Measurements of CP asymmetries in $D^0 \to \pi^+\pi^-\pi^0$ decays in Dalitz plot regions at *Belle II*

by

Emma R. Oxford

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
at
Carnegie Mellon University
Department of Physics
Pittsburgh, Pennsylvania

Advised by Professor Roy Briere

April 26, 2023

Abstract

We perform a search for charge-parity violation (CPV) in the density of events in a binned Dalitz plot of D^*-tagged $D^0 \rightarrow \pi^+\pi^-\pi^0$ decays. The data used for this analysis is $362\,\mathrm{fb}^{-1}$ of $\Upsilon(4S)$ data collected at *Belle II* from 2019–2022. We develop a novel binned Dalitz analysis method to determine local CPV; our method carefully accounts for nuisance asymmetries not due to CPV using the data itself rather than relying on simulations. All final asymmetry measurements are consistent with the no-CPV hypothesis. Our analysis method will scale well as *Belle II* collects more data, and we expect both the statistical and systematic uncertainties on the results to decrease if the analysis is redone on a larger dataset.

Acknowledgments

JAKE
ESMERALDA
CLAIRE
ROYBRIERE
HARRY

CHARM
JITENDRA
MARGI
BRADLEY
ENERGY
RENU

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Conservation laws are critical to the way physicists understand the universe. As the field of particle physics began to emerge in the mid-twentieth century, it was no exception. One particularly important concept for particle physicists was parity conservation, which holds that the laws of physics do not change under mirror symmetry: the mirror image of any physical process is also a valid physical process. This seemed obvious to particle physicists; it just made sense that this would be true. But in 1956, Lee and Yang pointed out that although it had been verified in the strong and electromagnetic interactions, parity conservation had never been confirmed in the weak interaction [1]. In 1957, in the famous cobalt-60 β-decay experiment, Wu and others found that parity was *not* conserved in the weak interaction [2]. (Lee and Yang won the 1957 Nobel Prize in Physics for their work. Madame Wu was notably denied this recognition, despite being nominated multiple times.)

The discovery of parity violation prompted something of a crisis in particle physics: did the weak interaction preserve any discrete symmetry? Very quickly, charge-parity (CP) conservation was proposed as a possibility [3]. CP conservation means that if you first charge-conjugate the particles involved in a physical process, *then* the mirror image is also a valid physical process. But this was also eventually found not to be a universal symmetry. The first observation of charge-parity violation (CPV) occurred in 1964 [4], and the search for additional examples as well as the development of a model to explain them has been a central part of particle physics ever since.

Our current understanding of CPV can be broken down into what is included in the Standard Model (SM) that was developed in the 1960s and 1970s and what must be explained another way. Even before experimental confirmation of quarks [5, 6] and the development of the CKM matrix [7, 8], CPV in the kaon sector had already been observed [4]. However, the original quark model only postulated the existence of three quark flavors – up, down, and strange – and the mathematics of this model could not account for CPV. In 1964, a fourth quark – charm – was postulated [9]; the phenomenology of the charm quark was further developed in 1970 [10]. In 1974, the charm quark was observed experimentally [11, 12], but a model with four quarks was still not sufficient to explain CPV. The CKM matrix called for a six-quark model, which would be able to account for CPV via a complex phase. The fifth quark – bottom, or beauty – was discovered in 1977 [13], and although the sixth quark remained elusive (the top quark was finally observed in 1995 at a much larger mass than anyone anticipated [14]), this was enough to convince particle physicists of the validity of

the quark model and CKM matrix and send them on a decades-long search for CPV in B mesons (the bottom sector) and D mesons (the charm sector).

CPV in the bottom sector was first observed in the early 2000s at BABAR [15] and *Belle* [16] in neutral B mesons. Additional measurements, plus observations of CPV in $B^\pm$ and B_s meson decays, came in the following decade [17–20]. These measurements are all consistent with the level of CPV predicted by the SM.

Despite this success, CPV in the charm sector remained elusive. Evidence for D^0-$\overline{D}^0$ mixing (one mechanism through which CPV could occur) was seen as early as 2007 [21–23], and the first observation of mixing in a single measurement was seen in 2013 [24], but these studies did not find evidence for CPV. CPV in the D meson system was finally observed in 2019 by LHCb [25]. This measurement is also consistent with the amount of CPV predicted by the SM.

But what drives physicists to study CPV is not just an attempt to confirm the SM, but rather a desire to prove it incorrect, or at least incomplete. The level of CPV predicted by the current CKM matrix, for example, is not large enough to explain the dominance of matter over antimatter in our universe [26]. There must therefore be sources of CPV coming from beyond-the-SM physics, or "New Physics."

Detecting CPV from New Physics is one of the main goals of the *Belle II* experiment's physics program [27]. Using 362 fb^{-1} of *Belle II* data collected at the $\Upsilon(4S)$ resonance, this analysis focuses on searching for CPV in one particular decay of the D^0 meson: D^*-tagged $D^0 \to \pi^+\pi^-\pi^0$ (charge-conjugate decays are implied throughout unless otherwise stated). This decay involves a $c \to d$ quark transition via the weak interaction, so based on the elements of the CKM matrix, it is suppressed relative to decays containing $c \to s$ transitions (i.e., D^0 decays with a K^- or $\overline{K}^0$ in the final state). This is easiest to see with the Wolfenstein parameterization of the CKM matrix. The general elements of the CKM matrix are:

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.1}$$

The Wolfenstein parameterization uses four real parameters: A, λ, ρ, and η. ρ and η are the CP-violating parameters. $\lambda \ll 1$, and the CKM elements are expanded in increasing powers of λ [28]:

$$V_{CKM} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \tag{1.2}$$

The elements along the diagonal, which represent the rates of $u \to d$, $c \to s$, and $t \to b$ quark transitions, are all $\mathcal{O}(1)$. But away from the diagonal, the power of λ increases. Of particular interest, V_{cd}, i.e., the $c \to d$ transition rate, is $\mathcal{O}(\lambda)$, so it is suppressed relative to $c \to s$.

A large amount of data is required to analyze suppressed decays, but they are good places to look for New Physics, as the lower amplitude of the SM processes means it should be easier to see any New Physics process. In a Cabibbo-favored (CF) decay, the amplitude of the SM processes is so large that even if CPV from New Physics were present, it might be very difficult to observe.

The other reason we consider specifically $D^0 \rightarrow \pi^+\pi^-\pi^0$ is the presence of the neutral π^0 meson. Reconstruction of π^0 mesons from $\pi^0 \rightarrow \gamma\gamma$ decays, using precise detection of photon showers in the electromagnetic calorimeter (ECL), is expected to be one of the strengths of *Belle II* [29, 30]. As we are still relatively early in *Belle II* data-taking, this analysis will test *Belle II*'s π^0 performance capabilities and inform future analyses of D meson decays containing π^0s.

In order to search for CPV, we develop a novel binned analysis method for the $D^0 \rightarrow \pi^+\pi^-\pi^0$ Dalitz plot. Importantly, this analysis method allows us to account for nuisance asymmetries (i.e., asymmetries caused by sources other than CPV) using the data rather than relying on simulations. It also accounts for peaking backgrounds, and it is broadly scalable and adaptable to larger amounts of data. Since our current data sample is only a small fraction of what *Belle II* will eventually collect, we set out to develop an analysis technique that could be used for later analyses at *Belle II* as well.

The structure of this thesis is as follows:

- Chapter 2 describes previous examples of Dalitz analyses in the literature and discusses their various uses. This chapter is supplemented by Appendix A and Appendix B, which go into more detail about the history of the development of Dalitz plots as well as the specific methodologies of different Dalitz analyses.

- Chapter 3 discusses the SuperKEKB accelerator and the *Belle II* detector. This analysis, like all physics analyses done at *Belle II*, relies on the excellent performance of the detector and the various subdetectors that comprise it.

- Chapter 4 describes the dataset used for the analysis in this thesis and details our procedure for event selection.

- Chapter 5 describes in detail our novel binned analysis method. This chapter is supplemented by Appendix C, which goes into more detail about additional analyses that may be performed to complement the results obtained with our method.

- Chapter 6 presents the results of our method when performed on simulated data (also called Monte-Carlo data or MC). Since this simulated data does *not* contain CPV, we should see asymmetry measurements consistent with zero. The results described in this chapter serve as a test of our methodology to confirm that we do not see false positives.

- Chapter 7 presents the results of our analysis on $362\,\mathrm{fb}^{-1}$ of *Belle II* data.

- Chapter 8 details the estimation of systematic uncertainties to apply to the results from chapter 7.

- Chapter 9 summarizes the procedure and results from this analysis and looks ahead to how these results could be improved upon and expanded in the future.

Chapter 2

Dalitz Analyses

This chapter discusses the purpose (section 2.1) and selected past uses (section 2.2) of Dalitz plots as well as the motivation for the analysis described in this thesis (section 2.3). For more detailed discussions of particular types of Dalitz analyses used to search for CPV, refer to chapter 5, Appendix A, and Appendix B.

2.1 Overview of Dalitz plots

The analysis presented in this thesis will use data binned in regions of the Dalitz plot to measure a local CP asymmetry. A Dalitz plot is a way to visualize a three-body decay, and Dalitz analyses have been used in particle physics for more than 70 years [31, 32]. The principle behind a Dalitz plot is as follows: if A is a particle of mass m_A and B, C, and D are spin-0 particles of masses m_B, m_C, and m_D, respectively (the masses of the three decay products need not be equal), then the decay $A \to B\,C\,D$ can be completely described by only two variables. It is common to use the invariant mass squared for two pairs of decay products (e.g., m_{BC}^2 for the invariant mass squared of the BC system and m_{CD}^2 for the invariant mass squared of the CD system) as Dalitz plot axes in modern analyses. There are two main reasons why these axes are a sensible choice. The first reason can be demonstrated by some simple four-vector algebra. Using the convention $\hat{p} = (E, \vec{p})$, four-momentum conservation gives:

$$
\begin{aligned}
\hat{p}_A &= \hat{p}_B + \hat{p}_C + \hat{p}_D, \\
\hat{p}_A^2 &= (\hat{p}_B + \hat{p}_C + \hat{p}_D)^2, \\
m_A^2 &= m_B^2 + m_C^2 + m_D^2 + 2\hat{p}_B \cdot \hat{p}_C + 2\hat{p}_C \cdot \hat{p}_D + 2\hat{p}_B \cdot \hat{p}_C.
\end{aligned}
\tag{2.1}
$$

If we let $\hat{p}_{BC}$ be the four-momentum of the BC system, then $\hat{p}_{BC} = \hat{p}_B + \hat{p}_C$, and it follows that:

$$
\begin{aligned}
\hat{p}_{BC}^2 &= (\hat{p}_B + \hat{p}_C)^2, \\
m_{BC}^2 &= m_B^2 + m_C^2 + 2\hat{p}_B \cdot \hat{p}_C, \\
2\hat{p}_B \cdot \hat{p}_C &= m_{BC}^2 - m_B^2 - m_C^2.
\end{aligned}
\tag{2.2}
$$

Using similar equations for $2\hat{p}_B \cdot \hat{p}_D$ and $2\hat{p}_C \cdot \hat{p}_D$, we can substitute Equation 2.2 into Equation 2.1, and (after rearranging) we have:

$$m_A^2 + m_B^2 + m_C^2 + m_D^2 = m_{BC}^2 + m_{CD}^2 + m_{BD}^2. \tag{2.3}$$

But the left-hand side of Equation 2.3 is just a constant. This leads to the constraint:

$$\text{constant} = m_{BC}^2 + m_{CD}^2 + m_{BD}^2. \tag{2.4}$$

Equation 2.4 is one justification for the choice of invariant pair masses squared as Dalitz plot axes: since the three invariant pair masses squared must always sum to a constant, if two of them appear on the x and y axes, the diagonals $x + y = \text{constant}$ represent the third. Resonances in a particular pair of decay products, say B and C, appear as stripes perpendicular to the m_{BC}^2 axis. Determining the masses of intermediate resonances is often a main goal of Dalitz analyses, and this choice of axes makes that very easy.

The other reason this choice of axes is particularly useful is because, for this three-body decay, the differential decay probability is [33]:

$$d\Gamma = \frac{1}{(2\pi)^3} \frac{1}{32m_A^3} |\mathcal{A}|^2 dm_{BC}^2 dm_{CD}^2. \tag{2.5}$$

The only non-constant term in the coefficients in Equation 2.5 is $\mathcal{A}$, the decay amplitude. Thus, any structure (i.e., non-uniform density) in a Dalitz plot is due to the variation in $\mathcal{A}$ [34]. Another way of saying this is that the area in the Dalitz plot is proportional to the phase space of the decay. Again, since understanding the intermediate resonances in a decay is often a goal of these analyses, this makes a Dalitz plot a particularly useful visualization of the decay process.

Figure 2.1 shows what the Dalitz plot of a decay with no structure (i.e., no resonances) would look like. Figure 2.2 shows the kinematic boundaries of a Dalitz plot, which are set by the invariant mass of the decaying particle, plus the configuration of the decay products around the edge of the plot.

The remaining sections of this chapter discuss how Dalitz analyses have been used to study D meson decays. More detail on the history of Dalitz plots can be found in Appendix A.

2.2 Dalitz analyses in practice

In the last two decades, Dalitz analyses have been used both to search for CPV and to determine amplitude models for resonant decays (see Appendix A for a discussion of amplitude models). They have been used by many different collaborations, at a wide range of energy scales, and on numerous different decays. The mode analyzed in this thesis, $D^0 \rightarrow \pi^+\pi^-\pi^0$, has been analyzed with several types of Dalitz analyses going back more than fifteen years.

2.2.1 $D^0 \rightarrow \pi^+\pi^-\pi^0$ analyses

The BABAR collaboration has published several analyses on $D^0 \rightarrow \pi^+\pi^-\pi^0$. In 2006, they published the most precise measurements (at the time) of the ratios $\frac{\Gamma(D^0 \rightarrow \pi^+\pi^-\pi^0)}{\Gamma(D^0 \rightarrow K^-\pi^+\pi^0)}$ and
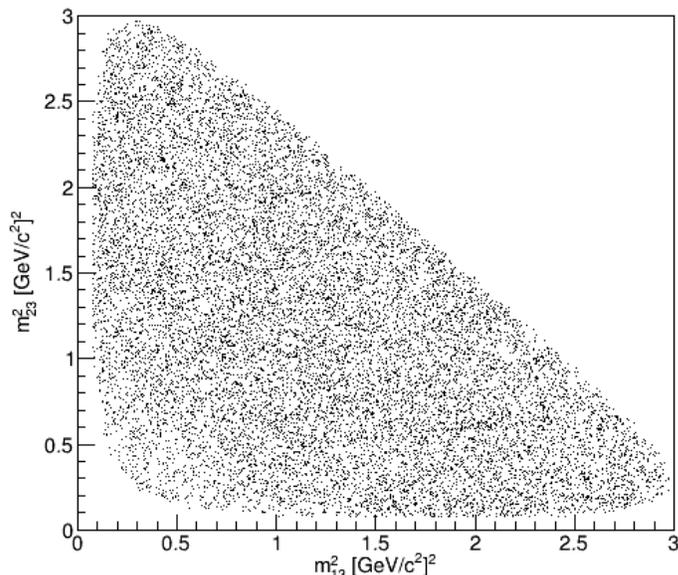
Figure 2.1: The area of a Dalitz plot reveals the phase-space structure of a decay. A uniform Dalitz plot, like this one, indicates that there are no resonances; resonances would appear as stripes along an axis.

$\frac{\Gamma(D^0 \to K^- K^+ \pi^0)}{\Gamma(D^0 \to K^- \pi^+ \pi^0)}$, calculated using Dalitz techniques [35]. In 2007, they published a Dalitz analysis of $D^0 \to \pi^+ \pi^- \pi^0$ from $B^\pm \to D^0 K^\pm$ decays using 315 fb^{-1} of $B\overline{B}$ data collected at or near the $\Upsilon(4S)$ resonance [36]. This analysis explicitly measured the SM CKM angle γ [33]. A theoretical follow-up to this paper analyzing the results in terms of isospin eigenstates was published the following year [37].

In 2008, BABAR published binned Dalitz analyses of $D^0 \to \pi^+ \pi^- \pi^0$ and $D^0 \to K^- K^+ \pi^0$ using a technique colloquially referred to as the "Miranda method" [38]. The Miranda method is a novel method for performing a binned analysis that involves measuring a significance in a particular bin of the Dalitz plot rather than a fractional asymmetry. Appendix B describes the Miranda method in detail, including its advantages over a traditional binned analysis and its limitations.

In 2008, *Belle II*'s predecessor, *Belle*, also published a Dalitz analysis of $D^0 \to \pi^+ \pi^- \pi^0$ using 532 fb^{-1} of $B\overline{B}$ data [39]. They measured the ratio $\frac{\Gamma(D^0 \to \pi^+ \pi^- \pi^0)}{\Gamma(D^0 \to K^- \pi^+ \pi^0)}$ as well as the CP asymmetry across the Dalitz plot. They used a binned analysis based on event counts instead of an amplitude analysis with an isobar model (see Appendix A for description) to reduce the model dependence of the CPV measurement. It is worth noting that this is *Belle*'s only analysis on this mode, and it was published before the full *Belle* dataset was available.

In 2015, LHCb published an analysis of this mode [40]. They analyzed the decay $D^0 \to \pi^+ \pi^- \pi^0$ using a technique called the "energy test," which was first described as a possible Dalitz-based technique in a 2011 paper by Williams [41]. The energy test is discussed in more detail in Appendix B.

None of these analyses found evidence for CPV. As mentioned in chapter 1, CPV in the
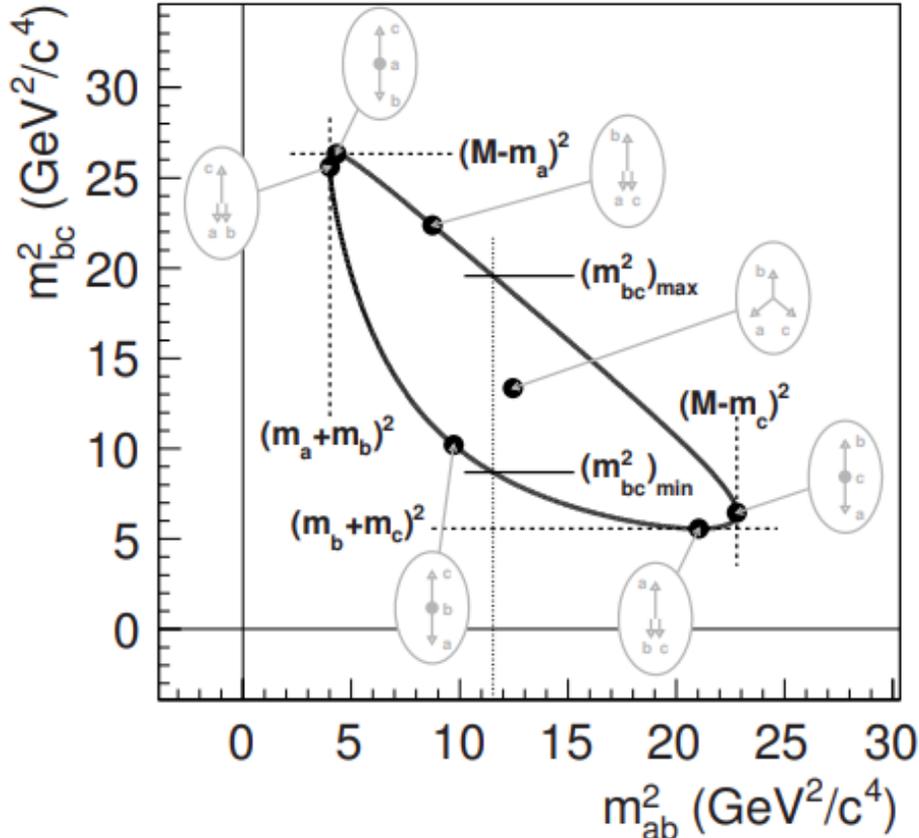
Figure 2.2: Kinematic boundaries of a Dalitz plot. The gray text demonstrates the relative orientation of the particles in the decay at these kinematic limits. The actual decay demonstrated in this plot is $B^0 \to \pi^- \overline{D}^0 K^+$, where $M = M_{B^0}$, $m_a = m_{\pi^-}$, $m_b = m_{\overline{D}^0}$, and $m_c = m_{K^+}$ [34].

charm sector was not observed until 2019, and it was found using two-body, not three-body, singly Cabibbo-suppressed (SCS) D^0 decays [25].

2.2.2 Other selected D decays

The decay $D^0 \to K^- \pi^+ \pi^0$ was studied with a Dalitz model by the CLEO II collaboration in 2001 [42]. They performed an amplitude analysis and calculated an integrated CP asymmetry. This analysis still provides the dominant contribution to the fit fraction $\frac{\Gamma(D^0 \to K^- \rho^+)}{\Gamma(D^0 \to K^- \pi^+ \pi^0)}$ reported by the Particle Data Group (PDG) [33]. The FOCUS collaboration published an isobar model of this mode in 2007 [43].

In 2002, an amplitude analysis of the $D^+ \to K^- \pi^+ \pi^+$ Dalitz plot was performed by the E791 collaboration at Fermilab [44]. This analysis was updated in 2006 [45] and still contributes to PDG averages for fit fractions of this decay [33].

The mode $D^+ \to K^+ K^- \pi^+$ was analyzed by CLEO-c with an amplitude analysis in 2008 [46] and by LHCb with a binned analysis (using the Miranda method) in 2011 [47]. LHCb also applied the Miranda method to a binned Dalitz analysis of $D^+ \to \pi^- \pi^+ \pi^+$ in 2014 [48].

Again, none of these analyses found evidence for CPV. However, they illustrate that Dalitz analyses are well-established in the study of charm physics. They further demonstrate the wide range of Dalitz analyses that exist. This one basic technique can be used with amplitude models, binned analyses (which have many variations), and the newer energy test method to extract a wealth of information from many different decays.

2.3 Motivation

Given that the mode $D^0 \to \pi^+\pi^-\pi^0$ has been studied with Dalitz analyses by three different collaborations multiple times over the last fifteen years, one may reasonably wonder why we propose to do so again in this thesis. There are three main reasons. First, *Belle II*'s predecessor, *Belle*, only published one analysis on this mode early in the experiment [39], so it would be useful to restart analysis of $D^0 \to \pi^+\pi^-\pi^0$ at *Belle II*. Second, *Belle II* is a relatively new experiment, and this analysis will provide proof of concept that charm decays with neutral daughters can be well-measured by this collaboration. Third, we have developed a novel, data-based methodology for performing a Dalitz analysis that we demonstrate in this thesis.

Chapter 3

Experimental Apparatus

This chapter discusses the accelerator at which *Belle II* is based (section 3.1) as well as the individual components of the *Belle II* detector (section 3.2). In section 3.3, we discuss sources of nuisance asymmetries within the detector.

3.1 SuperKEKB Accelerator

The *Belle II* experiment is run at the SuperKEKB accelerator located at the High Energy Accelerator Research Organization (KEK, abbreviation from the Japanese) in Tsukuba, Japan. SuperKEKB is an asymmetric electron-positron collider, with the electron (positron) beam at 7 GeV (4 GeV) [49]. A schematic is shown in Figure 3.1. The diagram includes the four experimental halls – Nikko, Fuji, Oho, and Tsukuba – located around the accelerator's circumference. The *Belle II* experiment is located in Tsukuba Hall, which contains the only interaction point of the two beam lines. The asymmetric beams allow for time-dependent CPV analyses in B meson decays [29]. The collider is run primarily at a center-of-mass (CM) energy of $\sqrt{s} = 10.58$ GeV, which is the mass of the $\Upsilon(4S)$ resonance, with some data also collected at other Υ resonances or off-resonance energies [29]. The $\Upsilon(4S)$ decays primarily into B^+B^- and $B^0\overline{B}^0$ pairs. However, the $e^+e^- \to c\bar{c}$ cross section is roughly equal to $e^+e^- \to \Upsilon(4S) \to B\overline{B}$, which makes charm physics accessible at SuperKEKB as well [29].

SuperKEKB began taking data on April 26, 2018 [50], and has already achieved multiple instantaneous luminosity records [51]. The design luminosity of the collider was originally reported as 8×10^{35} cm^{-2}s^{-1} [29], but in 2020 this was amended to 6×10^{35} cm^{-2}s^{-1} [52]. The target integrated luminosity remains $50\,\text{ab}^{-1}$ by the end of the *Belle II* experiment [29].

3.2 *Belle II* Detector

Belle II is the first next-generation (super) B-factory detector. While the original B-factory detectors, BaBar and *Belle*, probed CPV in the SM, *Belle II*, with higher luminosities and greater precision, will be able to probe CPV both in and beyond the SM [34]. A state-of-the-art detector is required to undertake this physics program. Figure 3.2 shows a top view cross section of the *Belle II* detector. The specific components of the detector, especially those most relevant to the work described in this thesis, are detailed in the subsections below.
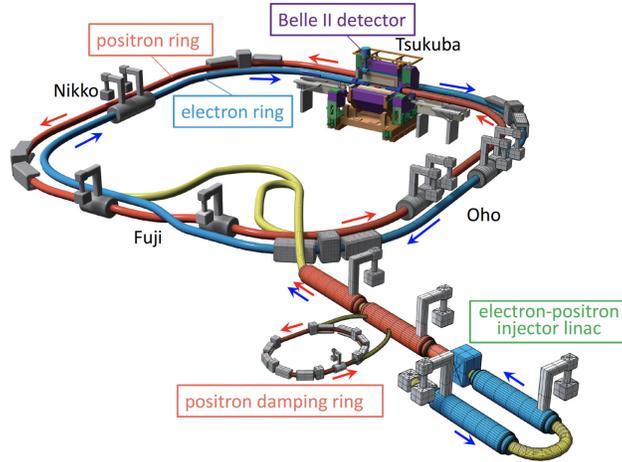
Figure 3.1: Schematic of the SuperKEKB accelerator [49].
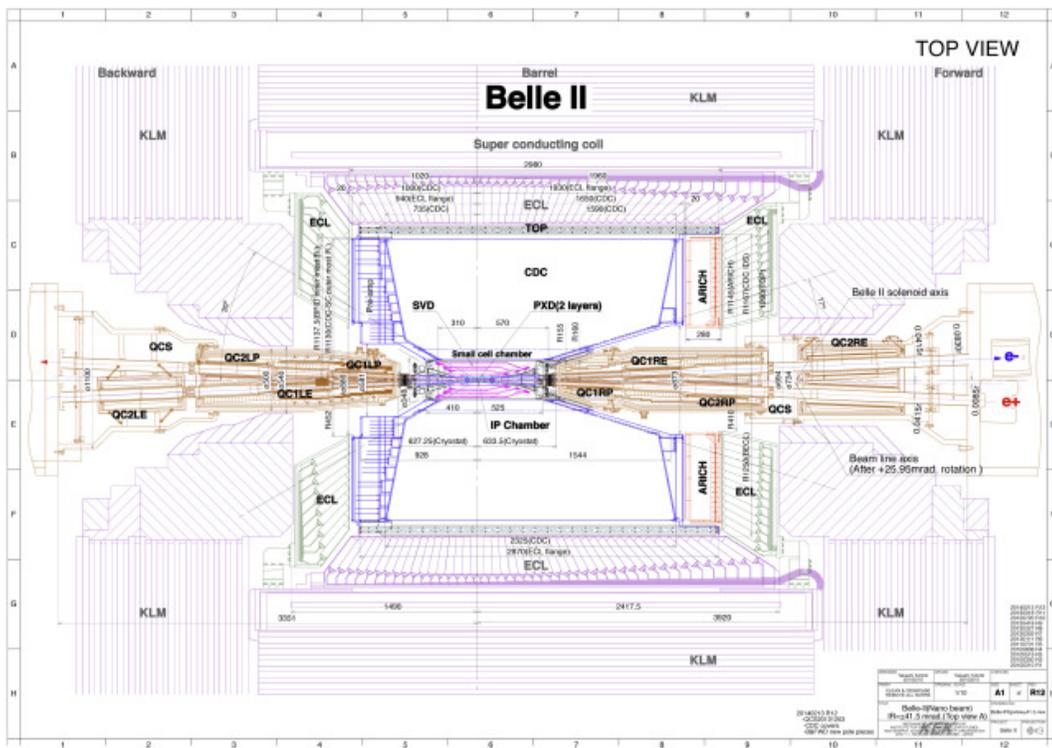
Figure 3.2: Top-view schematic of the *Belle II* detector. In this schematic, the z-axis (i.e., the beamline) points from the left, where the electron beam enters, to the right, where the positron beam enters from the opposite direction. The x-axis points up the page, and the remaining y-axis is found by the right-hand rule to point out of the page [53].

3.2.1 Vertex detector (VXD)

The vertex detector (VXD) is the first detector component outside of the beryllium beam pipe and is crucial to the vertex resolution (and hence, the entire physics program) of *Belle II*. It is made of six total layers and two sub-detectors: the inner silicon pixel vertex detector

(PXD) and the outer silicon strip vertex detector (SVD). The PXD is two layers, and the remaining four layers are the SVD, made up of double-sided silicon strip sensors [29].

The use of pixels in the PXD (as opposed to a six-layer strip detector) is in response to the nano-beam scheme employed by SuperKEKB. Figure 3.3 shows a schematic of the beam crossing; the beam pipe radius in the collision region is only 10 mm. While this is good for vertex reconstruction, the small radius increases the background in the collision region. The silicon strip sensors of the SVD would not be able to handle this larger background, hence the inclusion of the inner PXD layers [30].
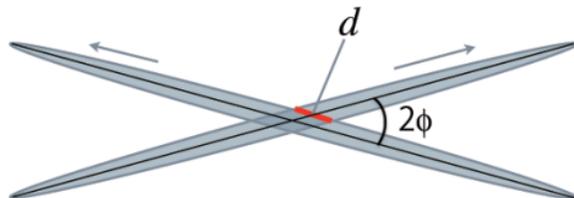
Figure 3.3: Nano-beam collision scheme at SuperKEKB. $d \approx \sigma_x^*/\phi$, where σ_x^* is the horizontal beam size at the interaction point (IP), and ϕ is half the crossing angle. At SuperKEKB, $\sigma_x^* = 10.2\,\mu\mathrm{m}$ and $\phi = 41.5\,\mathrm{mrad}$, which means $d \approx 0.25\,\mathrm{mm}$ [30].

3.2.2 Central drift chamber (CDC)

The central drift chamber (CDC) is the most important detector for charged particle tracking. It also plays a crucial role in particle identification (PID), especially for low-momentum tracks. The CDC at *Belle II* contains 14,336 sense wires in 56 layers inside a $\mathrm{He-C_2H_6}$ 50:50 gas mixture with an average $3\,\mathrm{cm}/\mu\mathrm{s}$ drift velocity. The sense wires in each layer are in either an axial or stereo orientation. Axial wires are aligned with the detector's z-axis, and thus with the solenoidal magnetic field inside the detector. In contrast, stereo wires are skewed with respect to the detector axis (imagine holding a bunch of uncooked spaghetti in both hands and then slightly twisting your hands in opposite directions: this is effectively the geometry of stereo wires) [29].

As a charged particle moves through the CDC, it ionizes atoms and molecules in the gas mixture. The ionization electrons drift toward the sense wires, since there is a voltage difference between the sense and field wires, and this permits the detection of the original charged particle. The charge, positive or negative, is determined by the curvature of the track. While both axial and stereo wires give information about the particle's r–ϕ position, the measurement angle of the stereo wires allows us to extract the particle's z position as well. Combining this information with the drift time measurement gives the vector momentum of the particle.

Particles of different masses will exhibit different amounts of ionization energy loss as they travel through the CDC gas. These characteristic curves, called dE/dx (i.e., energy loss per distance) curves, are used for PID within the CDC [30, 54]. While essentially all charged particles will have dE/dx information (and thus PID information) from the CDC, the helical

track shape of low-momentum particles (due to the Lorentz force from the solenoidal field) means they may spiral in on themselves and never reach the dedicated PID system in the TOP or ARICH. Additionally, the TOP and ARICH do not cover the backward endcap. In these cases it is particularly critical to have high-quality PID information from the CDC.

3.2.3   Particle identification systems (TOP and ARICH)

In the barrel and forward endcap regions, particles that leave the CDC enter either the TOP (barrel) or ARICH (endcap) detectors. TOP stands for time-of-propagation and is a special type of Cherenkov detector. Figure 3.4 gives the conceptual overview of the TOP, and Figure 3.5 describes the theory behind Cherenkov detection in general. The TOP detector's main purpose is to improve K/π separation, but it provides information about other charged particles as well [30].
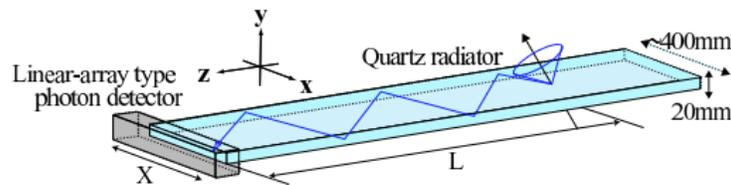
Figure 3.4: Schematic of the quartz radiator used in the TOP counter. Cherenkov photons are internally reflected in the quartz radiator, and the time-of-propagation is measured [30].
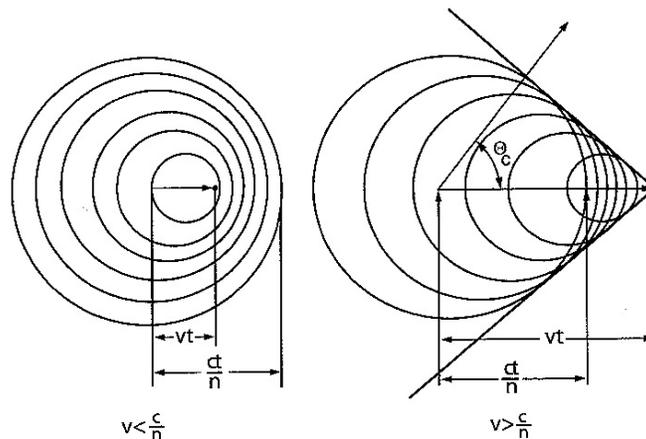
Figure 3.5: Schematic of the theory behind Cherenkov radiation. When a particle's velocity, v, is larger than the speed of light in the medium it travels in, c/n (where n = the refractive index of the medium), a Cherenkov cone with a characteristic angle Θ_C is produced. $\cos\Theta_C = c/(nv)$, so measuring Θ_C determines particle velocity. [55].

In the forward endcap, the aerogel ring-imaging Cherenkov detector (ARICH) provides particle identification (PID) information. Figure 3.6 shows a schematic of the ARICH. As the name implies, the ARICH uses an aerogel radiator, while the TOP uses a quartz radiator

[30]. The ARICH is also a more traditional Cherenkov detector, using ring-imaging of the Cherenkov cone to identify particles where the TOP is a time-of-propagation device (though the time-of-propagation is related to the Cherenkov angle).
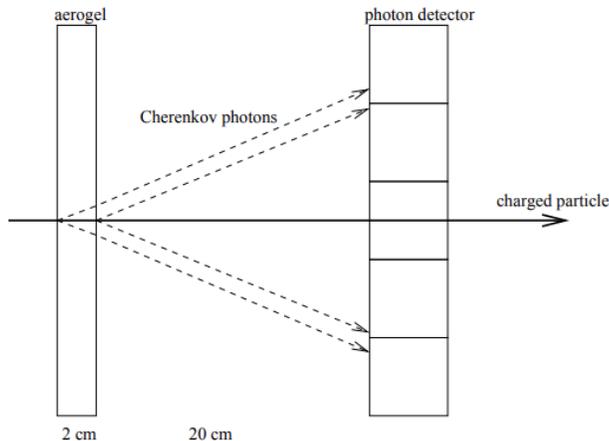
Figure 3.6: Principle behind the ARICH detector. As in the TOP, Cherenkov radiation is used, but rather than measuring the time-of-propagation, the ARICH images the ring of Cherenkov photons produced on the photon detector [30].

The reason for the different PID systems in different regions of the detector largely results from geometric constraints. Additionally, there is no PID system in the backward endcap. Due to the asymmetric beam energies, decays in *Belle II* are forward-boosted, and thus we expect the absence of a dedicated PID system in the backward endcap to have a small effect on the detector's capabilities. Overall, only about 10% of the CDC angular coverage is outside the TOP and ARICH coverage.

