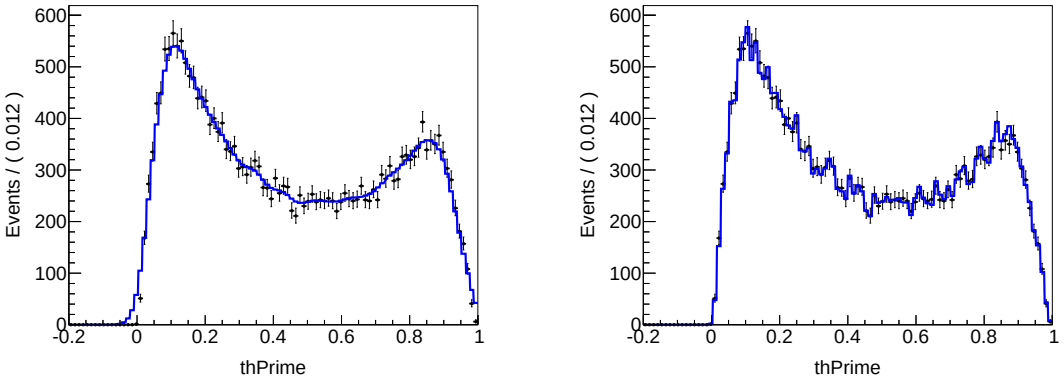
A.1. Continuum background



(a) Normal Dalitz plot

(b) Squared Dalitz plot

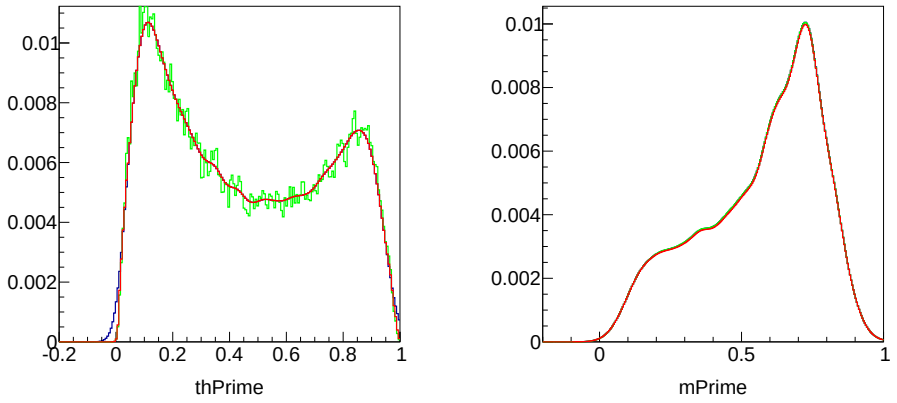
Figure A.1.: Comparison of the continuum background from M_{bc} sideband (selected with MC truth variables) of stream 0 of generic Monte Carlo in the normal and squared Dalitz plot representation ($B^0 \to D^- \pi^+ \pi^0$).



(a) High-bandwidth KDE

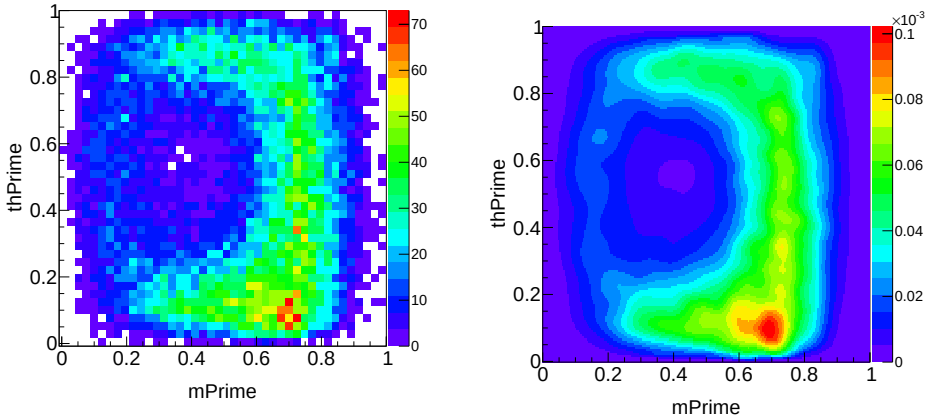
(b) Low-bandwidth KDE

Figure A.2.: Comparison of the High- and Low-bandwidth KDE in the projection to θ' in the off-resonance Monte Carlo sample ($B^0 \to D^- \pi^+ \pi^0$).



(a) (b)

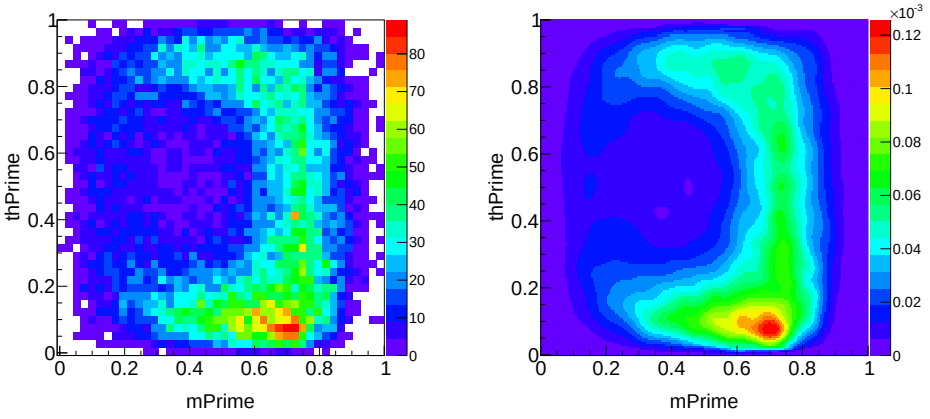
Figure A.3.: Result of the merging (red) of the low-bandwidth KDE (green) and high-bandwidth KDE (blue) (Continuum background from M_{bc} sideband, stream 1 of generic Monte Carlo, selected with MC truth variables) ($B^0 \to D^- \pi^+ \pi^0$).



(a) Continuum background in M_{bc} sideband

(b) Smoothed using the merged KDE

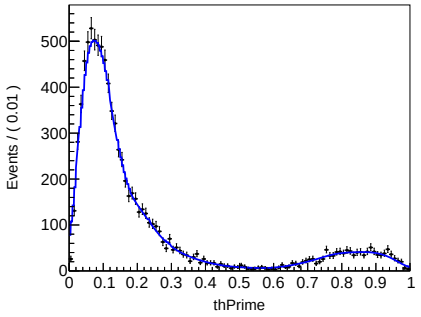
Figure A.4.: Comparison of continuum background in the M_{bc} sideband of stream 1 of generic Monte Carlo (selected with MC truth variables) and its smoothed distribution ($B^0 \to D^- \pi^+ \pi^0$).



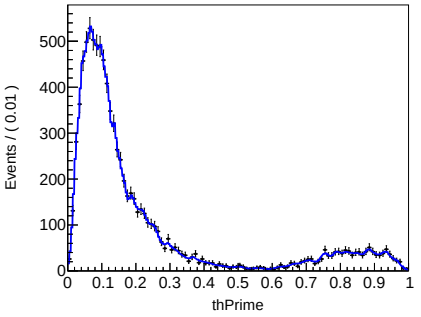
(a) Off-resonance data (b) Smoothed using the merged KDE

Figure A.5.: Comparison of off-resonance data and its smoothed distribution ($B^0 \to D^-\pi^+\pi^0$).

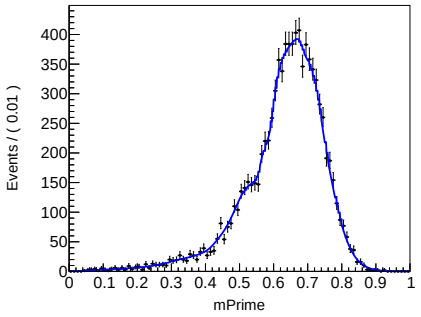
A.2. Combinatoric background



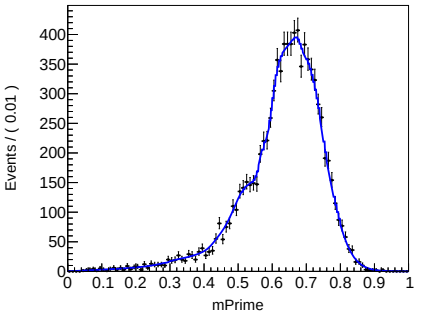
(a) High-bandwidth KDE - θ'



(b) Low-bandwidth KDE - θ'

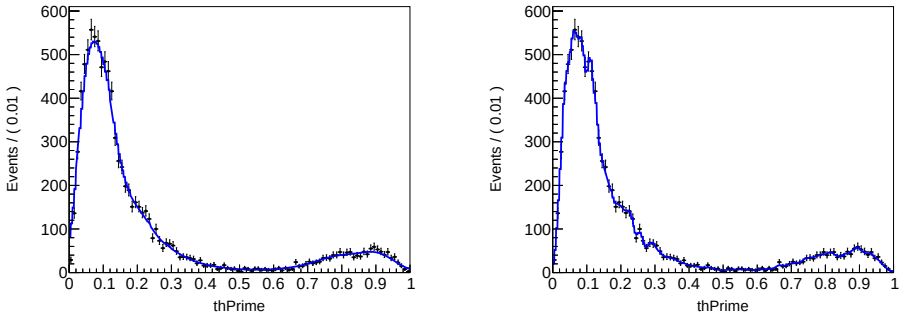


(c) High-bandwidth KDE - m'

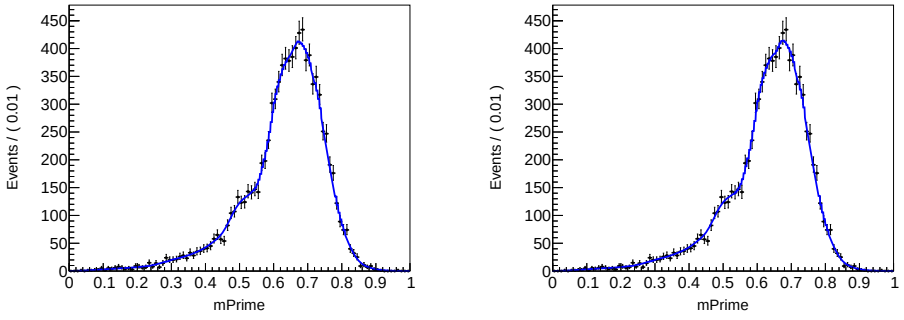


(d) Low-bandwidth KDE - m'

Figure A.6.: Comparison of the High- and Low-bandwidth KDE in the projection to θ' of combinatoric background in stream 0 of mixed generic Monte Carlo ($B^- \to D^0\pi^-\pi^0$).

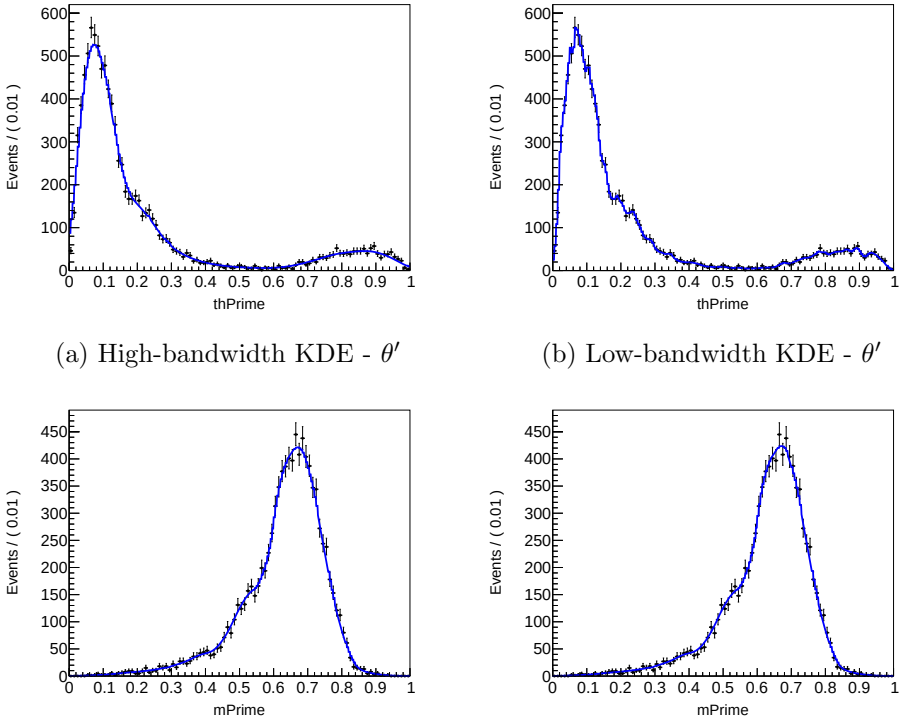


(a) High-bandwidth KDE - θ' (b) Low-bandwidth KDE - θ'



(c) High-bandwidth KDE - m' (d) Low-bandwidth KDE - m'

Figure A.7.: Comparison of the High- and Low-bandwidth KDE in the projection to θ' of combinatoric background in stream 1 of mixed generic Monte Carlo ($B^- \to D^0\pi^-\pi^0$).



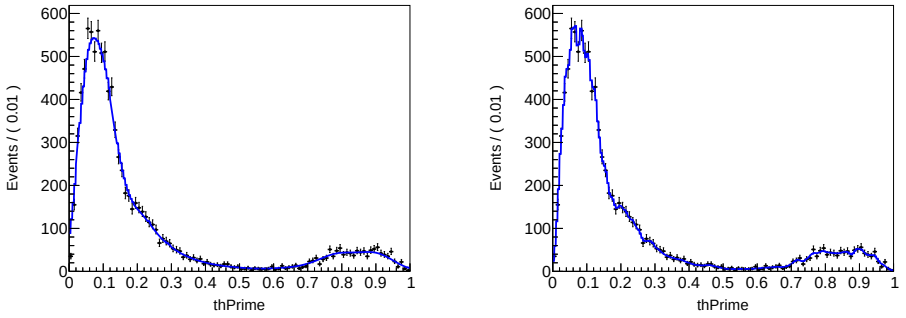
(a) High-bandwidth KDE - θ'

(b) Low-bandwidth KDE - θ'

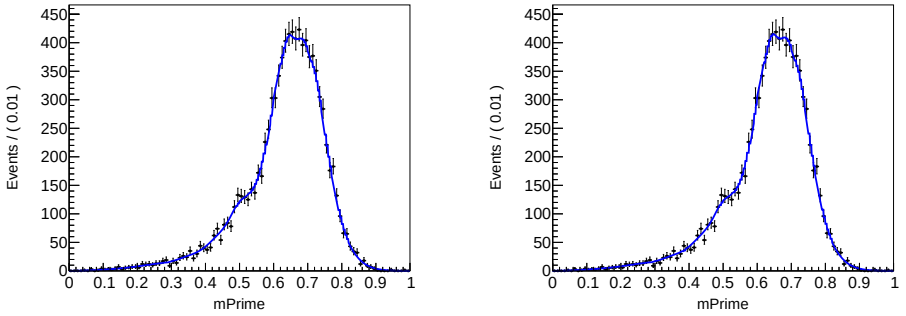
(c) High-bandwidth KDE - m'

(d) Low-bandwidth KDE - m'

Figure A.8.: Comparison of the High- and Low-bandwidth KDE in the projection to θ' of combinatoric background in stream 2 of mixed generic Monte Carlo ($B^- \to D^0\pi^-\pi^0$).

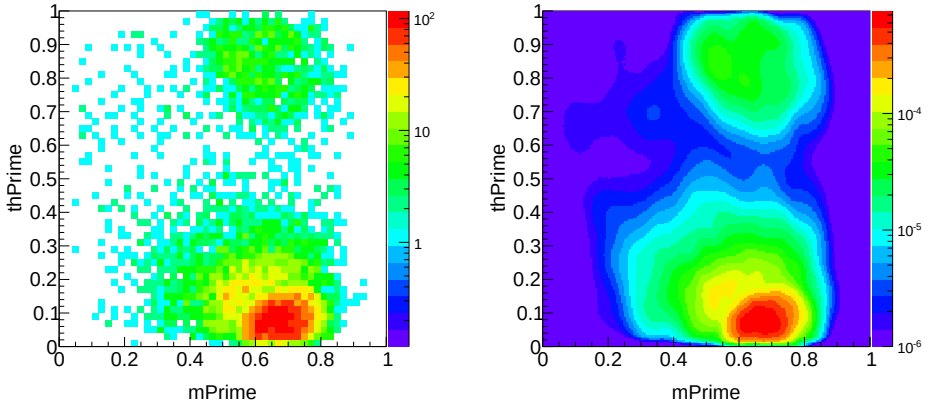


(a) High-bandwidth KDE - θ' (b) Low-bandwidth KDE - θ'



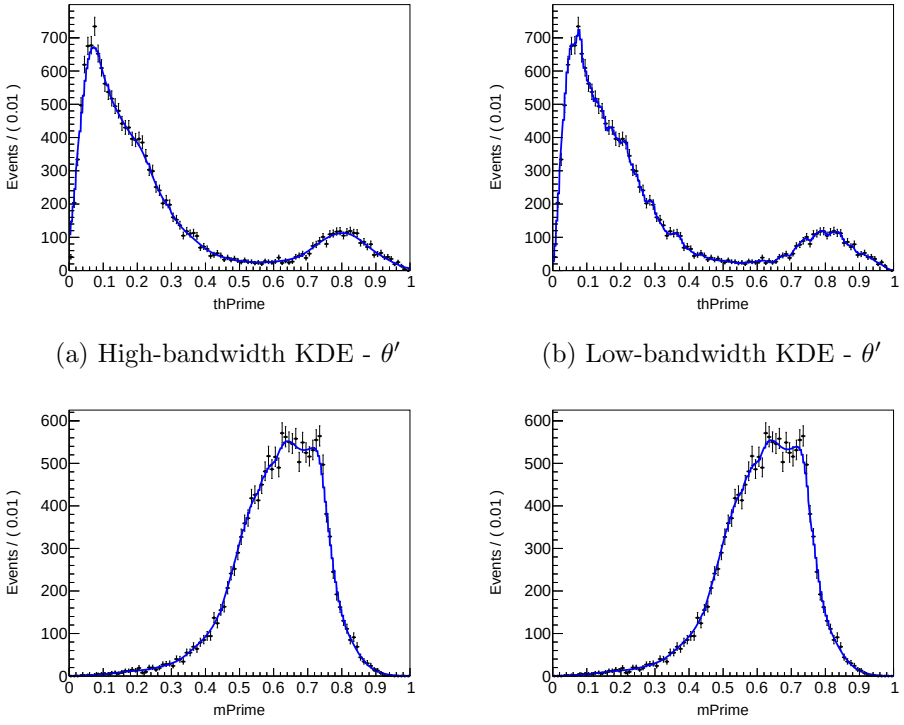
(c) High-bandwidth KDE - m' (d) Low-bandwidth KDE - m'

Figure A.9.: Comparison of the High- and Low-bandwidth KDE in the projection to θ' of combinatoric background in stream 3 of mixed generic Monte Carlo ($B^- \to D^0\pi^-\pi^0$).



(a) Combinatoric background in stream 0 of generic Monte Carlo (b) Smoothed using the merged KDE

Figure A.10.: Comparison of combinatoric background in stream 0 of generic Monte Carlo (selected with MC truth variables) and its smoothed distribution (Log scale).



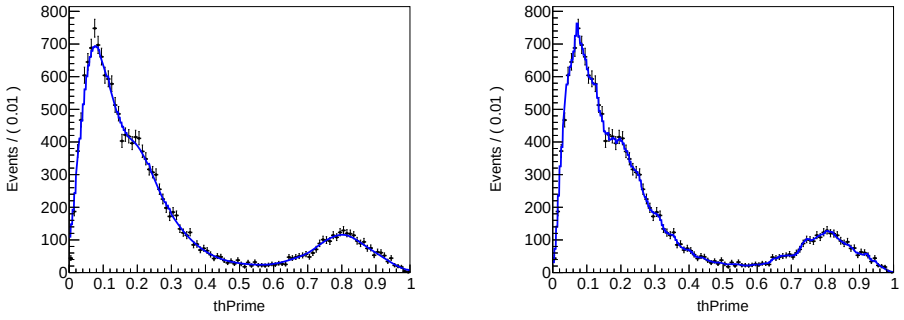
(a) High-bandwidth KDE - θ'

(b) Low-bandwidth KDE - θ'

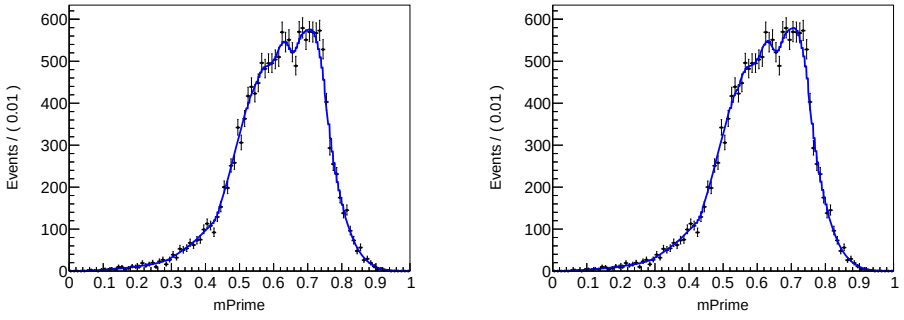
(c) High-bandwidth KDE - m'

(d) Low-bandwidth KDE - m'

Figure A.11.: Comparison of the High- and Low-bandwidth KDE in the projection to θ' of combinatoric background in stream 1 of charged generic Monte Carlo ($B^- \to D^0\pi^-\pi^0$).

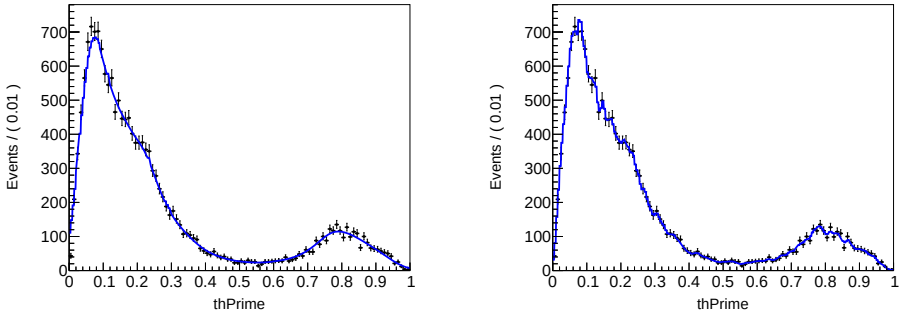


(a) High-bandwidth KDE - θ' (b) Low-bandwidth KDE - θ'

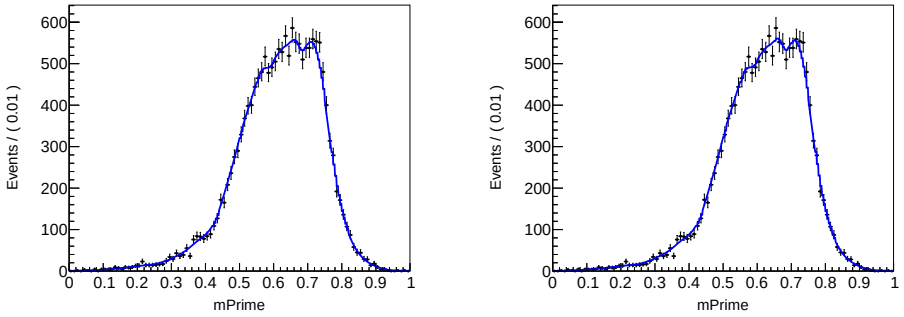


(c) High-bandwidth KDE - m' (d) Low-bandwidth KDE - m'

Figure A.12.: Comparison of the High- and Low-bandwidth KDE in the projection to θ' of combinatoric background in stream 2 of charged generic Monte Carlo ($B^- \to D^0\pi^-\pi^0$).

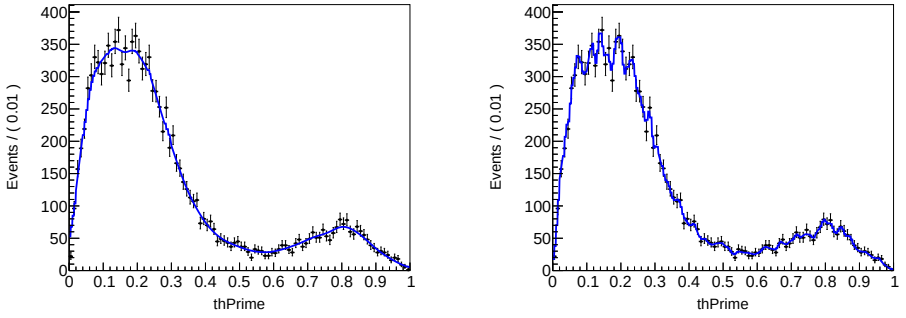


(a) High-bandwidth KDE - θ' (b) Low-bandwidth KDE - θ'

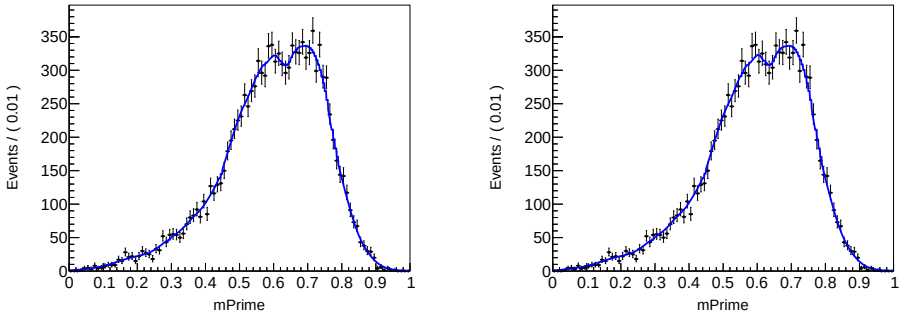


(c) High-bandwidth KDE - m' (d) Low-bandwidth KDE - m'

Figure A.13.: Comparison of the High- and Low-bandwidth KDE in the projection to θ' of combinatoric background in stream 3 of charged generic Monte Carlo ($B^- \to D^0\pi^-\pi^0$).

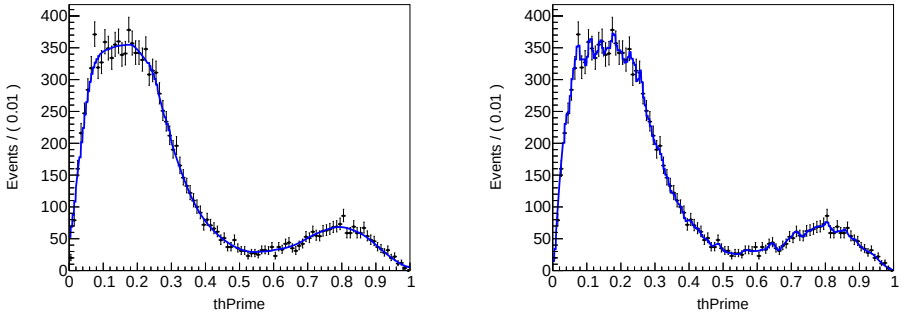


(a) High-bandwidth KDE - θ' (b) Low-bandwidth KDE - θ'

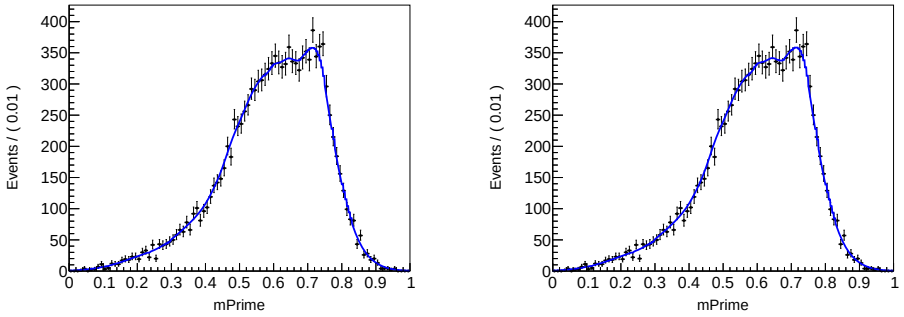


(c) High-bandwidth KDE - m' (d) Low-bandwidth KDE - m'

Figure A.14.: Comparison of the High- and Low-bandwidth KDE in the projection to θ' of combinatoric background in stream 0 of mixed generic Monte Carlo ($B^0 \to D^- \pi^+ \pi^0$).

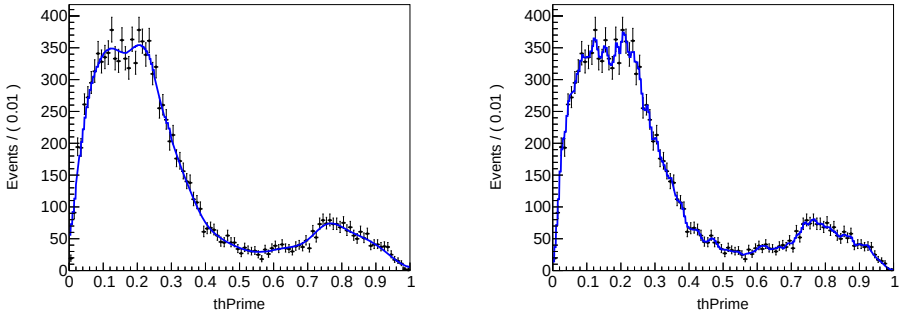


(a) High-bandwidth KDE - θ' (b) Low-bandwidth KDE - θ'

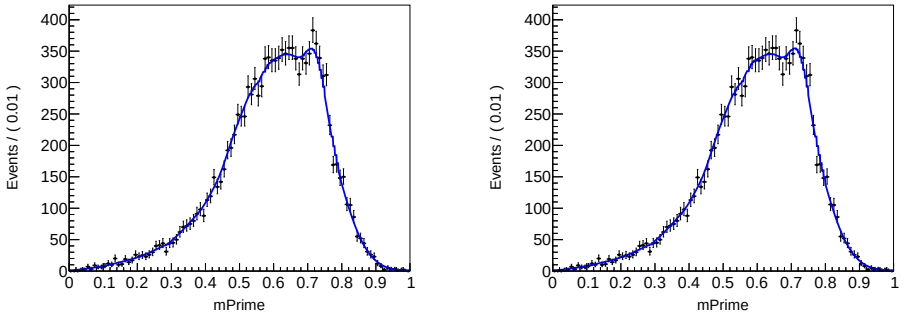


(c) High-bandwidth KDE - m' (d) Low-bandwidth KDE - m'

Figure A.15.: Comparison of the High- and Low-bandwidth KDE in the projection to θ' of combinatoric background in stream 1 of mixed generic Monte Carlo ($B^0 \to D^- \pi^+ \pi^0$).

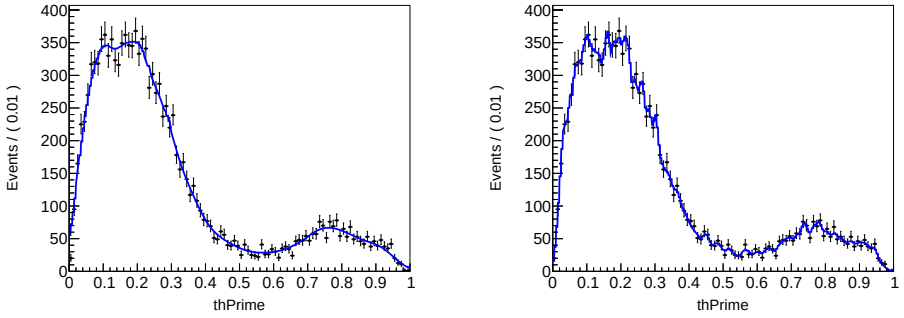


(a) High-bandwidth KDE - θ' (b) Low-bandwidth KDE - θ'

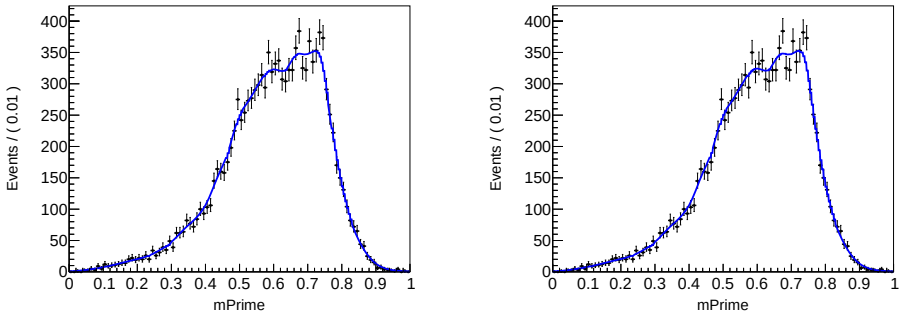


(c) High-bandwidth KDE - m' (d) Low-bandwidth KDE - m'

Figure A.16.: Comparison of the High- and Low-bandwidth KDE in the projection to θ' of combinatoric background in stream 2 of mixed generic Monte Carlo ($B^0 \to D^- \pi^+ \pi^0$).

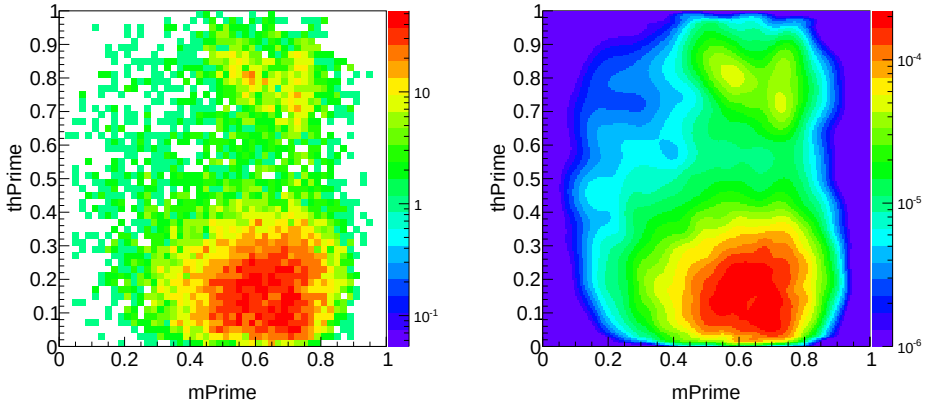


(a) High-bandwidth KDE - θ' (b) Low-bandwidth KDE - θ'



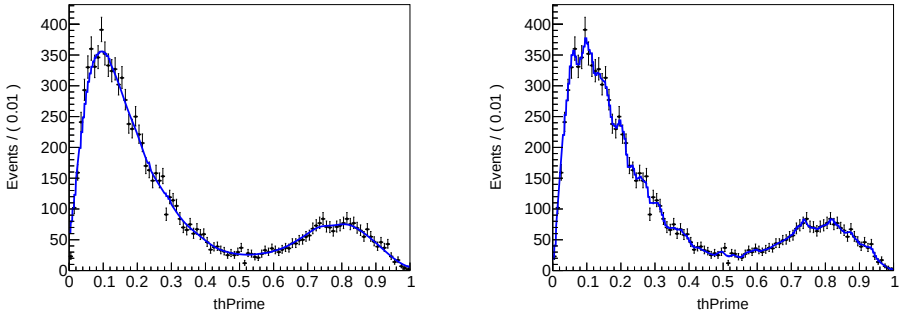
(c) High-bandwidth KDE - m' (d) Low-bandwidth KDE - m'

Figure A.17.: Comparison of the High- and Low-bandwidth KDE in the projection to θ' of combinatoric background in stream 3 of mixed generic Monte Carlo ($B^0 \to D^-\pi^+\pi^0$).