3.2.4   Electromagnetic calorimeter (ECL)

The electromagnetic calorimeter (ECL) is located outside the CDC and PID systems (see Figure 3.2) and covers all three detector regions (forward endcap, backward endcap, and barrel). The ECL contains 8,736 thallium-doped CsI crystals covering 90% of the solid angle in the CM system [29]. The main purpose of this detector is two-fold: 1) identify neutral decay products, in particular γ and π^0 particles, and 2) separate photons from electrons and hadrons. To accomplish this, in the ECL we look for localized large energy depositions called *clusters* or *showers*. Photons, electrons, and hadrons produce ECL showers with different characteristics, enabling separation between them. The ECL must also provide precise measurements of photon energy and angular position [30].

As in the VXD, the large (compared to *Belle*) detector backgrounds impact the ECL hardware and geometry. The CsI(Tl) crystals, which have been reused from *Belle*, exhibit scintillations with relatively long decay times, which increases the noise from background events. The detector electronics are equipped to handle this in the short term, but long-term, increased noise will degrade detector performance. Possible upgrades to address this, including replacing the CsI(Tl) crystals with pure CsI, are under consideration [29].

13

3.2.5 K_L^0 and muon detector (KLM)

The K_L^0 and muon detector (KLM) is the outermost subdetector in *Belle II*. The KLM covers all three detector regions and consists of 4.7 cm-thick iron plates alternating with glass electrode resistive plate chambers (RPCs).

An in-depth description of the KLM is beyond the scope of this document, since none of the particles relevant to this analysis are expected to reach it. The KLM is well-described elsewhere [30].

3.2.6 Other detector components

Besides the main subdetectors detailed above, there are several other components of the *Belle II* detector critical to its operation. Two important ones are the trigger system and the data acquisition system (DAQ).

The trigger system has two levels: the hardware-based low-level trigger (L1) and the software-based high-level trigger (HLT). The function of both triggers is to filter out beam background and retain only events of interest as well as to prescale common interactions like Bhabha scattering [29]. Filtering out hadronic events (events where a meson or baryon is produced) is relatively easy, as they "look" quite different to the detector compared to noise, beam backgrounds, and other events that are not of interest. This function couples into the DAQ, which would be overwhelmed by high rates without the trigger filtering out these events.

The DAQ is responsible for reading data from all *Belle II* subdetectors. It is designed to handle a 10 kHz maximum event rate from the HLT (reduced for offline storage from a 30 kHz trigger rate out of the L1), the value expected at peak luminosity [29, 53].

All of the hardware in the *Belle II* detector, including all of the front-end electronics in the DAQ, must be able to operate in a strong magnetic field, averaging around 1.5 T throughout the detector [30]. (For reference, the Earth's magnetic field is $\sim 50\,\mu$T.) This places some limitations on detector components and requires careful monitoring.

3.3 Sources of detector-induced asymmetries

Since the goal of the analysis described in this thesis is to measure an asymmetry due to CPV, the analysis method must account for asymmetries from other sources, so-called "nuisance asymmetries." It is worth discussing how the detector itself contributes to such nuisance asymmetries.

There are three nuisance asymmetries that we account for in our analysis (see chapter 5 for detailed discussions on each): the forward-backward production asymmetry, the slow pion reconstruction asymmetry, and the internal D^0 asymmetry. The production asymmetry comes from γ–Z^0 interference and higher order QED effects, but the other two nuisance asymmetries come from detector effects that lead to different reconstruction efficiencies for positive and negative tracks. There are two specific sources of detector asymmetries that contribute to this: the CP asymmetry of the detector material and the different behavior of positive and negative particles in the axial $\vec{B}$ field.

The CP asymmetry of the detector simply means the detector is made of matter, not equal parts matter and antimatter. Therefore, the detector is unavoidably not CP-symmetric. This means that the detector response is sensitive to differences in the cross sections of XN vs. $\overline{X}N$, where X and $\overline{X}$ are a particle-antiparticle pair, either of which may be produced by a decay at *Belle II*, and $N = p, n$, i.e., a nucleon in the detector. The differences in these cross sections are more prominent for kaons $(X = K)$, which our analysis does not have in the signal mode, but they are still present for pions $(X = \pi)$ as well. The overall effect is a difference in reconstruction efficiency for the two charges of pion.

The presence of an axial $\vec{B}$ field in the detector is a feature of the CDC that allows for charged-particle momentum measurements, but it must be noted that positive and negative particles do not behave the same way in this field: positively charged particles curl in one direction, and negatively charged particles curl in the opposite direction. However, particles of both charge types result in ionization electrons, which of course are always negative and always curl in the same direction. In effect, the axial $\vec{B}$ field breaks the symmetry of the drift cells in the r–ϕ projection. Like the CP asymmetry of the detector, this also leads to a difference in reconstruction efficiency for the two charges of pion. Calibration can remove much of this effect, but residual asymmetries lead to PID efficiency differences.

3.4   Summary

Overall, the components of *Belle II* come together to form a state-of-the-art detector. The analysis described in this document relies on the excellent performance of almost all *Belle II* subdetectors. The analysis methodology also takes into careful consideration the nuisance asymmetries caused by the detector itself, which must be accounted for in order to isolate CPV.

Chapter 4

Dataset and Event Selection

This chapter describes the *Belle II* dataset (section 4.1) and the selection criteria applied to it (section 4.2) for the Dalitz analysis described in this thesis. The event selection efficiency as a function of location in the Dalitz plot is also examined (section 4.3).

4.1 Dataset

We use $362\,\mathrm{fb}^{-1}$ of *Belle II* data collected at the $\Upsilon(4S)$ resonance for the analysis described in this thesis. The data was collected from 2019–2022, before the experiment entered its first long shutdown. Based on the way that data is processed at *Belle II*, the dataset is divided into two subsets: $187\,\mathrm{fb}^{-1}$ of data collected from March 2019–July 2021 (subset I) and $175\,\mathrm{fb}^{-1}$ of data collected from October 2021–June 2022 (subset II). Subset I is a reprocessing of older data with an updated software release (see section 4.2) and updated calibration constants. Subset II is a "prompt" processing of more recent data, performed with the same software release as subset I. Historically, prompt processings used preliminary calibrations that needed to be updated in later reprocessings; the difference between the prompt processing and the reprocessing of older data is very minor with current calibrations. Thus, beyond possible deterioration due to detector aging, the difference in quality between the two subsets is expected to be negligible, but this is checked explicitly and discussed further in chapter 8. Some plots in this chapter were made using only subset I; figures are clearly labeled to indicate when this is the case.

To study background events and to test the analysis procedure, we use simulated Monte-Carlo data (MC). Unless otherwise stated, the MC used here is $400\,\mathrm{fb}^{-1}$ of MC generated with the same background and detector conditions as the most recent processing of the data.

An advantage of examining MC is that we can separate correctly reconstructed signal events from background as well as identify the sources of background contamination. This is referred to as "truth-matching" or "truth-tagging." Several plots in this chapter use truth-matching. Plots of MC are clearly labeled to indicate if they show all reconstructed MC events or if truth-tagging has been used to isolate signal or background.

Figures included in this document are clearly labeled to indicate whether they show data or MC. All plots should be assumed to show the sum of D^0 and $\overline{D}^0$ events, unless it is explicitly stated that a plot is "charge-separated."

4.2   Event Selection

We use the *Belle II* Analysis Software Framework (`basf2`) [56] to process data and MC samples. Unless otherwise stated, the `basf2` version `light-2205-abys` is used for MC, and version `light-2210-devonrex` is used for data. (Each release was the most up-to-date version of the software at the time each analysis was performed; the practicle differences for our purposes are minimal.) The CERN data analysis tool `ROOT v6.24/06` is used to analyze the final selected events.

In order to reduce background while retaining events of interest, certain requirements are imposed on the particles in the decay chain. Additionally, two vertex fits and a best candidate selection procedure are performed. In the case of the real data, performance corrections are included on the charged track momenta and photon energies.[1] Table 4.1 shows the full list of selection criteria for $D^{*+} \rightarrow D^0 \ (\pi^+\pi^-\pi^0) \ \pi_s^+$ in both data and MC, and the following subsections explain these cuts in more detail. The goal of the selection criteria is to improve the signal-to-background ratio (equivalent to the purity) of the final selected events and to optimize the signal significance, defined as $s/\sqrt{s+b}$, where s = number of signal events and b = number of background events. Some basic optimization was done on key variables before the criteria listed in Table 4.1 were finalized.

4.2.1   Photon and π^0 reconstruction requirements

The π^0 in the final state is reconstructed from the decay $\pi^0 \rightarrow \gamma\gamma$, using ECL clusters to identify the photons. The branching fraction for $\pi^0 \rightarrow \gamma\gamma$ is 98.8% [33], so there is a small loss in efficiency at this stage. As noted in subsection 3.2.4, the ECL is subject to large detector backgrounds (e.g., fake photons from beam background radiation), so the cuts made here are to separate photons of interest from beam background photons as well as to distinguish photons from leptons and hadrons.

To remove beam background photons and detached hadronic shower fragments, we require the photon energy to be above $100\,\text{MeV}$. Most beam background photons and shower fragments are low-energy compared to the signal, so this loose cut removes much of the background without severely impacting the signal efficiency.

To distinguish photons from leptons and hadrons, we use several key variables. First, the cluster in question must be inside the CDC acceptance region. This region is slightly narrower than the ECL,[2] but this restriction allows us to determine whether the ECL cluster has a corresponding track in the CDC, in which case the cluster is due at least in part to a charged particle. Second, we look at the ratio of energies deposited in inner versus outer crystals of the ECL cluster via the variables `clusterE1E9` and `clusterE9E21`. Figure 4.1 shows a schematic of how these quantities are determined. If the crystal numbered 1 in this

---

[1]When data is processed at *Belle II*, care is taken to ensure particle information is calibrated correctly. However, some small biases may still occur. These biases are determined after the processing is complete, and the instructions for how to correct them are made available to analysts. These corrections to particle information applied after the central data processing are called "performance corrections."

[2]In terms of the cosine of the polar angle θ in the lab frame, the CDC acceptance region is $-0.866 < \cos\theta < 0.956$. The ECL acceptance region is $-0.907 < \cos\theta < -0.652$ (backward endcap) plus $-0.625 < \cos\theta < 0.846$ (barrel) plus $0.854 < \cos\theta < 0.977$ (forward endcap).

Particle	Selection criteria
γ	$E > 100\,\mathrm{MeV}$ created from ECL cluster in CDC geometric acceptance region `clusterE9E21` > 0.9 and `clusterE1E9` > 0.3
π^0	$105\,\mathrm{MeV}/c^2 < m(\gamma\gamma) < 150\,\mathrm{MeV}/c^2$
π_s^+	$dr < 0.5\,\mathrm{cm}$ and $|dz| < 2.0\,\mathrm{cm}$ in CDC geometric acceptance region
π^+, π^-	$dr < 0.5\,\mathrm{cm}$ and $|dz| < 2.0\,\mathrm{cm}$ in CDC geometric acceptance region $\mathrm{binaryPID}(\pi, K) > 0.4$
D^0	$1.7\,\mathrm{GeV}/c^2 < M(\pi^+\pi^-\pi^0) < 2.1\,\mathrm{GeV}/c^2$ `flightSig` > 0 (applied after D^{*+} vertex fits)
D^{*+}	$\Delta M < 0.160\,\mathrm{GeV}/c^2$ $p^* > 2.5\,\mathrm{GeV}/c$
Vertex fitting	Perform two fits with TreeFitter on full decay chain: 1) `ipConstraint` and `massConstraint` on π^0, require `chiProb`$_1 > 0.001$ 2) `ipConstraint` and `massConstraint` on π^0 and D^0, require fit not to fail but no other restrictions on `chiProb`$_2$
Other	$m(\pi^+\pi^-) < 485.0\,\mathrm{MeV}/c^2$ or $m(\pi^+\pi^-) > 502.5\,\mathrm{MeV}/c^2$ (K_S^0 veto) `BestCandidateSelection`, rank on χ^2-probability of second vertex fit Tracking and photon energy corrections applied to data only

Table 4.1: Finalized cuts for $D^{*+} \to D^0\,(\pi^+\pi^-\pi^0)\,\pi_s^+$ in both data and MC. Variable names are defined in-text.

figure is the center of an ECL cluster (i.e., the crystal where the most energy was deposited), then `clusterE1E9` $= E1/E9$, where $E1 = $ the energy deposited in crystal 1, and $E9 = $ the sum of the energies deposited in the crystals labeled 1–9 (i.e., the central 3×3 region). Similarly, `clusterE9E21` $= E9/E21$, where $E21 = $ the sum of the energies deposited in the crystals labeled 1–21 (i.e., all of the crystals in the schematic except for the outermost corners). Both `clusterE1E9` and `clusterE9E21` are always between 0 and 1. Most of the energy from a photon shower is deposited in the central crystals, which means both of these variables tend toward higher values for photons compared to hadrons. (Lepton showers also tend toward higher values for these variables, but as discussed above, we have other ways to distinguish between photons and leptons.)

Finally, a cut on the π^0 mass window requiring the invariant pair mass of the two photons, $m(\gamma\gamma)$, to be between $105\,\mathrm{MeV}/c^2$ and $150\,\mathrm{MeV}/c^2$ serves to remove incorrectly paired photons that did not come from a parent π^0. Figure 4.2 shows a plot of $m(\gamma\gamma)$. The resolution on the mass of the true π^0s is very wide, with an asymmetric low-side tail due to energy leakage out the back of the ECL crystals. This is why the cut on the mass is asymmetric about the π^0 mass peak, which is nominally $(134.9768 \pm 0.0005)\,\mathrm{MeV}/c^2$ [33]. The resolution on the reconstructed mass of the decaying particles, D^0 and D^{*+}, is improved by the vertex fits discussed in subsection 4.2.4, which both include a mass constraint on the π^0.

25	10	11	12	22
21	9	2	6	13
20	5	1	3	14
19	8	4	7	15
24	18	17	16	23
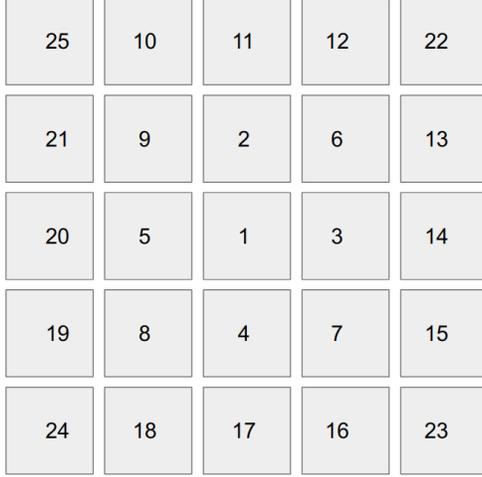
Figure 4.1: A top-down view of the ECL where each square represents a CsI(Th) crystal and the crystal numbered 1 is the center of an ECL cluster.
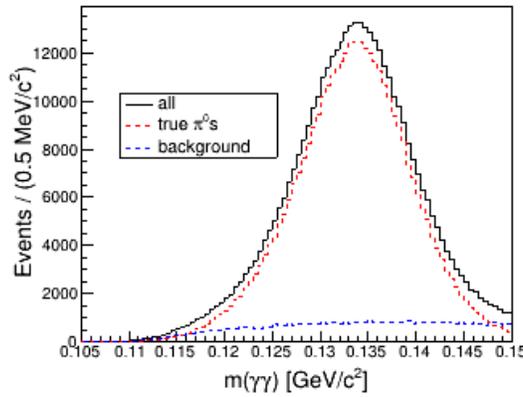
Figure 4.2: Invariant mass of the two photons used to reconstruct π^0 candidates in MC. The solid black line represents all candidates, while the dashed red line represents true (i.e., correctly reconstructed) π^0s and the dashed blue line represents background.

4.2.2 Charged pion reconstruction requirements

There are three charged pions in the final state for this decay: two from the Dalitz decay of the D^0 and one from $D^{*+} \to D^0 \, \pi_s^+$, which is used to tag the flavor of the D^0 ($\overline{D}^0$). The tagging pion is called the "slow pion" (or sometimes "soft pion") and usually labeled π_s^+ to distinguish it from the charged D^0 decay products, which we will collectively refer to here as "fast pions." As the moniker implies, the momentum of the slow pion tends to be lower than that of the fast pions. Figure 4.3 shows the spectra of the three charged pions in signal D^{*+} decays in MC. The slow pion spectrum tapers off below $0.6 \, \text{GeV}/c$, while the fast pion spectra have long tails that extend to almost $5 \, \text{GeV}/c$.

The quality cuts on the slow pion differentiate the pion track from detector noise by ensuring it originated from the interaction point (IP) and was reconstructed correctly. The IP
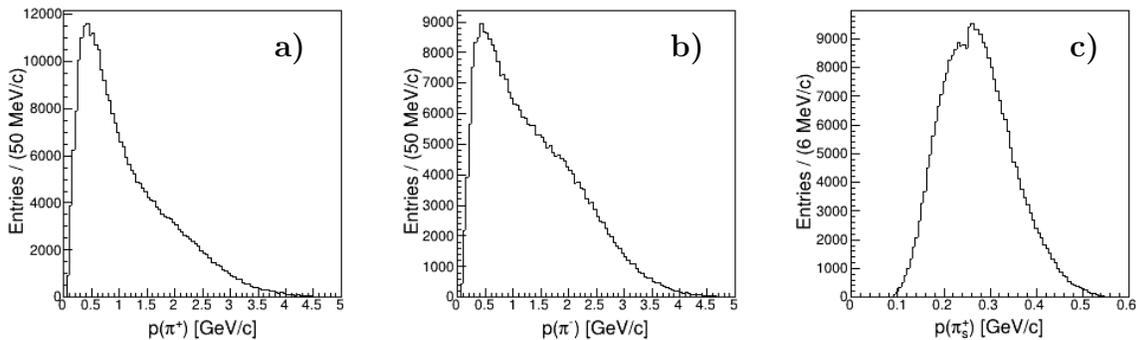
Figure 4.3: Charged pion spectra for signal events in MC. Note the different scale for images **a)** and **b)** (fast pions) as compared to **c)** (slow pion).

parameters, dr and dz, are the distances in the x-y plane and along the z-axis (respectively) from the particle's point of closest approach (POCA) to the IP (see Figure 3.2 for orientation of axes). We require $dr < 0.5\,\mathrm{cm}$ and $|dz| < 2.0\,\mathrm{cm}$. These are fairly loose cuts, so they do not have a large effect on signal efficiency, but they do remove background tracks that could not physically have been part of the D^{*+} decay. We also require that the slow pion track be within the CDC acceptance region. This is another loose cut that ensures we remove tracks with a misreconstructed polar angle.

For the fast pions, the same IP cuts and CDC acceptance requirement used for the slow pion are included, but there is an additional cut to remove background from $D^0 \to K^- X$ transitions. As discussed in chapter 1, final states containing K^- mesons are Cabibbo-favored for D^0 meson decays and consequently occur about an order of magnitude more often than final states containing only pions, so we need to remove decays in which a kaon has been misreconstructed as a pion. To achieve this, we use a particle ID (PID) variable that compares the likelihood that the particle is a pion to the likelihood that the particle is a kaon. Many different PID variables are available to identify any of six different charged particle types (electron, muon, pion, kaon, proton, or deuteron). The variable we use is explicitly defined as:

$$\mathrm{binaryPID}(\pi, K) = \frac{\mathcal{L}_\pi}{\mathcal{L}_\pi + \mathcal{L}_K} \tag{4.1}$$

where $\mathcal{L}_X$ indicates the likelihood of a particle being of type X. The likelihood has been calculated using information from all *Belle II* subdetectors except the SVD and PXD.[3] binaryPID(π, K) ranges from 0 to 1 (as do all other PID variables), with particles that are more likely to be pions tending toward higher values. The cut we use is binaryPID$(\pi, K) > 0.4$; this cut is relatively loose, but it removes much of the kaon background, which peaks near 0. Since the ultimate goal of this analysis is to measure an asymmetry, we need to avoid

---

[3]The PXD does not provide PID information and is excluded from PID variabels by default. At the time the analysis in this thesis was being completed, there was a bug in the `basf2` tracking software that meant SVD likelihoods were not computed for low-momentum particles in data. For consistency between data and MC, and to avoid having to make momentum-dependent PID cuts, we did not use SVD likelihoods in calculating PID.

charge asymmetries introduced by PID, which typically become more of an issue as the PID cut is tightened. Since both the D^0 and $\overline{D}^0$ decays contain a π^+ and a π^-, naïvely one might think any charge asymmetry in PID would cancel out. However, since the spectra of the pions are not identical (compare the left and center plots in Figure 4.3), PID asymmetries and detector asymmetries caused by the different spectra might couple together. The safest way to avoid these issues is to use a loose PID cut to reduce charge asymmetries.

There are three variables – `nCDCHits`, `pValue`, and globalPID – that are used for standard track quality and PID cuts in many other *Belle II* analyses but that we do not include here. `nCDCHits` is the number of wires in the CDC with a signal ("hit") from the particle track. `pValue` is the χ^2-probability from the track fit. globalPID for our purposes is defined as:

$$\text{globalPID}(\pi) = \frac{\mathcal{L}_\pi}{\mathcal{L}_\pi + \mathcal{L}_K + \mathcal{L}_e + \mathcal{L}_\mu + \mathcal{L}_p + \mathcal{L}_d} \tag{4.2}$$

so that we are effectively comparing the pion likelihood to all other particle likelihoods, not just the kaon likelihood. While finalizing the cuts for this analysis, we discovered that cuts on these three variables in particular introduce a large efficiency discrepancy between MC and data.

To some degree, a discrepancy between data and MC is acceptable for this analysis. Other than some systematic uncertainty calculations (see chapter 8), we are not using MC in calculations of asymmetries. However, we would like to use the MC to check the reasonableness of certain cuts and to test the analysis procedure. If there is a large discrepancy between data and MC, then we cannot be certain that conclusions drawn from the MC are valid. Since omitting these cuts does not significantly degrade the purity or signal significance, we do not include them.

4.2.3 D^0 and D^{*+} reconstruction requirements

After the final state particle lists have been created, we make loose cuts on the D^0 and D^{*+} particles to remove candidates made from tracks and π^0s that, while they may be real pions, are unlikely to have come from the decay of interest as well as to reduce the combinatorial background.

Firstly, we make a loose cut on the mass of the D^0 candidate, requiring that the invariant mass of the three pions, $M(\pi^+\pi^-\pi^0)$, be between $1.7\,\text{GeV}/c^2$ and $2.1\,\text{GeV}/c^2$. This cut will be substantially tightened when we define the signal region, but we leave it loose at this stage so that we can examine sideband regions if needed and to avoid biasing the vertex fits (see subsection 4.2.4).

Next, we make another loose cut on the variable $\Delta M = M(D^{*+}) - M(D^0)$, or in terms of the final state particles, $\Delta M = M(\pi^+\pi^-\pi^0\pi_s^+) - M(\pi^+\pi^-\pi^0)$. This is sometimes referred to as the "D^* trick." Since the D^{*+} is reconstructed from the same three particles used to reconstruct the D^0 plus one extra pion, the resolution on the difference between $M(D^{*+})$ and $M(D^0)$ is much narrower than the resolution on either of these masses individually. We cut on $\Delta M < 160\,\text{MeV}/c^2$ at this stage, again to preserve sidebands and to reduce bias from vertex fitting. The nominal value for ΔM is $(145.4258 \pm 0.0017)\,\text{MeV}/c^2$, and there is a physical lower bound at the rest mass of the slow pion, which is $(139.57039 \pm 0.00018)\,\text{MeV}/c^2$, so we do not include a lower bound explicitly in the cut [33].

Finally, the last cut we make before performing the vertex fits is on the center-of-mass (CM) momentum of the D^{*+}, $p^*(D^{*+})$, where the asterisk on the variable indicates it is measured in the CM frame. We cut at $p^*(D^{*+}) > 2.5\,\mathrm{GeV}/c$ to remove the high combinatorial background that occurs at lower momenta. This also has the effect of removing D^{*+} candidates coming from $B\bar{B}$ decays, since B mesons are produced at threshold at SuperKEKB. In principle, we would be happy to include signal decays from $B\bar{B}$, but they are not worth the high rate of combinatorics that would unavoidably be included with them.

There is one additional cut made on the D^0 candidate after the vertex fitting has been performed. To remove background from other $q\bar{q}$ processes ($s\bar{s}$, $u\bar{u}$, and $d\bar{d}$ fragmentation), we make a cut on the flight significance (`flightSig`) of the D^0 candidate. The flight significance is defined as the flight distance divided by the error on the flight distance:

$$\texttt{flightSig} = \frac{\texttt{flightDist}}{\texttt{flightDistErr}}. \tag{4.3}$$

Since the D^0 has a finite lifetime and flies a significant distance from the IP before it decays, `flightSig` must be positive (except for resolution smearing) for true D^0 mesons. The $s\bar{s}$, $u\bar{u}$, and $d\bar{d}$ processes do not exhibit this behavior, so their `flightSig` values are mostly symmetric about 0. Thus, the cut we include is `flightSig` > 0, which removes about half of the other $q\bar{q}$ background but has only a small effect on signal efficiency. It is worth noting, however, that this cut does not remove much $c\bar{c}$ background, since this background can also have a finite lifetime. We rely on other cuts (such as the particle ID cut discussed in subsection 4.2.2) to reduce $c\bar{c}$ background.

4.2.4 Vertex fitting

We perform two vertex fits on the full D^{*+} decay chain using TreeFitter. TreeFitter is a global fitting tool actively being worked on by members of *Belle II* [57, 58]. Its primary advantage, and what differentiates it from other fitters such as KFit, is that it fits an entire decay chain at once, eliminating the need to perform a separate fit for every intermediate particle. Based on the fit results, the fitter updates the kinematic variables of all decay products. Because intermediate particles are reconstructed from the four-vectors of their decay products, this also updates the invariant mass of the intermediate particles.

The first vertex fit requires the D^{*+} to originate at the IP, and it constrains the mass of the π^0 candidate to its nominal mass. The fit quality is determined by its χ^2-probability (`chiProb`$_1$), and so we also require `chiProb`$_1 > 0.001$ to remove events inconsistent with the decay topology. The results of this vertex fit are used to define the signal region (via cuts on $M(\pi^+\pi^-\pi^0)$ and ΔM) and to perform yield fits (on ΔM).

The second vertex fit requires the D^{*+} to originate at the IP, and it constrains the mass of both the π^0 candidate and the D^0 candidate to their nominal masses. The fit is required not to fail (i.e., `chiProb`$_2 \geq 0$), but no other cuts are made on fit quality. Because this fit mass-constrains the D^0, even a very loose cut on fit quality can have a large effect on signal efficiency. The results of this vertex fit are used to calculate the variables that define the Dalitz plot regions in the binned analysis because mass-constraining the D^0 candidate enforces the physical boundary of the Dalitz plot.

4.2.5 K_S^0 veto

The decay we are interested in for this analysis is the Dalitz decay $D^0 \to \pi^+\pi^-\pi^0$. However, there is another way for the parent D^0 to proceed to the same final state: through the channel $D^0 \to K_S^0\pi^0$, $K_S^0 \to \pi^+\pi^-$, which is an entirely different decay process and not part of the Dalitz decay we are interested in. Since the K_S^0 also has a finite lifetime, much of this background is removed by the vertex fitting (and, to a lesser extent, the track IP cuts), but a small amount remains and contaminates the Dalitz plot. To remove this residual K_S^0 background, we apply a K_S^0 veto in the form of a narrow cut on the $\pi^+\pi^-$ invariant mass, $m(\pi^+\pi^-)$, after both vertex fits. Figure 4.4 shows a plot of the $\pi^+\pi^-$ mass for this K_S^0 background in MC. The cut we apply is $m(\pi^+\pi^-) < 485.0\,\mathrm{MeV}/c^2$ or $m(\pi^+\pi^-) > 502.5\,\mathrm{MeV}/c^2$, which removes most of the K_S^0 peak. The few remaining events are fractionally $\mathcal{O}(10^{-4})$ of all retained events and are negligible compared to the signal. The cut region is asymmetric about the nominal K_S^0 mass ($497.611 \pm 0.013\,\mathrm{MeV}/c^2$ [33]) because the peak itself is asymmetric.
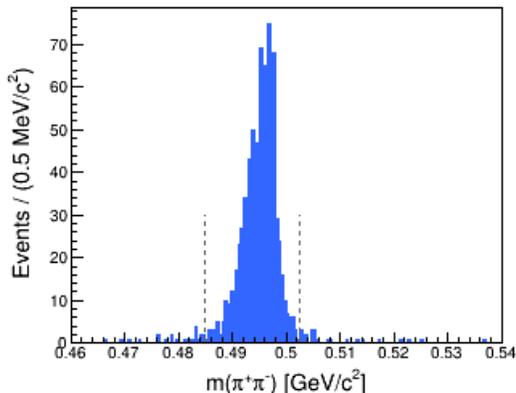
Figure 4.4: $K_S^0 \to \pi^+\pi^-$ background, identified using truth-matching in MC. The dashed vertical lines are the boundaries of the region removed with our K_S^0 veto. This plot was made with $200\,\mathrm{fb}^{-1}$ of MC analyzed with `basf2` version `light-2203-zeus`. All of the cuts are identical to those described in Table 4.1 except the K_S^0 veto is not yet included.

4.2.6 Best candidate selection

Even after all of the quality cuts and the fitting procedure described in the previous subsections, we are left with many events that contain multiple candidates of a given charge (D^{*+} or D^{*-}). Figure 4.5 illustrates the fraction of events this occurs for in MC, and Figure 4.6 shows the same in data. The relatively high fraction of multiple candidates is due to the combinatorics of the final state particles: if, for example, multiple π^0s are present in an event, it is possible that several of them passed our quality cuts and all of them could be consistent with a $D^0 \to \pi^+\pi^-\pi^0$ decay. Only one of them (at most) is, though, and so we need a way to pick the candidate most likely to contain the correct π^0. The same issue could arise for multiple charged pions in an event as well.
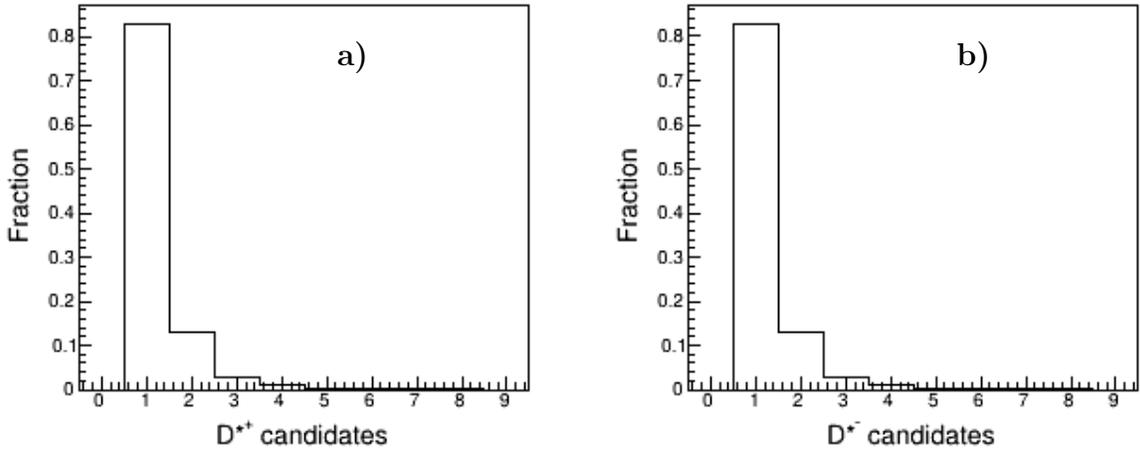
Figure 4.5: The percentage of events that contain a particular number of **a)** D^{*+} and **b)** D^{*-} candidates in MC. Notably, 82.9% (82.9%) of events that have *at least* one D^{*+} (D^{*-}) candidate contain *only* one D^{*+} (D^{*-}) candidate.
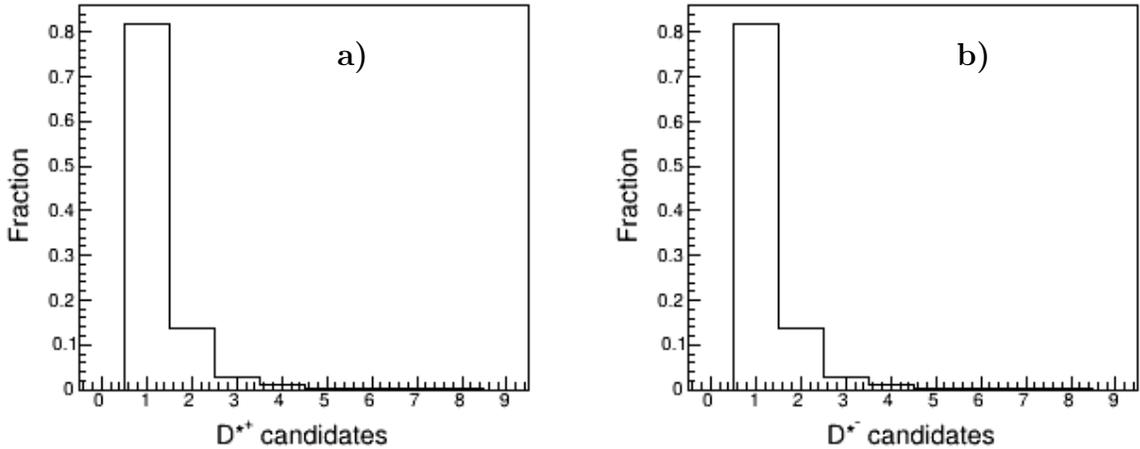
Figure 4.6: The percentage of events that contain a particular number of **a)** D^{*+} and **b)** D^{*-} candidates in data. Notably, 81.8% (81.9%) of events that have *at least* one D^{*+} (D^{*-}) candidate contain *only* one D^{*+} (D^{*-}) candidate. This is similar to what we see in MC; it is not unexpected for the percentage of single-candidate events to be slightly lower. These plots were made with a $187\,\text{fb}^{-1}$ subset of the full dataset.

The most effective way to pick the correct candidate is to rank all the candidates of a particular charge (i.e., treat D^{*+} and D^{*-} candidates separately, since one candidate of each charge per event is possible) in a single event on the χ^2-probability of the second vertex fit described in subsection 4.2.4 (`chiProb`$_2$). This is the fit that constrains the D^0 mass to its nominal value. Table 4.2 shows the relative efficiency of ranking on `chiProb`$_1$ vs. `chiProb`$_2$

for each charge and for different numbers of multiple candidates in MC. Figures 4.5 and 4.6 show MC and data behave similarly in terms of multiple candidates, so we expect the relative efficiencies to be similar, perhaps a bit lower, in data. Relative efficiency is calculated using truth-matching in MC, dividing the number of signal candidates that are rank 1 in a particular ranking by the total number of signal candidates being ranked. For example, there are 28,249 true D^{*+} signal candidates in events with exactly 2 D^{*+} candidates. Ranking the candidates with `chiProb`$_1$, 18,393 signal candidates are retained (i.e., are rank 1), for an efficiency of 65.1%. Using `chiProb`$_2$ instead, 25,026 signal candidates are retained, for an efficiency of 88.6%. For both charges and for all numbers of candidates (note the case of only 1 candidate per event is trivial), `chiProb`$_2$ is the more effective ranking variable.

Using `chiProb`$_2$, from the vertex fit that mass-constrains the D^0, in order to pick the best candidate has the downside of biasing the background toward the D^0 mass peak. However, it does not have the same impact on the ΔM peak, which is what we use for performing yield fits. Additionally, as shown in Table 4.2, it is much better at correctly identifying the signal candidate as compared to ranking on `chiProb`$_1$, so we accept this bias in the $M(\pi^+\pi^-\pi^0)$ background as a trade-off for keeping the signal efficiency as high as possible.

Particle	Number of candidates	Relative efficiency for BCS Method:	
		Rank on `chiProb`$_1$	Rank on `chiProb`$_2$
D^{*+}	1	100%	100%
	2	65.1%	88.6%
	3	51.2%	84.2%
	> 3	30.4%	58.3%
D^{*-}	1	100%	100%
	2	65.4%	88.4%
	3	50.8%	83.8%
	> 3	29.1%	57.7%

Table 4.2: Relative efficiency of ranking on `chiProb`$_1$ and `chiProb`$_2$ given charge and number of candidates in MC.