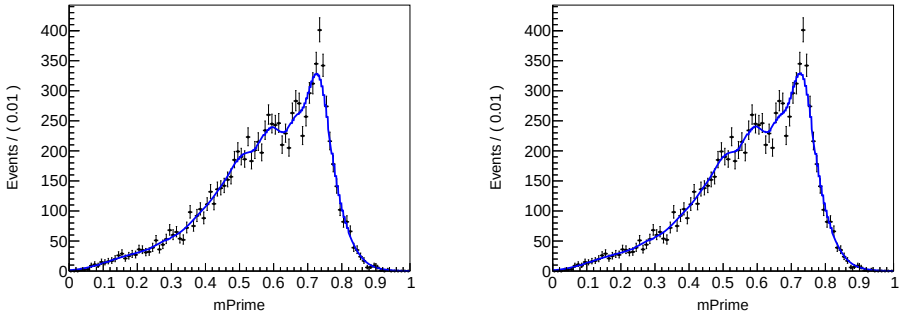


(a) Combinatoric background in stream 0 of generic Monte Carlo (b) merged KDE

Figure A.18.: Comparison of combinatoric background in stream 0 of mixed generic Monte Carlo (selected with MC truth variables) and its smoothed distribution (Log scale) ($B^0 \to D^-\pi^+\pi^0$).

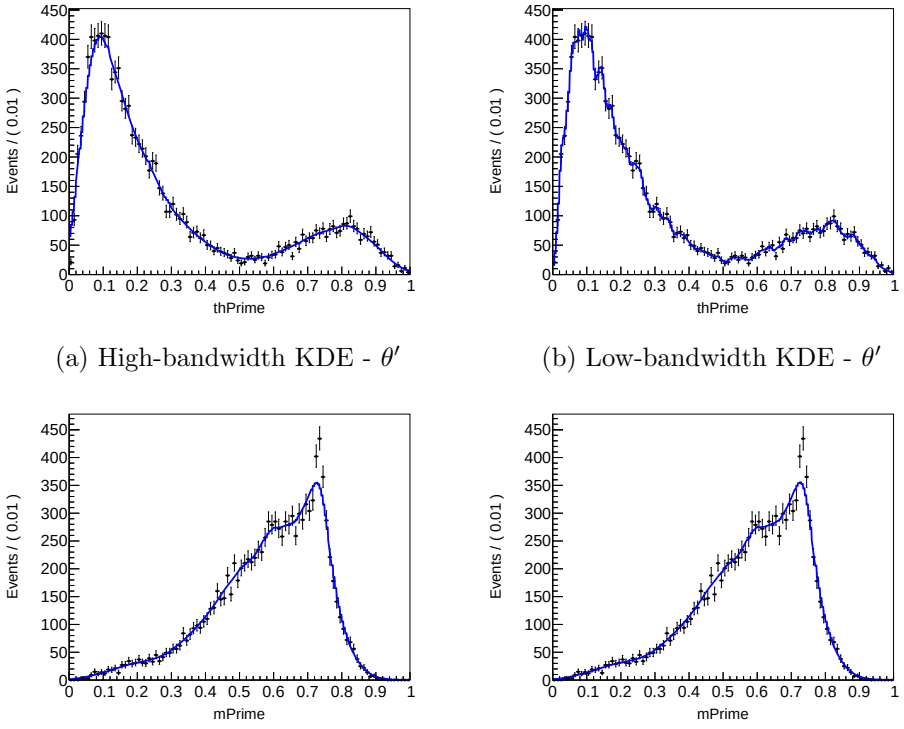


(a) High-bandwidth KDE - θ' (b) Low-bandwidth KDE - θ'



(c) High-bandwidth KDE - m' (d) Low-bandwidth KDE - m'

Figure A.19.: Comparison of the High- and Low-bandwidth KDE in the projection to θ' of combinatoric background in stream 0 of charged generic Monte Carlo ($B^0 \to D^- \pi^+ \pi^0$).



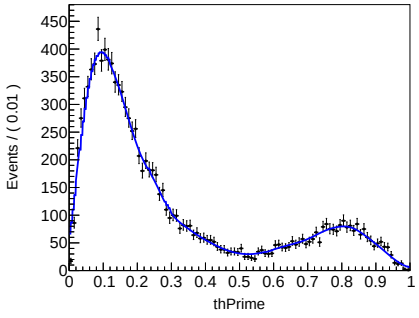
(a) High-bandwidth KDE - θ'

(b) Low-bandwidth KDE - θ'

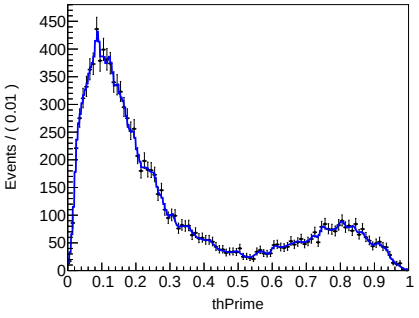
(c) High-bandwidth KDE - m'

(d) Low-bandwidth KDE - m'

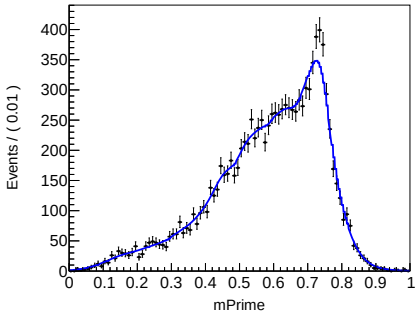
Figure A.20.: Comparison of the High- and Low-bandwidth KDE in the projection to θ' of combinatoric background in stream 1 of charged generic Monte Carlo ($B^0 \to D^- \pi^+ \pi^0$).



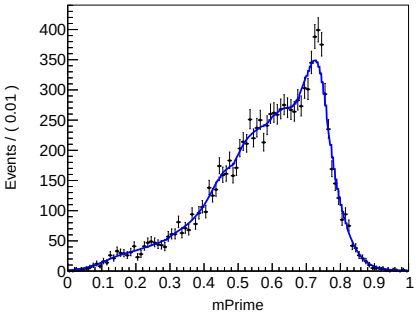
(a) High-bandwidth KDE - θ'



(b) Low-bandwidth KDE - θ'

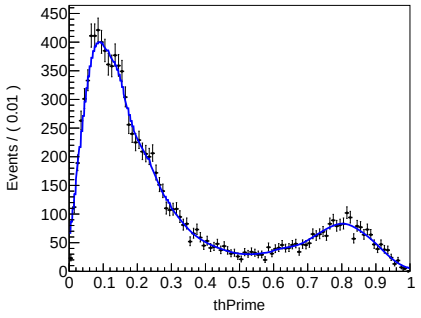


(c) High-bandwidth KDE - m'

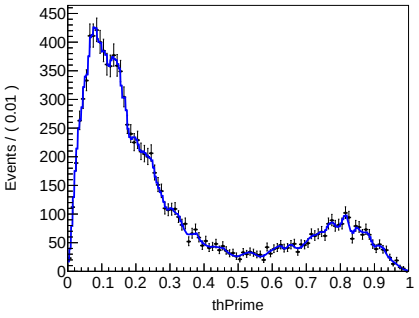


(d) Low-bandwidth KDE - m'

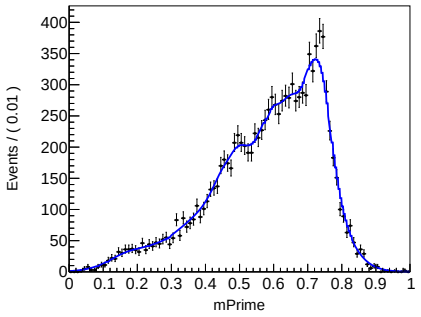
Figure A.21.: Comparison of the High- and Low-bandwidth KDE in the projection to θ' of combinatoric background in stream 2 of charged generic Monte Carlo ($B^0 \to D^- \pi^+ \pi^0$).



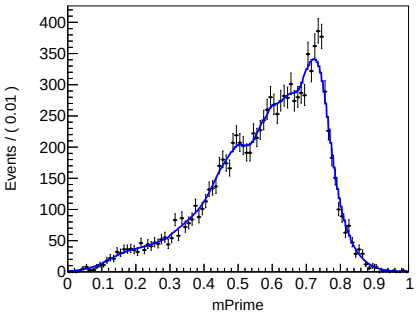
(a) High-bandwidth KDE - θ'



(b) Low-bandwidth KDE - θ'

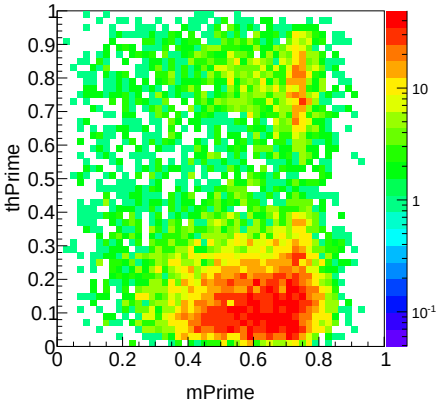


(c) High-bandwidth KDE - m'

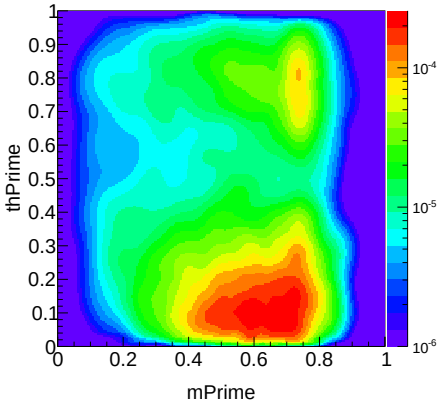


(d) Low-bandwidth KDE - m'

Figure A.22.: Comparison of the High- and Low-bandwidth KDE in the projection to θ' of combinatoric background in stream 3 of charged generic Monte Carlo ($B^0 \to D^-\pi^+\pi^0$).



(a) Combinatoric background in stream 0 of generic Monte Carlo



(b) Smoothed using the merged KDE

Figure A.23.: Comparison of combinatoric background in stream 0 of charged generic Monte Carlo (selected with MC truth variables) and its smoothed distribution (Log scale) ($B^0 \to D^-\pi^+\pi^0$).

B. Scatter plots of fits to single resonance in the normal Dalitz plot

B.1. Fit results for $B^- \to D^0\pi^-\pi^0$

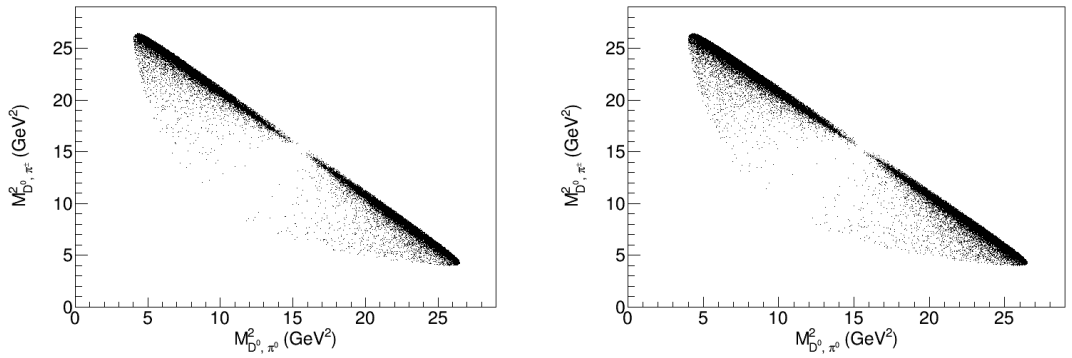


Figure B.24.: Scatter plots in the normal Dalitz plot showing the input sample (left) and the fit result (right) for the fit to truth matched ρ sample.

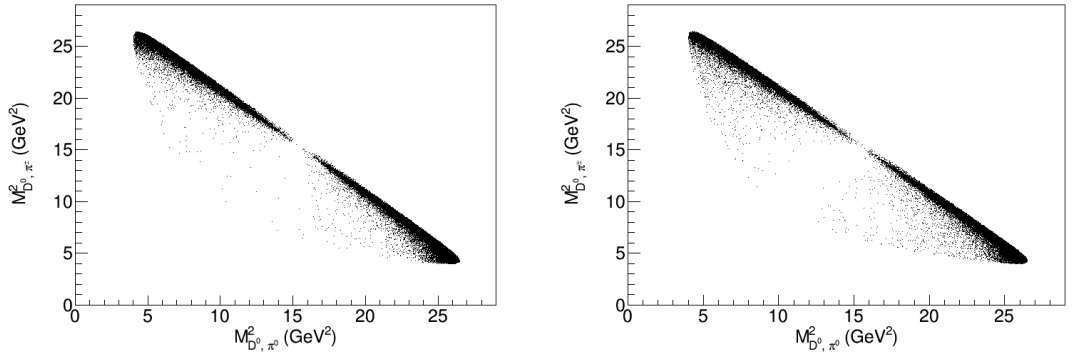


Figure B.25.: Scatter plots in the normal Dalitz plot showing the input sample (left) and the fit result (right) for the fit to the full ρ sample.

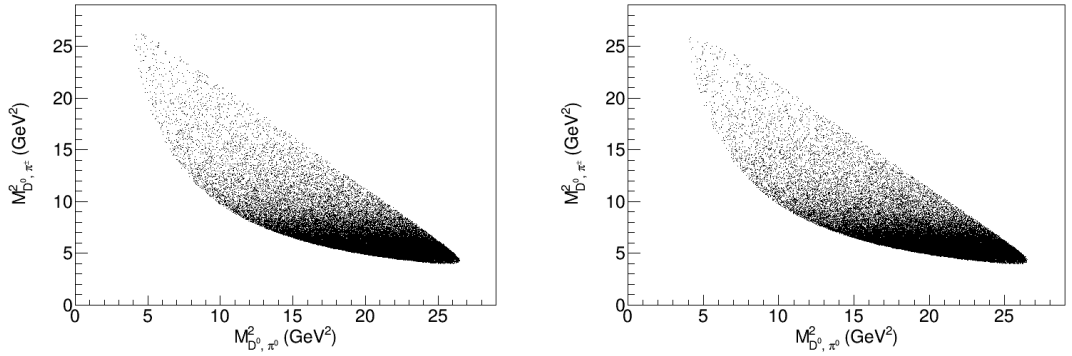


Figure B.26.: Scatter plots in the normal Dalitz plot showing the input sample (left) and the fit result (right) for the fit to truth matched $D_0^*(2400)^0$ sample.

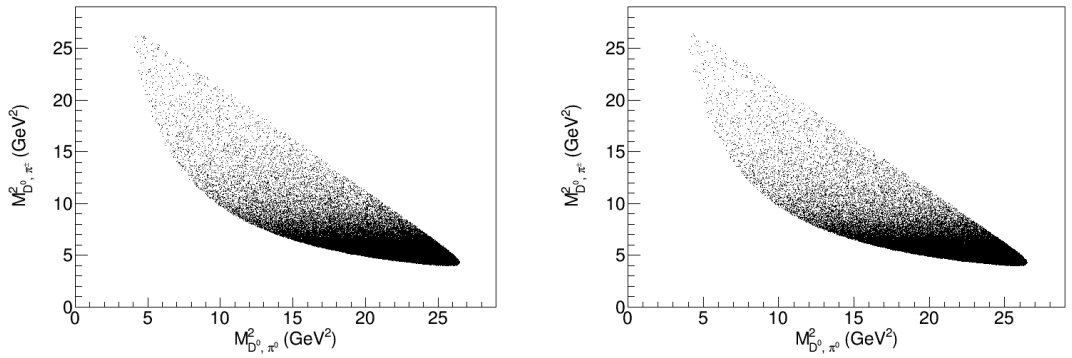


Figure B.27.: Scatter plots in the normal Dalitz plot showing the input sample (left) and the fit result (right)for the fit to the full $D_0^*(2400)^0$ sample.

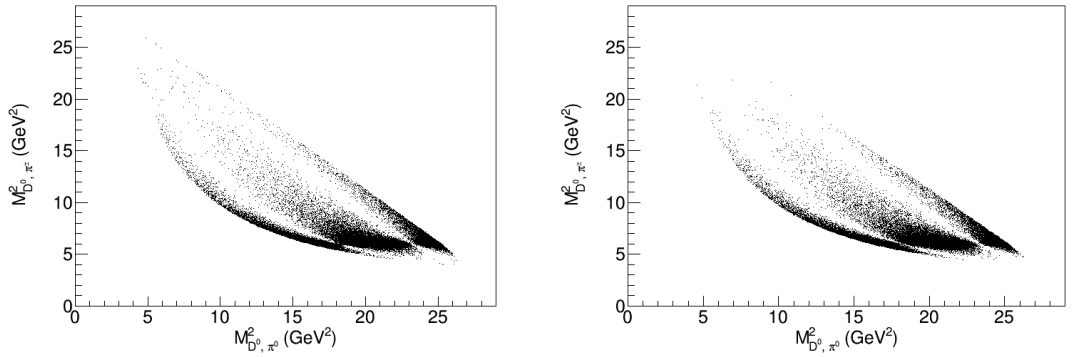


Figure B.28.: Scatter plots in the normal Dalitz plot showing the input sample (left) and the fit result (right) for the fit to truth matched $D_2^*(2462)^0$ sample.

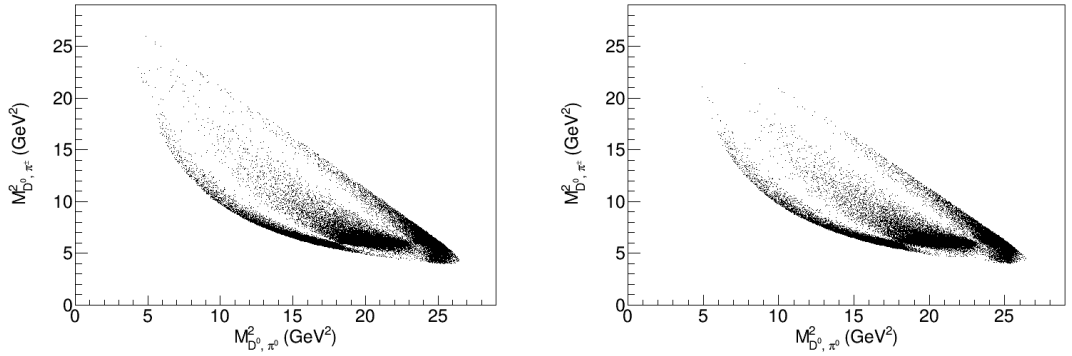


Figure B.29.: Scatter plots in the normal Dalitz plot showing the input sample (left) and the fit result (right) for the fit to the full $D_2^*(2462)^0$ sample.

B.2. Fit results for $B^0 \to D^- \pi^+ \pi^0$

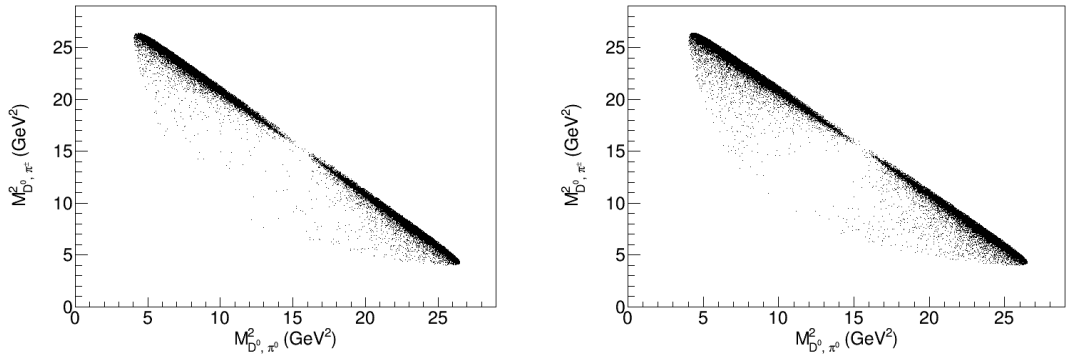


Figure B.30.: Scatter plots in the normal Dalitz plot showing the input sample (left) and the fit result (right) for the fit to truth matched ρ sample.

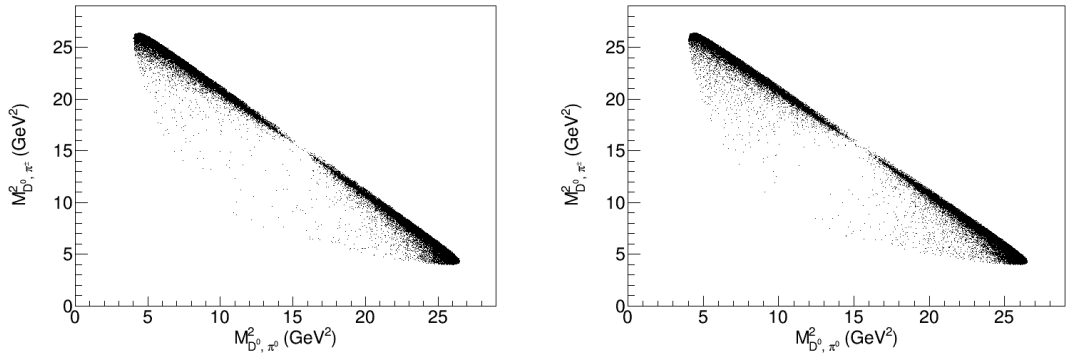


Figure B.31.: Scatter plots in the normal Dalitz plot showing the input sample (left) and the fit result (right) for the fit to the full ρ sample.

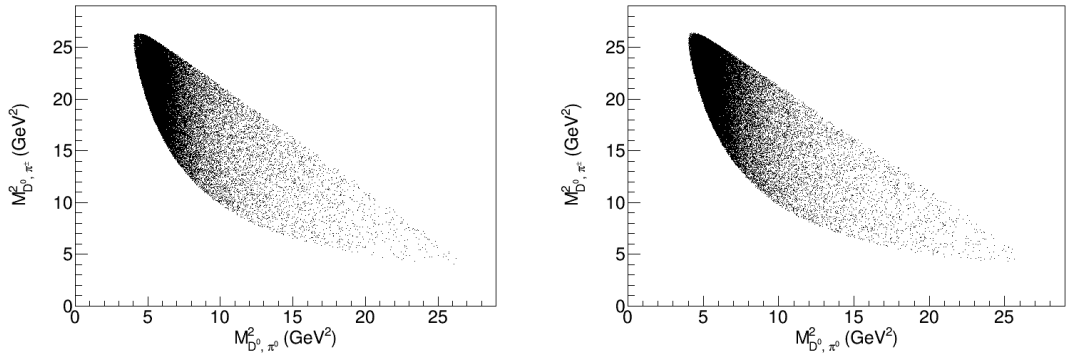


Figure B.32.: Scatter plots in the normal Dalitz plot showing the input sample (left) and the fit result (right) for the fit to truth matched $D_0^*(2400)^0$ sample.

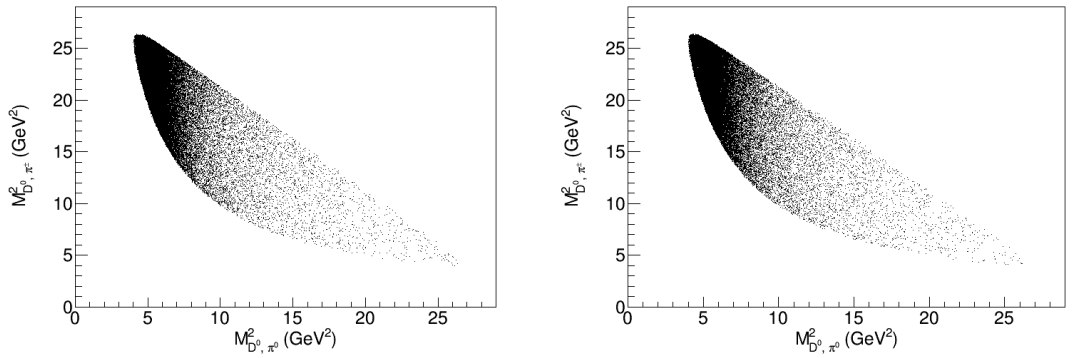


Figure B.33.: Scatter plots in the normal Dalitz plot showing the input sample (left) and the fit result (right) for the fit to the full $D_0^*(2400)^0$ sample.

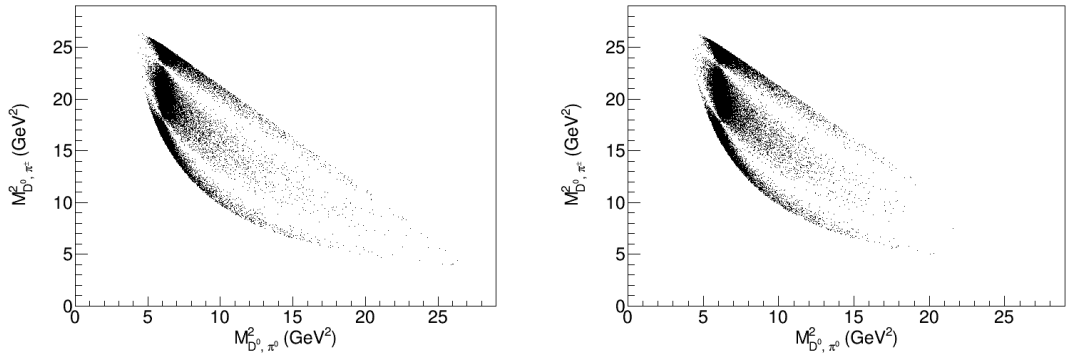


Figure B.34.: Scatter plots in the normal Dalitz plot showing the input sample (left) and the fit result (right) for the fit to truth matched $D_2^*(2462)^0$ sample.

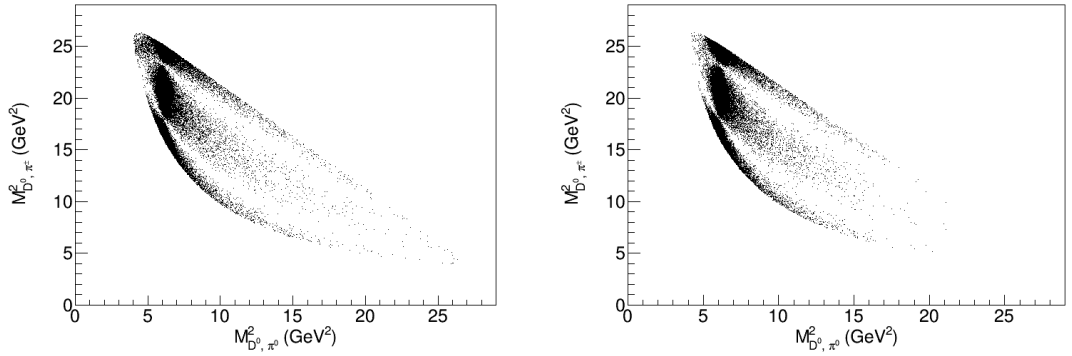


Figure B.35.: Scatter plots in the normal Dalitz plot showing the input sample (left) and the fit result (right) for the fit to the full $D_2^*(2462)^0$ sample.

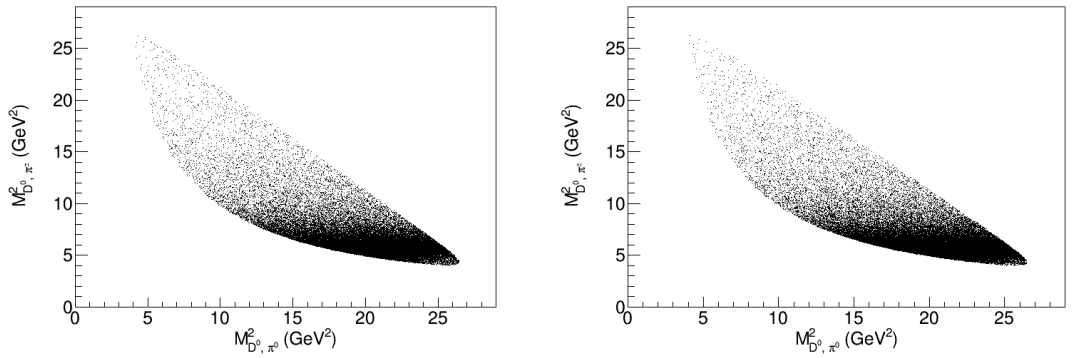


Figure B.36.: Scatter plots in the normal Dalitz plot showing the input sample (left) and the fit result (right) for the fit to truth matched $D_0^*(2400)^-$ sample.

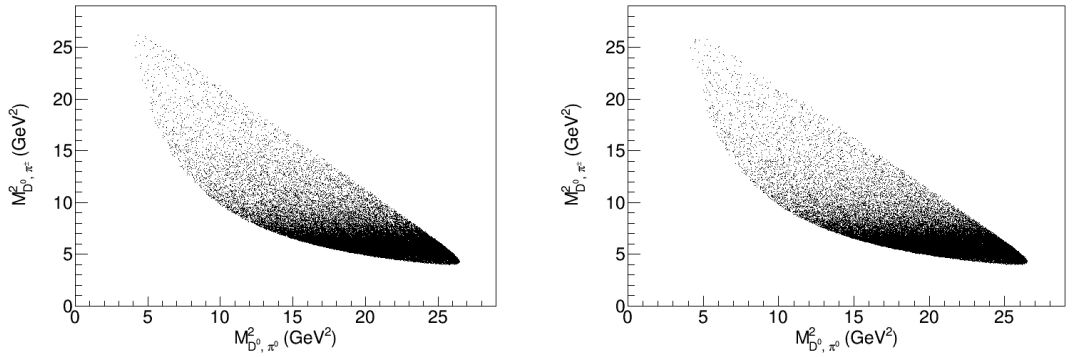


Figure B.37.: Scatter plots in the normal Dalitz plot showing the input sample (left) and the fit result (right) for the fit to the full $D_0^*(2400)^-$ sample.

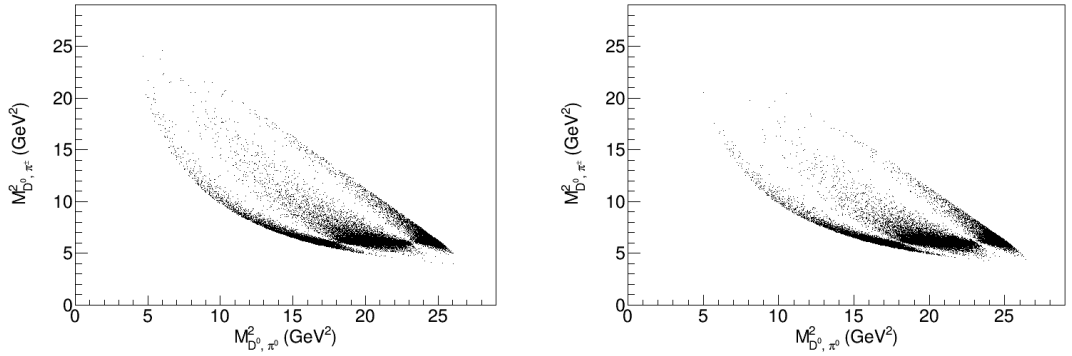


Figure B.38.: Scatter plots in the normal Dalitz plot showing the input sample (left) and the fit result (right) for the fit to truth matched $D_2^*(2462)^-$ sample.

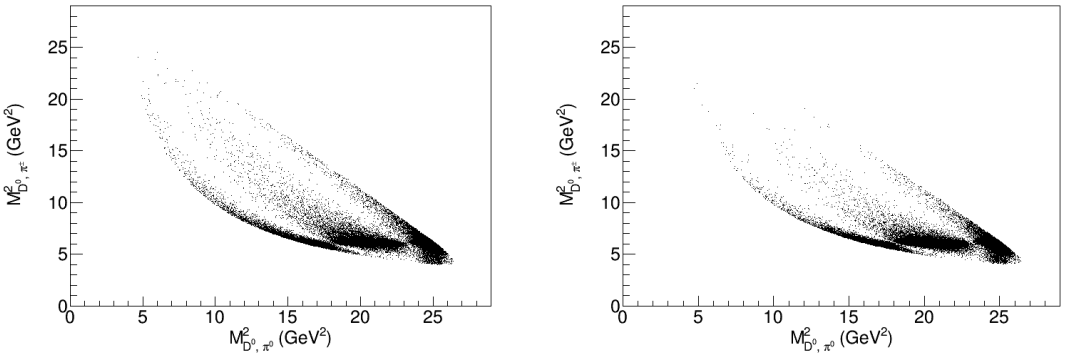


Figure B.39.: Scatter plots in the normal Dalitz plot showing the input sample (left) and the fit result (right) for the fit to the full $D_2^*(2462)^-$ sample.