4.2.7 Defining the signal region

Figure 4.7 shows plots of the $M(\pi^+\pi^-\pi^0)$ and ΔM distributions in MC for the remaining events once all the cuts described in the preceding sections are applied. Truth-matched signal and background distributions are plotted separately. The plots are "signal-enhanced," which means that each distribution is cut on the signal region of the other. The two signal regions are: $1.83\,\mathrm{GeV}/c^2 < M(\pi^+\pi^-\pi^0) < 1.89\,\mathrm{GeV}/c^2$ and $144.5\,\mathrm{MeV}/c^2 < \Delta M < 146.2\,\mathrm{MeV}/c^2$. The two cuts together define the overall signal region used when making Dalitz plots or plots of other relevant variables. Note that quantitatively, only the $M(\pi^+\pi^-\pi^0)$ signal region cut is particularly relevant; the asymmetry measurement performed in this analysis relies on a fit to the ΔM distribution, so we never explicitly cut on the ΔM variable when performing these calculations. The ΔM signal region is relevant only in that it helps to isolate signal decays and estimate signal efficiency when examining other distributions.
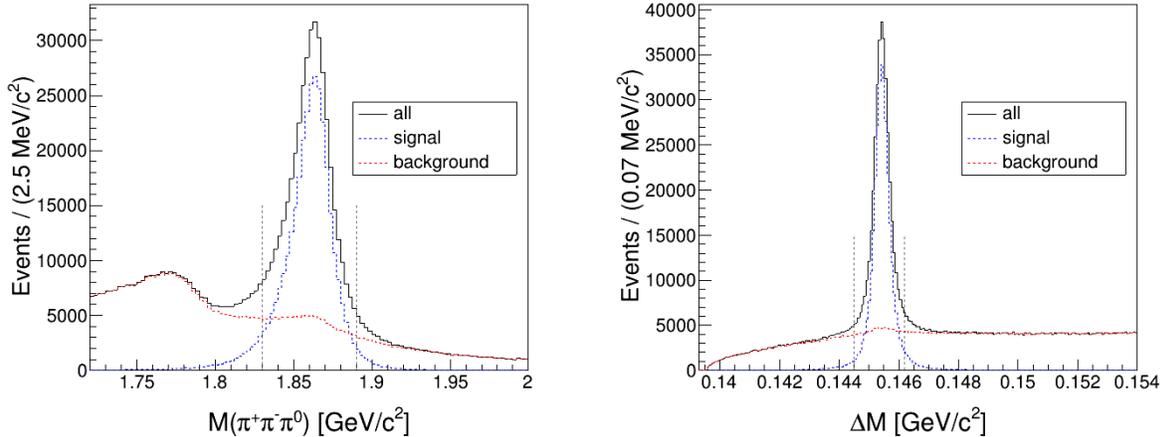
Figure 4.7: Signal-enhanced $M(\pi^+\pi^-\pi^0)$ and ΔM plots of candidates in MC retained after all selection criteria are applied. "Signal-enhanced" means each distribution is cut on the signal region of the other. The dotted vertical lines indicate the signal regions.

4.2.8 MC vs. data

Figure 4.8 shows the $M(\pi^+\pi^-\pi^0)$ and ΔM distributions in data overlaid with the distributions in MC. The contributions coming from different types of MC (i.e., $c\bar{c}$ processes vs. other types of background) are plotted separately. A discrepancy between the height of the signal peak in data and MC is visible in both plots, although the background is in very good agreement. Most of what we want to understand from MC is the behavior of the background, so this level of discrepancy is acceptable. More importantly, this discrepancy is observed by many different analyses of charm hadrons at *Belle II*, and we believe it is a feature of *Belle II* MC in general (probably because the fraction of c quarks forming D^* mesons is incorrect) and not an artifact introduced by this specific analysis.

4.3 Dalitz plots

We apply all of the selection criteria described in the previous subsections in order to isolate signal events to use in our binned Dalitz analysis. Using truth-matched MC events, the signal efficiency of the selection criteria is $(5.712 \pm 0.010)\%$ in the signal region. The purity in the signal region is 74%, and the signal significance, $s/\sqrt{s+b}$, is 475. Note that although efficiency and purity are independent of the size of the dataset, the significance is not.

Figure 4.9 (Figure 4.10) shows the Dalitz plot of events in the signal region in MC (data). The three ρ bands that dominate the phase space of the decay, including the interference between them, are clearly visible. To emphasize that the three ρ bands are not all equally populated, contour plots are shown in addition to scatterplots. Both D^0 and $\overline{D}^0$ events are included in both plots; for D^0, the plot is filled as indicated by the axis labels. For $\overline{D}^0$, $m(\pi^+\pi^0)^2$ and $m(\pi^-\pi^0)^2$ are interchanged and then the plot is filled (i.e., the axes are charge-conjugated for the $\overline{D}^0$ events).

26
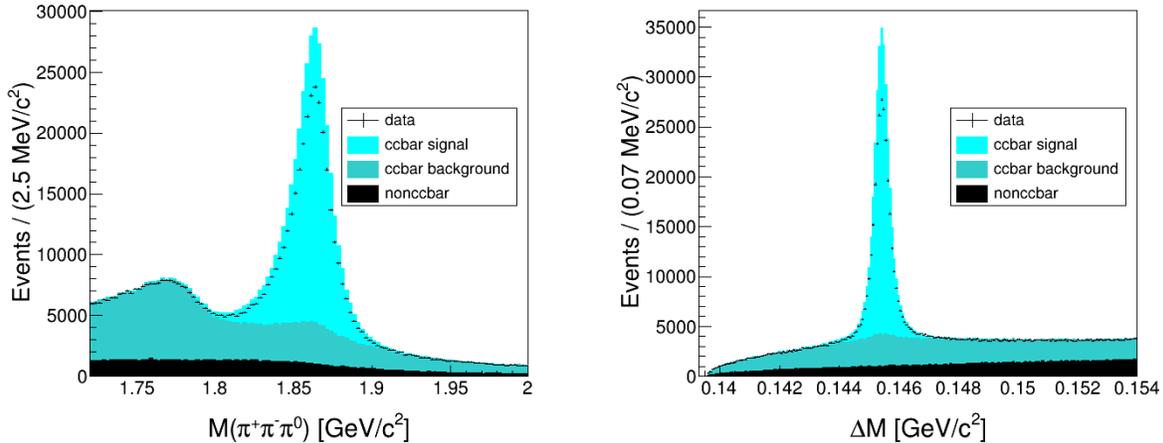
Figure 4.8: Signal-enhanced $M(\pi^+\pi^-\pi^0)$ and ΔM plots of candidates in MC and data.(See Figure 4.7 for a description of "signal-enhanced" and the signal region).

Comparing the contour plots, there is a resonance in the MC that is not nearly as prominent in the real data. This is attributed to the $f_2(1270) \to \pi^+\pi^-$ resonance having too large an amplitude in the software that generates the MC. This too-large amplitude is reported in the paper the MC parameters are taken from, but it is consistently not observed in data [36]. Figure 4.11 shows a comparison of the three Dalitz projections for MC and data; while the overall shapes are similar, there is some discrepancy coming from the incorrectly enhanced $f_2(1270)$. Since we only use MC for background studies and to determine systematic uncertainties, this discrepancy should not be problematic for this analysis.

4.3.1   Dalitz plot efficiency

We have shown that the cuts described in the previous section result in a clear signal peak with an overall signal efficiency of $(5.712 \pm 0.010)\%$. However, since we are performing a binned Dalitz analysis, it is not only the overall signal efficiency that is relevant: we are also interested in the signal efficiency as a function of position in the Dalitz plot. Figure 4.12 shows contour plots of the signal efficiency in a 10×10 grid across the Dalitz plot for both D^0 and $\bar{D}^0$ candidates. In both plots, the efficiency is largely consistent across the plot except at low $m(\pi^-\pi^0)^2$ and low $m(\pi^+\pi^0)^2$ (the lower left corner of the plot), where there is a noticeable decrease. Table 4.3 gives the actual signal efficiency for the D^0 and $\bar{D}^0$ candidates combined. (At this level of quantitative detail, separating the charges in MC is unlikely to map to what we expect to see in data.) The lower-left corner of the Dalitz plot, where a noticeable drop in efficiency occurs, corresponds to low-momentum π^0 events. As described in subsection 4.2.1, our photon energy cuts remove low-energy photons and hence many low-momentum π^0s, so we expect to see a drop in efficiency in this region.

This thesis does not efficiency-correct the Dalitz plot before calculating asymmetries. In principle, one should do this to ensure differences in efficiency are not obscuring real asymmetries. Future work will look into the best way to accomplish this.

27
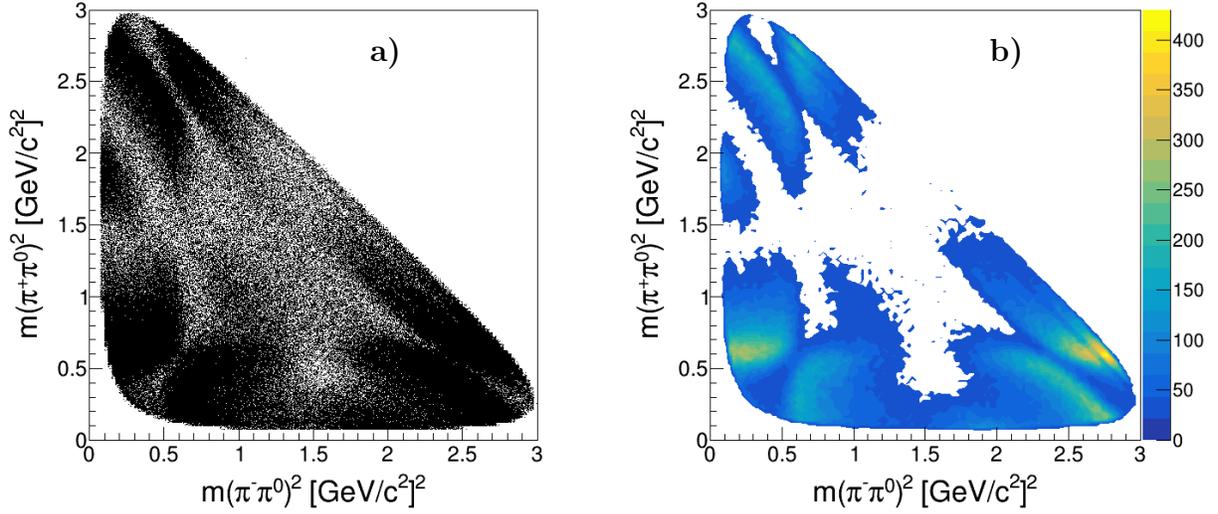
Figure 4.9: Dalitz **a)** scatterplot and **b)** contour plot of events in the signal region in MC. The axis labels are for the D^0 decay, but $\overline{D}^0$ events are included as well (with the charge-conjugate axes). The total number of events in the plot $(D^0 + \overline{D}^0)$ is $N = 410{,}978$.
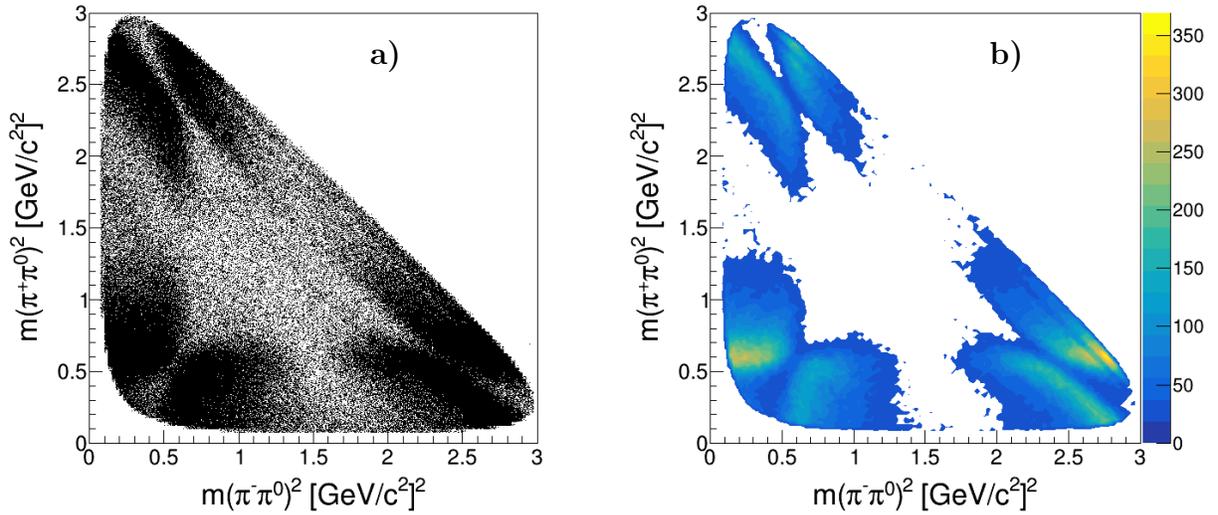
Figure 4.10: Dalitz **a)** scatterplot and **b)** contour plot of events in the signal region in data. The axis labels are for the D^0 decay, but $\overline{D}^0$ events are included as well (with the charge-conjugate axes). The total number of events in the plot $(D^0 + \overline{D}^0)$ is $N = 313{,}264$.
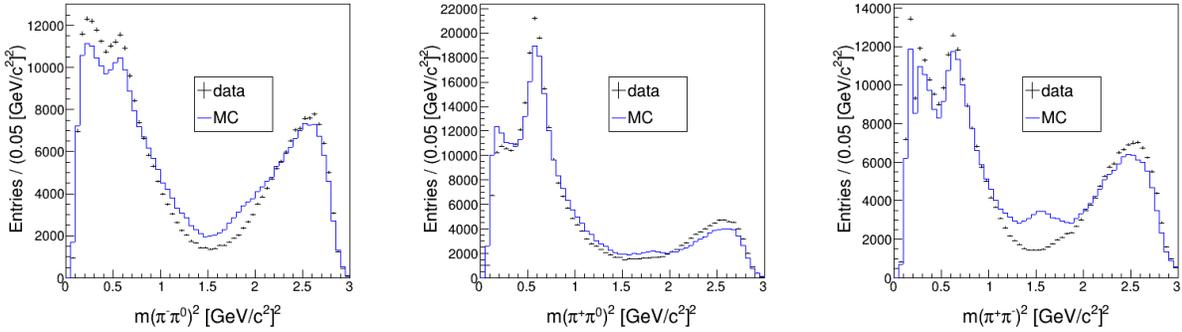
Figure 4.11: Comparison of projections onto the three Dalitz axes in MC and data. The MC has been scaled to the data luminosity and to account for the signal peak discrepancy observed in Figure 4.8
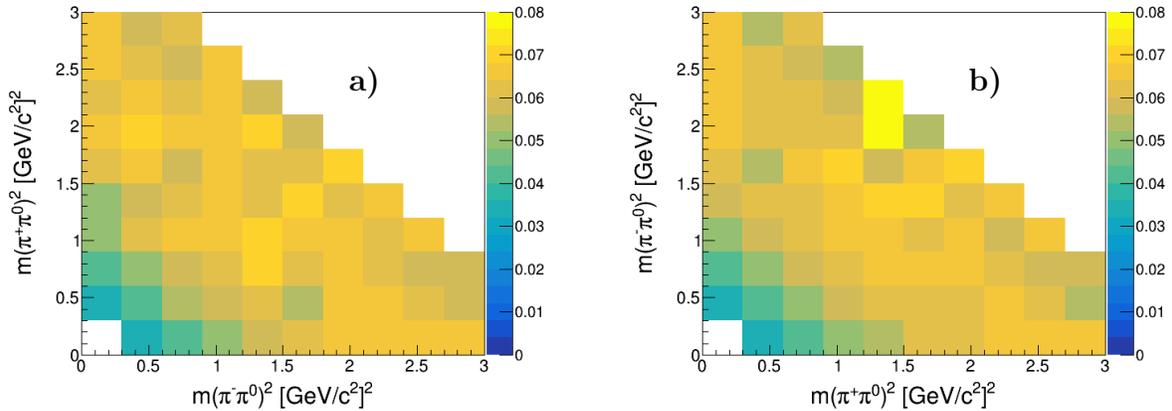
Figure 4.12: Plot of efficiency as a function of Dalitz region in MC for **a)** D^0 decays and **b)** $\overline{D}^0$ decays. Truth-matched signal MC is used to calculate efficiency.

6.632 ± 0.094	5.521 ± 0.093	6.08 ± 0.16							
6.565 ± 0.081	6.241 ± 0.060	5.790 ± 0.058	5.93 ± 0.28						
6.42 ± 0.20	6.422 ± 0.066	6.251 ± 0.092	6.15 ± 0.14	6.88 ± 0.70					
6.72 ± 0.12	6.55 ± 0.12	6.47 ± 0.16	6.34 ± 0.18	7.26 ± 0.53	5.46 ± 0.90				
6.26 ± 0.12	5.98 ± 0.23	6.34 ± 0.23	6.89 ± 0.23	6.10 ± 0.28	6.27 ± 0.23	6.89 ± 0.48			
5.33 ± 0.23	5.85 ± 0.19	6.28 ± 0.20	6.70 ± 0.19	6.56 ± 0.21	6.93 ± 0.16	6.26 ± 0.13	6.56 ± 0.33		
5.166 ± 0.093	6.011 ± 0.077	6.48 ± 0.17	6.68 ± 0.13	6.87 ± 0.18	6.46 ± 0.20	6.63 ± 0.10	5.896 ± 0.082	6.61 ± 0.21	
4.275 ± 0.038	5.070 ± 0.038	5.87 ± 0.11	6.240 ± 0.099	6.67 ± 0.15	6.33 ± 0.16	6.52 ± 0.11	6.260 ± 0.062	5.787 ± 0.041	5.92 ± 0.11
3.377 ± 0.044	4.296 ± 0.051	5.186 ± 0.043	5.918 ± 0.070	6.18 ± 0.14	5.80 ± 0.17	6.446 ± 0.080	6.481 ± 0.051	6.242 ± 0.046	5.482 ± 0.056
	3.315 ± 0.073	4.271 ± 0.047	4.999 ± 0.080	5.62 ± 0.11	6.29 ± 0.10	6.477 ± 0.085	6.614 ± 0.087	6.597 ± 0.058	6.517 ± 0.086

The y-axis is labeled $m(\pi^+\pi^0)^2\ [\,\mathrm{GeV}/c^2]^2$ and the x-axis is labeled $m(\pi^-\pi^0)^2\ [\,\mathrm{GeV}/c^2]^2$.

Table 4.3: Reconstruction efficiency (given in %) in a 10×10 ($0.3[\,\mathrm{GeV}/c^2]^2 \times 0.3[\,\mathrm{GeV}/c^2]^2$) binned Dalitz plot in MC. The efficiencies' positions in this table map to the Dalitz plot regions in Figure 4.12. While Figure 4.12 separates D^0 and $\overline{D}^0$ Dalitz plots to emphasize that the two plots are similar, this table gives the reconstruction efficiency for the combined plot. The axis labels are for the D^0 decay, but $\overline{D}^0$ events are included as well (with the charge-conjugate axes).

Chapter 5

Analysis Procedure

This chapter describes our novel binned analysis method used to analyze the dataset from chapter 4. In section 5.1, we define the different asymmetries, and in section 5.2, we describe the relevant considerations to ensure cancellation of all nuisance asymmetries in our search for CPV. We provide a summary in section 5.3.

5.1   Definitions

We investigate CPV by measuring an asymmetry in the number of D^0 vs. $\overline{D}^0$ candidates in bins of the Dalitz plot. However, sources other than CPV may contribute to such an asymmetry. This chapter discusses several different sources of asymmetries, one of which may be true CPV and the rest of which are nuisance asymmetries. Our methodology is based on cancelling the nuisance asymmetries using the data itself rather than relying on simulations. Table 5.1 provides a comprehensive list of the asymmetries discussed in this chapter and their definitions. Asymmetries without a superscript i for the bin index are treated as constant across the Dalitz plot modulo some systematic uncertainty. The asymmetry we will ultimately measure is the local CP asymmetry in each bin, $\mathcal{A}_{CP,\text{loc}}^i$.

The asymmetries are related as follows:

$$\mathcal{A}_{\text{bin}}^{*i} = \mathcal{A}_{\text{bin}}^i - \mathcal{A}_{D^0}^i = \mathcal{A}_{CP}^i + \mathcal{A}_{\text{FB}} + \mathcal{A}_{\pi_s} \tag{5.1}$$

$$\mathcal{A}_{\text{avg}} = \frac{\sum_i \left(\text{n}^i \mathcal{A}_{\text{bin}}^{*i} \right)}{\sum_i \text{n}^i}, \qquad \text{n}^i = \text{number of signal events in bin } i \tag{5.2}$$

$$\mathcal{A}_{CP,\text{loc}}^i = \mathcal{A}_{\text{bin}}^{*i} - \mathcal{A}_{\text{avg}} \tag{5.3}$$

Note that n^i includes both D^0 and $\overline{D}^0$ signal events. More detail is provided in section 5.2 as to how $\mathcal{A}_{\text{bin}}^{*i}$ is measured and how $\mathcal{A}_{CP,\text{loc}}^i$ is calculated.

Equation 5.3 gives the actual CP asymmetry in each bin relative to the Dalitz average, which can be demonstrated by some simple algebra. Substituting Equation 5.1 and

Equation 5.2 into Equation 5.3, we have:

$$\mathcal{A}^i_{CP,\text{loc}} = \mathcal{A}^i_{CP} + \mathcal{A}_{\text{FB}} + \mathcal{A}_{\pi_s} - \frac{\sum_i \left(\text{n}^i (\mathcal{A}^i_{CP} + \mathcal{A}_{\text{FB}} + \mathcal{A}_{\pi_s}) \right)}{\sum_i \text{n}^i}$$

$$= \mathcal{A}^i_{CP} + \mathcal{A}_{\text{FB}} + \mathcal{A}_{\pi_s} - \frac{\sum_i \left(\text{n}^i \mathcal{A}^i_{CP} \right)}{\sum_i \text{n}^i} - (\mathcal{A}_{\text{FB}} + \mathcal{A}_{\pi_s})$$

$$= \mathcal{A}^i_{CP} - \frac{\sum_i \left(\text{n}^i \mathcal{A}^i_{CP} \right)}{\sum_i \text{n}^i}$$

$$\mathcal{A}^i_{CP,\text{loc}} = \mathcal{A}^i_{CP} - \mathcal{A}_{CP,\text{avg}} \tag{5.4}$$

$\mathcal{A}_{\text{FB}}$ and $\mathcal{A}_{\pi_s}$ cancel (modulo some systematic uncertainty) when subtracting $\mathcal{A}_{\text{avg}}$ from $\mathcal{A}^{*i}_{\text{bin}}$, so $\mathcal{A}^i_{CP,\text{loc}}$ is indeed the difference between the real CP asymmetry in bin i ($\mathcal{A}^i_{CP}$) and the average CP asymmetry across the whole Dalitz plot.

Asymmetry	Definition
$\mathcal{A}^i_{CP}$	The asymmetry in bin i coming from real CPV in the decay. We cannot calculate this with only the information from the binned analysis.
$\mathcal{A}^i_{\text{bin}}$	The total asymmetry in bin i, consisting of real CPV, $\mathcal{A}^i_{CP}$, and the sum of three nuisance asymmetries, $\mathcal{A}_{\text{FB}}$, $\mathcal{A}_{\pi_s}$, and $\mathcal{A}^i_{D^0}$.
$\mathcal{A}^i_{D^0}$	The internal D^0 asymmetry in bin i coming from the fact that the Dalitz plot is not exactly symmetric.
$\mathcal{A}^{*i}_{\text{bin}}$	$\mathcal{A}^i_{\text{bin}}$ corrected for the internal D^0 asymmetry ($\mathcal{A}^i_{D^0}$).
$\mathcal{A}_{\text{FB}}$	The forward-backward production asymmetry in the detector, treated as bin-independent with the small bin dependence as a systematic effect.
$\mathcal{A}_{\pi_s}$	The slow pion reconstruction asymmetry in the detector, treated as bin-independent with the small bin dependence as a systematic effect.
$\mathcal{A}_{\text{avg}}$	The weighted average of $\mathcal{A}^{*i}_{\text{bin}}$ across all bins, i.e., the average asymmetry from all sources (except $\mathcal{A}^i_{D^0}$) across the whole Dalitz plot. The weighting is done according to the number of events in each bin.
$\mathcal{A}_{CP,\text{avg}}$	The weighted average of $\mathcal{A}^i_{CP}$ across all bins, i.e., the average real CPV across the whole Dalitz plot. We cannot calculate this with only the information from the binned analysis.
$\mathcal{A}^i_{CP,\text{loc}}$	The local CP asymmetry in bin i, i.e., the difference between $\mathcal{A}^i_{CP}$ and $\mathcal{A}_{CP,\text{avg}}$. This is what we ultimately want to measure in the binned analysis. $\mathcal{A}^i_{CP}$ and $\mathcal{A}_{CP,\text{avg}}$ may be found another way (see Appendix C).

Table 5.1: Definitions of asymmetries. i is the bin index in the Dalitz plot.

5.2 Binned analysis

To search for CPV in particular regions of the Dalitz plot, we first divide the plot in Figure 4.10 into six bins, as shown in Figure 5.1. Note that bins may be disjoint. These bins are symmetric across the diagonal line $m(\pi^+\pi^0)^2 = m(\pi^-\pi^0)^2$, henceforth referred to as

the "$m(\pi^{\pm}\pi^0)^2$ symmetry axis", which is the line bisecting the plot starting in the lower left corner. The symmetry of the bins allows us to correct for the internal D^0 asymmetry, $\mathcal{A}_{D^0}^i$ (defined in section 5.1 and further explained in subsection 5.2.3), when calculating the per-bin asymmetry. We perform fits to the ΔM variable in each bin (taking D^0 and $\overline{D}^0$ separately) to calculate the yield for each charge and use this yield to calculate an asymmetry. The subsections below explain this process in more detail.
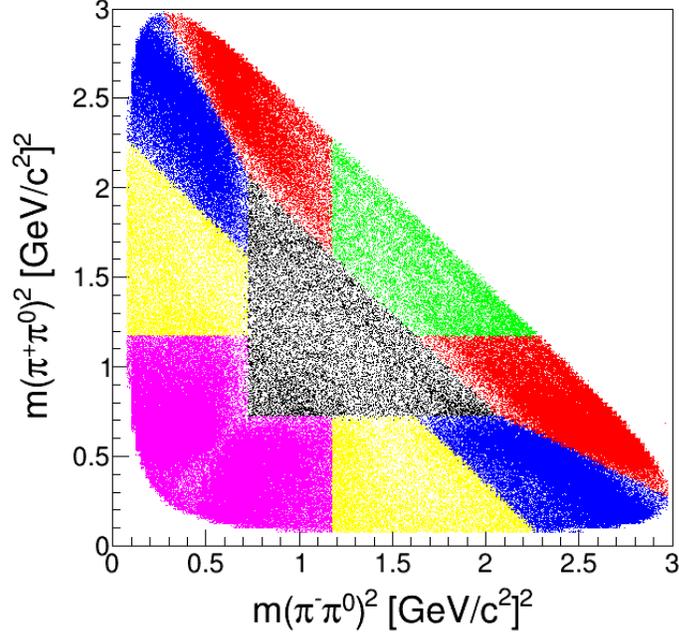
Figure 5.1: Binned version of the Dalitz plot in Figure 4.10. Each different color (black, red, green, blue, yellow, and magenta) indicates a different bin. The axis labels are for the D^0 decay, but $\overline{D}^0$ events are included as well (with the charge-conjugate axes).

5.2.1 Binning choice

Figure 5.1 shows the binned Dalitz plot used in this analysis. The number and orientation of the bins were informed by both practical and physical concerns.

The number of bins (6) is in part limited by the statistics of the plot. Since we perform fits to determine the signal yield and asymmetry in each bin (rather than sideband-subtracting and counting the number of signal events), the statistics per bin must be high enough to get a reasonable fit. Further, fits to the D^0 and $\overline{D}^0$ signal are performed separately and, as will be described in subsection 5.2.3, each bin is divided in half along the $m(\pi^{\pm}\pi^0)^2$ symmetry axis, so each bin must accommodate four separate fits. If subsequent versions of this analysis are performed on a significantly larger dataset, using more bins may be justified.

The bin orientation is informed by the structure of the Dalitz decay. As seen in Figure 4.10, the Dalitz plot is dominated by three interfering ρ resonances, each with a $\cos^2\theta$

dependence. The bins roughly follow this structure, with three bins in the signal-dense corner regions, two bins in the depleted edge regions, and one bin in the center. The orientation is also symmetric across the the $m(\pi^\pm\pi^0)^2$ symmetry axis to allow us to cancel $\mathcal{A}_{D^0}^i$.

It is worth addressing why we do not attempt to bin according to the strong phase. While binning in the strong phase makes sense when searching for CPV within the SM, we are interested in CPV from New Physics. CPV from New Physics is proportional to $\sin(\phi - \phi')$, where ϕ = the strong phase in the SM and ϕ' = the strong phase from New Physics. Therefore, it is not possible to optimize the binning without making assumptions about ϕ', which we know nothing about. If some expectations about New Physics (or other inputs relevant to binning) become clearer, subsequent versions of this analysis may use a different binning scheme.

5.2.2 Accounting for $\mathcal{A}_{\mathrm{FB}}$ and $\mathcal{A}_{\pi_s}$

The forward-backward production asymmetry, $\mathcal{A}_{\mathrm{FB}}$, and the slow pion asymmetry, $\mathcal{A}_{\pi_s}$, occur at the level of the $D^{*+} \to D^0 \pi_s^+$ decay and are approximately constant across the Dalitz plot. With perfect efficiency, they would be exactly constant, since we cannot violate causality, so asymmetries occurring at the level of the tagging decay cannot depend on position in the $D^0 \to \pi^+\pi^-\pi^0$ Dalitz plot. We do not have perfect efficiency, so some small Dalitz bin-dependence may exist. Systematics to account for the small differences in each bin will be discussed in chapter 8. $\mathcal{A}_{\mathrm{FB}}$ comes from γ–Z^0 interference and higher order QED effects, and $\mathcal{A}_{\pi_s}$ comes from the different detector efficiencies for positive and negative tracks. We do not expect these asymmetries to vary significantly across the Dalitz plot, but this can be checked via plots of kinematic variables correlated with $\mathcal{A}_{\mathrm{FB}}$ and $\mathcal{A}_{\pi_s}$ in each bin.

$\mathcal{A}_{\mathrm{FB}}$ is related to the center-of-mass (CM) angle of the D^* candidate: if plots of $\cos\theta_{D^*}^*$ (the asterisk on θ indicates the angle is measured in the CM frame) in each bin show that this variable behaves the same way across all bins, it is reasonable to conclude that $\mathcal{A}_{\mathrm{FB}}$ is constant across all bins and cancels out of the final asymmetry calculation, as shown in Equation 5.4. Similarly, $\mathcal{A}_{\pi_s}$ is related to the slow pion kinematics, specifically the momentum, p_{π_s}, and angle, $\cos\theta_{\pi_s}$. If these variables behave the same way across all Dalitz bins, it is reasonable to conclude that $\mathcal{A}_{\pi_s}$ is constant and cancels out of the final asymmetry calculation, as shown in Equation 5.4. Figures 5.3 – 5.8 show these normalized kinematic distributions. (Note that not all bins are equally populated; bins with fewer signal events exhibit larger fluctuations in these distributions.) Specifically:

- Figure 5.3 shows the $\cos\theta_{D^*}^*$ distribution in all MC events as well as in signal and background separately. While it appears in the case of all events that there is some disagreement between these distributions in the different bins, this disagreement disappears when only signal events are considered. We conclude that any discrepancy observed here is the result of varying signal-to-background ratios between the bins, as bins 1, 3 and 5 have a much higher signal-to-background ratio than bins 0, 2, and 4, and not due to differences in behavior within the signal.

- Figure 5.4 shows the $\cos\theta_{D^*}^*$ distribution in data. Based on what we observed in MC in Figure 5.3, we conclude that any discrepancy observed here is again the result of

varying signal-to-background ratios between the bins. Note that in data the signal-to-background ratio is worse in general, and especially in bins 0, 2, and 4, than it is in MC. This explains why the disagreement between the bins in Figure 5.4 is exaggerated as compared to Figure 5.3a.

- Figure 5.5 (Figure 5.7) shows the p_{π_s} ($\cos\theta_{\pi_s}$) distribution in all MC events as well as in signal and background separately. While it appears in the case of all events that there is some disagreement between these distributions in the different bins, this disagreement disappears when only signal events are considered. We conclude that any discrepancy observed here is the result of varying signal-to-background ratios between the bins, as bins 1, 3 and 5 have a much higher signal-to-background ratio than bins 0, 2, and 4, and not due to differences in behavior within the signal.

- Figure 5.6 (Figure 5.8) shows the p_{π_s} ($\cos\theta_{\pi_s}$) distribution in data. Based on what we observed in MC in Figure 5.5 (Figure 5.7), we conclude that any discrepancy observed here is again the result of varying signal-to-background ratios between the bins. Note that in data the signal-to-background ratio is worse in general, and especially in bins 0, 2, and 4, than it is in MC. This explains why the disagreement between the bins in Figure 5.6 (Figure 5.8) is exaggerated as compared to Figure 5.5a (Figure 5.7a).

Based on Figures 5.3 and 5.4, we conclude that it is reasonable to treat $\mathcal{A}_{\mathrm{FB}}$ as approximately constant across the Dalitz plot. Similarly, based on Figures 5.5, 5.6, 5.7, and 5.8, we conclude that it is reasonable to treat $\mathcal{A}_{\pi_s}$ as approximately constant across the Dalitz plot. Since global asymmetries cancel out when two particular asymmetry measurements are subtracted, we never explicitly determine $\mathcal{A}_{\mathrm{FB}}$ or $\mathcal{A}_{\pi_s}$ as part of this analysis. The cancellation of both $\mathcal{A}_{\mathrm{FB}}$ and $\mathcal{A}_{\pi_s}$ has some small systematic uncertainty, which we discuss in chapter 8.