C. Fit to fully simulated data

Fits to EvtGen sample with aproximately the expected contributions, D^* excluded ($B^- \to D^0\pi^-\pi^0$)

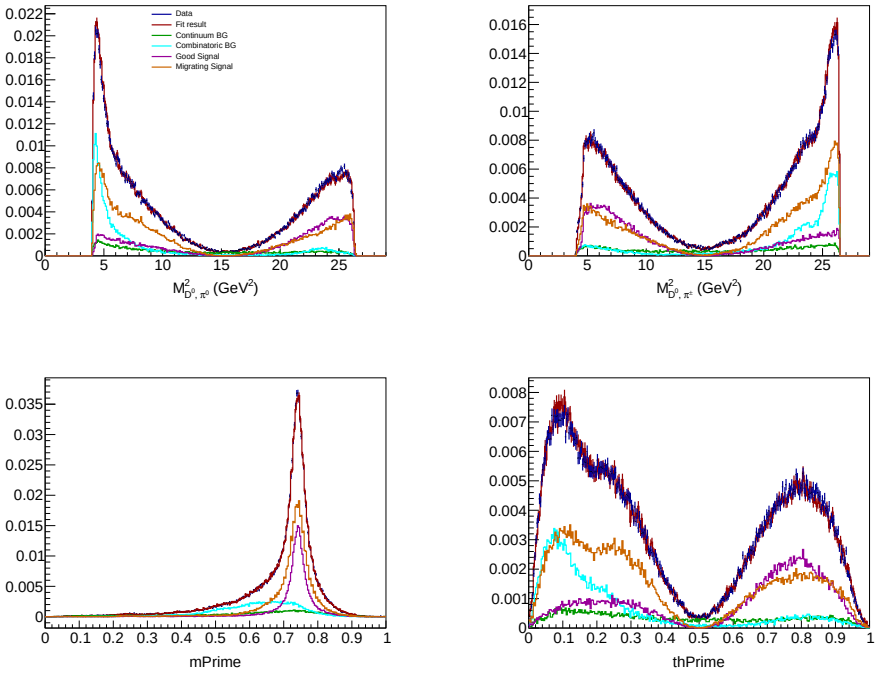


Figure C.40.: Projections to the normal and squared Dalitz plot variables of the fit to signal Monte Carlo including misreconstructed events.

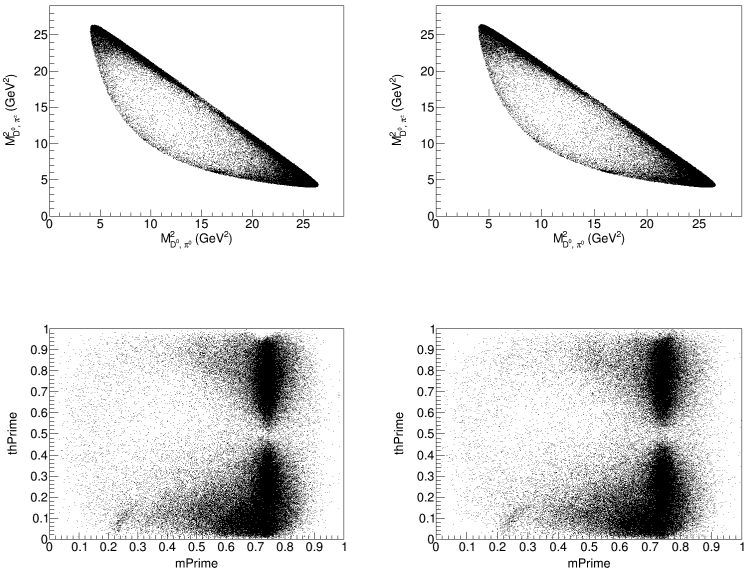


Figure C.41.: Scatter plots in the normal (top) and squared (bottom) Dalitz plot showing the input sample (left) and the fit result (right).

Fits to EvtGen sample with aproximately the expected contributions, D^* excluded ($B^0 \to D^- \pi^+ \pi^0$)

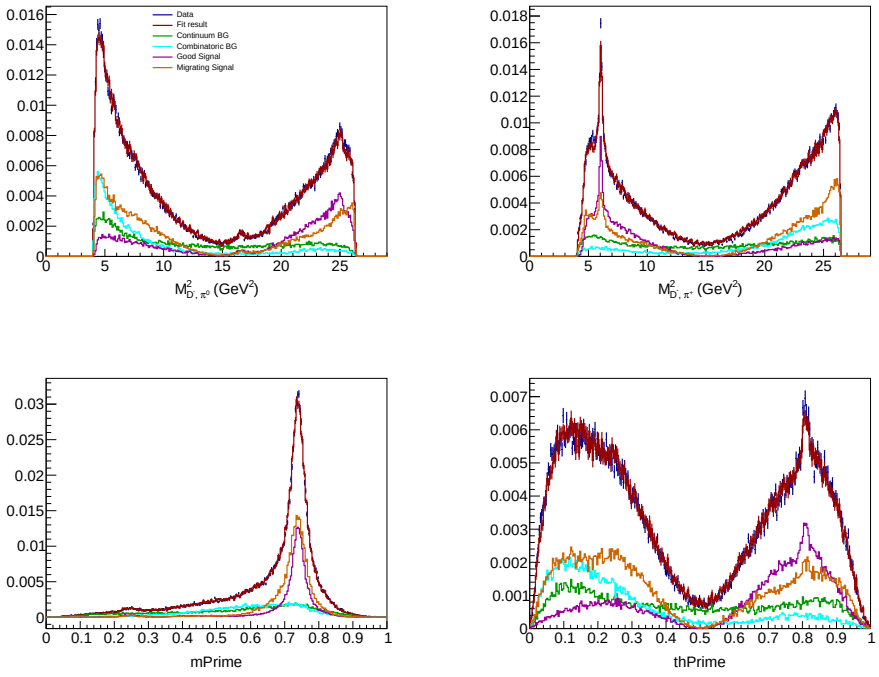


Figure C.42.: Projections to the normal and squared Dalitz plot variables of the fit to signal Monte Carlo including misreconstructed events.

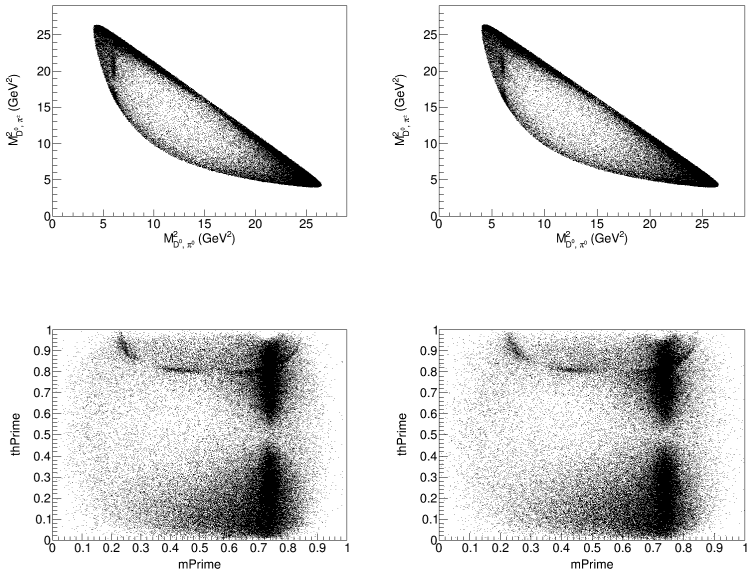


Figure C.43.: Scatter plots in the normal (top) and squared (bottom) Dalitz plot showing the input sample (left) and the fit result (right).

D. Results of ensemble test of random models prior to unblinding

D.1. Results for $B^- \to D^0\pi^-\pi^0$

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	300	0.070	-0.086 ± 0.055	0.86 ± 0.047
$D_0^*(2400)^0$ Y	300	0.027	0.019 ± 0.071	1.1 ± 0.064
$D_2^*(2462)^0$ X	300	-0.045	0.090 ± 0.064	0.99 ± 0.056
$D_2^*(2462)^0$ Y	300	-0.011	-0.082 ± 0.064	1.0 ± 0.052
NR	300	0.10	0.0098 ± 0.050	0.81 ± 0.044
D^*	300	0.27	0.24 ± 0.060	0.96 ± 0.052
Continuum BG	300	14 000	0.10 ± 0.057	0.88 ± 0.050
Combinatoric BG	300	25 000	-0.25 ± 0.049	0.80 ± 0.040
ρ X	-	-0.33	-	-
ρ Y	-	0.25	-	-

Table D.3.: Random model 2, $B^- \to D^0\pi^-\pi^0$

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	300	-0.072	-0.032 ± 0.059	0.93 ± 0.055
$D_0^*(2400)^0$ Y	300	-0.023	0.0068 ± 0.059	0.95 ± 0.048
$D_2^*(2462)^0$ X	300	-0.035	0.23 ± 0.053	0.83 ± 0.043
$D_2^*(2462)^0$ Y	300	0.051	0.095 ± 0.063	1.0 ± 0.052
NR	300	0.064	-0.060 ± 0.057	0.90 ± 0.048
D^*	300	0.26	0.22 ± 0.058	0.92 ± 0.046
Continuum BG	300	14 000	0.21 ± 0.053	0.84 ± 0.042
Combinatoric BG	300	25 000	-0.28 ± 0.050	0.81 ± 0.039
ρ X	-	-0.44	-	-
ρ Y	-	0.13	-	-

Table D.4.: Random model 3, $B^- \to D^0\pi^-\pi^0$

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	300	0.057	0.0094 ± 0.061	0.95 ± 0.050
$D_0^*(2400)^0$ Y	300	-0.044	-0.12 ± 0.064	1.0 ± 0.060
$D_2^*(2462)^0$ X	300	0.044	0.028 ± 0.061	0.96 ± 0.049
$D_2^*(2462)^0$ Y	300	0.041	-0.18 ± 0.065	1.0 ± 0.057
NR	300	0.036	-0.035 ± 0.065	1.0 ± 0.052
D^*	300	0.24	0.48 ± 0.081	1.1 ± 0.070
Continuum BG	300	14 000	0.12 ± 0.053	0.84 ± 0.050
Combinatoric BG	300	25 000	-0.23 ± 0.048	0.76 ± 0.041
ρ X	-	-0.27	-	-
ρ Y	-	0.62	-	-

Table D.5.: Random model 4, $B^- \to D^0\pi^-\pi^0$

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	300	-0.11	$-0.000\,15 \pm 0.066$	0.99 ± 0.055
$D_0^*(2400)^0$ Y	300	-0.014	-0.12 ± 0.069	1.1 ± 0.076
$D_2^*(2462)^0$ X	300	0.052	-0.045 ± 0.072	1.00 ± 0.069
$D_2^*(2462)^0$ Y	300	0.014	0.037 ± 0.064	0.93 ± 0.056
NR	300	0.082	-0.061 ± 0.061	0.96 ± 0.052
D^*	300	0.25	0.33 ± 0.061	0.97 ± 0.055
Continuum BG	300	14 000	0.15 ± 0.061	0.99 ± 0.051
Combinatoric BG	300	25 000	-0.27 ± 0.054	0.82 ± 0.047
ρ X	-	0.31	-	-
ρ Y	-	0.30	-	-

Table D.6.: Random model 5, $B^- \to D^0\pi^-\pi^0$

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	300	0.071	0.11 ± 0.058	0.91 ± 0.046
$D_0^*(2400)^0$ Y	300	0.041	-0.063 ± 0.062	0.96 ± 0.047
$D_2^*(2462)^0$ X	300	0.061	-0.15 ± 0.077	1.2 ± 0.069
$D_2^*(2462)^0$ Y	300	-0.016	-0.076 ± 0.069	1.0 ± 0.056
NR	300	0.027	0.070 ± 0.069	1.0 ± 0.054
D^*	300	0.27	0.37 ± 0.062	0.95 ± 0.053
Continuum BG	300	14 000	0.16 ± 0.059	0.91 ± 0.046
Combinatoric BG	300	25 000	-0.24 ± 0.045	0.74 ± 0.036
ρ X	-	0.23	-	-
ρ Y	-	0.60	-	-

Table D.7.: Random model 6, $B^- \to D^0\pi^-\pi^0$

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	300	0.025	0.054 ± 0.064	0.91 ± 0.061
$D_0^*(2400)^0$ Y	300	0.11	-0.089 ± 0.058	0.86 ± 0.049
$D_2^*(2462)^0$ X	300	0.017	0.025 ± 0.059	0.92 ± 0.048
$D_2^*(2462)^0$ Y	300	0.053	-0.069 ± 0.062	0.96 ± 0.047
NR	300	0.082	0.047 ± 0.066	1.0 ± 0.054
D^*	300	0.25	0.34 ± 0.060	0.96 ± 0.050
Continuum BG	300	14 000	0.096 ± 0.054	0.86 ± 0.042
Combinatoric BG	300	25 000	-0.19 ± 0.054	0.80 ± 0.043
ρ X	-	-0.54	-	-
ρ Y	-	0.17	-	-

Table D.8.: Random model 7, $B^- \to D^0\pi^-\pi^0$

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	300	0.075	-0.032 ± 0.069	1.1 ± 0.058
$D_0^*(2400)^0$ Y	300	0.053	0.050 ± 0.067	1.0 ± 0.051
$D_2^*(2462)^0$ X	300	0.014	-0.16 ± 0.065	1.0 ± 0.059
$D_2^*(2462)^0$ Y	300	-0.053	0.016 ± 0.063	1.0 ± 0.053
NR	300	0.094	0.012 ± 0.054	0.87 ± 0.048
D^*	300	0.26	0.31 ± 0.057	0.94 ± 0.049
Continuum BG	300	14 000	0.070 ± 0.053	0.85 ± 0.040
Combinatoric BG	300	25 000	-0.12 ± 0.050	0.79 ± 0.044
ρ X	-	0.46	-	-
ρ Y	-	0.054	-	-

Table D.9.: Random model 8, $B^- \to D^0\pi^-\pi^0$

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	300	0.046	-0.0064 ± 0.059	0.94 ± 0.060
$D_0^*(2400)^0$ Y	300	-0.085	0.099 ± 0.061	0.99 ± 0.051
$D_2^*(2462)^0$ X	300	0.051	0.057 ± 0.067	1.0 ± 0.066
$D_2^*(2462)^0$ Y	300	-0.047	-0.073 ± 0.063	0.96 ± 0.051
NR	300	0.019	0.085 ± 0.054	0.88 ± 0.046
D^*	300	0.26	0.42 ± 0.060	0.93 ± 0.048
Continuum BG	300	14 000	0.082 ± 0.060	0.88 ± 0.049
Combinatoric BG	300	25 000	-0.27 ± 0.050	0.80 ± 0.041
ρ X	-	-0.60	-	-
ρ Y	-	-0.16	-	-

Table D.10.: Random model 9, $B^- \to D^0\pi^-\pi^0$

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	300	-0.050	-0.073 ± 0.068	1.0 ± 0.064
$D_0^*(2400)^0$ Y	300	-0.075	0.13 ± 0.063	0.99 ± 0.052
$D_2^*(2462)^0$ X	300	-0.041	0.0039 ± 0.059	0.94 ± 0.046
$D_2^*(2462)^0$ Y	300	-0.024	0.052 ± 0.064	0.97 ± 0.055
NR	300	0.082	0.032 ± 0.066	1.0 ± 0.061
D^*	300	0.25	0.36 ± 0.062	0.99 ± 0.056
Continuum BG	300	14 000	0.24 ± 0.061	0.97 ± 0.051
Combinatoric BG	300	25 000	-0.35 ± 0.052	0.82 ± 0.045
ρ X	-	-0.16	-	-
ρ Y	-	-0.44	-	-

Table D.11.: Random model 10, $B^- \to D^0\pi^-\pi^0$

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	300	-0.046	-0.065 ± 0.062	0.98 ± 0.051
$D_0^*(2400)^0$ Y	300	-0.054	-0.16 ± 0.061	0.96 ± 0.065
$D_2^*(2462)^0$ X	300	0.046	0.033 ± 0.060	0.96 ± 0.049
$D_2^*(2462)^0$ Y	300	-0.033	-0.015 ± 0.061	0.96 ± 0.053
NR	300	0.075	0.10 ± 0.065	1.0 ± 0.057
D^*	300	0.27	0.34 ± 0.065	1.0 ± 0.058
Continuum BG	300	14 000	0.13 ± 0.054	0.86 ± 0.045
Combinatoric BG	300	25 000	-0.25 ± 0.051	0.83 ± 0.038
ρ X	-	-0.17	-	-
ρ Y	-	0.39	-	-

Table D.12.: Random model 11, $B^- \to D^0\pi^-\pi^0$

D.2. Results for $B^0 \to D^-\pi^+\pi^0$

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	236	0.064	-0.034 ± 0.070	0.96 ± 0.061
$D_0^*(2400)^0$ Y	236	-0.027	0.089 ± 0.072	1.0 ± 0.059
$D_2^*(2462)^0$ X	236	-0.20	0.095 ± 0.073	0.95 ± 0.067
$D_2^*(2462)^0$ Y	236	-0.16	-0.075 ± 0.070	0.94 ± 0.060
NR	236	0.046	0.058 ± 0.074	1.0 ± 0.073
D^*	236	0.35	0.23 ± 0.073	0.99 ± 0.062
Continuum BG	236	14 000	0.19 ± 0.061	0.81 ± 0.065
Combinatoric BG	236	25 000	-0.17 ± 0.056	0.80 ± 0.049
ρ X	-	0.17	-	-
ρ Y	-	0.77	-	-

Table D.13.: Random model 1, $B^0 \to D^-\pi^+\pi^0$

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	291	-0.16	0.022 ± 0.063	0.97 ± 0.064
$D_0^*(2400)^0$ Y	291	0.11	0.071 ± 0.062	0.97 ± 0.056
$D_2^*(2462)^0$ X	291	0.12	0.097 ± 0.067	1.00 ± 0.060
$D_2^*(2462)^0$ Y	291	0.24	-0.29 ± 0.070	1.1 ± 0.061
NR	291	0.26	-0.043 ± 0.070	1.0 ± 0.067
D^*	290	0.33	0.17 ± 0.060	0.94 ± 0.052
Continuum BG	291	14 000	0.14 ± 0.062	0.95 ± 0.051
Combinatoric BG	291	25 000	-0.13 ± 0.061	0.84 ± 0.054
ρ X	-	-0.52	-	-
ρ Y	-	0.77	-	-

Table D.14.: Random model 2, $B^0 \to D^-\pi^+\pi^0$

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	252	-0.11	0.074 ± 0.065	0.87 ± 0.059
$D_0^*(2400)^0$ Y	252	-0.042	-0.042 ± 0.075	1.1 ± 0.065
$D_2^*(2462)^0$ X	252	-0.14	0.040 ± 0.070	0.99 ± 0.065
$D_2^*(2462)^0$ Y	252	-0.17	0.073 ± 0.065	0.92 ± 0.058
NR	252	0.16	0.021 ± 0.072	1.0 ± 0.070
D^*	252	0.31	0.10 ± 0.067	0.98 ± 0.057
Continuum BG	252	14 000	0.18 ± 0.073	1.0 ± 0.058
Combinatoric BG	252	25 000	-0.18 ± 0.064	0.92 ± 0.053
ρ X	-	0.71	-	-
ρ Y	-	0.21	-	-

Table D.15.: Random model 3, $B^0 \to D^-\pi^+\pi^0$

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	258	-0.095	-0.0068 ± 0.066	0.99 ± 0.058
$D_0^*(2400)^0$ Y	258	0.018	0.032 ± 0.072	1.0 ± 0.063
$D_2^*(2462)^0$ X	258	-0.037	0.30 ± 0.064	0.90 ± 0.060
$D_2^*(2462)^0$ Y	258	0.20	-0.20 ± 0.065	0.95 ± 0.054
NR	258	0.13	-0.044 ± 0.066	0.97 ± 0.060
D^*	258	0.33	0.18 ± 0.069	0.99 ± 0.071
Continuum BG	258	14 000	0.065 ± 0.062	0.91 ± 0.060
Combinatoric BG	258	25 000	-0.096 ± 0.068	0.96 ± 0.062
ρ X	-	-0.46	-	-
ρ Y	-	0.66	-	-

Table D.16.: Random model 4, $B^0 \to D^-\pi^+\pi^0$

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	253	0.038	0.089 ± 0.063	0.88 ± 0.062
$D_0^*(2400)^0$ Y	253	0.069	-0.19 ± 0.061	0.89 ± 0.053
$D_2^*(2462)^0$ X	253	0.18	-0.14 ± 0.065	0.92 ± 0.057
$D_2^*(2462)^0$ Y	253	-0.20	-0.0055 ± 0.067	0.91 ± 0.065
NR	253	-0.032	-0.023 ± 0.060	0.86 ± 0.058
D^*	253	0.35	0.19 ± 0.073	1.00 ± 0.057
Continuum BG	253	14 000	0.18 ± 0.069	0.98 ± 0.065
Combinatoric BG	253	25 000	-0.20 ± 0.063	0.83 ± 0.055
ρ X	-	0.48	-	-
ρ Y	-	0.63	-	-

Table D.17.: Random model 5, $B^0 \to D^-\pi^+\pi^0$

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	292	-0.022	0.14 ± 0.060	0.95 ± 0.055
$D_0^*(2400)^0$ Y	292	-0.040	-0.020 ± 0.060	0.94 ± 0.053
$D_2^*(2462)^0$ X	292	0.17	0.13 ± 0.080	1.1 ± 0.061
$D_2^*(2462)^0$ Y	292	0.28	-0.12 ± 0.057	0.90 ± 0.044
NR	292	0.049	0.19 ± 0.070	1.0 ± 0.063
D^*	292	0.35	0.23 ± 0.060	0.94 ± 0.050
Continuum BG	292	14 000	0.12 ± 0.061	0.95 ± 0.050
Combinatoric BG	292	25 000	-0.21 ± 0.057	0.86 ± 0.046
ρ X	-	0.83	-	-
ρ Y	-	-0.74	-	-

Table D.18.: Random model 6, $B^0 \to D^-\pi^+\pi^0$

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	260	-0.095	-0.11 ± 0.082	1.1 ± 0.080
$D_0^*(2400)^0$ Y	260	-0.10	-0.014 ± 0.078	1.0 ± 0.072
$D_2^*(2462)^0$ X	260	-0.19	0.17 ± 0.068	0.97 ± 0.066
$D_2^*(2462)^0$ Y	260	-0.14	0.076 ± 0.076	1.1 ± 0.077
NR	260	-0.062	0.10 ± 0.060	0.83 ± 0.059
D^*	260	0.34	0.034 ± 0.062	0.90 ± 0.052
Continuum BG	260	14 000	0.21 ± 0.064	0.92 ± 0.055
Combinatoric BG	260	25 000	-0.15 ± 0.060	0.82 ± 0.052
ρ X	-	-0.68	-	-
ρ Y	-	0.55	-	-

Table D.19.: Random model 7, $B^0 \to D^-\pi^+\pi^0$

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	289	$-0.000\,78$	0.061 ± 0.068	1.1 ± 0.058
$D_0^*(2400)^0$ Y	289	0.18	-0.14 ± 0.066	1.0 ± 0.064
$D_2^*(2462)^0$ X	289	0.16	-0.21 ± 0.065	0.98 ± 0.065
$D_2^*(2462)^0$ Y	289	0.16	-0.17 ± 0.073	1.00 ± 0.085
NR	289	0.088	0.014 ± 0.072	1.00 ± 0.060
D^*	289	0.34	0.073 ± 0.067	0.95 ± 0.059
Continuum BG	289	14 000	0.24 ± 0.065	0.97 ± 0.056
Combinatoric BG	289	25 000	-0.20 ± 0.058	0.89 ± 0.045
ρ X	-	-0.99	-	-
ρ Y	-	-0.37	-	-

Table D.20.: Random model 8, $B^0 \to D^-\pi^+\pi^0$

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	298	0.059	0.062 ± 0.073	1.1 ± 0.060
$D_0^*(2400)^0$ Y	298	-0.068	-0.14 ± 0.057	0.91 ± 0.048
$D_2^*(2462)^0$ X	298	-0.22	-0.085 ± 0.053	0.84 ± 0.039
$D_2^*(2462)^0$ Y	298	-0.15	0.17 ± 0.057	0.91 ± 0.053
NR	298	0.084	0.086 ± 0.072	1.1 ± 0.059
D^*	298	0.32	0.23 ± 0.066	1.0 ± 0.057
Continuum BG	298	14 000	-0.012 ± 0.055	0.87 ± 0.048
Combinatoric BG	298	25 000	-0.082 ± 0.055	0.85 ± 0.045
ρ X	-	0.84	-	-
ρ Y	-	-0.76	-	-

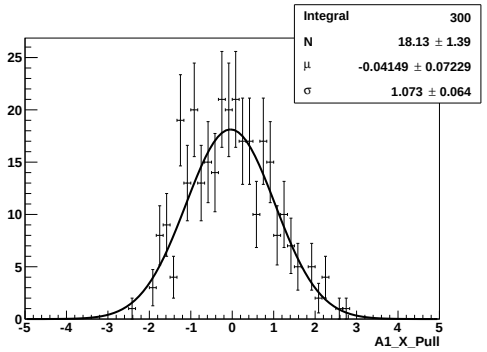
Table D.21.: Random model 9, $B^0 \to D^-\pi^+\pi^0$

Fit Parameter	Entries	True	Pull Mean	Pull Width
$D_0^*(2400)^0$ X	285	0.092	0.069 ± 0.069	1.0 ± 0.060
$D_0^*(2400)^0$ Y	285	-0.18	0.12 ± 0.066	1.00 ± 0.060
$D_2^*(2462)^0$ X	285	-0.18	0.0070 ± 0.068	1.0 ± 0.061
$D_2^*(2462)^0$ Y	285	0.033	0.042 ± 0.067	1.0 ± 0.064
NR	285	0.098	-0.048 ± 0.066	1.0 ± 0.053
D^*	285	0.33	0.037 ± 0.060	0.91 ± 0.058
Continuum BG	285	14 000	0.17 ± 0.054	0.82 ± 0.047
Combinatoric BG	285	25 000	-0.040 ± 0.061	0.85 ± 0.051
ρ X	-	0.96	-	-
ρ Y	-	-0.14	-	-

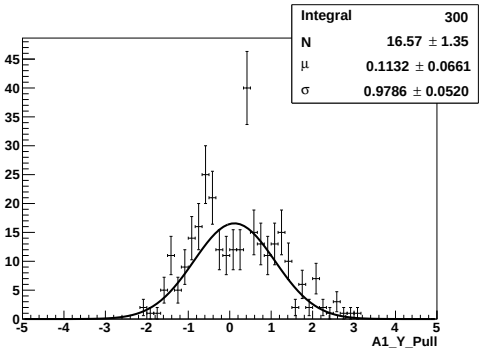
Table D.22.: Random model 10, $B^0 \to D^- \pi^+ \pi^0$

E. Results of ensemble tests of the data model

E.1. Pull distributions for $B^- \to D^0\pi^-\pi^0$

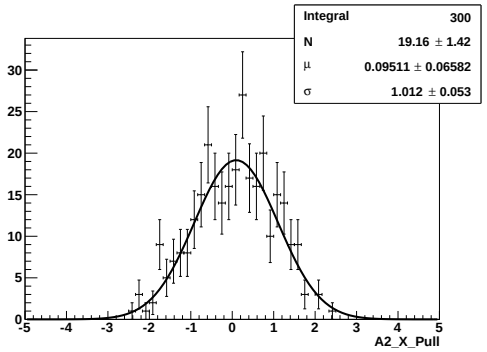


(a) Real part

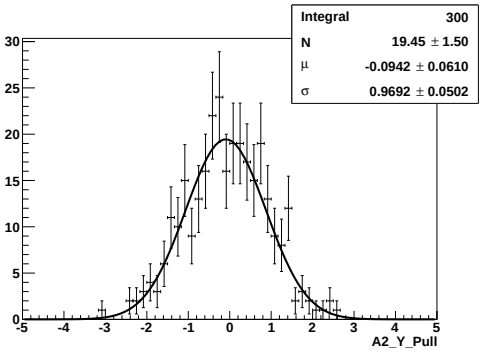


(b) Imaginary part

Figure E.44.: Pull distribution of the $D_0^*(2400)^0$ contribution



(a) Real part



(b) Imaginary part

Figure E.45.: Pull distribution of the $D_2^*(2462)^0$ contribution

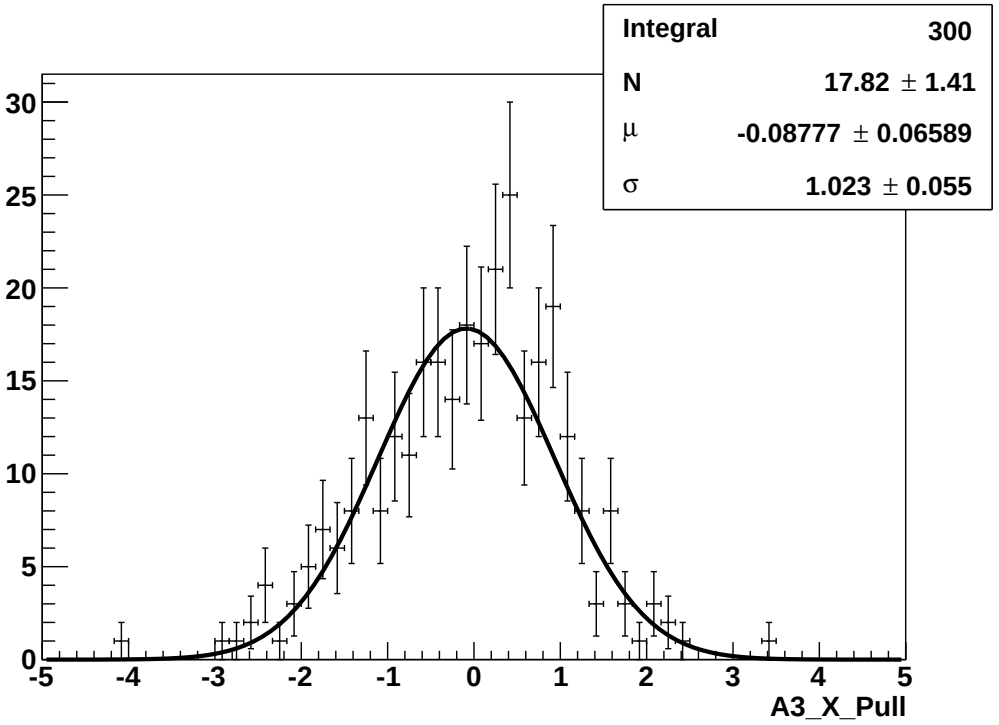
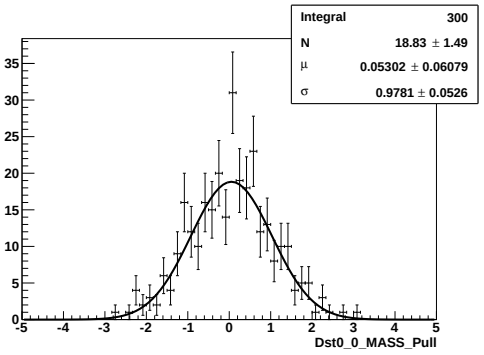
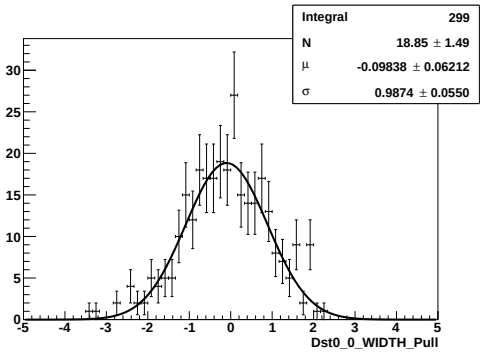


Figure E.46.: Pull distribution of the non resonant contribution

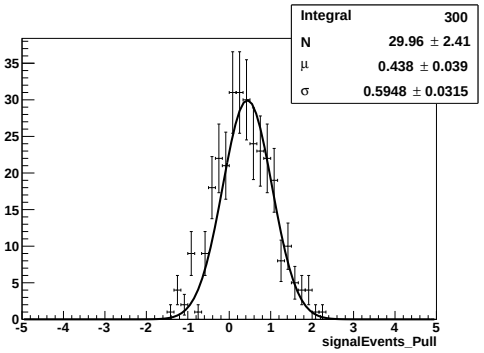


(a) $D_0^*(2400)^0$ mass

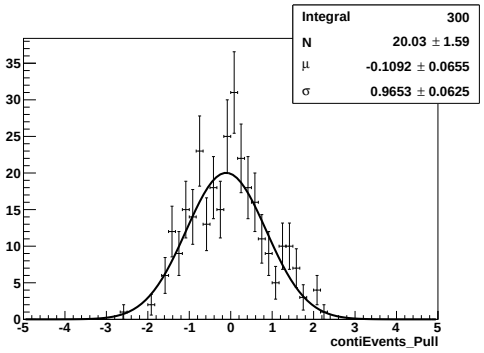


(b) $D_0^*(2400)^0$ width

Figure E.47.: Pull distribution of the $D_2^*(2462)^0$ $D_0^*(2400)^0$ parameters

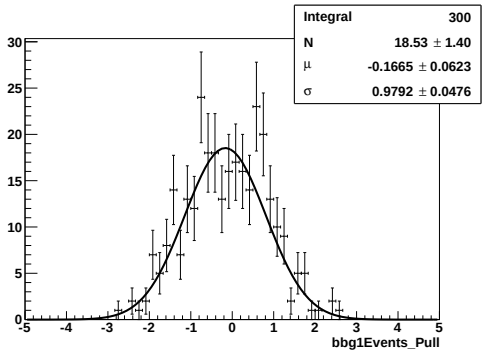


(a) Signal yield (constrained)

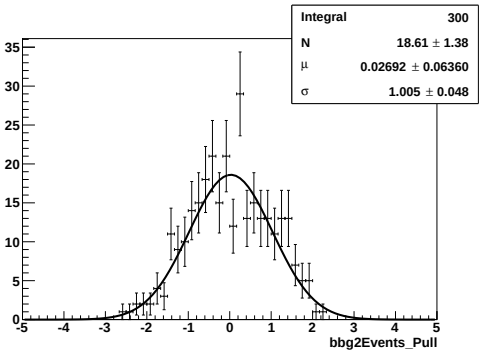


(b) Continuum background

Figure E.48.: Pull distribution of yields

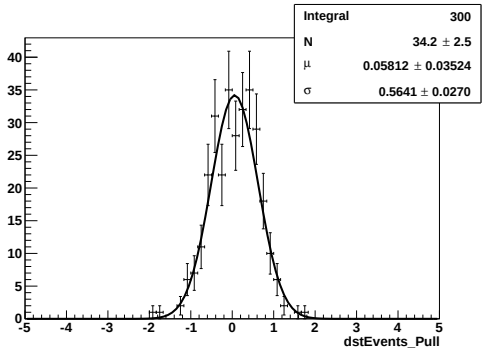


(a) Mixed combinatoric background

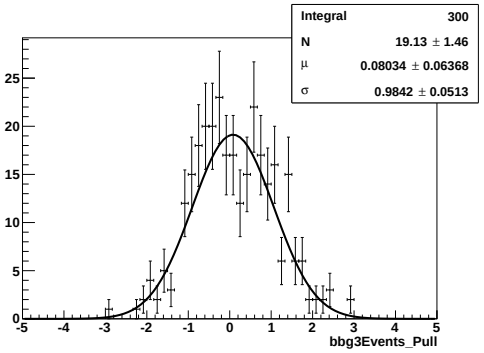


(b) Charged combinatoric background

Figure E.49.: Pull distribution of yields



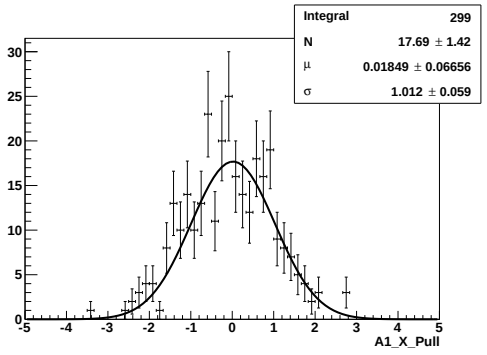
(a) Good signal



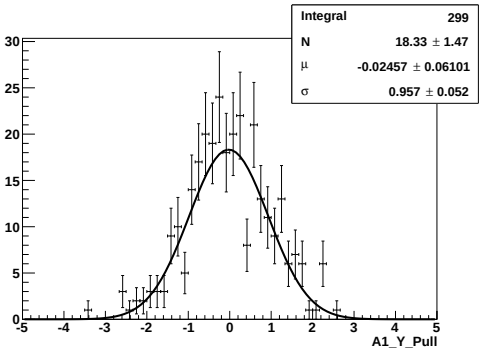
(b) Bad signal

Figure E.50.: Pull distribution of the D^* contribution (constrained).