5.2.3 Accounting for $\mathcal{A}_{D^0}^i$

Unlike $\mathcal{A}_{\mathrm{FB}}$ and $\mathcal{A}_{\pi_s}$, which are approximately constant across the Dalitz plot, $\mathcal{A}_{D^0}^i$ must be accounted for separately in each bin. Even though the final state we are interested in, $\pi^+\pi^-\pi^0$, is self-conjugate, so both the D^0 and $\bar{D}^0$ candidates contain one π^+ and one π^-, the Dalitz plot and subregions of it are *not* self-conjugate. The Dalitz structure is intrinsically related to the momenta of the decay products at any point on the plot, and in an arbitrary region of the Dalitz plot, the π^+ and π^- particles will not have the same momentum spectrum. Thus, when charge-conjugating, the π^+ from the D^0 decay has a different momentum spectrum than the π^+ from the $\bar{D}^0$ decay (and similarly for the π^-). This then couples into the detector effects which cause positive and negative particles to be detected at different rates (the same issue that causes $\mathcal{A}_{\pi_s}$) and leaves us with an intrinsic D^0 asymmetry in each bin, which we call $\mathcal{A}_{D^0}^i$. Note that the way in which the spectra differ can change bin-to-bin, which is why this is not a global asymmetry but is bin-dependent. This asymmetry can be accounted for by choosing a binning structure that is symmetric across the $m(\pi^\pm\pi^0)^2$ symmetry axis, which we have done (see Figure 5.1). However, this does not completely fix the issue, as two other concerns remain. First, since we have bins of a finite size, as opposed to infinitesimally small bins, it is possible that even within a symmetric bin, the *position* of events may not be exactly symmetric across the $m(\pi^\pm\pi^0)^2$

symmetry axis. This becomes more of a concern as the bins get larger, but six bins is enough to mitigate this problem considerably. Second, the *number* of events on one side of the symmetry axis as compared to the other may be significantly different. This is due to the underlying, non-self-conjugate structure of the Dalitz decay. For example, the fit fraction (see Appendix A for a description of fit fractions) for the resonance $D^0 \to \rho^+\pi^-$ ($\overline{D}^0 \to \rho^-\pi^+$) is $(1.01 \pm 0.04)\%$, while the fit fraction for the charge-conjugate resonance $D^0 \to \rho^-\pi^+$ ($\overline{D}^0 \to \rho^+\pi^-$) is $(0.515 \pm 0.025)\%$ [33]. The former occurs almost twice as frequently as the latter, so a rate difference across the Dalitz plot should be expected. This is also clearly visible in the contour plots in Figure 4.9 and Figure 4.10.
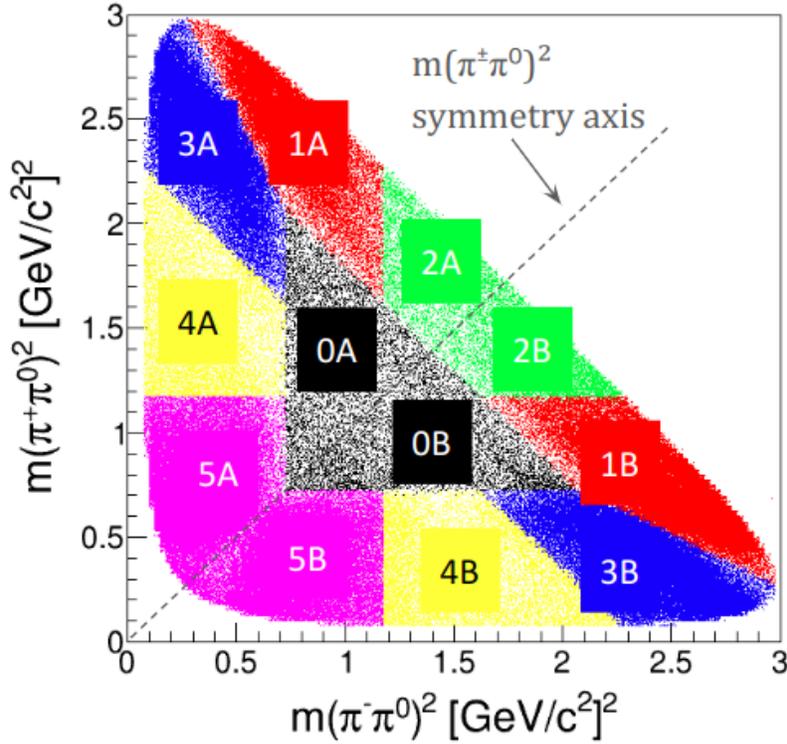
Figure 5.2: Figure 5.1 with additional labels for the $m(\pi^\pm\pi^0)^2$ symmetry axis and the upper (A) and lower (B) bin regions.

We must go one step further to account for this normalization effect. In addition to choosing bins that are symmetric across the $m(\pi^\pm\pi^0)^2$ symmetry axis, we also split each bin into the region above the axis (A) and the region below the axis (B), illustrated in Figure 5.2. We calculate an asymmetry separately in region A ($\mathcal{A}_A^i$) and region B ($\mathcal{A}_B^i$), and then we take an unweighted average of these asymmetries to determine $\mathcal{A}_{\text{bin}}^{*i}$:

$$\mathcal{A}_{\text{bin}}^{*i} = \mathcal{A}_{\text{bin}}^i - \mathcal{A}_{D^0}^i = \frac{\mathcal{A}_A^i + \mathcal{A}_B^i}{2} \qquad (5.5)$$

To demonstrate that this works (i.e., this method effectively corrects for $\mathcal{A}_{D^0}^i$), note that an equivalent method would be to weight the number of events so that region A and region B contributed the same number of events to the asymmetry calculation and then to calculate

$\mathcal{A}_{\text{bin}}^{*i}$ directly. Therefore, if we plot the π^+ and π^- p and $\cos\theta$ distributions in each bin, with the number of events weighted so that the same number are coming from above and below the axis, we expect the distributions to be the same. Figures 5.9 – 5.16 demonstrate this procedure. In each plot, the number above the plot corresponds to the bin number, as defined in Figure 5.2; the solid line is the momentum or cosine of the π^+ (π^-) in the D^0 ($\overline{D}^0$) decay, and the dashed line is the momentum or cosine of the π^- (π^+) in the D^0 ($\overline{D}^0$) decay. In weighted plots, events in region B of each bin are weighted by N_A^i/N_B^i, where $N_A^i =$ the number of events in region A of bin i and $N_B^i =$ the number of events in region B of bin i. The specific plots are:

- Figure 5.9 (Figure 5.13) shows the **unweighted** p_π $(\cos\theta_\pi^*)$ distribution in truth-tagged signal MC. Although the cosine distributions are in relatively good agreement even without weighting, there is a noticeable difference in the π^+ and π^- momenta, especially in bins 1 and 3.

- Figure 5.10 (Figure 5.14) shows the **weighted** p_π $(\cos\theta_\pi^*)$ distribution in truth-tagged signal MC. The difference in the spectra that we observe in Figure 5.9 has largely disappeared.

- Figure 5.11 (Figure 5.15) shows the **weighted** p_π $(\cos\theta_\pi^*)$ distribution in all MC. Even in the presence of background, the weighting procedure brings the spectra in each bin into much better agreement.

- Figure 5.12 (Figure 5.16) shows the **weighted** p_π $(\cos\theta_\pi^*)$ distribution in data. We cannot look at the signal by itself in data, but as shown by comparing plots of signal MC to plots of all MC, even in the presence of background, the weighting procedure brings the spectra in each bin into much better agreement.

Based on Figures 5.9 – 5.16, we conclude that most of the disagreement between the fast pion kinematics in a particular bin is due to the rate difference above and below the $m(\pi^\pm\pi^0)^2$ symmetry axis and that our method of taking an unweighted average of $\mathcal{A}_A^i$ and $\mathcal{A}_B^i$ effectively cancels $\mathcal{A}_{D^0}^i$ (modulo systematic uncertainties that are discussed in chapter 8).

5.2.4   Performing yield fits

To determine $\mathcal{A}_A^i$, $\mathcal{A}_B^i$, and $\mathcal{A}_{\text{bin}}^{*i}$ for each bin i, we perform a yield fit to a plot of the ΔM variable in the $M(\pi^+\pi^-\pi^0)$ signal region. Each bin requires four fits to determine signal yield: 1) fit to D^0 in region A, 2) fit to $\overline{D}^0$ in region A, 3) fit to D^0 in region B, and 4) fit to $\overline{D}^0$ in region B. Table 5.2 defines the relevant signal parameters for each of these four fits.

Each fit is performed with a binned, log-likelihood fit. The signal function is a Johnson's S_U-distribution or, in the case of a lower statistics bin, a single Gaussian. A Johnson's S_U-distribution is similar to a Gaussian but with an asymmetric tail; in addition to the mean and width parameters, μ and σ, the Johnson's S_U-distribution has two additional shape parameters, γ and δ [59, 60]. The background function is the sum of 1/2-power and 3/2-power threshold functions with a single shape parameter, α, that gives the ratio of the 1/2-power and 3/2-power functions. (This is referred to as just a "a 3/2-power threshold

Variable	Definition
n_A^i (σ_A^i)	The number (error on the number) of D^0 signal events in region A of bin i.
$\bar{n}_A^i$ $(\bar{\sigma}_A^i)$	The number (error on the number) of $\overline{D}^0$ signal events in region A of bin i.
n_B^i (σ_B^i)	The number (error on the number) of D^0 signal events in region B of bin i.
$\bar{n}_B^i$ $(\bar{\sigma}_B^i)$	The number (error on the number) of $\overline{D}^0$ signal events in region B of bin i.

Table 5.2: Definitions of yield parameters for the four fits performed in each Dalitz bin. Note that in Equation 5.2, $\mathtt{n}^i = n_A^i + \bar{n}_A^i + n_B^i + \bar{n}_B^i$.

function" going forward.) In the case of peaking background, a single Gaussian is added to the background function.

All fit functions are normalized so that the signal and background both integrate to 1 over the region of the fit. Two normalization parameters (one for signal, one for background) are included directly in the fit, and the yield and its associated error is simply related to the signal normalization parameter. Specifically, if $I_{A,\mathrm{sig}}^i$ is the signal normalization parameter for the D^0 fit in region A of bin i, then:

$$n_A^i = I_{A,\mathrm{sig}}^i / w \qquad (5.6)$$

and similarly for $\bar{n}_A^i$ and region B. w is the bin width in the ΔM histogram being fit. In every case for our nominal analysis, $w = 0.07\,\mathrm{MeV}/c^2$.

The specific fit functions used for the bins in this analysis in MC (data) are discussed in chapter 6 (chapter 7), and plots of fits and residuals are included as well.

We explored the possibility of using simultaneous fitting for the D^0 and $\overline{D}^0$ distributions in a particular region. This involves constraining the background shape or signal shape parameters (or both) to be the same for both the D^0 and $\overline{D}^0$ fits. The main motivation for this would be to lower the relative error on the final asymmetry measurements. When tested on MC, constraining the background shape parameter α improved the statistical uncertainty on $\mathcal{A}_{CP,\mathrm{loc}}^i$ by 5–10%. Constraining the signal width σ did slightly better, improving the statistical uncertainty by 6–16%. However, the relative error is not the only consideration. We also need good evidence that it is reasonable to expect a particular parameter to be the same in the D^0 and $\overline{D}^0$ distributions. Especially since we are trying to measure an asymmetry, forcing a parameter that is not the same in reality to be the same in the fits might reduce the statistical uncertainty but induce a fake asymmetry in the process. The signal width σ is potentially susceptible to this: different detector resolutions for positive and negative particles could lead to real differences in σ between the D^0 and $\overline{D}^0$ distributions. In light of these concerns and observing only a modest improvement in the uncertainties, we opted against simultaneous fitting. Future versions of this analysis may revisit simultaneous fitting or change the fit parameterization in other ways if further investigation indicates this is warranted.

5.2.5   Calculating $\mathcal{A}^i_{CP,\text{loc}}$

Once we have obtained the yield parameters in Table 5.2 from the fits described in subsection 5.2.4, we can calculate $\mathcal{A}^i_{CP,\text{loc}}$. First, we calculate $\mathcal{A}^i_{\text{A}}$, $\mathcal{A}^i_{\text{B}}$, $\mathcal{A}^{*i}_{\text{bin}}$, and their associated errors for each bin i:

$$\mathcal{A}^i_{\text{A}} \pm \sigma_{\mathcal{A}^i_{\text{A}}} = \frac{n^i_A - \bar{n}^i_A}{n^i_A + \bar{n}^i_A} \pm \frac{2\sqrt{(n^i_A\bar{\sigma}^i_A)^2 + (\bar{n}^i_A\sigma^i_A)^2}}{(n^i_A + \bar{n}^i_A)^2} \tag{5.7}$$

$$\mathcal{A}^i_{\text{B}} \pm \sigma_{\mathcal{A}^i_{\text{B}}} = \frac{n^i_B - \bar{n}^i_B}{n^i_B + \bar{n}^i_B} \pm \frac{2\sqrt{(n^i_B\bar{\sigma}^i_B)^2 + (\bar{n}^i_B\sigma^i_B)^2}}{(n^i_B + \bar{n}^i_B)^2} \tag{5.8}$$

$$\mathcal{A}^{*i}_{\text{bin}} \pm \sigma^{*i}_{\text{bin}} = \frac{\mathcal{A}^i_{\text{A}} + \mathcal{A}^i_{\text{B}}}{2} \pm \frac{\sqrt{\sigma^2_{\mathcal{A}^i_{\text{A}}} + \sigma^2_{\mathcal{A}^i_{\text{B}}}}}{2} \tag{5.9}$$

We can then calculate $\mathcal{A}^i_{CP,\text{loc}}$ using Equation 5.2 and Equation 5.3. We avoid correlated errors by rearranging the sum in Equation 5.2 as follows:

$$\mathcal{A}^i_{CP,\text{loc}} = \frac{\mathcal{A}^i_{\text{A}} + \mathcal{A}^i_{\text{B}}}{2} - \frac{\sum_j \left(\frac{\mathcal{A}^j_{\text{A}} + \mathcal{A}^j_{\text{B}}}{2} \mathbb{n}^j \right)}{\sum_j \mathbb{n}^j} \tag{5.10}$$

$$= \frac{\mathcal{A}^i_{\text{A}} + \mathcal{A}^i_{\text{B}}}{2} - \frac{(\mathcal{A}^i_{\text{A}} + \mathcal{A}^i_{\text{B}})\,\mathbb{n}^i}{2\sum_j \mathbb{n}^j} - \frac{\sum_{j\neq i} (\mathcal{A}^j_{\text{A}} + \mathcal{A}^j_{\text{B}})\,\mathbb{n}^j}{2\sum_j \mathbb{n}^j} \tag{5.11}$$

$$= \frac{(\mathcal{A}^i_{\text{A}} + \mathcal{A}^i_{\text{B}})\left(\sum_j \mathbb{n}^j\right) - (\mathcal{A}^i_{\text{A}} + \mathcal{A}^i_{\text{B}})\,\mathbb{n}^i}{2\sum_j \mathbb{n}^j} - \frac{\sum_{j\neq i} (\mathcal{A}^j_{\text{A}} + \mathcal{A}^j_{\text{B}})\,\mathbb{n}^j}{2\sum_j \mathbb{n}^j} \tag{5.12}$$

$$= \frac{(\mathcal{A}^i_{\text{A}} + \mathcal{A}^i_{\text{B}})\left(\left(\sum_j \mathbb{n}^j\right) - \mathbb{n}^i\right)}{2\sum_j \mathbb{n}^j} - \frac{\sum_{j\neq i} (\mathcal{A}^j_{\text{A}} + \mathcal{A}^j_{\text{B}})\,\mathbb{n}^j}{2\sum_j \mathbb{n}^j} \tag{5.13}$$

$$\boxed{\mathcal{A}^i_{CP,\text{loc}} = \frac{(\mathcal{A}^i_{\text{A}} + \mathcal{A}^i_{\text{B}})\left(\sum_{j\neq i} \mathbb{n}^j\right) - \sum_{j\neq i} (\mathcal{A}^j_{\text{A}} + \mathcal{A}^j_{\text{B}})\,\mathbb{n}^j}{2\sum_j \mathbb{n}^j}} \tag{5.14}$$

Propagating the errors from Equation 5.7 and Equation 5.8, and ignoring the error on $\mathbb{n}^i$, we have:

$$\boxed{\sigma^i_{CP,\text{loc}} = \frac{1}{2\sum_j \mathbb{n}^j} \sqrt{\left(\sigma^2_{\mathcal{A}^i_{\text{A}}} + \sigma^2_{\mathcal{A}^i_{\text{B}}}\right)\left(\sum_{j\neq i} \mathbb{n}^j\right)^2 + \sum_{j\neq i}\left(\sigma^2_{\mathcal{A}^j_{\text{A}}} + \sigma^2_{\mathcal{A}^j_{\text{B}}}\right)(\mathbb{n}^j)^2}} \tag{5.15}$$

5.3   Summary

The binned analysis method described in this chapter fully accounts for the three nuisance asymmetries $\mathcal{A}_{\pi_s}$, $\mathcal{A}_{\text{FB}}$, and $\mathcal{A}^i_{D^0}$ in a data-based way, and it allows for a direct calculation of the per-bin asymmetries across the Dalitz plot even in the presence of a non-negligible number of background events. This method is also scalable, if more data is used in the

future, and allows for changes to the procedure, such as finer binning or a different fitting method, if a larger dataset warrants them.

It is worth noting that this analysis does *not* capture any real CP asymmetries that are constant across the Dalitz plot. A complementary analysis to this one could check for CPV at this global level via a Dalitz-integrated analysis (i.e., one that is independent of the structure of the Dalitz plot). Appendix C describes what such an analysis could look like.
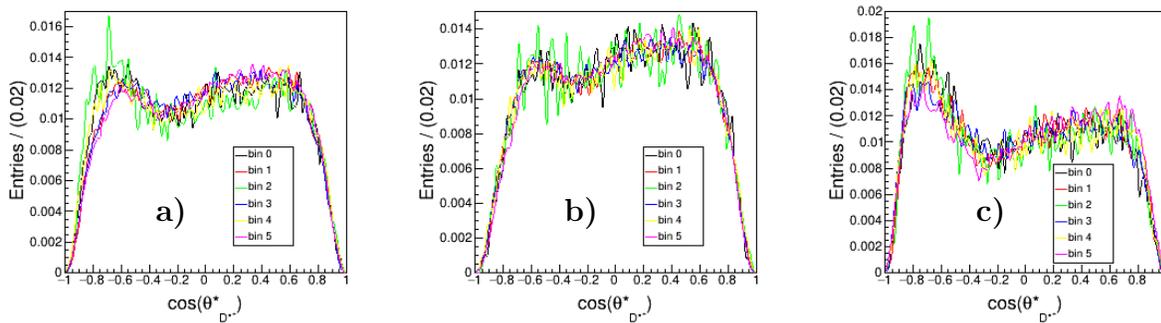
Figure 5.3: Comparison of $\cos\theta_{D*}^*$ in Dalitz plot bins in MC in **a)** all events, **b)** signal events, and **c)** background events.
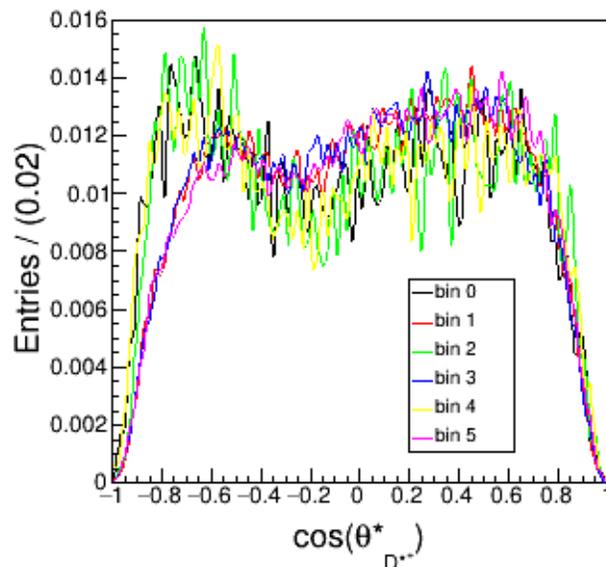
Figure 5.4: Comparison of $\cos\theta_{D*}^*$ in Dalitz plot bins in data. This plot was made with a $187\,\mathrm{fb}^{-1}$ subset of the full dataset.
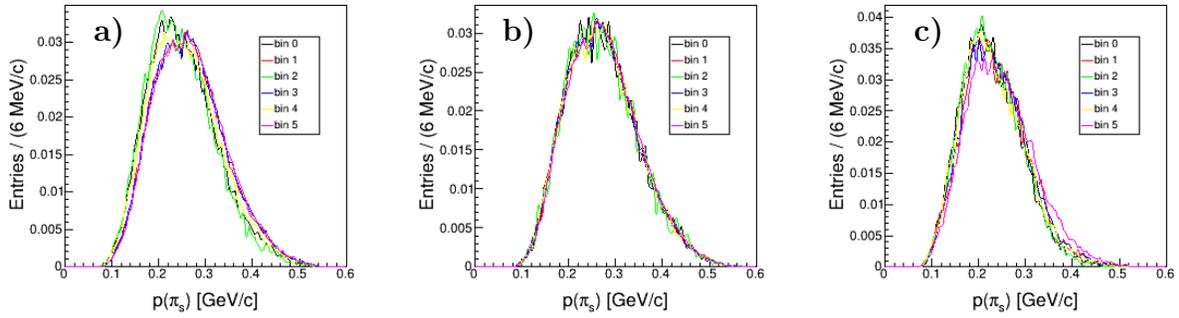
Figure 5.5: Comparison of p_{π_s} in Dalitz plot bins in MC in **a)** all events, **b)** signal events, and **c)** background events.
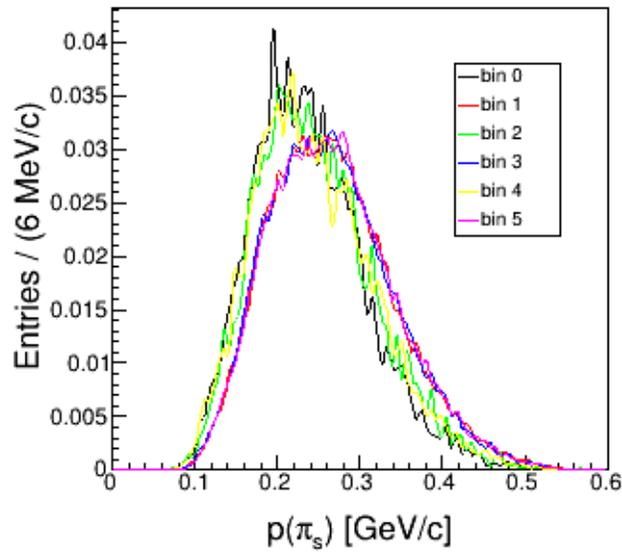
Figure 5.6: Comparison of p_{π_s} in Dalitz plot bins in data. This plot was made with a $187\,\mathrm{fb}^{-1}$ subset of the full dataset.
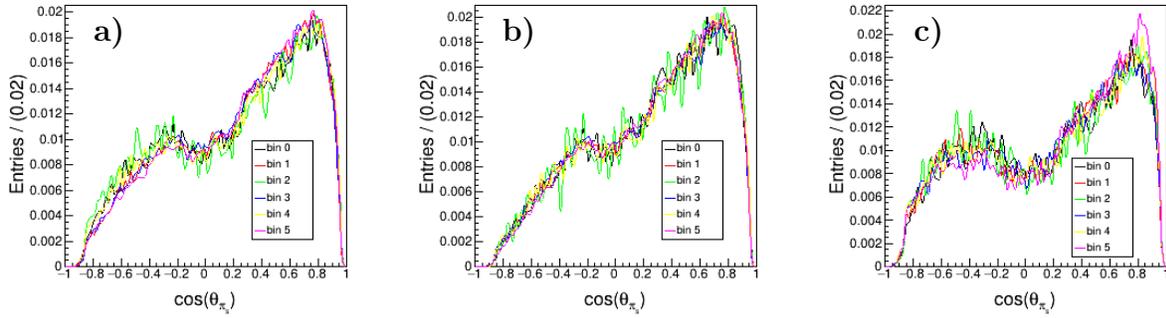
Figure 5.7: Comparison of $\cos\theta_{\pi_s}$ in Dalitz plot bins in MC in **a)** all events, **b)** signal events, and **c)** background events.
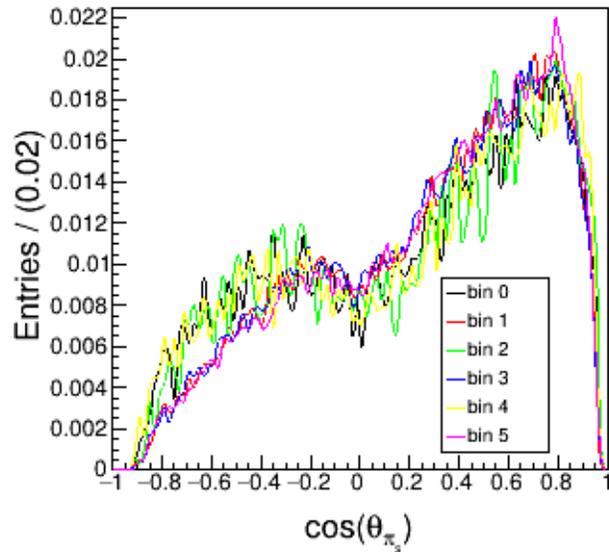
Figure 5.8: Comparison of $\cos\theta_{\pi_s}$ in Dalitz plot bins in data. This plot was made with a $187\,\text{fb}^{-1}$ subset of the full dataset.
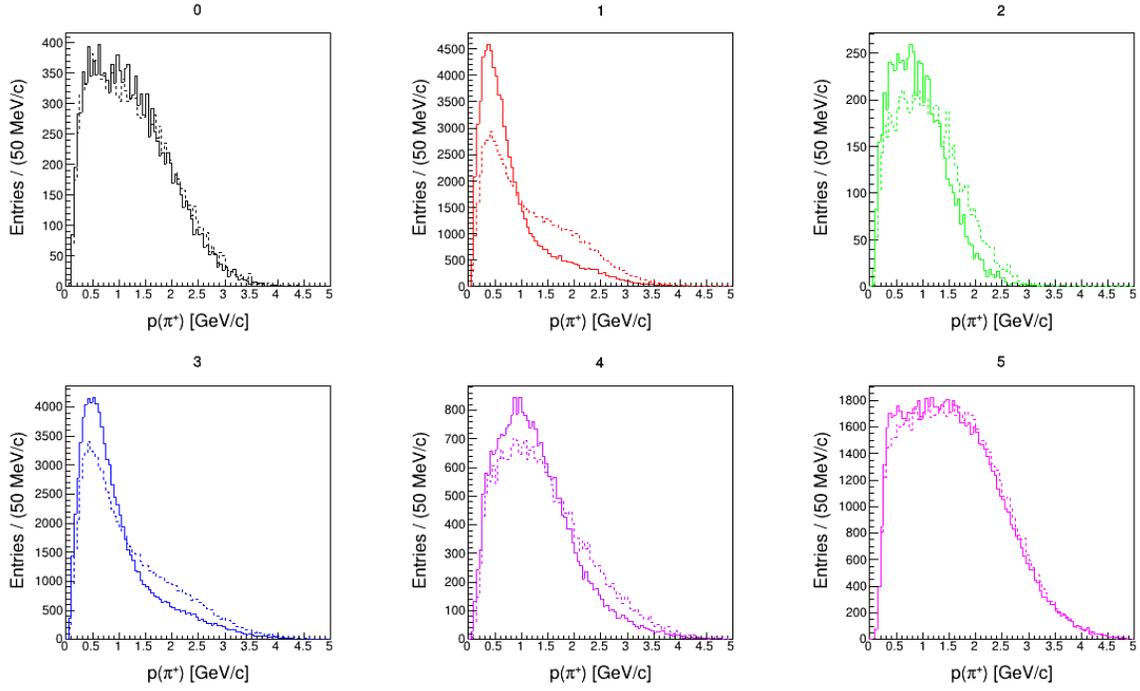
Figure 5.9: Unweighted pion p in truth-tagged signal MC. The number above each plot corresponds to the bin number, as defined in Figure 5.2. In each plot, the solid line is the momentum of the π^+ (π^-) in the D^0 $(\overline{D}^0)$ decay, and the dashed lines is the momentum of the π^- (π^+) in the D^0 $(\overline{D}^0)$ decay.
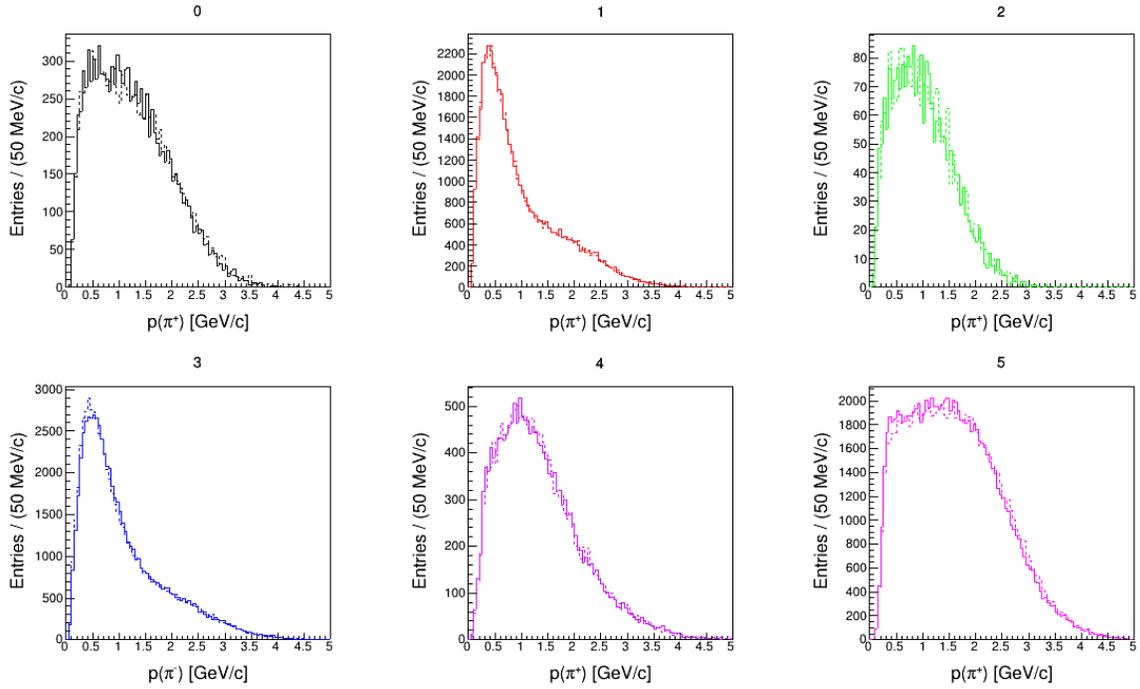
Figure 5.10: Weighted pion p in truth-tagged signal MC. The number above each plot corresponds to the bin number, as defined in Figure 5.2. In each plot, the solid line is the momentum of the π^+ (π^-) in the D^0 $(\overline{D}^0)$ decay, and the dashed lines is the momentum of the π^- (π^+) in the D^0 $(\overline{D}^0)$ decay. Events in region B of each bin are weighted by N_A^i/N_B^i, where N_A^i = the number of events in region A of bin i and N_B^i = the number of events in region B of bin i.
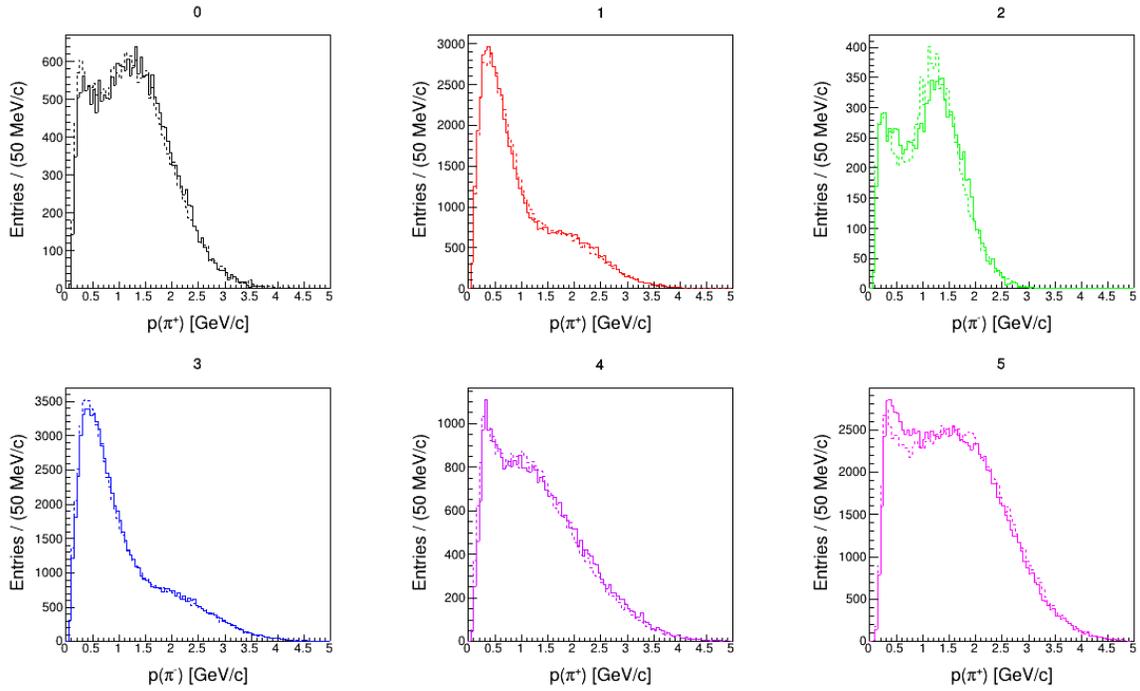
Figure 5.11: Weighted pion p in all MC. The number above each plot corresponds to the bin number, as defined in Figure 5.2. In each plot, the solid line is the momentum of the π^+ (π^-) in the D^0 ($\overline{D}^0$) decay, and the dashed lines is the momentum of the π^- (π^+) in the D^0 ($\overline{D}^0$) decay. Events in region B of each bin are weighted by N_A^i/N_B^i, where $N_A^i =$ the number of events in region A of bin i and $N_B^i =$ the number of events in region B of bin i.
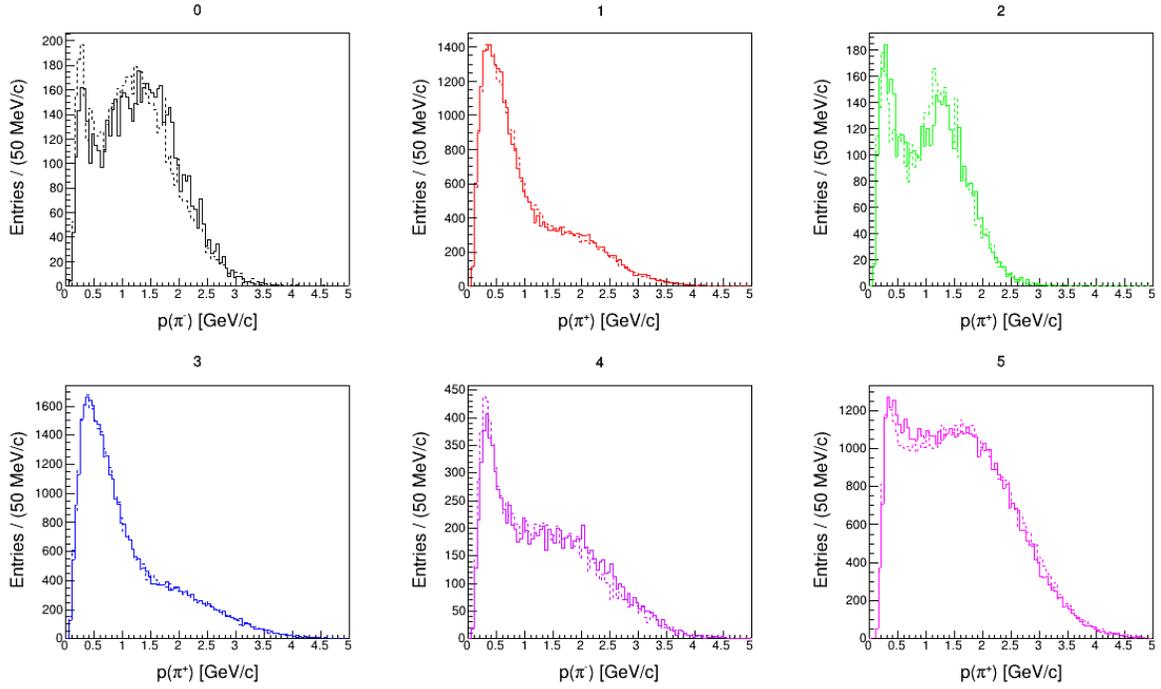
Figure 5.12: Weighted pion p in data. The number above each plot corresponds to the bin number, as defined in Figure 5.2. In each plot, the solid line is the momentum of the π^+ (π^-) in the D^0 ($\overline{D}^0$) decay, and the dashed line is the momentum of the π^- (π^+) in the D^0 ($\overline{D}^0$) decay. Events in region B of each bin are weighted by N_A^i/N_B^i, where $N_A^i =$ the number of events in region A of bin i and $N_B^i =$ the number of events in region B of bin i. This plot was made with a $187\,\mathrm{fb}^{-1}$ subset of the full dataset.
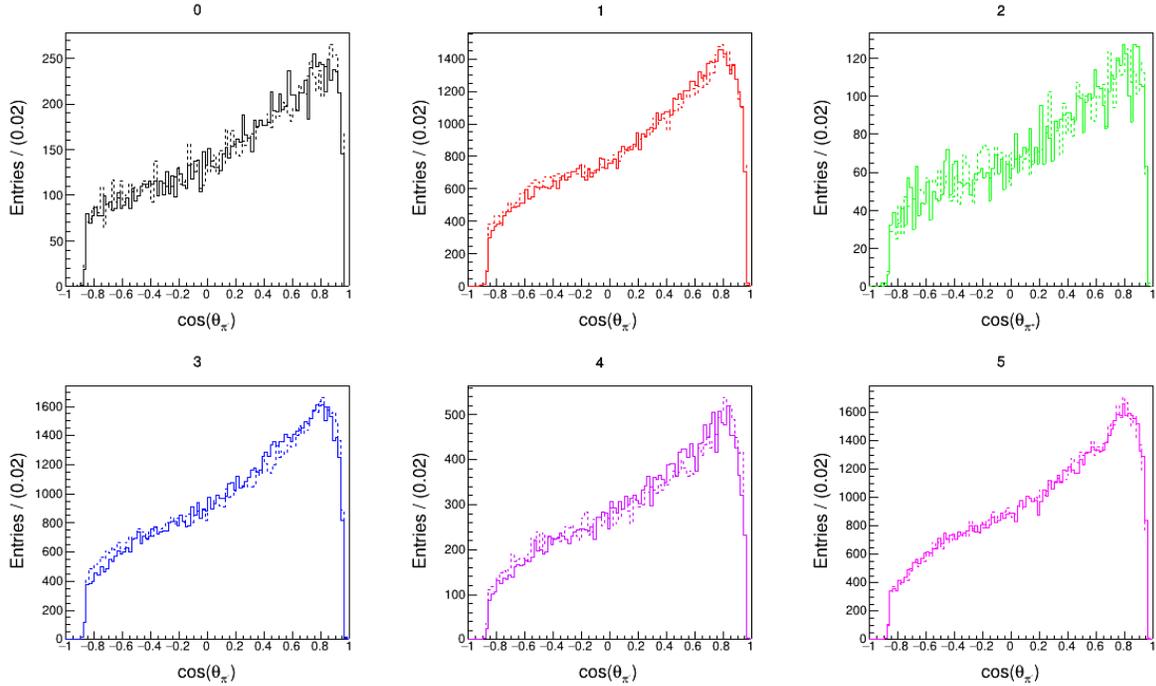
Figure 5.13: Unweighted pion $\cos\theta$ distributions in truth-tagged signal MC. The number above each plot corresponds to the bin number, as defined in Figure 5.2. In each plot, the solid line is the cosine of the π^+ (π^-) in the D^0 ($\overline{D}^0$) decay, and the dashed lines is the cosine of the π^- (π^+) in the D^0 ($\overline{D}^0$) decay.

47
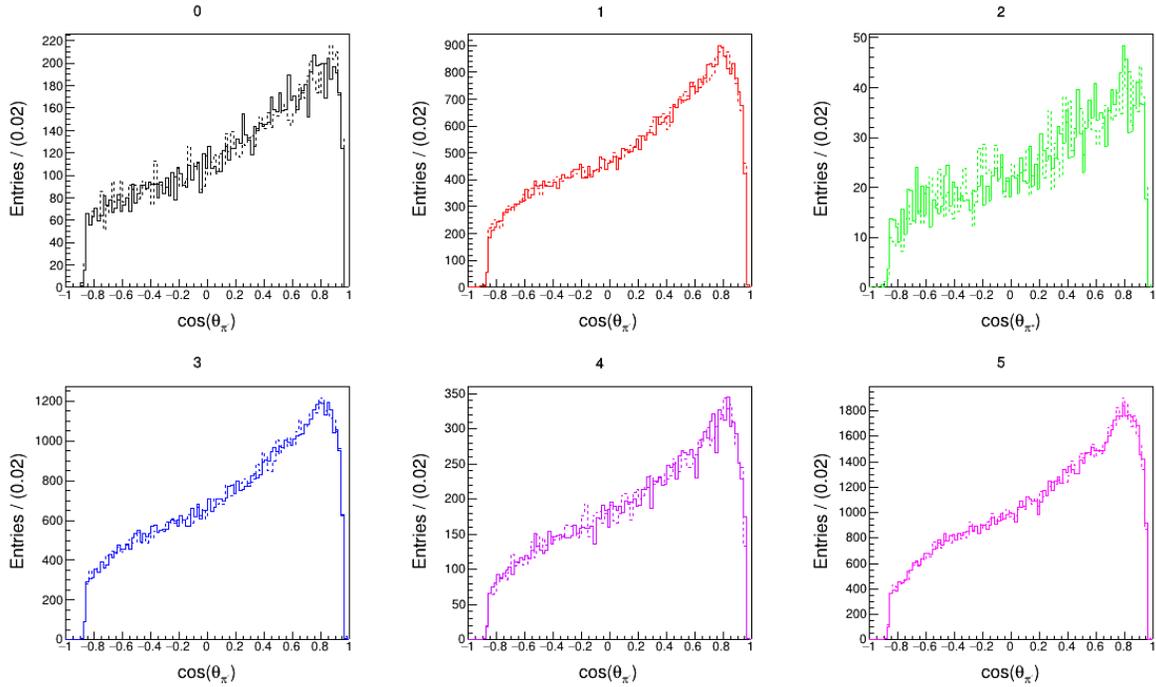
Figure 5.14: Weighted pion $\cos\theta$ distributions in truth-tagged signal MC. The number above each plot corresponds to the bin number, as defined in Figure 5.2. In each plot, the solid line is the cosine of the π^+ (π^-) in the D^0 ($\overline{D}^0$) decay, and the dashed lines is the cosine of the π^- (π^+) in the D^0 ($\overline{D}^0$) decay. Events in region B of each bin are weighted by N_A^i/N_B^i, where N_A^i = the number of events in region A of bin i and N_B^i = the number of events in region B of bin i.
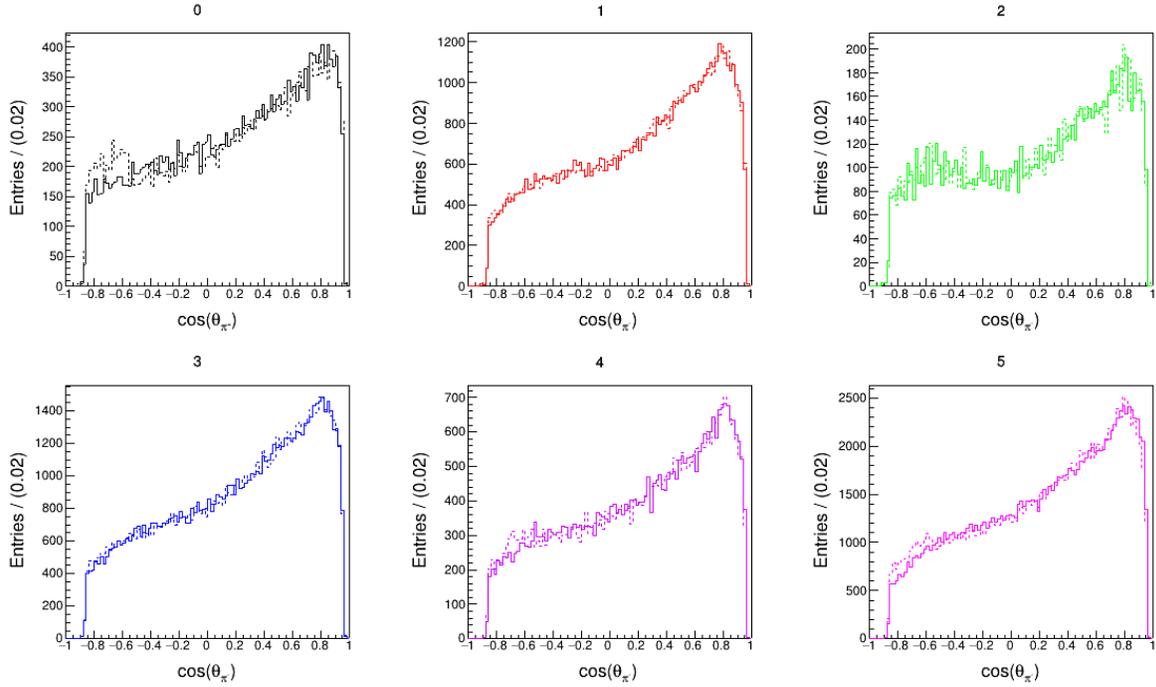
Figure 5.15: Weighted pion $\cos\theta$ distributions in all MC. The number above each plot corresponds to the bin number, as defined in Figure 5.2. In each plot, the solid line is the cosine of the π^+ (π^-) in the D^0 ($\overline{D}^0$) decay, and the dashed line is the cosine of the π^- (π^+) in the D^0 ($\overline{D}^0$) decay. Events in region B of each bin are weighted by N_A^i/N_B^i, where $N_A^i =$ the number of events in region A of bin i and $N_B^i =$ the number of events in region B of bin i.
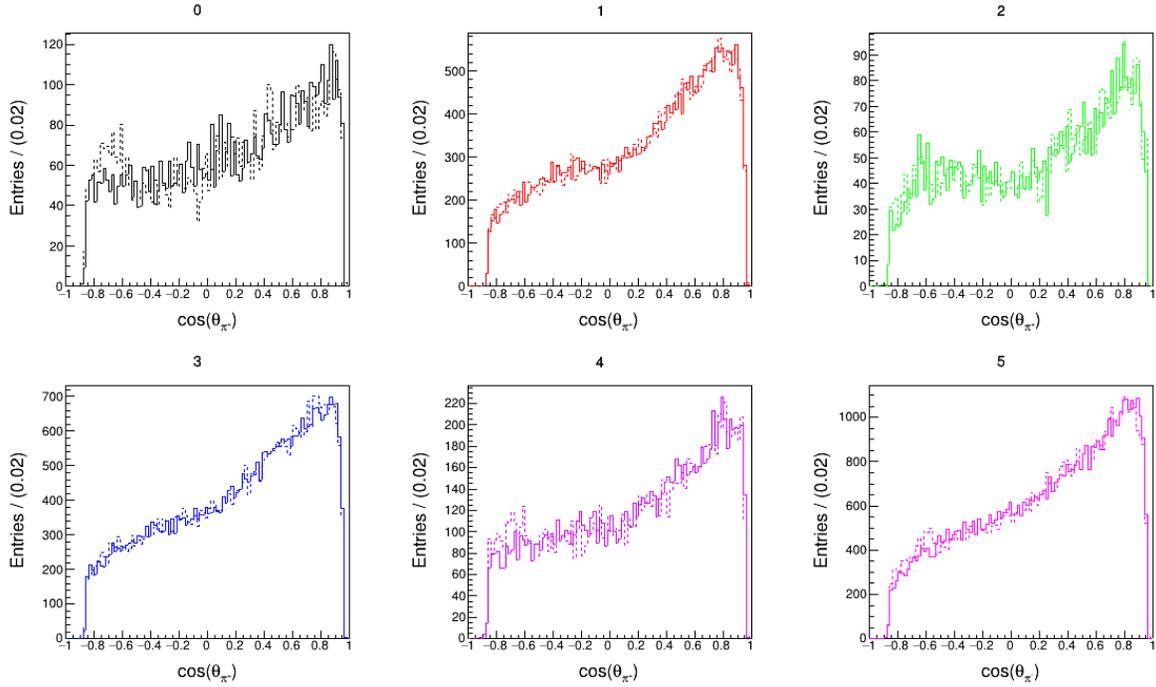
Figure 5.16: Weighted pion $\cos\theta$ distributions in data. The number above each plot corresponds to the bin number, as defined in Figure 5.2. In each plot, the solid line is the cosine of the π^+ (π^-) in the D^0 $(\overline{D}^0)$ decay, and the dashed line is the cosine of the π^- (π^+) in the D^0 $(\overline{D}^0)$ decay. This plot was made with a $187\,\mathrm{fb}^{-1}$ subset of the full dataset. Events in region B of each bin are weighted by N_A^i/N_B^i, where $N_A^i =$ the number of events in region A of bin i and $N_B^i =$ the number of events in region B of bin i.

Chapter 6

Methodology Test with Monte Carlo

This chapter presents the results of performing the analysis described in chapter 5 on the Monte Carlo data (MC) described in chapter 4.

Since we know that in MC there is no CP asymmetry coming from CPV within the decay process, all CP asymmetries present in the MC are nuisance asymmetries, and we should observe $\mathcal{A}_{CP,\mathrm{loc}}^i$ consistent with zero in all bins. Performing the full analysis on MC thus validates our analysis procedure: if we do not find $\mathcal{A}_{CP,\mathrm{loc}}^i$ to be consistent with zero in all bins, this would indicate that our procedure does not correctly account for all nuisance asymmetries.

6.1 Fit functions

In each region (A or B) of each bin (1–6) (see Figure 5.2 for visualization), two fits were performed (one for D^0 and one for $\overline{D}^0$) for a total of 24 fits. Table 6.1 (Table 6.2) summarizes the fit functions used in region A (B) to determine signal yield. In most cases, a Johnson's S_U-distribution was used as the signal function, and a 3/2-power threshold function was used as the background. The exceptions to this were 1) in region A of bin 2, a single Gaussian was used as the signal function because the statistics were too low to accommodate a function with more parameters (a 3/2-power threshold function was still used as the background), and 2) in regions A and B of bin 5, a single Gaussian was added to the 3/2-power threshold function to accommodate the peaking background in that region (a Johnson's S_U-distribution was still used as the signal). In each region, the same function was used for D^0 and $\overline{D}^0$ fits.

i	Signal function	Background function
0	Johnson's S_U-distribution	
1		
2	Single Gaussian	3/2-power threshold function
3	Johnson's S_U-distribution	
4		
5		Single Gaussian and 3/2-power threshold function

Table 6.1: Fit functions for fits to D^0 and $\overline{D}^0$ ΔM plots in MC in **region A** for all six bins.

i	Signal function	Background function
0		
1		
2	Johnson's S_U-distribution	3/2-power threshold function
3		
4		
5		Single Gaussian and 3/2-power threshold function

Table 6.2: Fit functions for fits to D^0 and $\overline{D}^0$ ΔM plots in MC in **region B** for all six bins.

6.2 Fit results

The fits to D^0 and $\overline{D}^0$ in each of the twelve Dalitz regions are shown in Figures 6.1 – 6.6. The raw yields and fit quality (i.e., χ^2/n.d.f.) are listed in Table 6.3 (Table 6.4) for region A (B). The number of degrees of freedom (n.d.f.) in each case is the number of bins in the ΔM histogram being fit minus the number of fit parameters. The number of bins is always 210, and the number of fit parameters varies between 5 and 10 depending on the fit function being used. The full list of signal parameters for each fit is given in Tables 6.6 – 6.11.

We observe much greater precision in the raw asymmetries for bins 1 and 3 as compared to bins 0, 2, and 4. This is due to the higher statistics in the odd-numbered bins. The precision in bin 5 is an outlier because the fit includes several additional parameters to account for peaking background present in this bin; the additional parameters that the fit must accommodate cause the precision on the asymmetry to decrease compared to bins 1 and 3. The fit quality is good, with the χ^2/n.d.f. falling between 0.86 and 1.48 for all fits, with an average value of 1.10. The distributions of the residuals (shown in Figures 6.1 – 6.6) show no obvious structure.

A	D^0		$\overline{D}^0$		$\mathcal{A}_\mathrm{A}^i$ (%)
i	χ^2/n.d.f.	n_A^i	χ^2/n.d.f.	$\bar{n}_A^i$	
0	0.86	2809 ± 97	0.92	2803 ± 102	0.1 ± 2.5
1	1.06	12956 ± 164	1.01	12639 ± 162	1.24 ± 0.90
2	1.16	531 ± 66	1.15	611 ± 52	-7.0 ± 7.5
3	1.26	17976 ± 179	1.22	17519 ± 177	1.29 ± 0.71
4	1.19	4360 ± 126	1.06	4769 ± 141	-4.5 ± 2.1
5	1.00	25280 ± 346	1.18	24483 ± 401	1.6 ± 1.1

Table 6.3: Fit results and raw asymmetries from fits to D^0 and $\overline{D}^0$ ΔM plots in **region A** for all six bins. Errors are statistical only.

6.3 Asymmetries

Using equation 5.3 and equations 5.10 – 5.15, we calculate the local CP asymmetry in each bin i from the raw asymmetries in Table 6.3 and Table 6.4. The results are shown in Table 6.5.

B	D^0		$\overline{D}^0$		$\mathcal{A}_{\mathrm{B}}^i$ (%)
i	χ^2/n.d.f.	n_B^i	χ^2/n.d.f.	$\bar{n}_B^i$	
0	1.12	4567 ± 125	1.02	4328 ± 115	2.7 ± 1.9
1	1.48	31178 ± 239	1.26	30369 ± 229	1.31 ± 0.54
2	1.13	2937 ± 89	0.95	2866 ± 86	1.2 ± 2.1
3	1.12	32007 ± 235	1.15	31482 ± 234	0.83 ± 0.52
4	1.05	9981 ± 160	1.05	9861 ± 157	0.6 ± 1.1
5	1.07	20668 ± 348	1.01	19726 ± 49	2.3 ± 1.2

Table 6.4: Fit results and raw asymmetries from fits to D^0 and $\overline{D}^0$ ΔM plots in **region B** for all six bins. Errors are statistical only.

i	$\mathcal{A}_{\mathrm{A}}^i$ (%)	$\mathcal{A}_{\mathrm{B}}^i$ (%)	$\mathcal{A}_{\mathrm{bin}}^{*i}$ (%)	$\mathbf{\mathcal{A}_{\mathbf{CP,loc}}^i}$ (%)
0	0.1 ± 2.5	2.7 ± 1.9	1.4 ± 1.6	$\mathbf{0.4 \pm 1.5}$
1	1.24 ± 0.90	1.31 ± 0.54	1.28 ± 0.52	$\mathbf{0.24 \pm 0.39}$
2	-7.0 ± 7.5	1.2 ± 2.1	-2.9 ± 3.9	$\mathbf{-3.9 \pm 3.8}$
3	1.29 ± 0.71	0.83 ± 0.52	1.06 ± 0.44	$\mathbf{0.02 \pm 0.31}$
4	-4.5 ± 2.1	0.6 ± 1.1	-1.9 ± 1.2	$\mathbf{-3.0 \pm 1.1}$
5	1.6 ± 1.1	2.3 ± 1.2	1.97 ± 0.81	$\mathbf{0.93 \pm 0.59}$

Table 6.5: Final values for $\mathcal{A}_{CP,\mathrm{loc}}^i$ in each Dalitz bin. Errors are statistical only. The average asymmetry across the plot is $\mathcal{A}_{\mathrm{avg}} = (1.03 \pm 0.33)\%$. $\mathcal{A}_{CP,\mathrm{loc}}^i$ is calculated by subtracting $\mathcal{A}_{\mathrm{avg}}$ from $\mathcal{A}_{\mathrm{bin}}^{*i}$ for each bin.

As with $\mathcal{A}_{\mathrm{A}}^i$ and $\mathcal{A}_{\mathrm{B}}^i$, $\mathcal{A}_{CP,\mathrm{loc}}^i$ is much more precise in regions of higher statistics. This is to be expected, given the use of a weighted average to determine $\mathcal{A}_{\mathrm{avg}}$ (see Equation 5.2 for definition). Weighting $\mathcal{A}_{\mathrm{bin}}^{*i}$ by the number of events in bin i when calculating $\mathcal{A}_{\mathrm{avg}}$ means we preserve the high precision measurements in high-statistics bins; by contrast, using an unweighted average would dilute these more precise measurements by effectively assigning disproportionate weight to the low-statistics bins.

We observe $\mathcal{A}_{CP,\mathrm{loc}}^i$ to be consistent with zero in all bins, which is what we expect for MC where no real CPV is present. Note we have not included systematic uncertainties on these measurements. We can test the level of agreement by calculating a χ^2/n.d.f., using:

$$\chi^2/\mathrm{n.d.f.} = \frac{1}{6} \sum \left(\frac{\mathcal{A}_{CP,\mathrm{loc}}^i}{\sigma_{CP,\mathrm{loc}}^i} \right)^2 \tag{6.1}$$

where 6 is the number of bins (and hence the number of values of $\mathcal{A}_{CP,\mathrm{loc}}^i$). For the results in Table 6.5, we find χ^2/n.d.f. $= 1.91$ (we expect χ^2/n.d.f. $= 1$ for a "good" result). For only 6 measurements, this is a reasonable χ^2/n.d.f., indicating good agreement with $\mathcal{A}_{CP,\mathrm{loc}}^i = 0$ across the Dalitz plot.
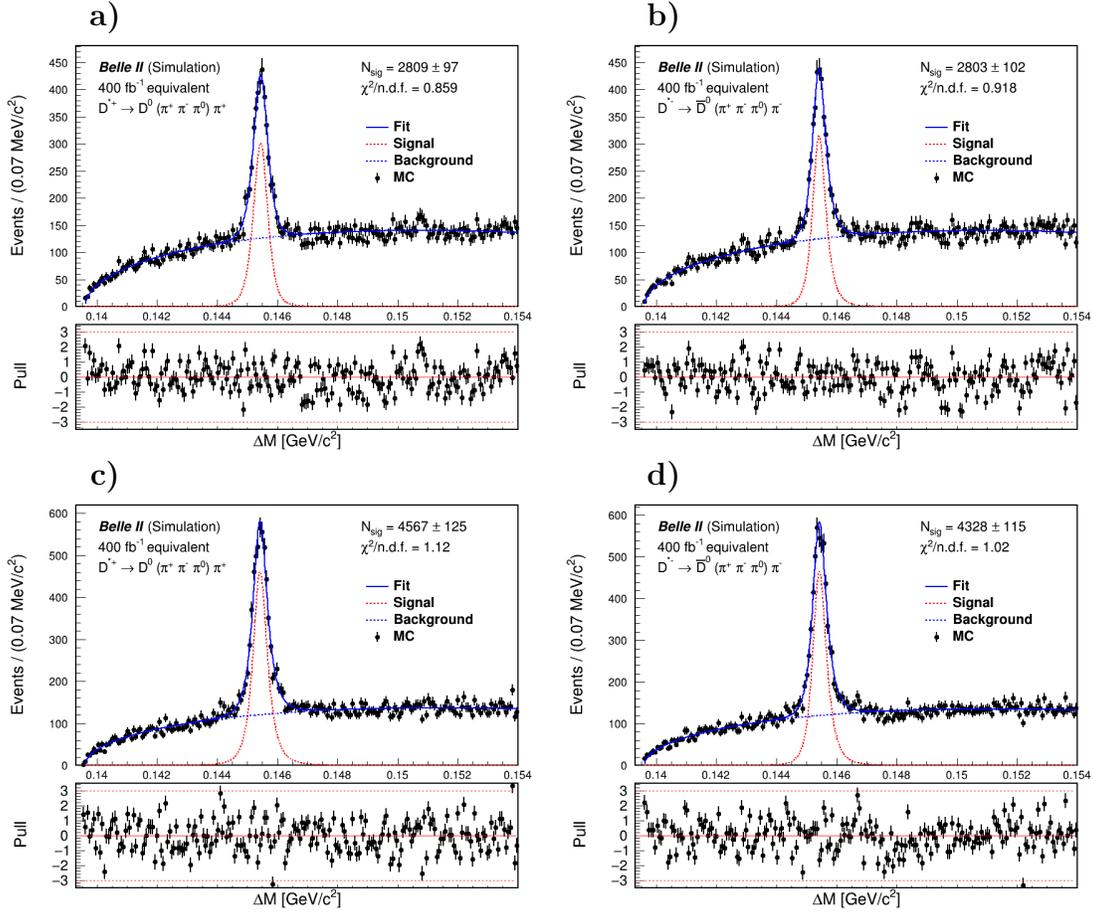
Figure 6.1: Fits in MC for **a)** D^0 in bin 0, region A, **b)** $\overline{D}^0$ in bin 0, region A, **c)** D^0 in bin 0, region B, and **d)** $\overline{D}^0$ in bin 0, region B.

Param.	A		B	
	D^0	$\overline{D}^0$	D^0	$\overline{D}^0$
I^0_{sig}	0.1966 ± 0.0068	0.1962 ± 0.0071	0.3197 ± 0.0087	0.3029 ± 0.0080
μ^0 (MeV/c^2)	145.455 ± 0.041	145.379 ± 0.032	145.401 ± 0.023	145.383 ± 0.024
σ^0 (MeV/c^2)	0.407 ± 0.088	0.324 ± 0.056	0.321 ± 0.035	0.321 ± 0.037
γ^0	0.08 ± 0.17	-0.19 ± 0.14	-0.109 ± 0.089	-0.23 ± 0.10
δ^0	1.57 ± 0.33	1.31 ± 0.23	1.17 ± 0.13	1.25 ± 0.14

Table 6.6: Signal parameters in MC for all fits in bin 0. I^i_{sig} is the normalization parameter from which the yields are calculated (see Equation 5.6).

54
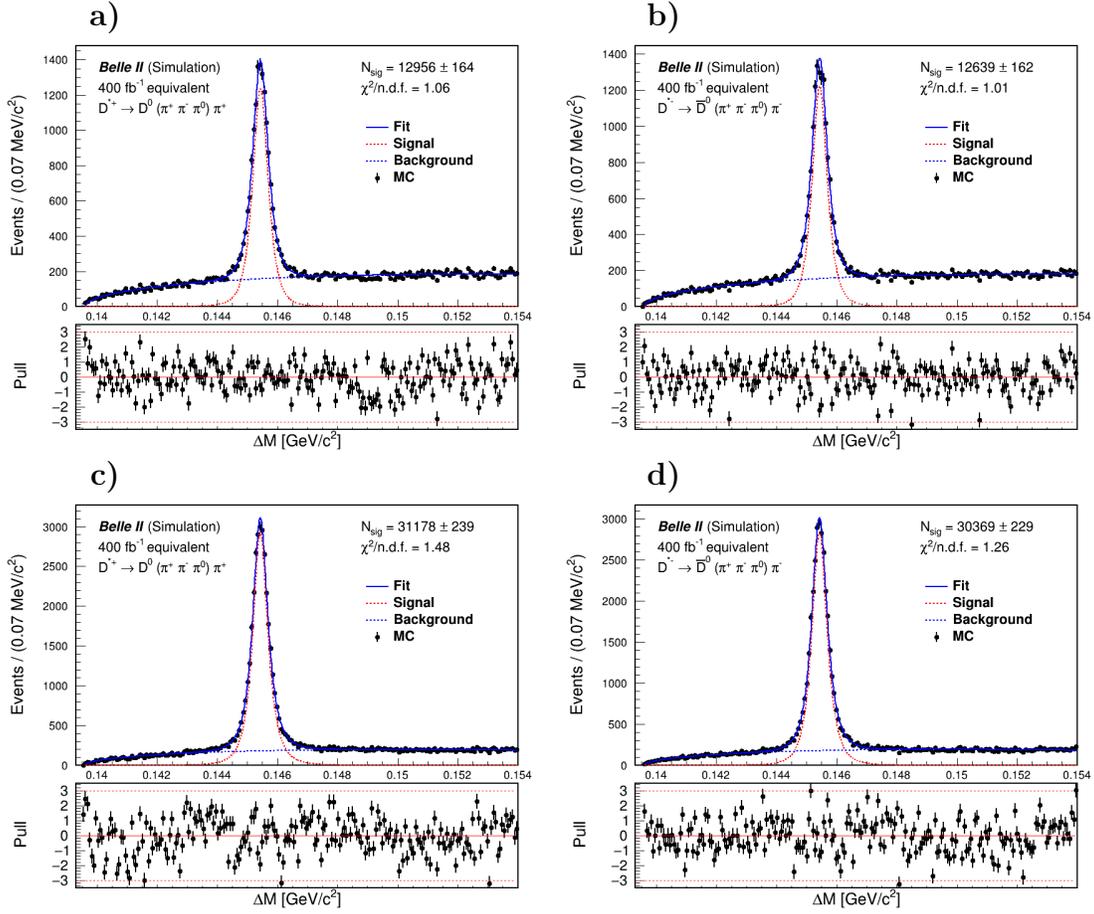
Figure 6.2: Fits in MC for **a)** D^0 in bin 1, region A, **b)** $\overline{D}^0$ in bin 1, region A, **c)** D^0 in bin 1, region B, and **d)** $\overline{D}^0$ in bin 1, region B.

	A		B	
Param.	D^0	$\overline{D}^0$	D^0	$\overline{D}^0$
I^1_{sig}	0.907 ± 0.011	0.885 ± 0.011	2.182 ± 0.017	2.126 ± 0.016
μ^1 (MeV/c^2)	145.435 ± 0.012	145.427 ± 0.012	145.417 ± 0.006	145.415 ± 0.006
σ^1 (MeV/c^2)	0.374 ± 0.022	0.361 ± 0.021	0.3204 ± 0.0099	0.340 ± 0.011
γ^1	0.009 ± 0.044	-0.026 ± 0.044	-0.057 ± 0.021	-0.056 ± 0.022
δ^1	1.279 ± 0.073	1.250 ± 0.069	1.081 ± 0.031	1.139 ± 0.033

Table 6.7: Signal parameters in MC for all fits in bin 1. I^i_{sig} is the normalization parameter from which the yields are calculated (see Equation 5.6).
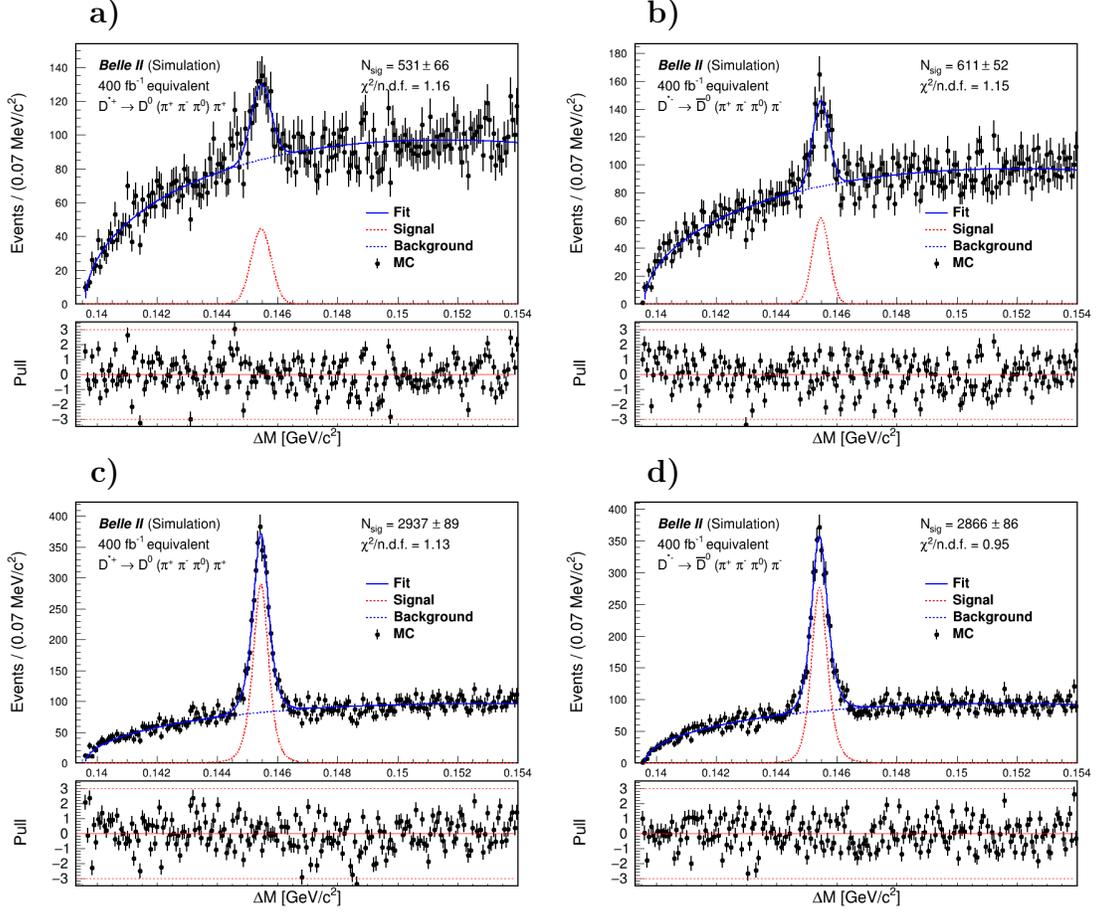
Figure 6.3: Fits in MC for **a)** D^0 in bin 2, region A, **b)** $\overline{D}^0$ in bin 2, region A, **c)** D^0 in bin 2, region B, and **d)** $\overline{D}^0$ in bin 2, region B.

Param.	A		B	
	D^0	$\overline{D}^0$	D^0	$\overline{D}^0$
I_{sig}^2	0.0372 ± 0.0046	0.0427 ± 0.0036	0.2056 ± 0.0062	0.2006 ± 0.0060
μ^2 (MeV/c^2)	145.458 ± 0.039	145.469 ± 0.026	145.462 ± 0.036	145.385 ± 0.043
σ^2 (MeV/c^2)	0.333 ± 0.055	0.277 ± 0.025	0.421 ± 0.072	0.457 ± 0.080
γ^2	–	–	0.06 ± 0.13	-0.17 ± 0.16
δ^2	–	–	1.49 ± 0.25	1.59 ± 0.27

Table 6.8: Signal parameters in MC for all fits in bin 2. μ and σ are shape parameters for both Gaussian and Johnson's S_U signal functions. δ and γ are parameters only for Johnson's S_U-distributions. I_{sig}^i is the normalization parameter from which the yields are calculated (see Equation 5.6).
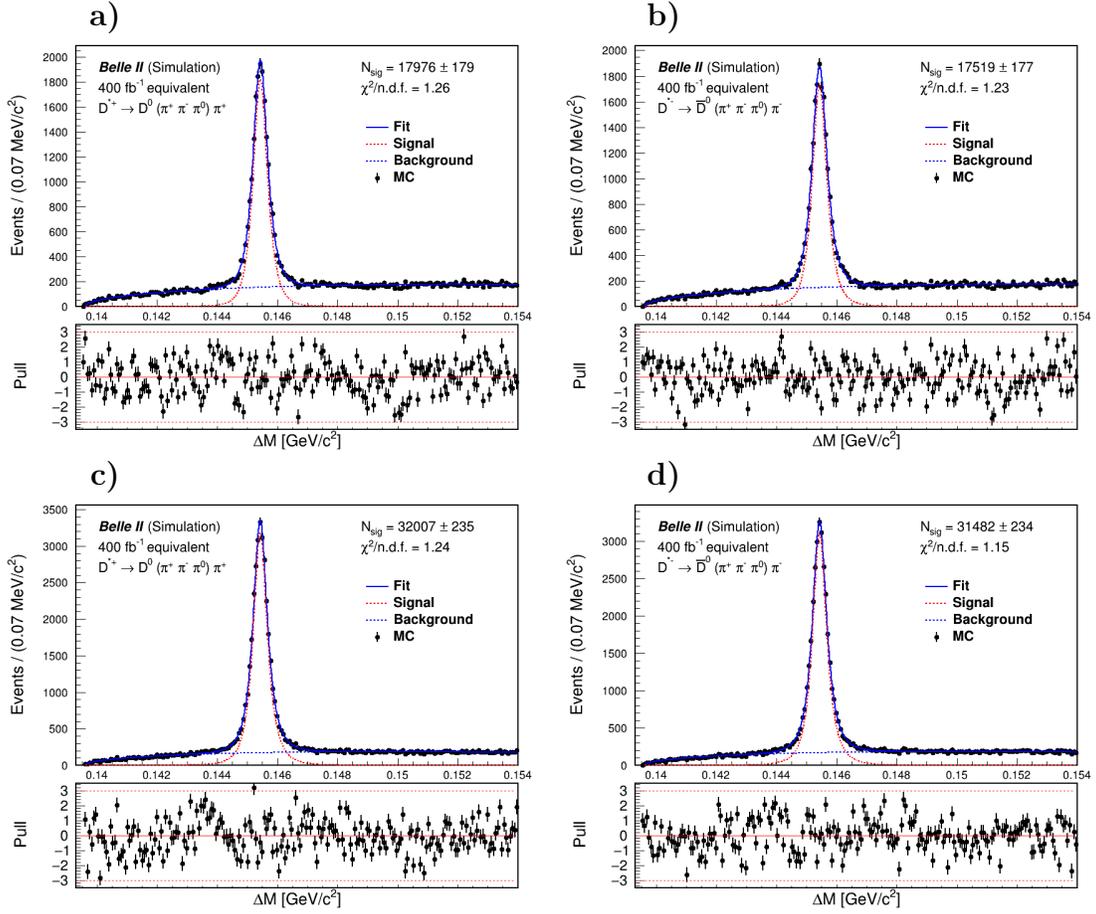
Figure 6.4: Fits in MC for **a)** D^0 in bin 3, region A, **b)** $\overline{D}^0$ in bin 3, region A, **c)** D^0 in bin 3, region B, and **d)** $\overline{D}^0$ in bin 3, region B.