E.2. Pull distributions for $B^0 \to D^- \pi^+ \pi^0$

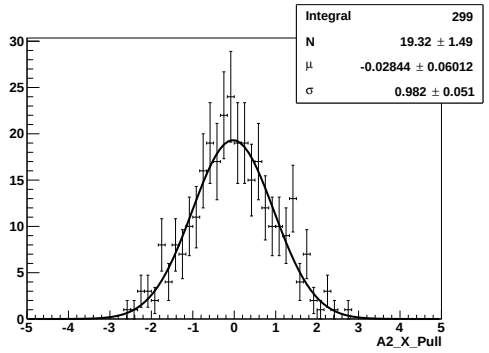


(a) Real part

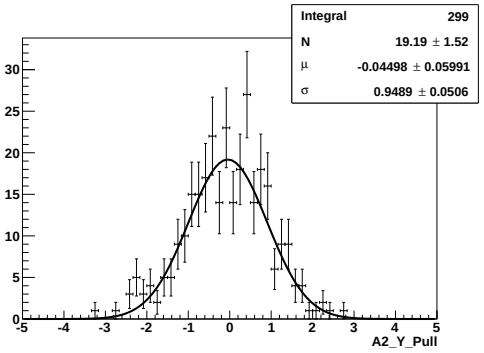


(b) Imaginary part

Figure E.51.: Pull distribution of the $D_0^*(2400)^0$ contribution

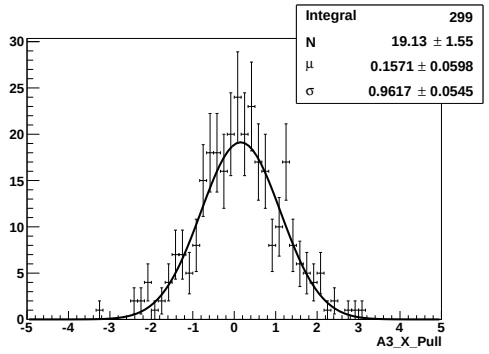


(a) Real part

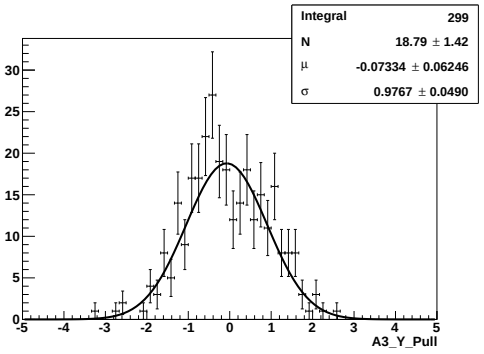


(b) Imaginary part

Figure E.52.: Pull distribution of the $D_2^*(2462)^0$ contribution

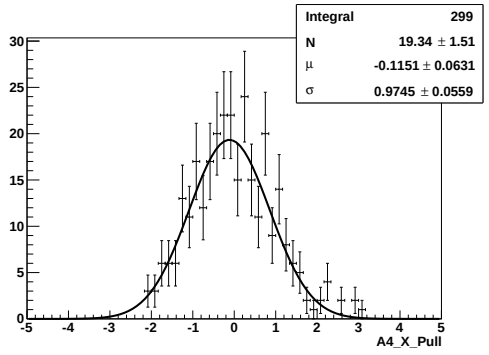


(a) Real part

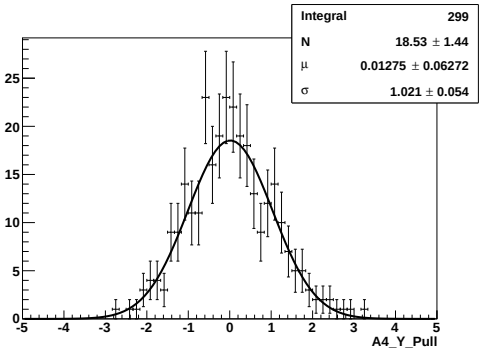


(b) Imaginary part

Figure E.53.: Pull distribution of the $D_0^*(2400)^-$ contribution



(a) Real part



(b) Imaginary part

Figure E.54.: Pull distribution of the $D_2^*(2462)^-$ contribution

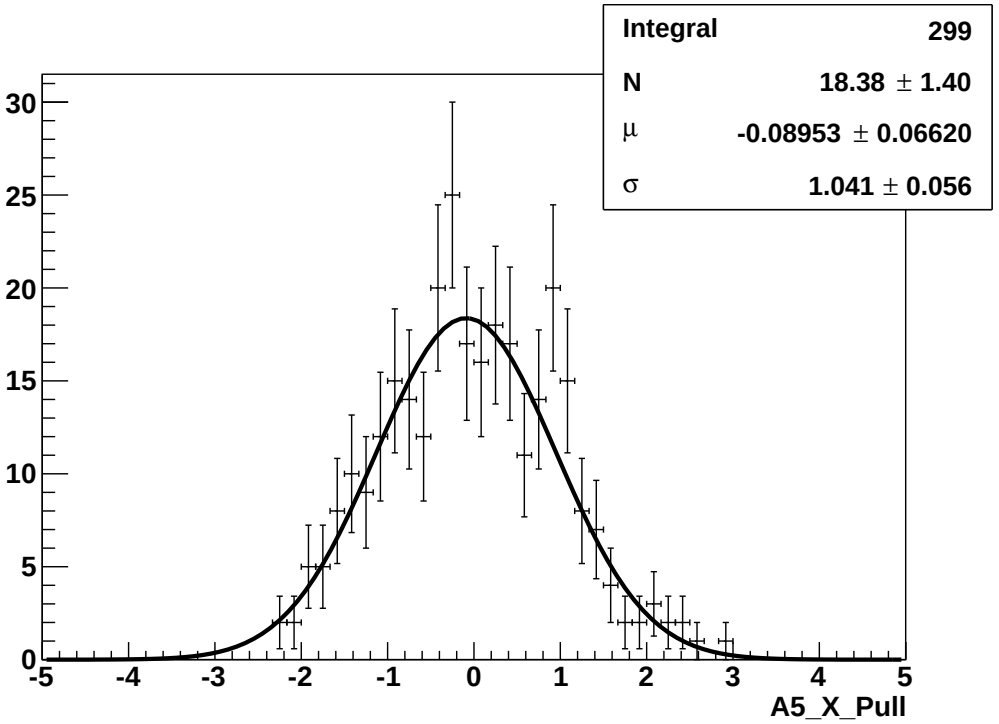
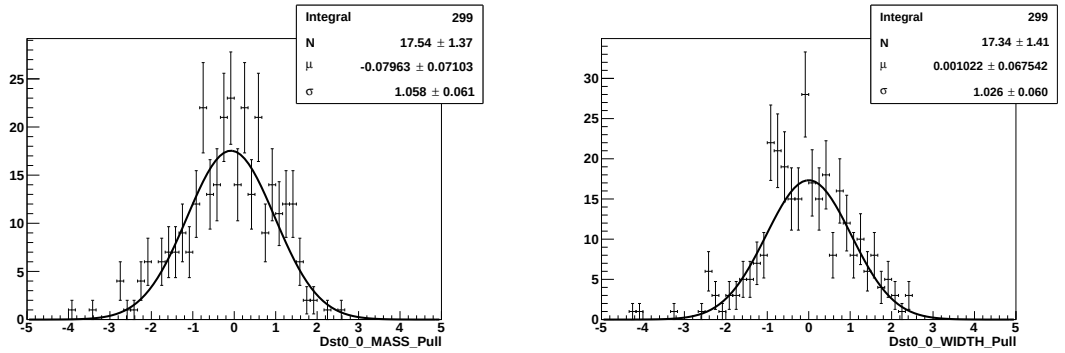
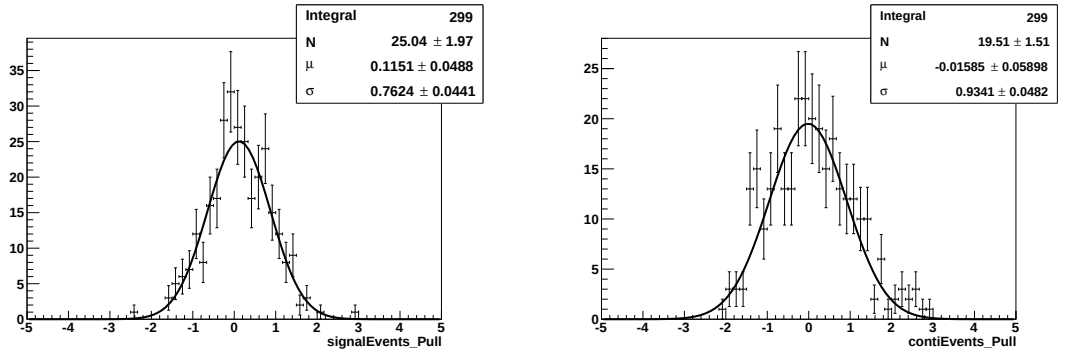


Figure E.55.: Pull distribution of the non resonant contribution



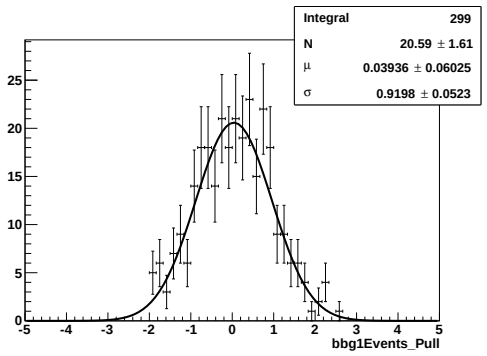
(a) $D_0^*(2400)^0$ mass (b) $D_0^*(2400)^0$ width

Figure E.56.: Pull distribution of the $D_2^*(2462)^0$ $D_0^*(2400)^0$ parameters

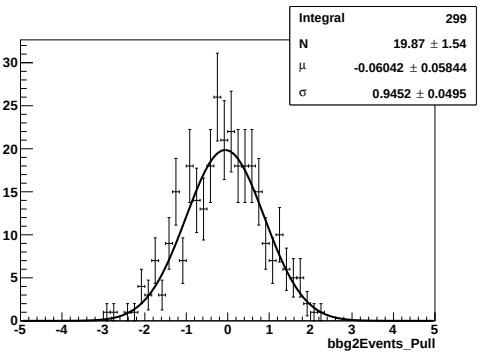


(a) Signal yield (constrained) (b) Continuum background

Figure E.57.: Pull distribution of yields

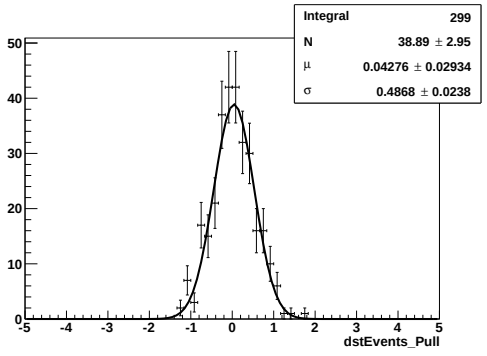


(a) Mixed combinatoric background

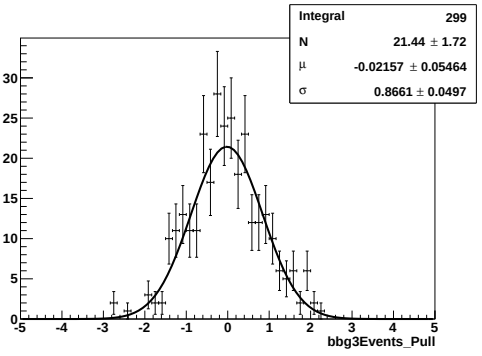


(b) Charged combinatoric background

Figure E.58.: Pull distribution of yields



(a) Good signal



(b) Bad signal

Figure E.59.: Pull distribution of the D^* contribution (constrained).

E.3. Variation of migrating signal fraction

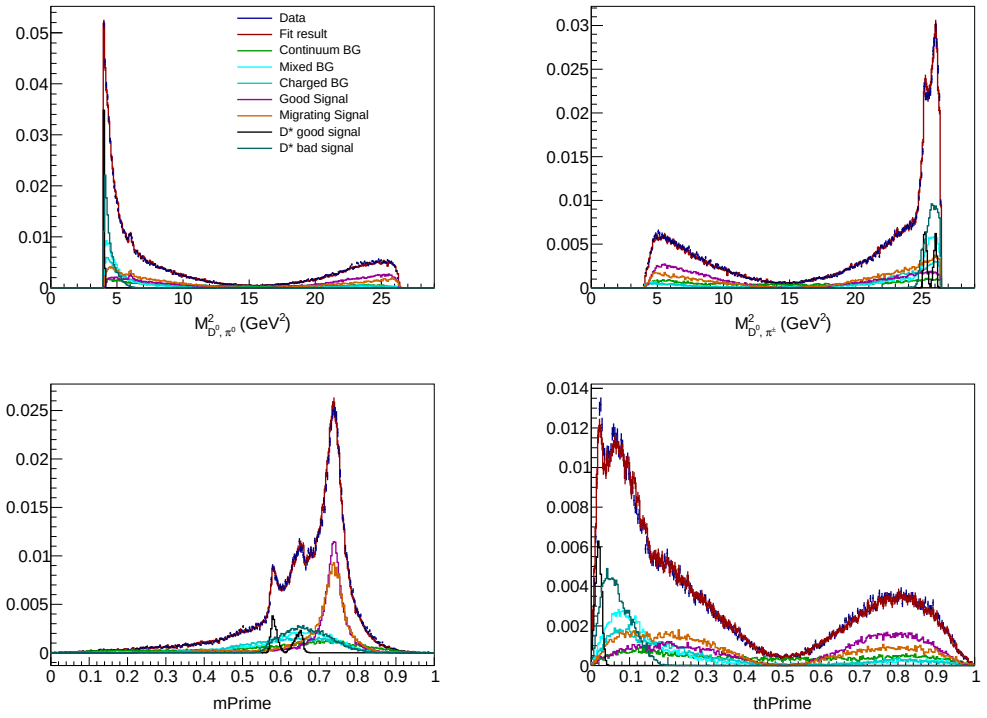


Figure E.60.: Projections to the normal and squared Dalitz plot variables of the fit to data, migrating signal decreased by 20% ($B^- \to D^0\pi^-\pi^0$).

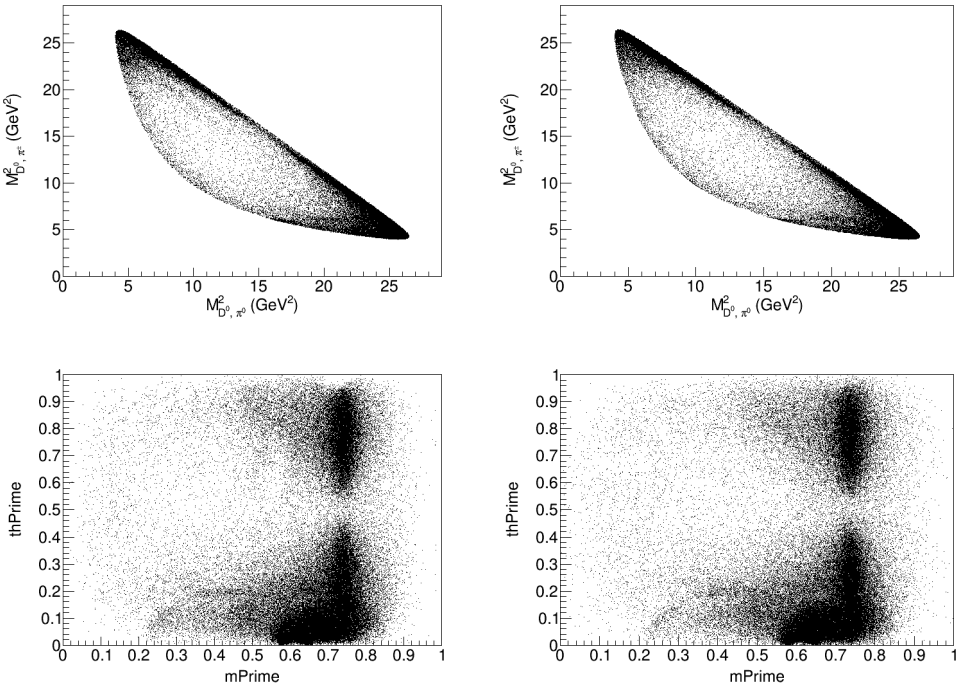


Figure E.61.: Scatter plots in the normal (top) and squared (bottom) Dalitz plot showing the input sample (left) and the fit result (right), migrating signal idecreased by 20% ($B^- \to D^0\pi^-\pi^0$).