	A		B	
Param.	D^0	$\overline{D}^0$	D^0	$\overline{D}^0$
I^3_{sig}	1.258 ± 0.013	1.226 ± 0.012	2.241 ± 0.016	2.204 ± 0.016
μ^3 (MeV/c^2)	145.418 ± 0.009	145.413 ± 0.009	145.417 ± 0.005	145.417 ± 0.005
σ^3 (MeV/c^2)	0.339 ± 0.016	0.339 ± 0.015	0.2976 ± 0.0087	0.2975 ± 0.0088
γ^3	-0.058 ± 0.033	-0.080 ± 0.033	-0.052 ± 0.019	-0.062 ± 0.019
δ^3	1.234 ± 0.053	1.201 ± 0.049	1.062 ± 0.028	1.051 ± 0.028

Table 6.9: Signal parameters in MC for all fits in bin 3. I^i_{sig} is the normalization parameter from which the yields are calculated (see Equation 5.6).
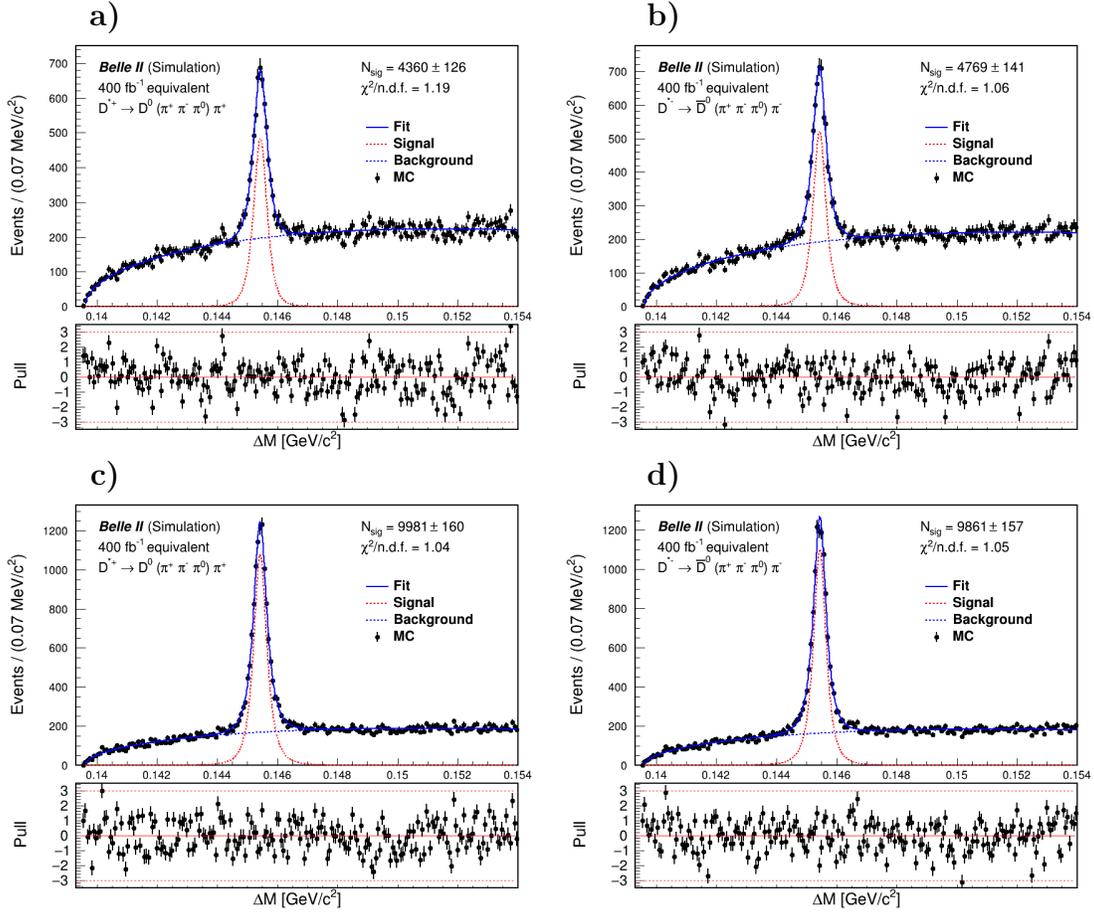
Figure 6.5: Fits in MC for **a)** D^0 in bin 4, region A, **b)** $\overline{D}^0$ in bin 4, region A, **c)** D^0 in bin 4, region B, and **d)** $\overline{D}^0$ in bin 4, region B.

Param.	A		B	
	D^0	$\overline{D}^0$	D^0	$\overline{D}^0$
I^4_{sig}	0.3052 ± 0.0088	0.3338 ± 0.0099	0.699 ± 0.011	0.690 ± 0.011
μ^4 (MeV/c^2)	145.444 ± 0.026	145.431 ± 0.020	145.414 ± 0.011	145.443 ± 0.010
σ^4 (MeV/c^2)	0.339 ± 0.050	0.291 ± 0.035	0.283 ± 0.019	0.268 ± 0.017
γ^4	0.08 ± 0.11	0.018 ± 0.088	-0.062 ± 0.045	0.040 ± 0.045
δ^4	1.35 ± 0.20	1.14 ± 0.14	1.094 ± 0.069	1.073 ± 0.064

Table 6.10: Signal parameters in MC for all fits in bin 4. I^i_{sig} is the normalization parameter from which the yields are calculated (see Equation 5.6).
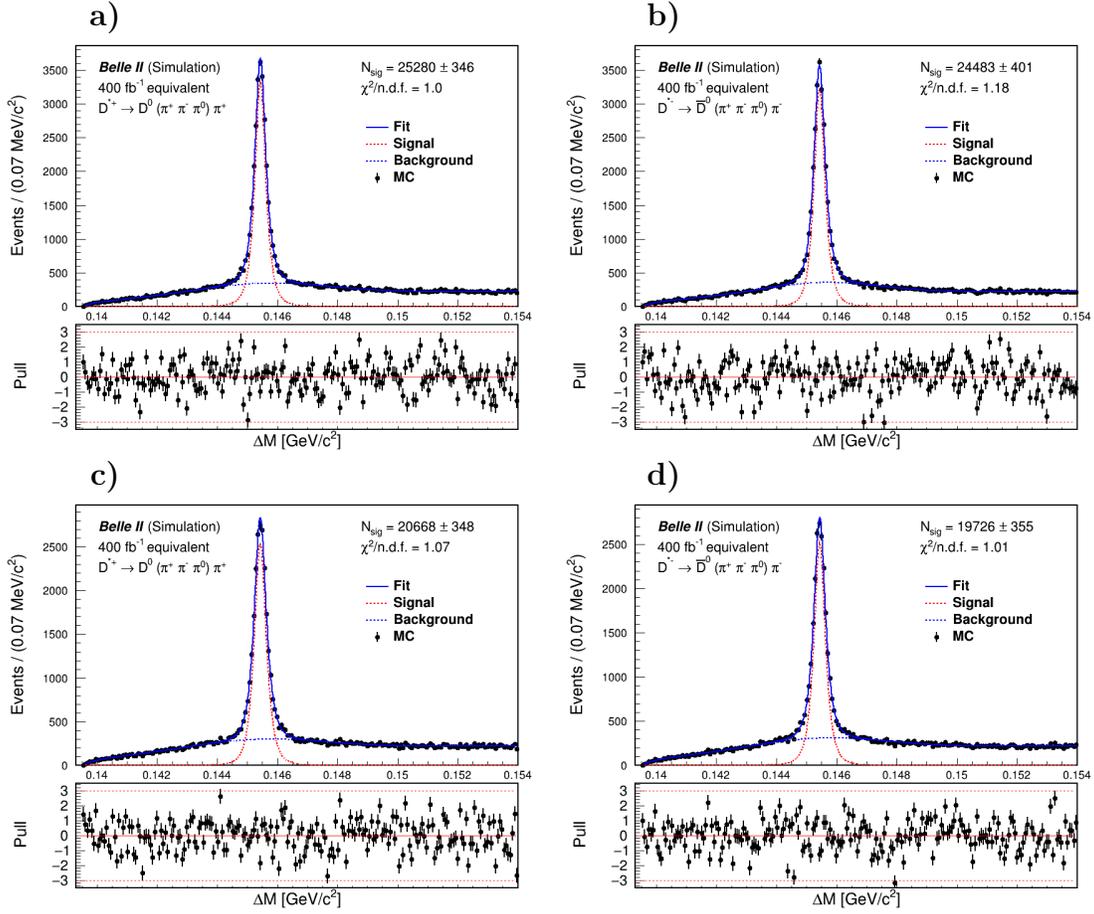
Figure 6.6: Fits in MC for **a)** D^0 in bin 5, region A, **b)** $\overline{D}^0$ in bin 5, region A, **c)** D^0 in bin 5, region B, and **d)** $\overline{D}^0$ in bin 5, region B.

	A		B	
Param.	D^0	$\overline{D}^0$	D^0	$\overline{D}^0$
I^5_{sig}	1.770 ± 0.024	1.714 ± 0.028	1.447 ± 0.024	1.381 ± 0.025
μ^5 (MeV/c^2)	145.426 ± 0.006	145.416 ± 0.006	145.418 ± 0.008	145.431 ± 0.007
σ^5 (MeV/c^2)	0.244 ± 0.011	0.253 ± 0.013	0.284 ± 0.016	0.261 ± 0.015
γ^5	-0.017 ± 0.027	-0.073 ± 0.030	-0.054 ± 0.035	-0.006 ± 0.033
δ^5	1.152 ± 0.055	1.190 ± 0.067	1.251 ± 0.076	1.182 ± 0.074

Table 6.11: Signal parameters in MC for all fits in bin 5. I^i_{sig} is the normalization parameter from which the yields are calculated (see Equation 5.6).

Chapter 7

Results

This chapter presents the results of performing the analysis described in chapter 5 on the *Belle II* dataset described in chapter 4.

7.1   Fit functions

In each region (A or B) of each bin (1–6) (see Figure 5.2 for visualization), two fits were performed (one for D^0 and one for $\overline{D}^0$) for a total of 24 fits. Table 7.1 (Table 7.2) summarizes the fit functions used in region A (B) to determine signal yield. In most cases, a Johnson's S_U-distribution was used as the signal function, and a 3/2-power threshold function was used as the background. The exceptions to this were 1) in region A of bins 0, 2, and 4, a single Gaussian was used as the signal function because the statistics were too low to accommodate a function with more parameters (a 3/2-power threshold function was still used as the background), and 2) in regions A and B of bin 5, a single Gaussian was added to the 3/2-power threshold function to accommodate the peaking background in that region (a Johnson's S_U-distribution was still used as the signal). In each region, the same function was used for D^0 and $\overline{D}^0$ fits.

i	Signal function	Background function
0	Single Gaussian	
1	Johnson's S_U-distribution	
2	Single Gaussian	3/2-power threshold function
3	Johnson's S_U-distribution	
4	Single Gaussian	
5	Johnson's S_U-distribution	Single Gaussian and 3/2-power threshold function

Table 7.1: Fit functions for fits to D^0 and $\overline{D}^0$ ΔM plots in data in **region A** for all six bins.

7.2   Fit results

The fits to D^0 and $\overline{D}^0$ in each of the twelve Dalitz regions are shown in Figures 7.1 – 7.6. The raw yields and fit quality (i.e., χ^2/n.d.f.) are listed in Table 7.3 (Table 7.4) for region A

i	Signal function	Background function
0		
1		
2	Johnson's S_U-distribution	3/2-power threshold function
3		
4		
5		Single Gaussian and 3/2-power threshold function

Table 7.2: Fit functions for fits to D^0 and $\overline{D}^0$ ΔM plots in data in **region B** for all six bins.

(B). The number of bins is always 210, and the number of fit parameters varies between 5 and 10 depending on the fit function being used. The full list of signal parameters for each fit are listed in Tables 7.6 – 7.11.

We observe much greater precision in the raw asymmetries for bins 1 and 3 as compared to bins 0, 2, and 4. This is due to the higher statistics in the odd-numbered bins. The precision in bin 5 is an outlier because the fit includes several additional parameters to account for peaking background present in this bin; the additional parameters that the fit must accommodate cause the precision on the asymmetry to decrease slightly compared to bins 1 and 3. The fit quality is not unreasonable, with the χ^2/n.d.f. falling between 0.96 and 1.40 for all fits, with an average value of 1.14. The distributions of the residuals (shown in Figures 7.1 – 7.6) show no obvious structure.

A	D^0		$\overline{D}^0$		$\mathcal{A}_{\mathrm{A}}^i$ (%)
i	χ^2/n.d.f.	n_A^i	χ^2/n.d.f.	$\bar{n}_A^i$	
0	1.26	561 ± 57	1.11	394 ± 57	17.4 ± 8.6
1	1.18	12074 ± 163	1.13	11565 ± 156	2.15 ± 0.93
2	1.05	485 ± 56	1.18	376 ± 43	12.7 ± 8.0
3	1.14	15888 ± 171	1.20	15811 ± 179	0.24 ± 0.78
4	1.15	1124 ± 75	1.08	1151 ± 76	-1.2 ± 4.7
5	0.90	21177 ± 367	1.05	20363 ± 344	2.0 ± 1.2

Table 7.3: Fit results and raw asymmetries from fits to D^0 and $\overline{D}^0$ ΔM plots in **region A** for all six bins. Errors are statistical only.

7.3 Asymmetries

Using equation 5.3 and equations 5.10 – 5.15, we calculate the local CP asymmetry in each bin i from the raw asymmetries in Table 7.3 and Table 7.4. The results are shown in Table 7.5.

As with $\mathcal{A}_{\mathrm{A}}^i$ and $\mathcal{A}_{\mathrm{B}}^i$, $\mathcal{A}_{CP,\mathrm{loc}}^i$ is much more precise in regions of higher statistics. This is to be expected, given the use of a weighted average to determine $\mathcal{A}_{\mathrm{avg}}$ (see Equation 5.2 for definition). Weighting $\mathcal{A}_{\mathrm{bin}}^{*i}$ by the number of events in bin i when calculating $\mathcal{A}_{\mathrm{avg}}$ means we preserve the high precision measurements in high-statistics bins; by contrast, using an

B	D^0		$\overline{D}^0$		$\mathcal{A}_B^i$ (%)
i	χ^2/n.d.f.	n_B^i	χ^2/n.d.f.	$\bar{n}_B^i$	
0	0.99	985 ± 83	0.96	923 ± 127	3.2 ± 8.1
1	1.10	24443 ± 218	1.40	24066 ± 226	0.78 ± 0.65
2	1.30	1676 ± 79	1.05	1756 ± 88	-2.3 ± 3.4
3	1.20	23228 ± 208	1.35	23107 ± 221	0.26 ± 0.66
4	1.18	3066 ± 154	1.16	2580 ± 120	8.6 ± 3.3
5	1.15	16632 ± 276	1.09	15719 ± 332	2.8 ± 1.3

Table 7.4: Fit results and raw asymmetries from fits to D^0 and $\overline{D}^0$ ΔM plots in **region B** for all six bins. Errors are statistical only.

i	$\mathcal{A}_A^i$ (%)	$\mathcal{A}_B^i$ (%)	$\mathcal{A}_{\text{bin}}^{*i}$ (%)	$\mathcal{A}_{\textbf{CP,loc}}^{\textbf{i}}$ (%)
0	17.4 ± 8.6	3.2 ± 8.1	10.4 ± 5.9	$\mathbf{8.8 \pm 5.8}$
1	2.15 ± 0.93	0.78 ± 0.65	1.47 ± 0.58	$\mathbf{-0.14 \pm 0.40}$
2	12.7 ± 8.0	-2.3 ± 3.4	5.2 ± 4.4	$\mathbf{3.6 \pm 4.3}$
3	0.24 ± 0.78	0.26 ± 0.66	0.25 ± 0.51	$\mathbf{-1.35 \pm 0.34}$
4	-1.2 ± 4.7	8.6 ± 3.3	3.7 ± 2.9	$\mathbf{2.1 \pm 2.8}$
5	2.0 ± 1.2	2.8 ± 1.3	2.39 ± 0.90	$\mathbf{0.78 \pm 0.62}$

Table 7.5: Final values for $\mathcal{A}_{CP,\text{loc}}^i$ in each Dalitz bin. Errors are statistical only. The average asymmetry across the plot is $\mathcal{A}_{\text{avg}} = (1.60 \pm 0.40)\%$. $\mathcal{A}_{CP,\text{loc}}^i$ is calculated by subtracting $\mathcal{A}_{\text{avg}}$ from $\mathcal{A}_{\text{bin}}^{*i}$ for each bin.

unweighted average would dilute these more precise measurements by effectively assigning disproportionate weight to the low-statistics bins.

The results appear to be consistent with zero in all bins except bin 3, where the central value is almost 4σ from zero. However, one should also carefully consider possible systematic uncertainties, which we discuss in chapter 8, before assessing significance.
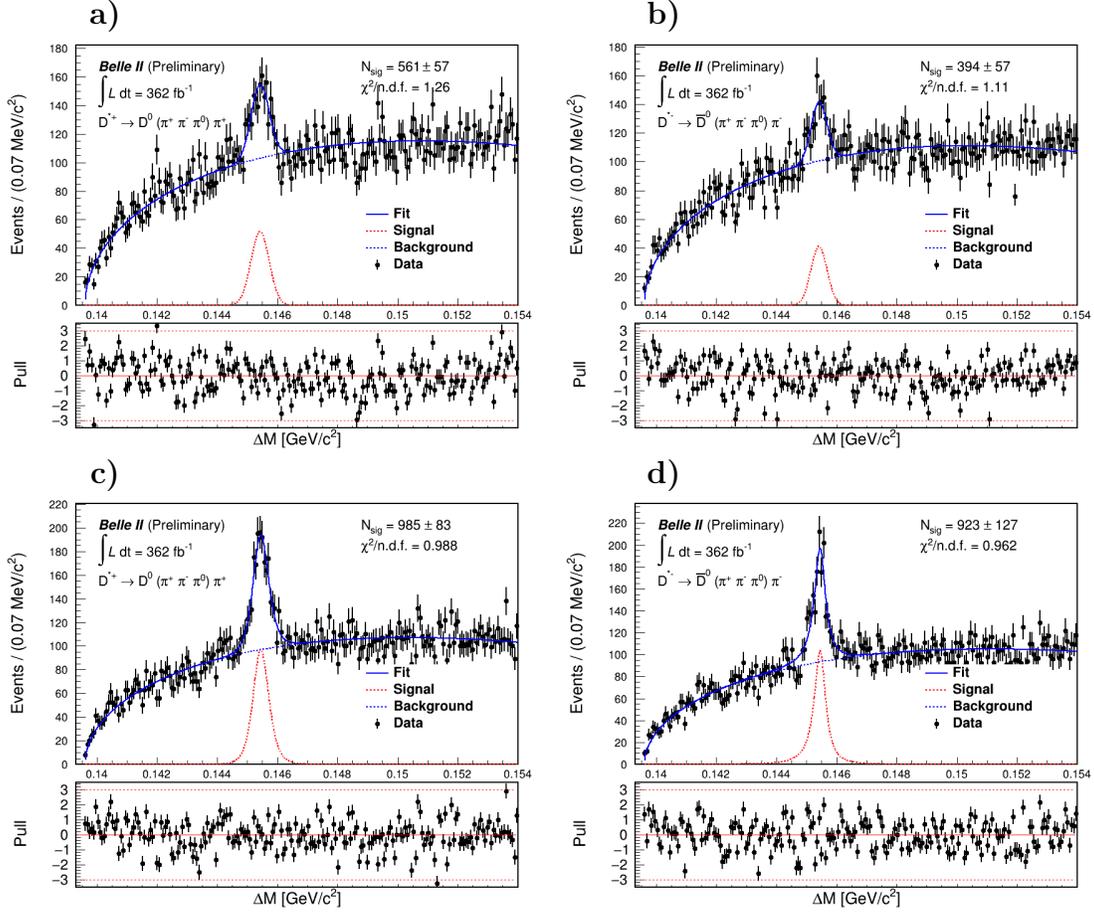
Figure 7.1: Fits for **a)** D^0 in bin 0, region A, **b)** $\overline{D}^0$ in bin 0, region A, **c)** D^0 in bin 0, region B, and **d)** $\overline{D}^0$ in bin 0, region B.

	A		B	
Param.	D^0	$\overline{D}^0$	D^0	$\overline{D}^0$
I^0_{sig}	0.0393 ± 0.0040	0.0276 ± 0.0040	0.0690 ± 0.0058	0.0646 ± 0.0089
μ^0 (MeV/c^2)	145.421 ± 0.035	145.408 ± 0.040	145.41 ± 0.13	145.460 ± 0.042
σ^0 (MeV/c^2)	0.303 ± 0.031	0.268 ± 0.048	0.49 ± 0.29	0.207 ± 0.080
γ^0	–	–	-0.14 ± 0.49	0.13 ± 0.22
δ^0	–	–	1.7 ± 1.0	0.84 ± 0.36

Table 7.6: Signal parameters for all fits in bin 0. μ and σ are shape parameters for both Gaussian and Johnson's S_U signal functions. δ and γ are parameters only for Johnson's S_U-distributions. I^i_{sig} is the normalization parameter from which the yields are calculated (see Equation 5.6).
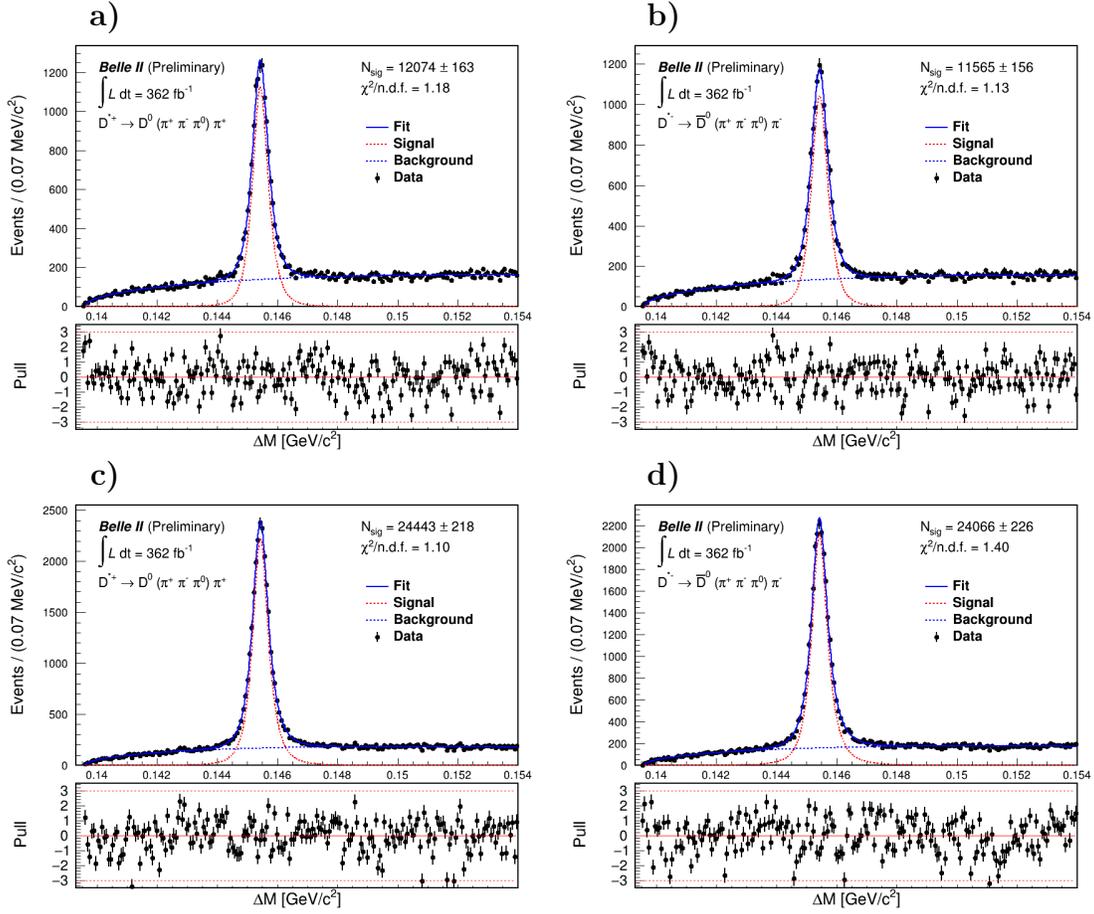
Figure 7.2: Fits for **a)** D^0 in bin 1, region A, **b)** $\overline{D}^0$ in bin 1, region A, **c)** D^0 in bin 1, region B, and **d)** $\overline{D}^0$ in bin 1, region B.

Param.	A		B	
	D^0	$\overline{D}^0$	D^0	$\overline{D}^0$
I^1_{sig}	0.845 ± 0.011	0.810 ± 0.011	1.711 ± 0.015	1.685 ± 0.016
μ^1 (MeV/c^2)	145.410 ± 0.013	145.411 ± 0.015	145.416 ± 0.007	145.408 ± 0.007
σ^1 (MeV/c^2)	0.367 ± 0.023	0.403 ± 0.026	0.447 ± 0.012	0.326 ± 0.012
γ^1	-0.092 ± 0.044	-0.086 ± 0.049	-0.087 ± 0.025	-0.097 ± 0.024
δ^1	1.227 ± 0.073	1.304 ± 0.082	1.101 ± 0.038	1.030 ± 0.035

Table 7.7: Signal parameters for all fits in bin 1. I^i_{sig} is the normalization parameter from which the yields are calculated (see Equation 5.6).
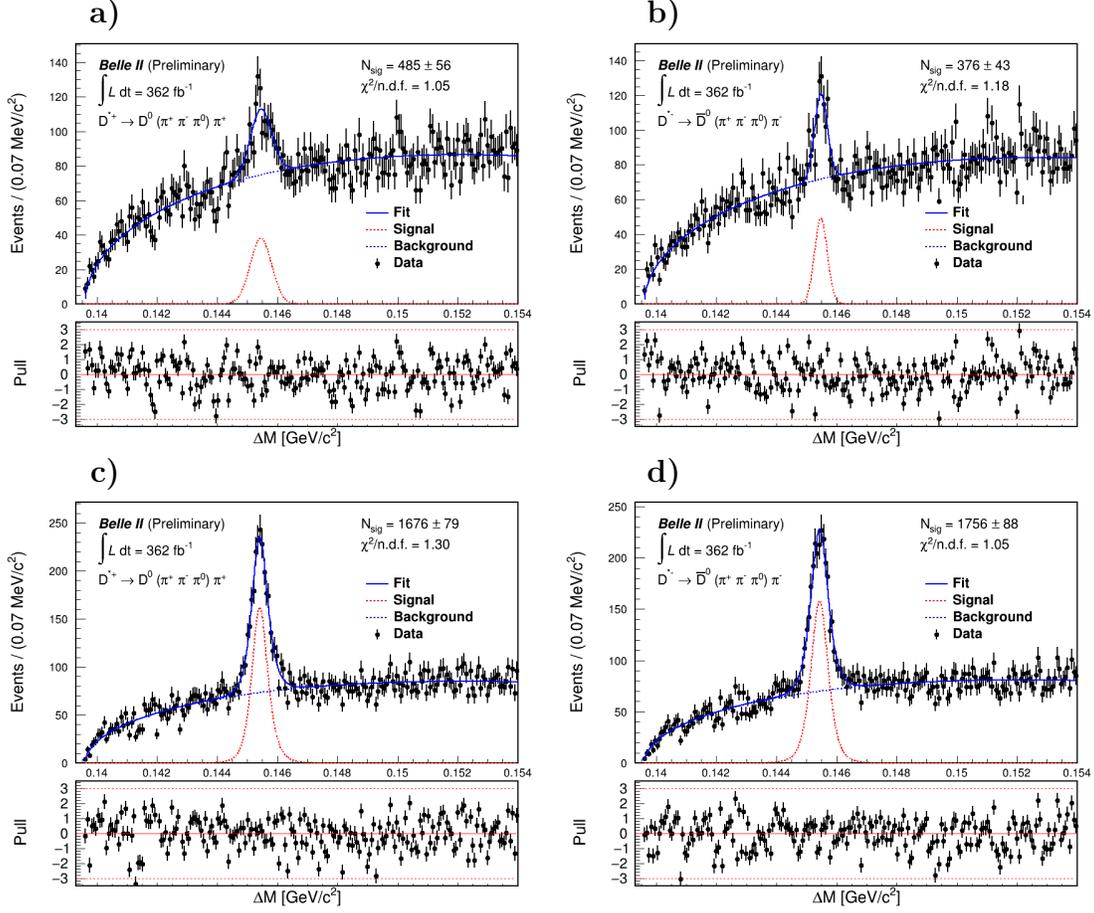
Figure 7.3: Fits for **a)** D^0 in bin 2, region A, **b)** $\overline{D}^0$ in bin 2, region A, **c)** D^0 in bin 2, region B, and **d)** $\overline{D}^0$ in bin 2, region B.

Param.	A		B	
	D^0	$\overline{D}^0$	D^0	$\overline{D}^0$
I^2_{sig}	0.0340 ± 0.0039	0.0264 ± 0.0030	0.1173 ± 0.0055	0.1229 ± 0.0061
μ^2 (MeV/c^2)	145.449 ± 0.043	145.475 ± 0.028	145.379 ± 0.062	145.425 ± 0.078
σ^2 (MeV/c^2)	0.355 ± 0.045	0.214 ± 0.029	0.44 ± 0.13	0.51 ± 0.16
γ^2	−	−	-0.14 ± 0.23	-0.01 ± 0.27
δ^2	−	−	1.54 ± 0.45	1.64 ± 0.52

Table 7.8: Signal parameters for all fits in bin 2. μ and σ are shape parameters for both Gaussian and Johnson's S_U signal functions. δ and γ are parameters only for Johnson's S_U-distributions. I^i_{sig} is the normalization parameter from which the yields are calculated (see Equation 5.6).
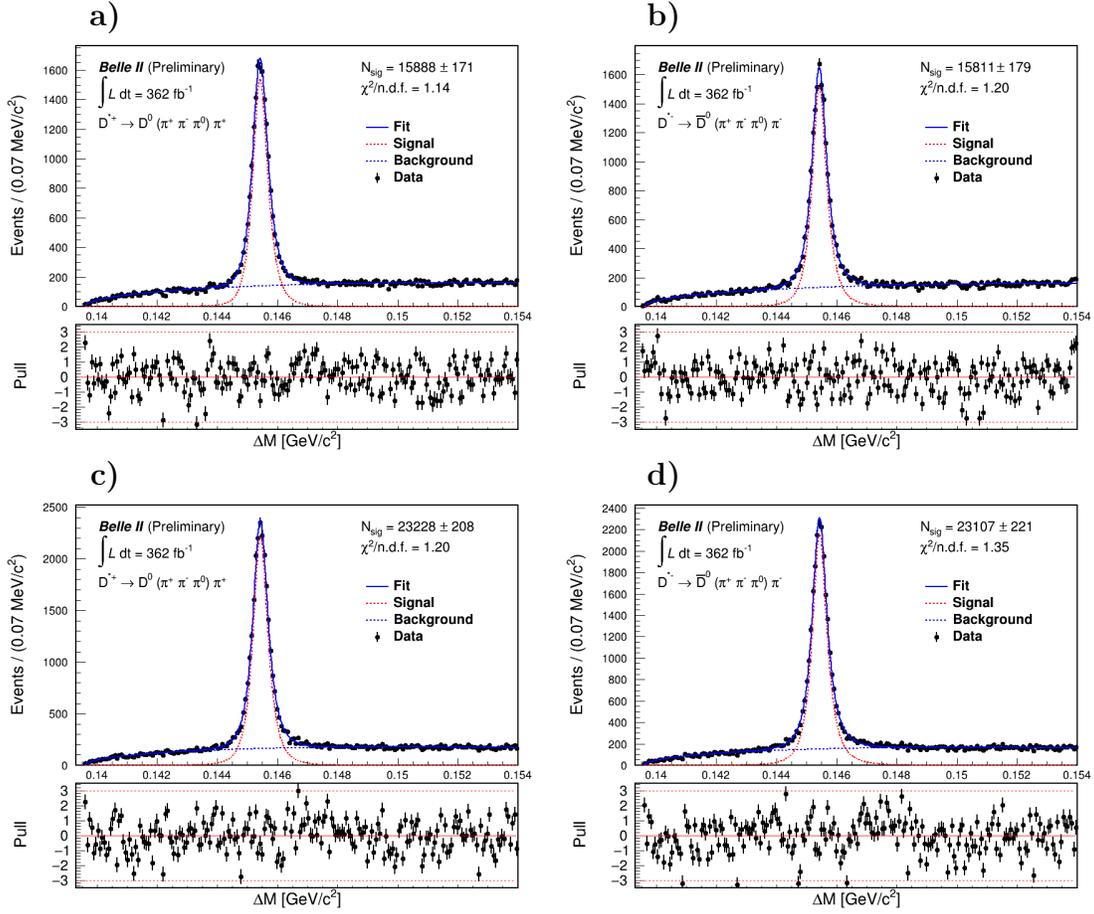
Figure 7.4: Fits for **a)** D^0 in bin 3, region A, **b)** $\overline{D}^0$ in bin 3, region A, **c)** D^0 in bin 3, region B, and **d)** $\overline{D}^0$ in bin 3, region B.

	A		B	
Param.	D^0	$\overline{D}^0$	D^0	$\overline{D}^0$
I^3_{sig}	1.112 ± 0.012	1.107 ± 0.013	1.626 ± 0.015	1.617 ± 0.015
μ^3 (MeV/c^2)	145.394 ± 0.010	145.401 ± 0.009	145.417 ± 0.007	145.413 ± 0.007
σ^3 (MeV/c^2)	0.355 ± 0.017	0.311 ± 0.014	0.326 ± 0.012	0.298 ± 0.011
γ^3	-0.149 ± 0.036	-0.103 ± 0.031	-0.078 ± 0.025	-0.093 ± 0.023
δ^3	1.237 ± 0.057	1.074 ± 0.045	1.116 ± 0.039	0.999 ± 0.034

Table 7.9: Signal parameters for all fits in bin 3. I^i_{sig} is the normalization parameter from which the yields are calculated (see Equation 5.6).
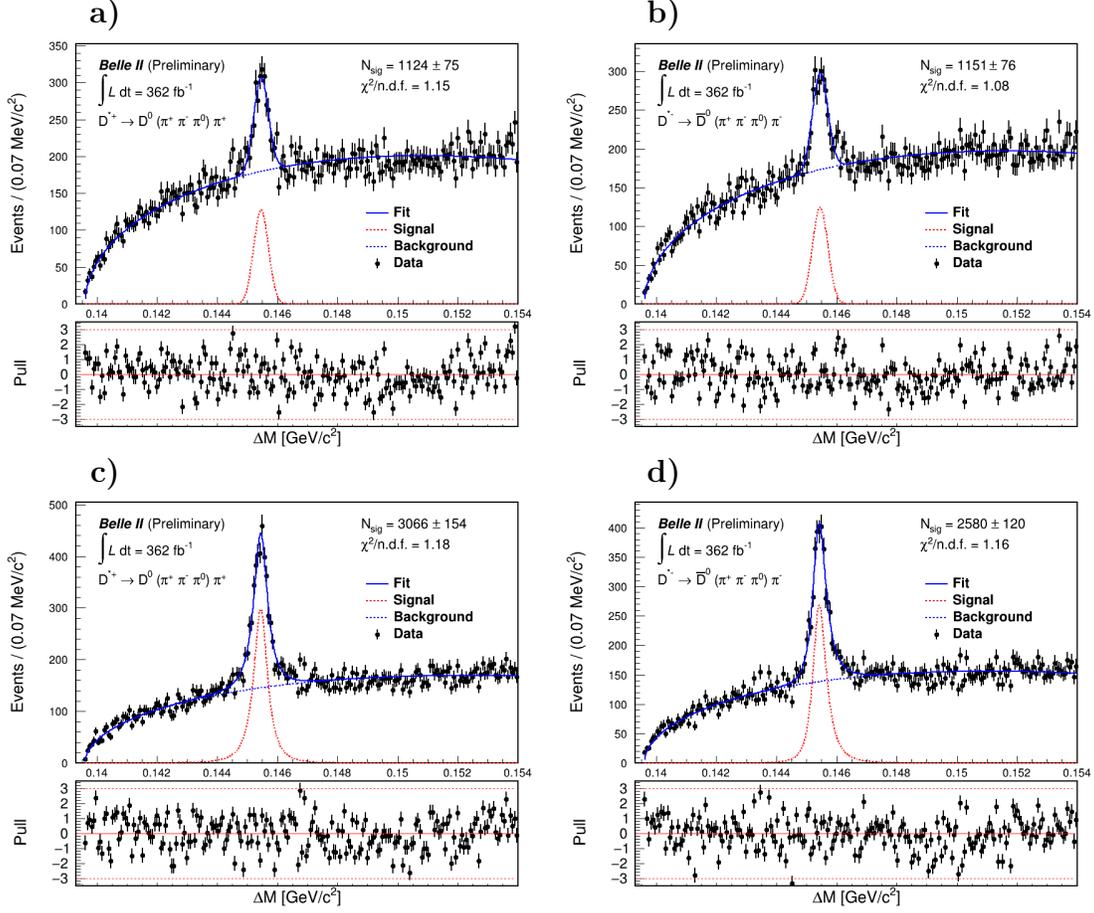
Figure 7.5: Fits for **a)** D^0 in bin 4, region A, **b)** $\overline{D}^0$ in bin 4, region A, **c)** D^0 in bin 4, region B, and **d)** $\overline{D}^0$ in bin 4, region B.

	A		B	
Param.	D^0	$\overline{D}^0$	D^0	$\overline{D}^0$
I_{sig}^4	0.0787 ± 0.0052	0.0805 ± 0.0053	0.215 ± 0.011	0.1806 ± 0.0084
μ^4 (MeV/c^2)	145.455 ± 0.017	145.448 ± 0.018	145.451 ± 0.0026	145.380 ± 0.030
σ^4 (MeV/c^2)	0.247 ± 0.020	0.258 ± 0.021	0.268 ± 0.038	0.273 ± 0.045
γ^4	–	–	0.06 ± 0.11	-0.27 ± 0.13
δ^4	–	–	0.93 ± 0.15	1.03 ± 0.17

Table 7.10: Signal parameters for all fits in bin 4. μ and σ are shape parameters for both Gaussian and Johnson's S_U signal functions. δ and γ are parameters only for Johnson's S_U-distributions. I_{sig}^i is the normalization parameter from which the yields are calculated (see Equation 5.6).
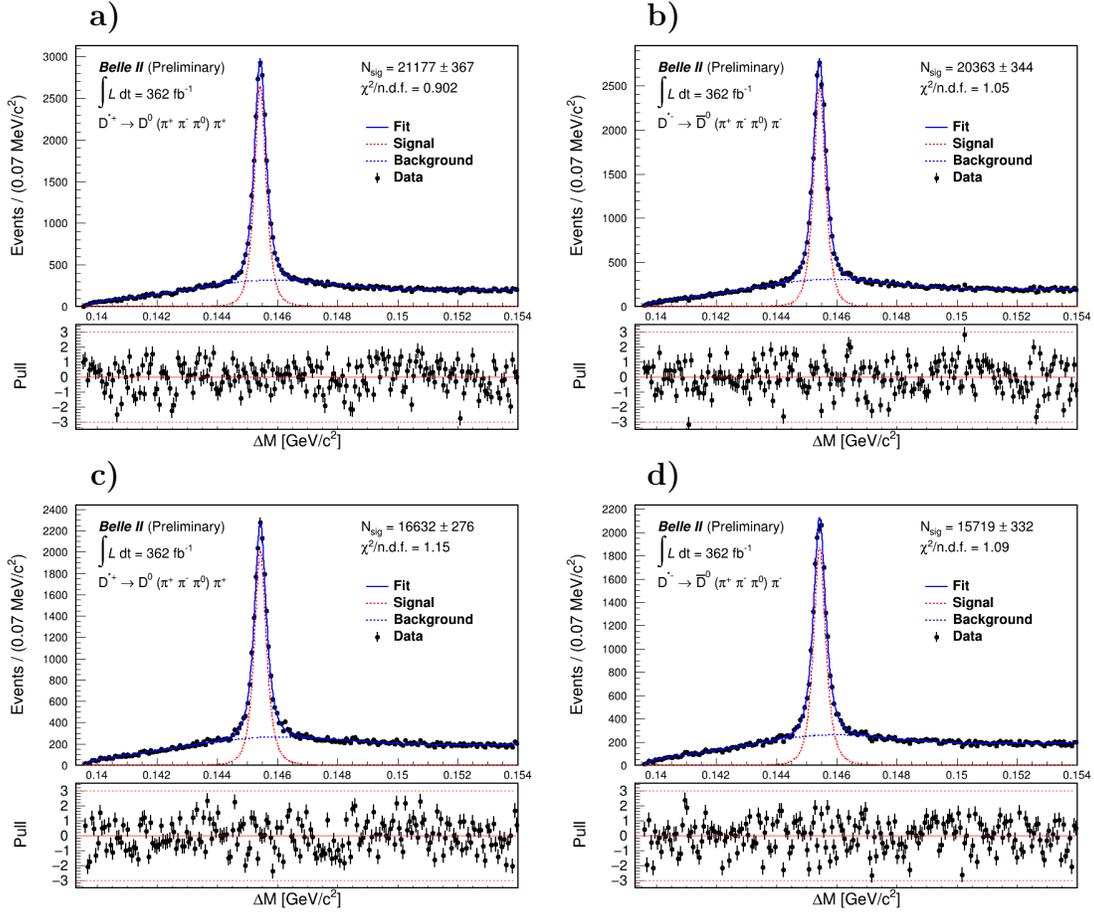
Figure 7.6: Fits for **a)** D^0 in bin 5, region A, **b)** $\overline{D}^0$ in bin 5, region A, **c)** D^0 in bin 5, region B, and **d)** $\overline{D}^0$ in bin 5, region B.

| Param. | A | | B | |
	D^0	$\overline{D}^0$	D^0	$\overline{D}^0$
I^5_{sig}	1.482 ± 0.026	1.425 ± 0.024	1.164 ± 0.019	1.100 ± 0.023
μ^5 (MeV/c^2)	145.417 ± 0.007	145.427 ± 0.008	145.419 ± 0.008	145.415 ± 0.009
σ^5 (MeV/c^2)	0.276 ± 0.016	0.290 ± 0.017	0.267 ± 0.015	0.275 ± 0.018
γ^5	-0.053 ± 0.33	-0.008 ± 0.035	-0.063 ± 0.036	-0.079 ± 0.039
δ^5	1.233 ± 0.077	1.265 ± 0.080	1.157 ± 0.069	1.174 ± 0.082

Table 7.11: Signal parameters for all fits in bin 5. I^i_{sig} is the normalization parameter from which the yields are calculated (see Equation 5.6).

Chapter 8

Systematics

The error on the asymmetries presented in chapter 7 is entirely statistical. We also want to include systematic uncertainties on the final asymmetry values. We identified four main sources of systematic uncertainty: 1) imperfect cancellation of $\mathcal{A}_{\mathrm{FB}}$ in Equation 5.4, 2) imperfect cancellation of $\mathcal{A}_{\pi_s}$ in Equation 5.4, 3) imperfect cancellation of $\mathcal{A}_{D^0}^i$ from choosing bins that are symmetric across the $m(\pi^{\pm}\pi^0)^2$ symmetry axis, and 4) systematic effects from the fits used to obtain the signal yield. This chapter details our analysis of each of these systematic effects in sections 8.2–8.5. In section 8.6, some consistency checks that were performed are described. Finally, in section 8.7, the results from Table 7.5 are presented with systematic uncertainties included.

To calculate the uncertainty from the imperfect cancellation of $\mathcal{A}_{\mathrm{FB}}$, $\mathcal{A}_{\pi_s}$, and $\mathcal{A}_{D^0}^i$, we used $400\,\mathrm{fb}^{-1}$ of truth-tagged signal MC (see chapter 4 for full description). We noted in chapter 5 that the relevant distributions for calculating these asymmetries can have different shapes for signal and background (see Figures 5.3–5.16), and since we cannot select only signal events in the data, we rely on truth-tagging in the MC. Since these distributions exhibit relatively good agreement between data and MC and we are conservative in our estimates of systematic uncertainties, the truth-tagged signal MC events are an acceptable proxy for the signal distributions in data.

8.1 Notation

In sections 8.2, 8.3, and 8.4, we use j as a general bin index in a histogram (typically angle or momentum). Note this is distinct from the use of i in previous chapters as a bin index in the Dalitz plot. We also use Dalitz bin indices in this chapter, and we will be explicit about the difference between the two when there might be confusion.

Consider α as an arbitrary asymmetry. n_j ($\bar{n}_j$) is the number of particles of one type (the charge-conjugate type) in bin j. $n = \sum n_j$ and $\bar{n} = \sum \bar{n}_j$. The total asymmetry over all bins is:

$$\alpha = \frac{\sum n_j - \sum \bar{n}_j}{\sum n_j + \sum \bar{n}_j} = \frac{n - \bar{n}}{n + \bar{n}} \tag{8.1}$$

Note that this is *not* a simple sum of the bin asymmetries, $\alpha_j = \frac{n_j - \bar{n}_j}{n_j + \bar{n}_j}$, since the sums are not global but separate in numerator and denominator. In terms of the bin asymmetries,

the total asymmetry α can be recast as:

$$\alpha = \frac{1}{n + \bar{n}} \left[\sum \alpha_j (n_j + \bar{n}_j) \right].$$ (8.2)

We can also define normalized bin weights, $w_j = \frac{n_j + \bar{n}_j}{n + \bar{n}}$, so that α can be further simplified to:

$$\alpha = \sum w_j \alpha_j.$$ (8.3)

Thus, the global asymmetry is a weighted average of the bin asymmetries.

8.2   Cancellation of $\mathcal{A}_{\mathrm{FB}}$

Our justification for the cancellation of $\mathcal{A}_{\mathrm{FB}}$ is the strong agreement between the distribution of $\cos\theta^*_{D^*}$ in each bin, as seen in Figures 5.3 and 5.4. However, there is still some small bin dependence in these distributions for which we now calculate a systematic uncertainty. We start with the general form for the asymmetry:

$$\mathcal{A}_{\mathrm{FB}} = \mathcal{A}_M \cos\theta^*$$ (8.4)

a linear form where $\mathcal{A}_M$ is the maximum asymmetry, occurring at $\cos\theta^* = \pm 1$. We will assume $\mathcal{A}_M \approx 15\%$.

Using the result from section 8.1, with $\alpha = \mathcal{A}_{\mathrm{FB}}$ and $x \equiv \cos\theta^*$ for convenience:

$$\mathcal{A}_{\mathrm{FB}} = \sum w_j (\mathcal{A}_M x_j) = \mathcal{A}_M \sum w_j x_j = \mathcal{A}_M x_{avg}$$ (8.5)

where $x_{avg} \equiv \sum w_j x_j$ is the average value of x based on the distribution of data given by the weights w_j.

We can easily get x_{avg} in the signal MC for the global $\cos\theta^*_{D^*}$ distribution and also for each of the six Dalitz bins. The difference in means between the global distribution and the distribution in a particular Dalitz bin scaled by $\mathcal{A}_M$ is the expected shift in the nuisance asymmetry due to bin dependence.

The values of $\mathcal{A}_{\mathrm{FB}}$ (for the global distribution and the six separate bins) are small since the distributions are fairly symmetric, making x_{avg} quite small (< 0.04). The changes in these $\mathcal{A}_{\mathrm{FB}}$ are also small.

We also consider the standard error on the means, $\sigma_i / \sqrt{\mathrm{n}^i}$ where σ_i is the RMS of the $\cos\theta^*_{D^*}$ distribution and n^i is the total number of events in Dalitz bin i. This error is small for the high-statistics bins, but not for low-statistics ones. The error on the global mean – $\sigma / \sqrt{N}$, where $N = \sum \mathrm{n}^i$ – is also taken into account.

For each bin, we first sum in quadrature the error on the mean in a particular Dalitz bin and the error on the global mean, ignoring correlations (which is conservative). We then linearly sum this with the difference of the binned vs. global mean. Finally, this total is multiplied by $\mathcal{A}_M = 0.15$ to give the systematic on the production asymmetry for each Dalitz bin:

$$\text{Syst. uncert. from } \mathcal{A}_{\mathrm{FB}} \text{ for bin } i = \mathcal{A}_M * \left(\text{Difference of means} + \sqrt{\frac{\sigma_i^2}{\mathrm{n}^i} + \frac{\sigma^2}{N}} \right).$$ (8.6)

Table 8.1 shows the numbers used in the calculations for each Dalitz bin. It also compares the systematic uncertainty from this calculation to the statistical error in each bin. The largest systematic we observe is 0.18% in bin 5, which is still smaller than the statistical error in that bin (0.62%).

In the future, with higher statistics throughout the Dalitz plot, the error on the mean will become smaller, and one could correct for the varying means to reduce the systematic uncertainty.

i	Difference of means	$\sigma_i/\sqrt{n^i}$	$\sigma/\sqrt{N}$	Syst. uncert. (%)	Stat. uncert. (%)
0	$+0.00492$	0.00419		**0.14**	5.8
1	-0.00265	0.00173		**0.069**	0.40
2	-0.00122	0.00599	0.00088	**0.11**	4.3
3	-0.00890	0.00161		**0.16**	0.34
4	$+0.00224$	0.00297		**0.080**	2.8
5	$+0.01014$	0.00160		**0.18**	0.62

Table 8.1: Systematic uncertainty due to imperfect $\mathcal{A}_{\mathrm{FB}}$ cancellation in each bin, determined from truth-tagged signal MC. The statistical errors for the results in data (see section 7.3) are shown for comparison.

8.3 Cancellation of $\mathcal{A}_{\pi_s}$

Our justification for the cancellation of $\mathcal{A}_{\pi_s}$ is the strong agreement between the distributions of p_{π_s} and $\cos\theta_{\pi_s}$ in each bin, as seen in Figures 5.5 – 5.8. However, there is still some small bin dependence in these distributions for which we now calculate a systematic uncertainty to include on our result.

Unlike $\mathcal{A}_{\mathrm{FB}}$, we do not have a general form for $\mathcal{A}_{\pi_s}$. However, we know it is dependent on p_{π_s} and $\cos\theta_{\pi_s}$, and ultimately we only need an expression for the change between $\mathcal{A}_{\pi_s}$ in a specific Dalitz bin and $\mathcal{A}_{\pi_s}$ across the whole Dalitz plot. We will call this change $\Delta^i\mathcal{A}_{\pi_s}$ in Dalitz bin i. Referring back to Equation 8.3, we can write:

$$\Delta^i\mathcal{A}_{\pi_s} = \sum_j w_j^i \beta(x_j) - \sum_j w_j \beta(x_j) \tag{8.7}$$

where i is a Dalitz bin index and j is a bin index in the histogram $x \equiv p_{\pi_s}$ or $x \equiv \cos\theta_{\pi_s}$. The weights w_j^i and w_j are calculated for both p_{π_s} and $\cos\theta_{\pi_s}$ distributions. The superscript i indicates w_j^i is calculated in a particular Dalitz bin; w_j (no superscript) are the weights for the global distribution. $\beta(x)$ is the slow pion detection efficiency asymmetry as a function of either p_{π_s} or $\cos\theta_{\pi_s}$. Note that $\beta(x)$ is not dependent on the Dalitz bin, only on x. The reason $\mathcal{A}_{\pi_s}$ ultimately has some residual bin dependence is because of the difference in the weights: $w_j^i \neq w_j$ for an arbitrary value of j. We can define a weighted average of the detection efficiency as $\beta_{avg} = \sum_j w_j \beta(x_j)$ and use this new variable to rewrite Equation 8.7:

$$\Delta^i\mathcal{A}_{\pi_s} = \sum_j w_j^i \left(\beta(x_j) - \beta_{avg}\right) - \sum_j w_j \left(\beta(x_j) - \beta_{avg}\right) = \sum_j (w_j^i - w_j)(\beta(x_j) - \beta_{avg}) \tag{8.8}$$

We have simply added 0 to the right side of Equation 8.7, in the form of $-\beta_{avg} \left(\sum_j w_j^i - \sum_j w_j \right)$. This is 0 because the w_j^i and w_j both sum to 1, and β_{avg} is a constant that can be pulled out of the sums. We treat p_{π_s} and $\cos \theta_{\pi_s}$ separately and calculate an error on $\Delta^i \mathcal{A}_{\pi_s}$ for both distributions. The resulting uncertainties are added in quadrature to obtain the final systematic uncertainty on this calculation. We could consider both variables simultaneously by looking at 2-D distributions, but the lower statistics resulting from having many more bins in a 2-D plot cause large fluctuations that are not actual differences. As a conservative estimate, we take 5% as a scale for $\beta(x_j) - \beta_{avg}$. For the weights, we take the RMS of the actual $w_j^i - w_j$ distributions for 25 bins in p_{π_s} and $\cos \theta_{\pi_s}$ separately.

The choice of 25 bins is based on the following considerations: for a large number of bins, there will be large statistical fluctuations; for too few bins, one may integrate over structure in the weights or asymmetries. Using 25 bins allows for fairly detailed structures to be seen without being overwhelmed by statistical fluctuations.

For the final systematic uncertainty calculation, we multiply the RMS in the weight difference by the asymmetry scale (5%) and by the number of bins (25). If $\mathrm{RMS}_1^i =$ the RMS of the $w_j^i - w_j$ distribution in p_{π_s} for bin i, and $\mathrm{RMS}_2^i =$ the RMS of the $w_j^i - w_j$ distribution in $\cos \theta_{\pi_s}$ for bin i, then:

$$\text{Syst. uncert. from } \mathcal{A}_{\pi_s} \text{ for bin } i = (0.05 \times 25) \sqrt{\left(\mathrm{RMS}_1^i\right)^2 + \left(\mathrm{RMS}_2^i\right)^2}. \qquad (8.9)$$

This assumes complete correlation of the signs of $w_j^i - w_j$ and the asymmetries, which is quite conservative. Table 8.2 shows the numbers used in the calculation of this systematic asymmetry for each Dalitz bin. It also compares the systematic uncertainty from this calculation to the statistical error in each bin. The largest RMS and largest systematic we observe are 0.00240 and 0.42% (respectively), both in bin 2, which has the lowest statistics of any bin. This is a fairly conservative estimate and is still much smaller than the statistical error in that bin (4.3%).

i	RMS_1^i	RMS_2^i	**Syst. uncert. (%)**	Stat. uncert. (%)
0	0.00145	0.00168	**0.28**	5.8
1	0.00073	0.00064	**0.12**	0.40
2	0.00231	0.00240	**0.42**	4.3
3	0.00072	0.00092	**0.15**	0.34
4	0.00114	0.00105	**0.19**	2.8
5	0.00124	0.00126	**0.22**	0.62

Table 8.2: Systematic uncertainty due to imperfect $\mathcal{A}_{\pi_s}$ cancellation in each bin, determined from truth-tagged signal MC. The statistical errors for the results in data (see section 7.3) are shown for comparison.

8.4 Cancellation of $\mathcal{A}_{D^0}^i$

Unlike $\mathcal{A}_{\mathrm{FB}}$ and $\mathcal{A}_{\pi_s}$, which are largely global asymmetries that we cancel out by subtraction in Equation 5.4, $\mathcal{A}_{D^0}^i$ is a bin-specific asymmetry that we cancel out by choosing bins that are

symmetric across the $m(\pi^{\pm}\pi^0)^2$ symmetry axis and then taking an unweighted average of the asymmetries above and below the axis in each bin (see subsection 5.2.3 for full explanation). Our justification for why this method corrects for $\mathcal{A}_{D^0}^i$ is the strong agreement between the weighted distributions of the momentum and cosine of the like-sign pion as compared to the unlike-sign pion in each bin, as seen in Figures 5.9 – 5.16. However, there is still some small disagreement between these distributions for which we now calculate a systematic uncertainty to include on our result.

The method for calculating this systematic is the same as for calculating the error on the $\mathcal{A}_{\pi_s}$ cancellation described in the previous section. The major difference is that instead of comparing a distribution within a specific Dalitz bin to the global distribution, we compare two different distributions in a specific bin: p_{like} vs. p_{unlike} and $\cos\theta_{\mathrm{like}}$ vs. $\cos\theta_{\mathrm{unlike}}$. We once again choose 25 bins in the momentum and cosine histograms, but we take 3% (instead of 5%) as the asymmetry scale.

Table 8.3 shows the numbers used in the calculation of this systematic uncertainty for each Dalitz bin. It also compares the systematic uncertainty from this calculation to the statistical error in each bin. The largest RMS and largest systematic we observe are 0.00443 and 0.43% (respectively), both in bin 2, which has the lowest statistics of any bin. This is a fairly conservative estimate and is still much smaller than the statistical error in that bin (4.3%).

i	RMS_1^i	RMS_2^i	**Syst. uncert. (%)**	Stat. uncert. (%)
0	0.00239	0.00255	**0.26**	5.8
1	0.00073	0.00114	**0.10**	0.40
2	0.00374	0.00443	**0.43**	4.3
3	0.00265	0.00085	**0.21**	0.34
4	0.00162	0.00142	**0.16**	2.8
5	0.00206	0.00088	**0.17**	0.62

Table 8.3: Systematic uncertainty due to imperfect $\mathcal{A}_{D^0}^i$ cancellation in each bin, determined from truth-tagged signal MC. The statistical errors for the results in data (see section 7.3) are shown for comparison.

8.5   Yield systematics

A bias in our fitting method (described in subsection 5.2.4 and section 7.1) could introduce systematic effects into our calculations of $\mathcal{A}_{CP,\mathrm{loc}}^i$. To investigate the scale of these effects, we vary the fitting and yield calculation methods in several ways and redo the analysis on data with these new methods.

Each yield systematic requires 24 new fits. We perform four different systematic variations, for a total of 96 new fits. Rather than show all 96 of these plots, which will not add much to the information already given in the tables below, we choose region A of bin 1 as an example. Figure 8.3 (Figure 8.4) shows the four variations on the D^0 ($\bar{D}^0$) fit in that region. The results in other regions are similar.

8.5.1 Vary signal shape

The first variation is to change the signal function in fits using a Johnson's S_U-distribution. We change the signal function to either a single or double Gaussian for each of these fits (depending on which variation still provides good fit quality) and redo the analysis. Table 8.4 (Table 8.5) summarizes the new fit functions used in region A (B) to determine signal yield. Note that in region A of bins 0, 2, and 4, we could not modify the signal function because a single Gaussian was already being used. This is the simplest possible signal function, used in these regions because the statistics are not high enough to support additional parameters. Any variation would require introducing more parameters to the signal function and was not feasible in these three cases.

i	Signal function	Background function
0	Single Gaussian	
1	**Double Gaussian**	
2	Single Gaussian	3/2-power threshold function
3	**Double Gaussian**	
4	Single Gaussian	
5	**Double Gaussian**	Single Gaussian and 3/2-power threshold function

Table 8.4: Fit functions for fits to D^0 and $\overline{D}^0$ ΔM plots in data in **region A** for all six bins. Functions in **bold** have been changed from those described in Table 7.1 as a systematic check.

i	Signal function	Background function
0	**Single Gaussian**	
1	**Double Gaussian**	
2	**Single Gaussian**	3/2-power threshold function
3	**Double Gaussian**	
4	**Single Gaussian**	
5	**Double Gaussian**	Single Gaussian and 3/2-power threshold function

Table 8.5: Fit functions for fits to D^0 and $\overline{D}^0$ ΔM plots in data in **region B** for all six bins. Functions in **bold** have been changed from those described in Table 7.2 as a systematic check.

The raw yields and fit quality (i.e., χ^2/n.d.f.) from this variation on the nominal analysis are listed in Table 8.6 (Table 8.7) for region A (B). The final asymmetries, including a comparison to the nominal result from section 7.3, are shown in Table 8.8.

8.5.2 Vary background shape

The second variation is to change the background function in the fits. We replace the 3/2-power threshold function with a 5/2-power threshold function in all fits, which adds one additional background shape parameter to each fit.

A	D^0		$\overline{D}^0$		$\mathcal{A}^i_A$ (%)
i	χ^2/n.d.f.	n^i_A	χ^2/n.d.f.	$\bar{n}^i_A$	
0	1.26	561 ± 57	1.11	394 ± 57	17.4 ± 8.6
1	1.19	12293 ± 186	1.13	11653 ± 166	2.7 ± 1.0
2	1.05	485 ± 56	1.18	376 ± 43	12.7 ± 8.0
3	1.10	15974 ± 178	1.25	15678 ± 174	0.94 ± 0.79
4	1.15	1124 ± 75	1.08	1151 ± 76	-1.2 ± 4.7
5	0.94	20535 ± 308	1.00	20669 ± 320	-0.3 ± 1.1

Table 8.6: Fit results and raw asymmetries from fits to D^0 and $\overline{D}^0$ ΔM plots in **region A** for all six bins with the signal function changed as a systematic.

B	D^0		$\overline{D}^0$		$\mathcal{A}^i_B$ (%)
i	χ^2/n.d.f.	n^i_B	χ^2/n.d.f.	$\bar{n}^i_B$	
0	0.99	945 ± 60	0.97	799 ± 55	8.4 ± 4.6
1	1.22	24255 ± 215	1.39	24045 ± 226	0.43 ± 0.65
2	1.32	1596 ± 62	1.07	1659 ± 62	-1.9 ± 2.7
3	1.20	23251 ± 212	1.33	23296 ± 234	-0.10 ± 0.68
4	1.32	2622 ± 83	1.29	2310 ± 80	6.3 ± 2.3
5	1.08	16628 ± 254	1.07	15332 ± 287	4.1 ± 1.2

Table 8.7: Fit results and raw asymmetries from fits to D^0 and $\overline{D}^0$ ΔM plots in **region B** for all six bins with the signal function changed as a systematic.

i	$\mathcal{A}^i_A$ (%)	$\mathcal{A}^i_B$ (%)	$\mathcal{A}^{*i}_{\mathrm{bin}}$ (%)	$\mathcal{A}^i_{CP,\mathrm{loc}}$ (%)	Diff. from nominal
0	17.4 ± 8.6	8.4 ± 4.6	12.9 ± 4.9	11.4 ± 4.8	**2.6**
1	2.7 ± 1.0	0.43 ± 0.65	1.55 ± 0.61	0.05 ± 0.43	**0.192**
2	12.7 ± 8.0	-1.9 ± 2.7	5.4 ± 4.2	3.9 ± 4.1	**0.3**
3	0.94 ± 0.79	-0.10 ± 0.68	0.42 ± 0.52	-1.08 ± 0.35	**0.27**
4	-1.2 ± 4.7	6.3 ± 2.3	2.6 ± 2.6	1.1 ± 2.5	**−1.0**
5	-0.3 ± 1.1	4.1 ± 1.2	1.86 ± 0.81	0.36 ± 0.56	**−0.43**

Table 8.8: Final values for $\mathcal{A}^i_{CP,\mathrm{loc}}$ in each Dalitz bin with the signal function varied as a systematic. The average asymmetry across the plot is $\mathcal{A}_{\mathrm{avg}} = (1.50 \pm 0.37)\%$. $\mathcal{A}^i_{CP,\mathrm{loc}}$ is calculated by subtracting $\mathcal{A}_{\mathrm{avg}}$ from $\mathcal{A}^{*i}_{\mathrm{bin}}$ for each bin. The rightmost column is the difference between the central value of $\mathcal{A}^i_{CP,\mathrm{loc}}$ in this table and $\mathcal{A}^i_{CP,\mathrm{loc}}$ from Table 7.5.

The raw yields and fit quality (i.e., χ^2/n.d.f.) from this variation on the nominal analysis are listed in Table 8.9 (Table 8.10) for region A (B). The final asymmetries, including a comparison to the nominal result from section 7.3, are show in Table 8.11.

8.5.3 Widen $M(\pi^+\pi^-\pi^0)$ signal region

The third variation is to widen the $M(\pi^+\pi^-\pi^0)$ signal region, defined in subsection 4.2.7. This effectively increases the amount of background under the ΔM peak, so it tests how

A		D^0		$\overline{D}^0$		$\mathcal{A}^i_{\mathrm{A}}$ (%)
i	χ^2/n.d.f.	n^i_A	χ^2/n.d.f.	$\bar{n}^i_A$		
0	1.19	610 ± 61	1.07	434 ± 64		16.9 ± 8.7
1	1.17	12188 ± 175	1.14	11600 ± 164		2.5 ± 1.0
2	1.02	520 ± 60	1.16	399 ± 45		13.7 ± 7.9
3	1.15	15862 ± 177	1.20	15909 ± 191		-0.15 ± 0.82
4	1.05	1193 ± 79	1.03	1209 ± 81		-0.7 ± 4.7
5	0.87	20919 ± 368	1.03	20122 ± 386		1.9 ± 1.3

Table 8.9: Fit results and raw asymmetries from fits to D^0 and $\overline{D}^0$ ΔM plots in **region A** for all six bins with the background function changed as a systematic.

B		D^0		$\overline{D}^0$		$\mathcal{A}^i_{\mathrm{B}}$ (%)
i	χ^2/n.d.f.	n^i_B	χ^2/n.d.f.	$\bar{n}^i_B$		
0	0.97	1061 ± 108	0.96	1076 ± 305		-1 ± 15
1	1.10	24357 ± 224	1.41	24090 ± 239		0.55 ± 0.68
2	1.30	1731 ± 191	1.05	1805 ± 100		-2.1 ± 3.8
3	1.19	23150 ± 214	1.34	23003 ± 229		0.32 ± 0.68
4	1.12	3354 ± 220	1.10	2812 ± 185		8.8 ± 4.6
5	1.09	16448 ± 525	1.07	15568 ± 195		2.7 ± 1.7

Table 8.10: Fit results and raw asymmetries from fits to D^0 and $\overline{D}^0$ ΔM plots in **region B** for all six bins with the background function changed as a systematic.

i	$\mathcal{A}^i_{\mathrm{A}}$ (%)	$\mathcal{A}^i_{\mathrm{B}}$ (%)	$\mathcal{A}^{*i}_{\mathrm{bin}}$ (%)	$\mathcal{A}^i_{CP,\mathrm{loc}}$ (%)	Diff. from nominal
0	16.9 ± 8.7	-1 ± 15	8.1 ± 8.7	6.5 ± 8.6	**-2.3**
1	2.5 ± 1.0	0.55 ± 0.68	1.51 ± 0.61	-0.05 ± 0.43	**0.09**
2	13.7 ± 7.9	-2.1 ± 3.8	5.8 ± 4.4	4.3 ± 4.3	**0.7**
3	-0.15 ± 0.82	0.32 ± 0.68	0.09 ± 0.53	-1.48 ± 0.36	**-0.13**
4	-0.7 ± 4.7	8.8 ± 4.6	4.1 ± 3.3	2.5 ± 3.1	**0.4**
5	1.9 ± 1.3	2.7 ± 1.7	2.3 ± 1.1	0.78 ± 0.75	**-0.01**

Table 8.11: Final values for $\mathcal{A}^i_{CP,\mathrm{loc}}$ in each Dalitz bin with the background function varied as a systematic. The average asymmetry across the plot is $\mathcal{A}_{\mathrm{avg}} = (1.56 \pm 0.45)\%$. $\mathcal{A}^i_{CP,\mathrm{loc}}$ is calculated by subtracting $\mathcal{A}_{\mathrm{avg}}$ from $\mathcal{A}^{*i}_{\mathrm{bin}}$ for each bin. The rightmost column is the difference between the central value of $\mathcal{A}^i_{CP,\mathrm{loc}}$ in this table and $\mathcal{A}^i_{CP,\mathrm{loc}}$ from Table 7.5.

sensitive our yield fits are to the background level. None of the fit functions from Table 7.1 or Table 7.2 are changed, but the $M(\pi^+\pi^-\pi^0)$ signal region over which the ΔM distribution is plotted is widened from $1.83\,\mathrm{GeV}/c^2 < M(\pi^+\pi^-\pi^0) < 1.89\,\mathrm{GeV}/c^2$ to $1.815\,\mathrm{GeV}/c^2 < M(\pi^+\pi^-\pi^0) < 1.905\,\mathrm{GeV}/c^2$ (i.e., the width is increased by 50%). Figure 8.1 shows the global $M(\pi^+\pi^-\pi^0)$ and ΔM plots with this new signal region in MC; compare to Figure 4.7. Figure 8.2 shows the data–MC comparisons of the mass plots for this wider signal region; compare to Figure 4.8. As in subsection 4.2.7, we observe good data–MC agreement between the background levels in both plots but a discrepancy in the signal peak.
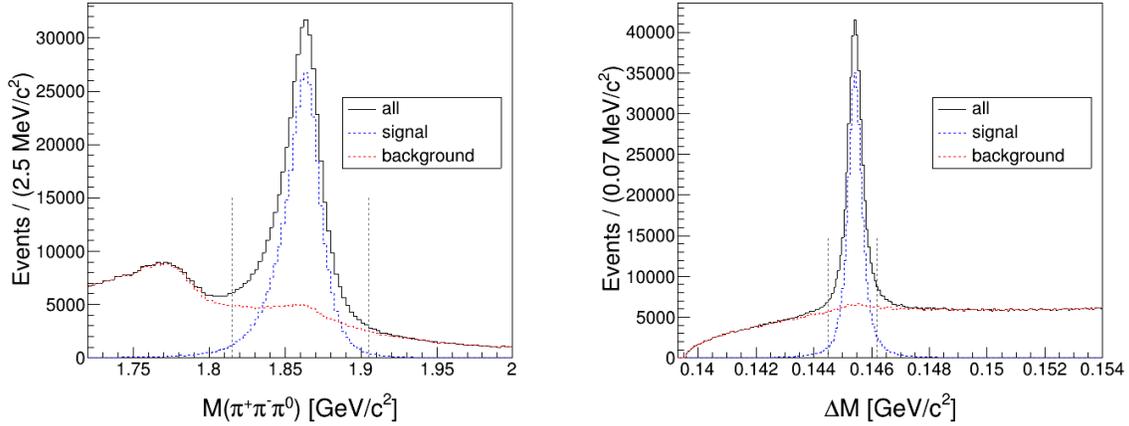
Figure 8.1: Signal-enhanced $M(\pi^+\pi^-\pi^0)$ and ΔM plots of candidates in MC retained after all selection criteria are applied. The dotted vertical lines indicate the signal regions. The $M(\pi^+\pi^-\pi^0)$ signal region is 50% wider than in Figure 4.7, which effectively increases the background in the ΔM plot.
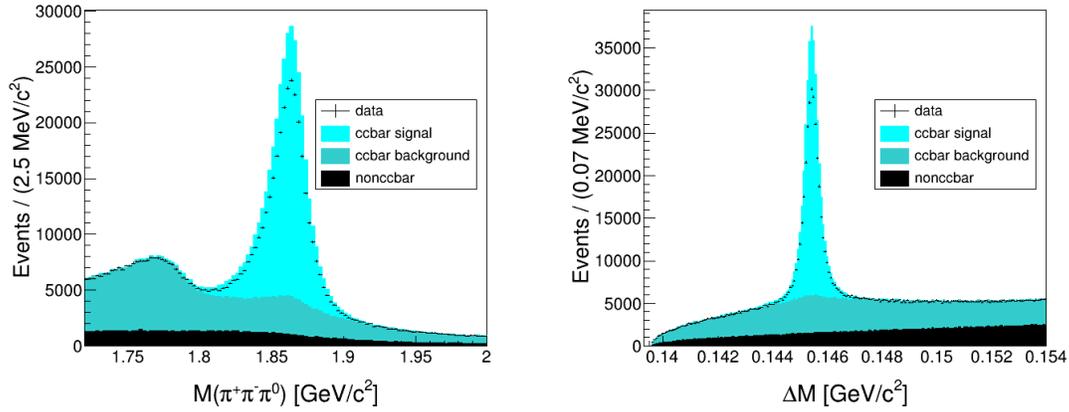
Figure 8.2: Signal-enhanced $M(\pi^+\pi^-\pi^0)$ and ΔM plots of candidates in MC and data over a wider $M(\pi^+\pi^-\pi^0)$ signal region.