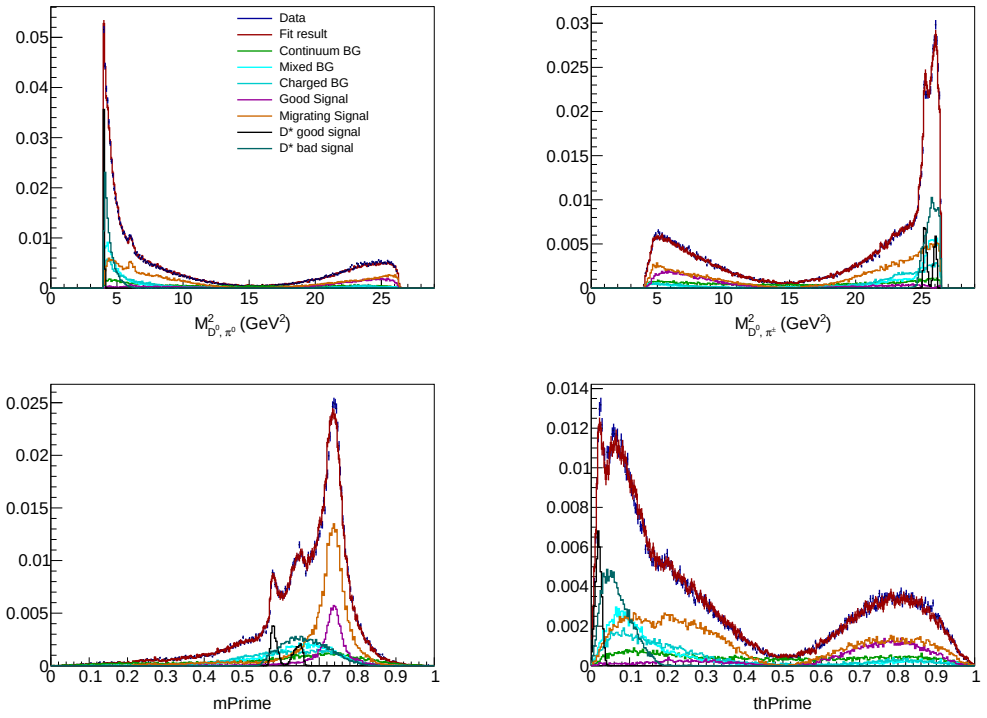


Figure E.62.: Projections to the normal and squared Dalitz plot variables of the fit to data, migrating signal increased by 20% ($B^- \to D^0\pi^-\pi^0$).

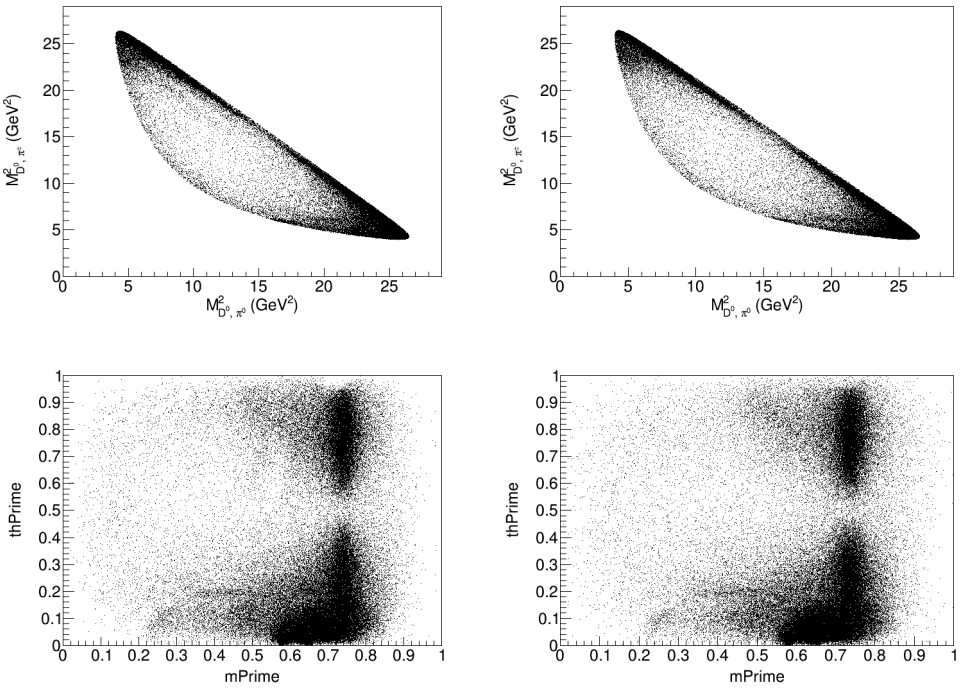


Figure E.63.: Scatter plots in the normal (top) and squared (bottom) Dalitz plot showing the input sample (left) and the fit result (right), migrating signal increased by 20% ($B^- \to D^0\pi^-\pi^0$).

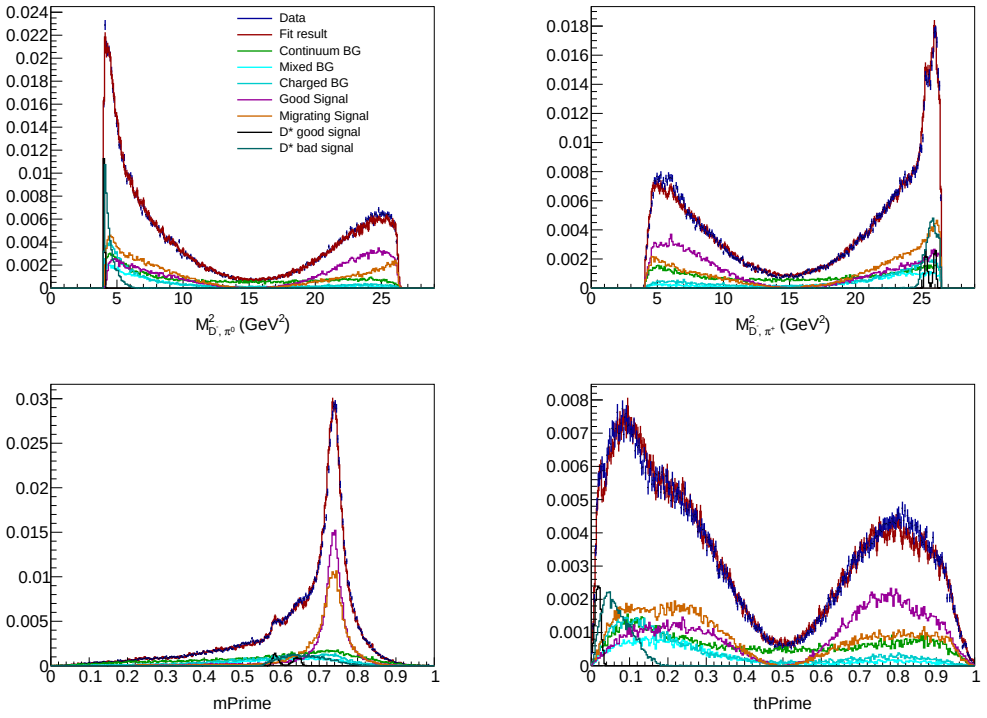


Figure E.64.: Projections to the normal and squared Dalitz plot variables of the fit to data, migrating signal decreased by 20% ($B^0 \to D^-\pi^+\pi^0$).

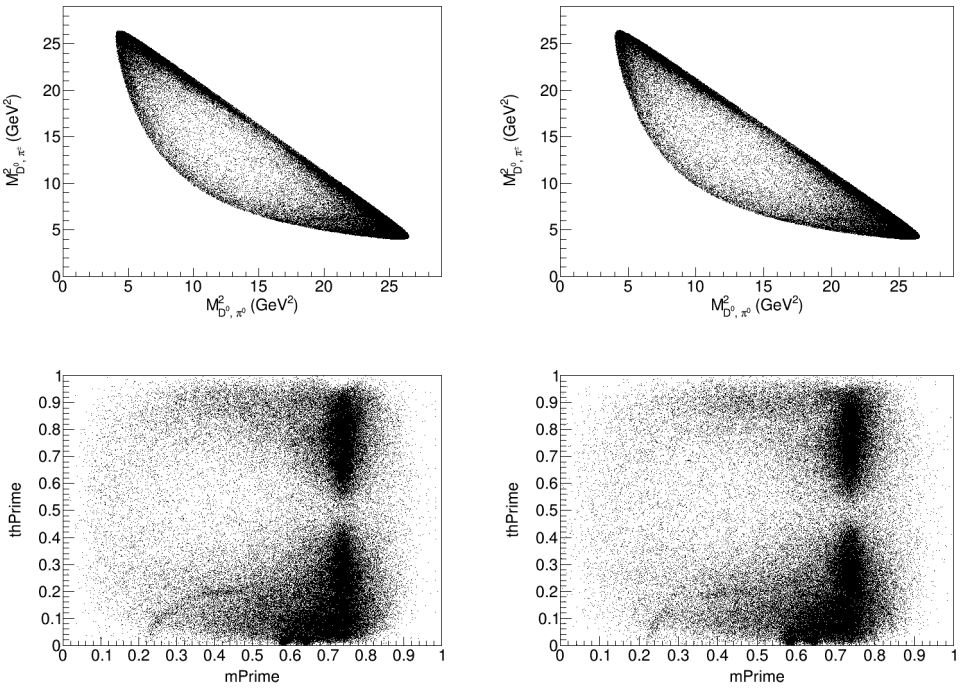


Figure E.65.: Scatter plots in the normal (top) and squared (bottom) Dalitz plot showing the input sample (left) and the fit result (right), migrating signal idecreased by 20% ($B^0 \to D^-\pi^+\pi^0$).

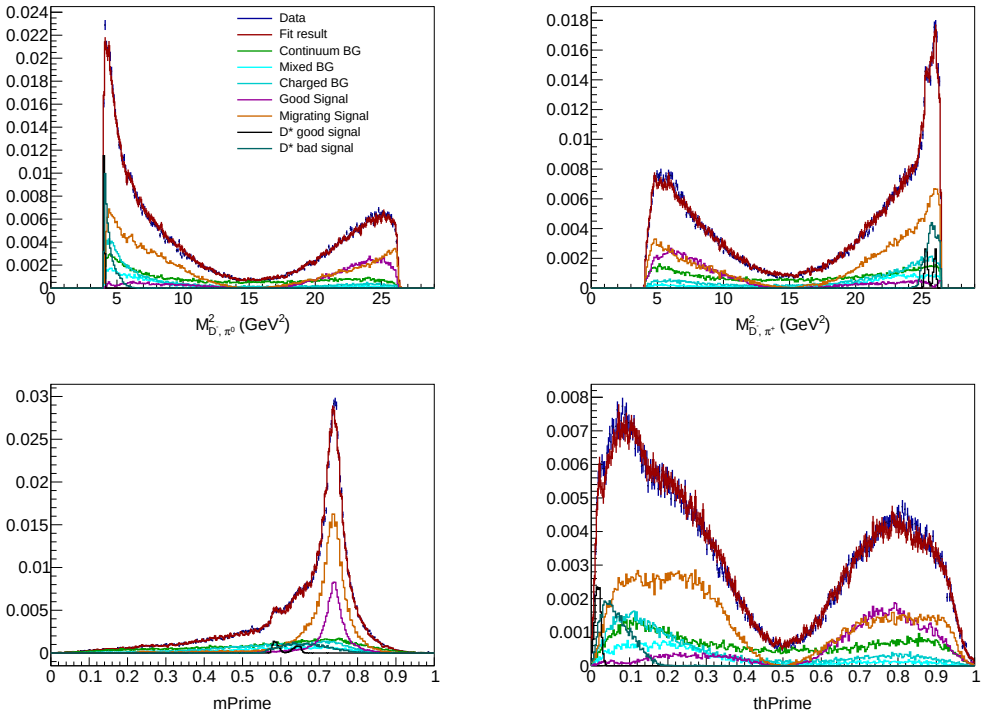


Figure E.66.: Projections to the normal and squared Dalitz plot variables of the fit to data, migrating signal increased by 20% ($B^0 \to D^-\pi^+\pi^0$).

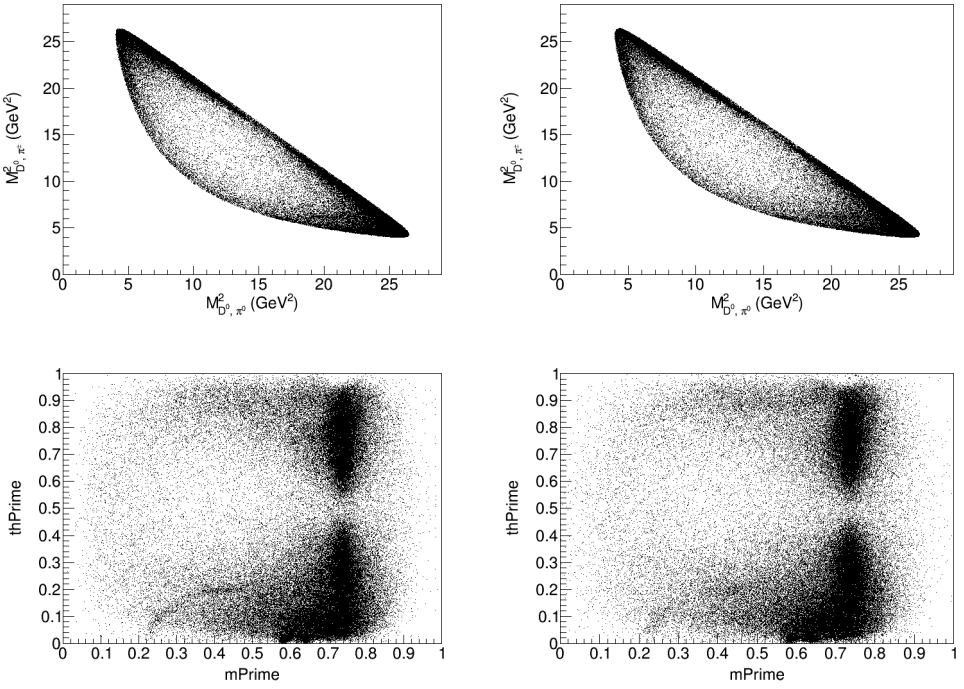


Figure E.67.: Scatter plots in the normal (top) and squared (bottom) Dalitz plot showing the input sample (left) and the fit result (right), migrating signal increased by 20% ($B^0 \to D^-\pi^+\pi^0$).

Danksagung

Mein Dank geht an erster Stelle an Prof. Michael Feindt, nicht nur für die Möglichkeit, diese Analyse überhaupt durchführen zu können, sondern auch für den fachlichen Beistand und Hilfe bei jedwedem Problem. Prof. Günter Quast danke ich für die trotz hohem Promotionsaufkommen bereitwillige Übernahme des Korreferats und seine wertvollen Kommentare.

Ohne die großartige Unterstützung der Post-Docs Pablo Goldenzweig, Thomas Kuhr, Martin Heck und Anže Zupanc wäre diese Arbeit nicht möglich gewesen. Hervorheben möchte ich davon Thomas, der immer wieder gute Ideen hat, und Pablo, der mir nicht nur zur späten Stunde über Skype half, Probleme zu lösen, sondern es auch verstand, mich zu motivieren und mir über Rückschläge hinwegzuhelfen. Vielen Dank dafür!

Ein besonderer Dank geht an die B-Gruppe, für außerordentliche Hilfsbereitschaft und Teamgeist und die angenehmste Arbeitsatmosphäre, die man sich vorstellen kann. Insbesondere Johannes Grygier (der alles weiß), Michael Ziegler (der immer mitgefiebert hat), Matthias Huschle (der sehr viel zugehört hat) und Bastian Kronenbitter (der immer wusste, was zu tun ist) haben mir sehr viel geholfen, wofür ich mich ganz herzlich bedanke!

Auch dem EKP-Admin-Team gebührt Dank, sie halten auch unter Einsatz ihrer Freizeit das ganze *Geraffel* am laufen. Der ekpplusctl danke ich für außergewöhnliches Durchhaltevermögen ganz ohne CPU-Kühlung.

Meinen Eltern und Schwiegereltern danke ich für die fortwährende Unterstüzung, sei es in Form von Naturalien, fahrbaren Untersätzen oder aber in Form von mir entgegengebrachtem Vertrauen.

Meine Frau Nicole war insbesondere in der Schlussphase eine große Hilfe für mich. Sie bot mir immer Rückhalt und unterstützte mich trotz eigener Prüfungen immer tatkräftig. Dafür und für alles andere bedanke ich mich herzlich bei ihr!