The raw yields and fit quality (i.e., χ^2/n.d.f.) from this variation on the nominal analysis are listed in Table 8.12 (Table 8.13) for region A (B). The final asymmetries, including a comparison to the nominal result from section 7.3, are show in Table 8.14.

8.5.4 Use finer ΔM binning

The fourth and final yield systematic we perform is to use a finer binning (i.e., increase the number of bins) in the ΔM histograms (this is different from the Dalitz plot bins – those remain unchanged). Since the yield fits we perform are binned log-likelihood fits, if the binning in the ΔM histograms is too wide, it could impact fit quality, especially in terms of

A	D^0		$\overline{D}^0$		$\mathcal{A}_{\mathrm{A}}^i$ (%)
i	$\chi^2/\mathrm{n.d.f.}$	n_A^i	$\chi^2/\mathrm{n.d.f.}$	$\bar{n}_A^i$	
0	1.28	600 ± 69	0.99	427 ± 64	16.8 ± 9.2
1	1.05	13629 ± 193	1.12	12917 ± 184	2.7 ± 1.0
2	1.12	508 ± 60	1.15	443 ± 53	6.8 ± 8.4
3	1.06	17356 ± 198	1.16	17304 ± 209	0.15 ± 0.83
4	0.97	1131 ± 86	1.17	1205 ± 90	-3.2 ± 5.3
5	1.08	22093 ± 390	1.11	20529 ± 221	3.7 ± 1.0

Table 8.12: Fit results and raw asymmetries from fits to D^0 and $\overline{D}^0$ ΔM plots in **region A** for all six bins with $M(\pi^+\pi^-\pi^0)$ signal region widened as a systematic.

B	D^0		$\overline{D}^0$		$\mathcal{A}_{\mathrm{B}}^i$ (%)
i	$\chi^2/\mathrm{n.d.f.}$	n_B^i	$\chi^2/\mathrm{n.d.f.}$	$\bar{n}_B^i$	
0	1.04	1052 ± 85	1.00	1020 ± 148	1.5 ± 8.3
1	1.07	27805 ± 261	1.35	27440 ± 277	0.66 ± 0.69
2	1.10	1940 ± 105	1.01	2019 ± 113	-2.0 ± 3.9
3	1.16	25832 ± 245	1.10	25629 ± 264	0.39 ± 0.70
4	1.12	3169 ± 166	1.14	2666 ± 145	8.6 ± 3.7
5	1.29	17282 ± 322	1.02	16326 ± 401	2.8 ± 1.5

Table 8.13: Fit results and raw asymmetries from fits to D^0 and $\overline{D}^0$ ΔM plots in **region B** for all six bins with $M(\pi^+\pi^-\pi^0)$ signal region widened as a systematic.

i	$\mathcal{A}_{\mathrm{A}}^i$ (%)	$\mathcal{A}_{\mathrm{B}}^i$ (%)	$\mathcal{A}_{\mathrm{bin}}^{*i}$ (%)	$\mathcal{A}_{CP,\mathrm{loc}}^i$ (%)	Diff. from nominal
0	16.8 ± 9.2	1.5 ± 8.3	9.2 ± 6.2	7.4 ± 6.1	-1.4
1	2.7 ± 1.0	0.66 ± 0.69	1.67 ± 0.61	-0.14 ± 0.42	**0.00**
2	6.8 ± 8.4	-2.0 ± 3.9	2.4 ± 4.6	0.6 ± 4.5	**-3.0**
3	0.15 ± 0.83	0.39 ± 0.70	0.27 ± 0.54	-1.54 ± 0.36	-0.19
4	-3.2 ± 5.3	8.6 ± 3.7	2.7 ± 3.3	0.9 ± 3.2	-1.2
5	3.7 ± 1.0	2.8 ± 1.5	3.26 ± 0.93	1.45 ± 0.66	**0.66**

Table 8.14: Final values for $\mathcal{A}_{CP,\mathrm{loc}}^i$ in each Dalitz bin with $M(\pi^+\pi^-\pi^0)$ signal region widened as a systematic. The average asymmetry across the plot is $\mathcal{A}_{\mathrm{avg}} = (1.81 \pm 0.41)\%$. $\mathcal{A}_{CP,\mathrm{loc}}^i$ is calculated by subtracting $\mathcal{A}_{\mathrm{avg}}$ from $\mathcal{A}_{\mathrm{bin}}^{*i}$ for each bin. The rightmost column is the difference between the central value of $\mathcal{A}_{CP,\mathrm{loc}}^i$ in this table and $\mathcal{A}_{CP,\mathrm{loc}}^i$ from Table 7.5.

the signal width. To check that these binning effects are not present, we decrease the bin width from $0.07\,\mathrm{MeV}/c^2$ to $0.05\,\mathrm{MeV}/c^2$ and repeat the analysis.

The raw yields and fit quality (i.e., $\chi^2/\mathrm{n.d.f.}$) from this variation on the nominal analysis are listed in Table 8.15 (Table 8.16) for region A (B). The final asymmetries, including a comparison to the nominal result from section 7.3, are show in Table 8.17.

A	D^0		$\overline{D}^0$		$\mathcal{A}_A^i$ (%)
i	χ^2/n.d.f.	n_A^i	χ^2/n.d.f.	$\bar{n}_A^i$	
0	1.25	555 ± 58	1.08	396 ± 58	16.7 ± 8.7
1	1.08	12087 ± 163	1.03	11562 ± 156	2.22 ± 0.95
2	1.07	486 ± 56	1.15	380 ± 43	12.2 ± 8.0
3	1.07	15094 ± 172	1.08	15803 ± 179	0.32 ± 0.78
4	1.06	1129 ± 75	0.98	1154 ± 76	-1.1 ± 4.7
5	0.97	21099 ± 355	1.04	20333 ± 352	1.8 ± 1.2

Table 8.15: Fit results and raw asymmetries from fits to D^0 and $\overline{D}^0$ ΔM plots in **region A** for all six bins with the number of ΔM histogram bins increased as a systematic.

B	D^0		$\overline{D}^0$		$\mathcal{A}_B^i$ (%)
i	χ^2/n.d.f.	n_B^i	χ^2/n.d.f.	$\bar{n}_B^i$	
0	0.96	989 ± 83	1.02	930 ± 130	3.1 ± 8.1
1	1.04	24439 ± 217	1.19	24076 ± 226	0.75 ± 0.65
2	1.23	1678 ± 79	0.96	1758 ± 87	-2.3 ± 3.4
3	1.12	23240 ± 209	1.32	23137 ± 222	0.22 ± 0.66
4	1.09	3070 ± 154	1.01	2573 ± 120	8.8 ± 3.4
5	1.12	16563 ± 288	1.10	15783 ± 342	2.4 ± 1.4

Table 8.16: Fit results and raw asymmetries from fits to D^0 and $\overline{D}^0$ ΔM plots in **region B** for all six bins with the number of ΔM histogram bins increased as a systematic.

i	$\mathcal{A}_A^i$ (%)	$\mathcal{A}_B^i$ (%)	$\mathcal{A}_{\text{bin}}^{*i}$ (%)	$\mathcal{A}_{CP,\text{loc}}^i$ (%)	Diff. from nominal
0	16.7 ± 8.7	3.1 ± 8.1	9.9 ± 6.0	8.4 ± 5.9	**-0.4**
1	2.22 ± 0.95	0.75 ± 0.65	1.48 ± 0.58	-0.05 ± 0.40	**0.095**
2	12.2 ± 8.0	-2.3 ± 3.4	5.0 ± 4.3	3.4 ± 4.2	**-0.2**
3	0.32 ± 0.78	0.22 ± 0.66	0.27 ± 0.51	-1.26 ± 0.35	**0.09**
4	-1.1 ± 4.7	8.8 ± 3.4	3.9 ± 2.9	2.3 ± 2.7	**0.2**
5	1.8 ± 1.2	2.4 ± 1.4	2.13 ± 0.92	0.60 ± 0.64	**-0.19**

Table 8.17: Final values for $\mathcal{A}_{CP,\text{loc}}^i$ in each Dalitz bin with the number of ΔM histogram bins increased as a systematic. The average asymmetry across the plot is $\mathcal{A}_{\text{avg}} = (1.53 \pm 0.40)\%$. $\mathcal{A}_{CP,\text{loc}}^i$ is calculated by subtracting $\mathcal{A}_{\text{avg}}$ from $\mathcal{A}_{\text{bin}}^{*i}$ for each bin. The rightmost column is the difference between the central value of $\mathcal{A}_{CP,\text{loc}}^i$ in this table and $\mathcal{A}_{CP,\text{loc}}^i$ from Table 7.5.

8.5.5 Summary

Table 8.18 compares the results from the four yield systematics to the nominal result. Table 8.19 explicitly compares the rightmost columns from Tables 8.8, 8.11, 8.14, and 8.17 and gives the final systematic uncertainty: the differences in the central values of $\mathcal{A}_{CP,\text{loc}}^i$ between the nominal analysis and the four variations in each bin are added in quadrature. In all cases, this is a conservative estimate of the error from our fitting method, since it does not take into account the direction of the difference in the central values.

| i | $\mathcal{A}^i_{CP,\text{loc}}$ (%) | | | | |
	Nominal result	Vary signal function	Vary bkg. function	Wider signal region	Smaller bins
0	**8.8 ± 5.8**	11.4 ± 4.8	6.5 ± 8.6	7.4 ± 6.1	8.4 ± 5.9
1	**-0.14 ± 0.40**	0.05 ± 0.43	-0.05 ± 0.43	-0.14 ± 0.42	-0.05 ± 0.40
2	**3.6 ± 4.3**	3.9 ± 4.1	4.3 ± 4.3	0.6 ± 4.5	3.4 ± 4.2
3	**-1.35 ± 0.34**	-1.08 ± 0.35	-1.48 ± 0.36	-1.54 ± 0.36	-1.26 ± 0.35
4	**2.1 ± 2.8**	1.1 ± 2.5	2.5 ± 3.1	0.9 ± 3.2	2.3 ± 2.7
5	**0.78 ± 0.62**	0.36 ± 0.56	0.78 ± 0.75	1.45 ± 0.66	0.60 ± 0.64

Table 8.18: Comparison of $\mathcal{A}^i_{CP,\text{loc}}$ from systematic variations in yield calculation methods. Uncertainties are all statistical only.

| i | $\mathcal{A}^i_{CP,\text{loc}}$ (%) | Diff. from nominal $\mathcal{A}^i_{CP,\text{loc}}$ (%) | | | | |
	Nominal result	Vary sig. function	Vary bkg. function	Wider sig. region	Smaller bins	Syst. uncert. (%)
0	**8.8 ± 5.8**	$+2.6$	-2.3	-1.4	-0.4	**3.8**
1	**-0.14 ± 0.40**	$+0.19$	$+0.09$	0.00	$+0.095$	**0.23**
2	**3.6 ± 4.3**	$+0.3$	$+0.7$	-3.0	-0.2	**3.1**
3	**-1.35 ± 0.34**	$+0.27$	$+0.13$	$+0.19$	$+0.09$	**0.37**
4	**2.1 ± 2.8**	-1.0	$+0.4$	-1.2	$+0.2$	**1.6**
5	**0.78 ± 0.62**	-0.43	$+0.01$	$+0.66$	-0.19	**0.81**

Table 8.19: Calculation of systematic uncertainty from variations in yield calculation methods. Uncertainties on the nominal results are statistical only.

8.6 Consistency checks

We perform several checks on our method for which a systematic uncertainty is not assigned. These studies are to check for aberrations in our data sample rather than to assess systematic uncertainties in our fitting method.

8.6.1 Analyze two pieces of dataset separately

As described in section 4.1, the $362\,\text{fb}^{-1}$ used for the analysis in this thesis can be naturally split into two subsets, one that is $187\,\text{fb}^{-1}$ and the other that is $175\,\text{fb}^{-1}$. We performed the analysis procedure from chapter 5 on each of these subsets separately and compared the values for $\mathcal{A}^i_{CP,\text{loc}}$ to the result in Table 7.5. The comparisons are shown in Table 8.20. As expected, the asymmetries in the two smaller subsets differ. The variation in the central values is greater in bins with lower statistics, and the error on the central values is higher in the two smaller subsets than in the combined sample. Only in bin 4 is the difference between the central values in the two subsets not consistent with zero; this may be due to the low statistics in that bin which made it difficult to achieve a reasonable fit in the two smaller subsets independently.
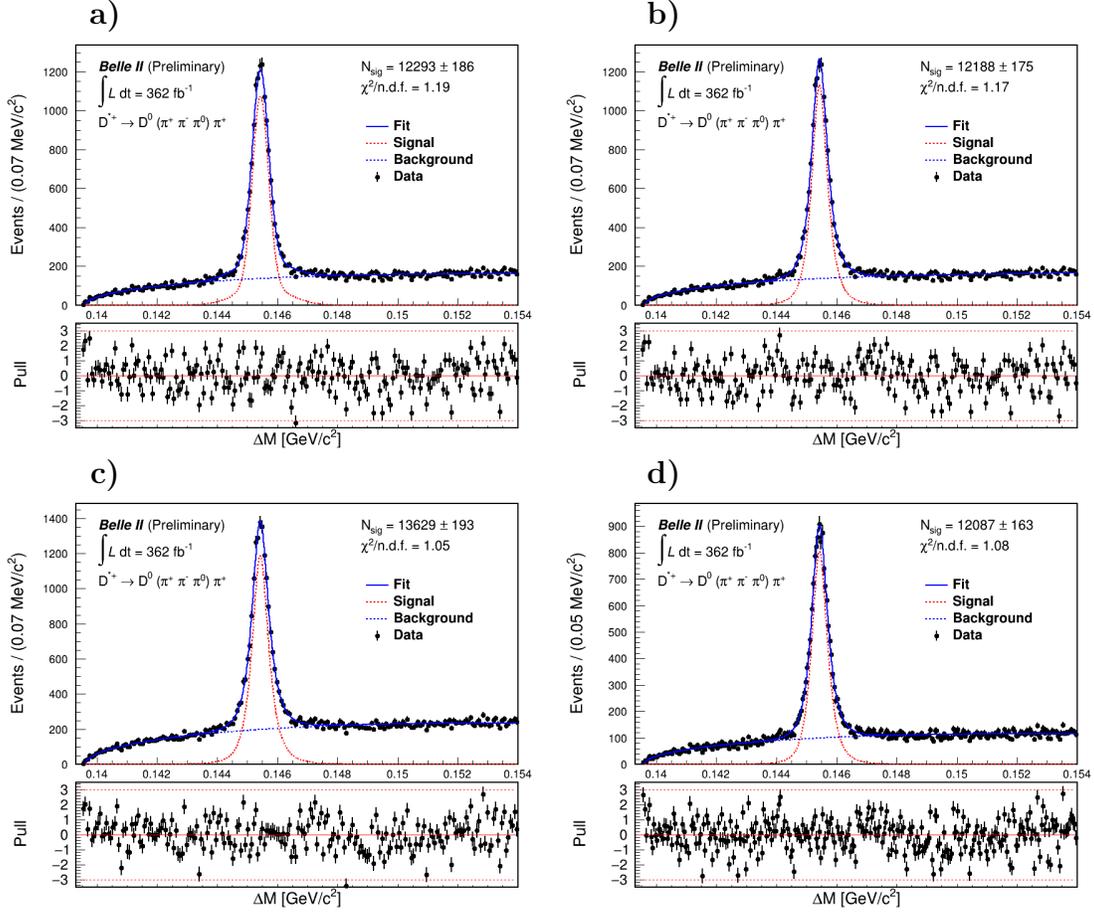
Figure 8.3: D^0 fits in bin 1, region A for **a)** double Gaussian signal function (instead of Johnson's S_U-distribution), **b)** 5/2-power threshold background (instead of 3/2), **c)** wider $M(\pi^+\pi^-\pi^0)$ signal region, and **d)** $0.05\,\mathrm{MeV}/c^2$ ΔM bin width (instead of $0.07\,\mathrm{MeV}/c^2$).

| i | $\mathcal{A}^i_{CP,\mathrm{loc}}$ (%) | | | Difference |
	Nominal result	Result from $187\,\mathrm{fb}^{-1}$ subset	Result from $175\,\mathrm{fb}^{-1}$ subset	
0	$\mathbf{8.8 \pm 5.8}$	14.3 ± 6.0	8.5 ± 7.9	5.8 ± 9.9
1	$\mathbf{-0.14 \pm 0.40}$	-0.58 ± 0.55	0.08 ± 0.59	-0.66 ± 0.81
2	$\mathbf{3.6 \pm 4.3}$	-2.9 ± 6.0	10.7 ± 5.7	-13.6 ± 8.3
3	$\mathbf{-1.35 \pm 0.34}$	-1.32 ± 0.48	-1.45 ± 0.50	0.13 ± 0.69
4	$\mathbf{2.1 \pm 2.8}$	11.4 ± 4.1	-8.8 ± 4.2	20.2 ± 5.9
5	$\mathbf{0.78 \pm 0.62}$	0.30 ± 0.82	1.48 ± 0.99	-1.2 ± 1.3

Table 8.20: Comparison of $\mathcal{A}^i_{CP,\mathrm{loc}}$ in $362\,\mathrm{fb}^{-1}$ dataset vs. $187\,\mathrm{fb}^{-1}$ and $175\,\mathrm{fb}^{-1}$ subsets. The rightmost column is the difference between the asymmetries in each of the subsets.
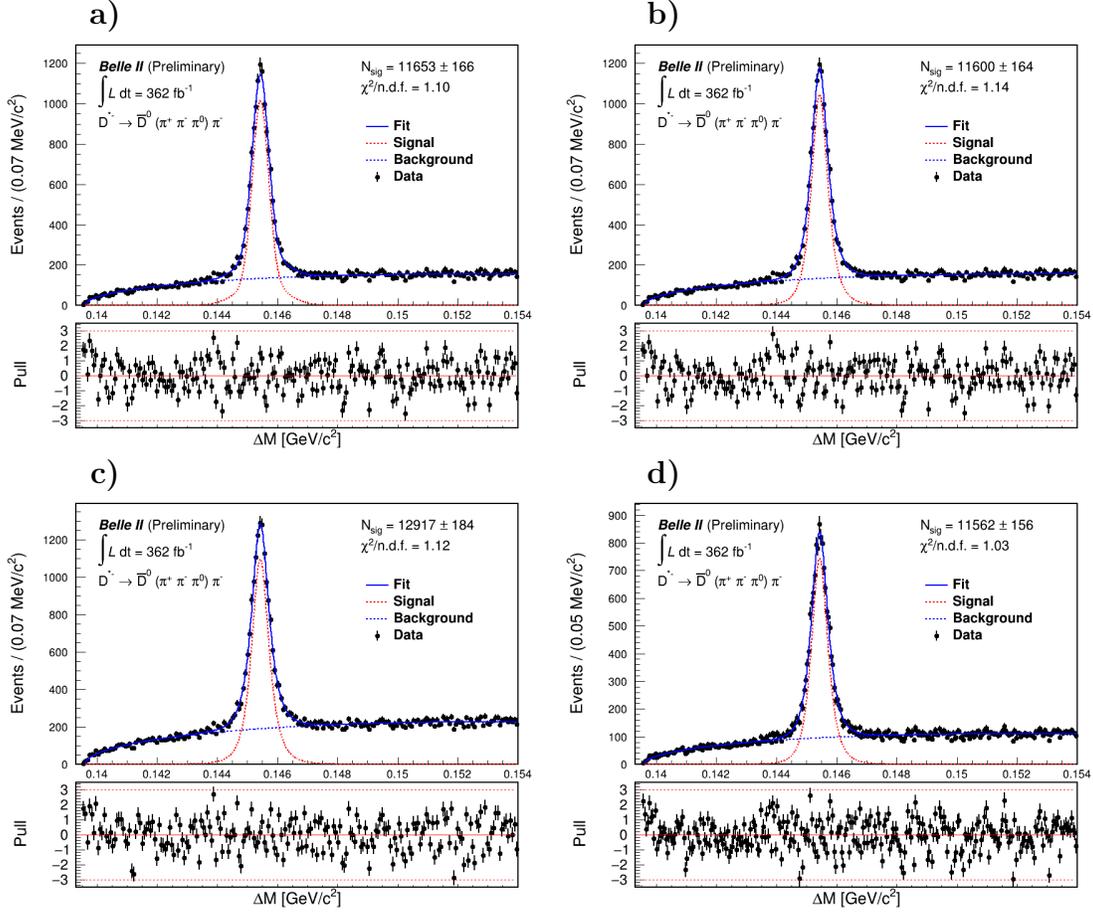
Figure 8.4: $\overline{D}^0$ fits in bin 1, region A for **a)** double Gaussian signal function (instead of Johnson's S_U-distribution), **b)** 5/2-power threshold background (instead of 3/2), **c)** wider $M(\pi^+\pi^-\pi^0)$ signal region, and **d)** $0.05\,\mathrm{MeV}/c^2$ ΔM bin width (instead of $0.07\,\mathrm{MeV}/c^2$).

8.6.2 Randomly assign tagged charge

Other than performing the analysis on MC (results shown in chapter 6), another way to verify that our method does not give false positive results is, rather than separate the signal candidates based on the charge of the slow pion (the "tagged charge"), separate them randomly. Effectively, instead of performing fits to D^0 and $\overline{D}^0$ candidates separately, we perform fits to one randomly selected half of the dataset and the corresponding half. Assigning the tagged charge this way should remove *all* asymmetries, nuisance asymmetries and potentially real CPV, from all parts of the analysis. We expect to see $\mathcal{A}_A^i$, $\mathcal{A}_B^i$, $\mathcal{A}_{bin}^{*i}$, $\mathcal{A}_{avg}$, and $\mathcal{A}_{CP,loc}^i$ all consistent with zero.

As a proxy for randomness, we assign candidates based on whether their event number is odd or even. Event number is a preassigned property of our data and is obviously not random, but importantly there are almost exactly the same number of even event numbers as odd ones, and event number is not correlated with tagged charge. Additionally, using a preassigned variable has a key advantage over generating a new random number, which is that candidates will be assigned to the same subset every time.

Table 8.21 shows $\mathcal{A}_A^i$, $\mathcal{A}_B^i$, $\mathcal{A}_{\mathrm{bin}}^{*i}$, $\mathcal{A}_{\mathrm{avg}}$, and $\mathcal{A}_{CP,\mathrm{loc}}^i$ from performing the analysis on this randomly assigned dataset. As expected, all asymmetries are consistent with zero. Note in particular that here $\mathcal{A}_{\mathrm{avg}} = (-0.50 \pm 0.41)\%$, while in the nominal analysis $\mathcal{A}_{\mathrm{avg}} = (1.60 \pm 0.40)\%$, which is 4σ from zero. $\mathcal{A}_{\mathrm{avg}}$ includes multiple nuisance asymmetries: we would not expect it to be zero under normal circumstances, but randomly assigning the tagged charge should remove all nuisance asymmetries. Since we observe $\mathcal{A}_{\mathrm{avg}}$ and all other asymmetries to be consistent with zero, we conclude that all nuisance asymmetries have been removed, and we take this check as further evidence that our method does not give false positive results.

i	$\mathcal{A}_A^i$ (%)	$\mathcal{A}_B^i$ (%)	$\mathcal{A}_{\mathrm{bin}}^{*i}$ (%)	$\mathcal{A}_{CP,\mathrm{loc}}^i$ (%)
0	-4.5 ± 8.2	7.1 ± 7.8	1.3 ± 5.6	1.8 ± 5.6
1	0.22 ± 0.95	-0.43 ± 0.65	-0.11 ± 0.58	0.39 ± 0.40
2	-9.9 ± 8.0	-0.5 ± 3.4	-5.2 ± 4.3	-4.7 ± 4.3
3	-0.99 ± 0.78	-0.57 ± 0.65	-0.78 ± -0.51	-0.28 ± 0.34
4	-2.8 ± 4.6	-1.2 ± 3.4	-2.0 ± 2.9	-1.5 ± 2.8
5	-0.4 ± 1.2	-0.08 ± 1.55	-0.23 ± 0.98	0.27 ± 0.68

Table 8.21: Results from randomly assigning tagged charge. $\mathcal{A}_{\mathrm{avg}} = (-0.50\pm0.41)\%$. $\mathcal{A}_{CP,\mathrm{loc}}^i$ is calculated by subtracting $\mathcal{A}_{\mathrm{avg}}$ from $\mathcal{A}_{\mathrm{bin}}^{*i}$ for each bin.

8.6.3 Vary yield calculation method

In our nominal analysis, we calculate the signal yield by using the signal normalization parameter from the fit (see subsection 5.2.4 for description of relevant parameters and calculation). Another way to calculate signal yield would be to use the total number of events in the histogram minus the background yield, calculated from the background function in the fit. Table 8.22 shows the results using this variation on the yield calculation method. The central values move almost undetectably from the nominal analysis, indicating our fits are correctly modeling the entire histogram range. The difference in the central values could be taken as a systematic uncertainty, but the difference is negligible.

i	$\mathcal{A}_{CP,\mathrm{loc}}^i$ (%)	
	Nominal result	Result from changing calculation method
0	$\mathbf{8.8 \pm 5.8}$	8.7 ± 5.1
1	$\mathbf{-0.14 \pm 0.40}$	-0.14 ± 0.10
2	$\mathbf{3.6 \pm 4.3}$	3.6 ± 3.3
3	$\mathbf{-1.35 \pm 0.34}$	-1.353 ± 0.082
4	$\mathbf{2.1 \pm 2.8}$	2.1 ± 2.2
5	$\mathbf{0.78 \pm 0.62}$	0.79 ± 0.51

Table 8.22: Comparison of $\mathcal{A}_{CP,\mathrm{loc}}^i$ in nominal results versus using the histogram integral and background function to calculate yield.

8.7   Results with systematic uncertainties

To obtain one number per bin for the systematic uncertainty, we add in quadrature the uncertainties determined in sections 8.2, 8.3, 8.4, and 8.5. Table 8.23 summarizes the sources of systematic uncertainty and the final result, using $\sigma_{[\text{source}]}$ as a shorthand for "the systematic uncertainty on our results associated with a particular source." Table 8.24 presents the results previously shown in Table 7.5 but with full statistical and systematic uncertainties.

i	$\sigma_{\mathcal{A}_{\text{FB}}}$	$\sigma_{\mathcal{A}_{\pi_s}}$	$\sigma_{\mathcal{A}_{D^0}^i}$	σ_{yield}	Total syst. uncert.
0	0.14	0.28	0.26	3.8	**3.8**
1	0.069	0.12	0.10	0.23	**0.29**
2	0.11	0.42	0.43	3.1	**3.2**
3	0.16	0.15	0.21	0.37	**0.48**
4	0.080	0.19	0.16	1.6	**1.8**
5	0.18	0.22	0.17	0.81	**0.88**

Table 8.23: Summary of all sources of systematic uncertainties. The right-most column is the sum in quadrature of the previous four columns.

i	$\mathcal{A}_{\text{CP,loc}}^i$ (%)
0	$\mathbf{8.8 \pm 5.8 \pm 3.8}$
1	$\mathbf{-0.14 \pm 0.40 \pm 0.29}$
2	$\mathbf{3.6 \pm 4.3 \pm 3.2}$
3	$\mathbf{-1.35 \pm 0.34 \pm 0.48}$
4	$\mathbf{2.1 \pm 2.8 \pm 1.8}$
5	$\mathbf{0.78 \pm 0.62 \pm 0.88}$

Table 8.24: Final values for $\mathcal{A}_{CP,\text{loc}}^i$ in each Dalitz bin. The first uncertainty is statistical, the second is systematic. See Table 7.5 for detailed calculations of central values.

Taking into account systematic as well as statistical uncertainties, all results are consistent with zero. We have used conservative estimates for the systematic uncertainties. In most bins, we observe the systematic uncertainty to be lower than the statistical one. Only in bins 3 and 5 is the systematic uncertainty higher. Many of the systematic uncertainties are based on data such that they would decrease with a larger sample size. For example, changes in the central value as a result of modifying the yield calculation method (currently the dominant contribution to the systematic uncertainty) are smaller for bins 1, 3, and 5, i.e., the higher statistics bins. This suggests that in future iterations of this analysis done with higher statistics (and perhaps a different binning choice to maximize the signal-to-background ratio in every bin) the systematic uncertainty in each bin would be lower.

Chapter 9

Summary and Outlook

This final chapter will summarize the main points of this thesis and look ahead to how the analysis we performed could be expanded in the future.

In chapter 1 and chapter 2, we provided an overview of the field of particle physics and its decades-long search for charge-parity violation (CPV) as well as a discussion of how Dalitz analyses have been used in the past. Searching for CPV beyond the Standard Model (SM) is one of the main research goals of the *Belle II* physics program [27], and Dalitz analyses like the one described in this thesis are one avenue for these studies. More detail on Dalitz analyses is provided in Appendix A.

In chapter 3, we described the SuperKEKB accelerator and the *Belle II* detector with which the data used for this analysis was collected. SuperKEKB is an upgrade to the KEKB accelerator, and *Belle II* is similarly an upgrade to *Belle*. The physics program that *Belle II* aims to undertake, including the analysis described in this thesis, relies on the excellent performance of the *Belle II* detector and its many components.

In chapter 4 and chapter 5, we discussed our procedure for event selection on the signal mode $D^0 \to \pi^+\pi^-\pi^0$ and described in detail our analysis methodology. The methodology allows us to calculate local CPV in a particular region of a Dalitz plot while carefully accounting for nuisance asymmetries. This method uses yield fits to allow for asymmetry measurements even in the presence of non-negligible or peaking background. Dalitz analysis techniques used in the past are often dealing with relatively low background levels, so they do not need to use fits to account for them. They are also sometimes cavalier about how nuisance asymmetries, particularly the internal D^0 asymmetry, may bias their measurements. Additionally, many types of Dalitz analyses are unavoidably model-dependent or rely heavily on simulated data for some aspect of their results. Our method is entirely data-driven except for some calculations of systematic uncertainties, and thus it complements the work done by other experiments. Appendix A and Appendix B discuss other types of Dalitz analyses in detail and contrast them with the methodology used in this thesis.

In chapter 6, we tested our analysis methodology on simulated Monte-Carlo data (MC) to verify that we did not see false-positive results. Since *Belle II* MC contains no CPV in the decay $D^0 \to \pi^+\pi^-\pi^0$, all asymmetries in the MC are nuisance asymmetries, and our final result for the asymmetry due to CPV should be consistent with zero. As shown in Table 6.5, all results in MC are consistent with the no-CPV hypothesis, indicating that our method effectively cancels all nuisance asymmetries and does not lead to false positives.

In chapter 7 we performed our analysis on $362\,\mathrm{fb}^{-1}$ of *Belle II* data collected at the $\Upsilon(4S)$ resonance between 2019 and 2022, and in chapter 8 we calculated the systematic uncertainties to include on the results. Table 8.24 shows the final asymmetry results with both statistical and systematic uncertainties included. All asymmetries are consistent with the no-CPV hypothesis.

Given the relatively small dataset we were working with and the fact that other collaborations working with similarly sized datasets have never observed CPV in this mode [38–40], it was very unlikely we would be able to achieve high enough sensitivity to observe CPV with this analysis at this stage in *Belle II* data-taking. However, the purpose of this analysis is not to be the final word on $D^0 \to \pi^+\pi^-\pi^0$ studies at *Belle II*: the purpose is to start the analysis of this mode with early *Belle II* data and to develop a methodology that can be adapted and scaled up as the experiment continues. The goal of the experiment is to collect $50\,\mathrm{ab}^{-1}$ of data by the end of its run [29]; this is a more than 100-fold increase in the size of the data sample compared to this thesis. Both the statistical and systematic uncertainty associated with the final asymmetry measurements in this analysis are expected to decrease as the amount of data increases. Additionally, in Appendix C, we propose some possible techniques for analyses that could supplement the one in this thesis by providing information about average CPV that is not localized to a particular Dalitz bin.

The *Belle II* experiment is still relatively new, having been taking physics data only since 2019. It has published several physics results already, but the fact remains that, for some analyses, more data is required before a reasonable result can be achieved. That is the case for the analysis described in this thesis. A larger dataset could improve the precision on our results considerably, but with the current dataset, we were still able to develop a novel, data-based Dalitz plot analysis method that will scale up as *Belle II* collects more data and be a valuable part of the *Belle II* physics program.

Appendix A

Dalitz Plots

A.1 History

The eponymous R.H. Dalitz pioneered the technique now known as Dalitz plot analysis in the 1950s [31, 32]. Although modern analyses use the invariant mass squared of pairs of decay products as axes, Dalitz's original methodology used a ternary plot with the dimensionless variables x and y. Consider a three-body decay of a particle A with mass m_A. B, C, and D are spin-0 particles of masses m_B, m_C, and m_D, respectively. The kinetic energies of B, C, and D are written T_B, T_C, and T_D. Dalitz then defined x and y as:

$$x = \frac{\sqrt{3}(T_B - T_C)}{E}$$
$$y = \frac{2T_D - T_C - T_B}{E} \tag{A.1}$$

where E is the energy released in the decay. If the rest frame of the parent A is used to calculate the kinetic energies of the daughters, then:

$$T_B = \frac{m_A^2 + m_B^2 - m_{CD}^2}{2m_A} - m_B \tag{A.2}$$

and similarly for T_C and T_D. Equation A.2 is derived from treating the decay $A \to B\ C\ D$ as a two-body decay into B (m_B) and the CD system (m_{CD}).

Equation A.2 is a Lorentz invariant quantity, even though it is calculated in a specific rest frame, and therefore it and Equation A.1 can be used even for relativistic analyses. However, if the invariant masses squared of pairs of decay products are used as axes instead, then it is easier to determine the mass of any intermediate resonance, which is usually a major goal of Dalitz analyses [34]. This is the main reason analysts moved away from Dalitz's original ternary plot and towards plots using invariant pair masses squared as the axes.

Figure A.1 shows the first Dalitz plot, used to model the 3π decay of a kaon, then called a τ meson. Since the kinetic energies of the particles in the decay Dalitz was considering could not exceed $50\,\mathrm{MeV}$, he did not use a relativistic description of the decay. It was Fabri who published the relativistic corrections in 1954 [34, 61]. Researchers continued to use Dalitz's technique mostly to analyze τ meson (kaon) decays for almost two decades after Dalitz's initial paper. Only in the 1970s, did Dalitz techniques become more widely used.
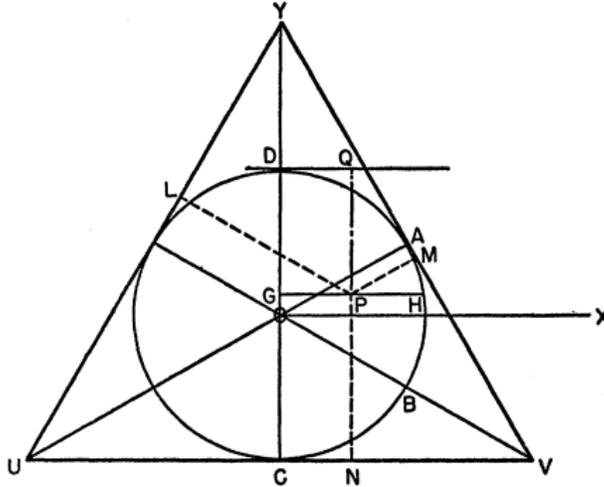
Figure A.1: The first description of a Dalitz plot, using the variables x and y to visualize a kaon (at the time known as a τ meson) decay [31, 32].

A.2 Types of Dalitz analyses

As mentioned in chapter 2, several types of Dalitz analyses are common in studies of multi-body decays. The two "traditional" types of Dalitz analyses are a binned analysis [39] and an amplitude analysis (also called an amplitude model or isobar model) [36, 38, 42, 44, 46]. (Two "non-traditional" types of Dalitz analyses are discussed in Appendix B.)

A binned analysis is in principle straightforward: the Dalitz plot is divided into many, usually $\mathcal{O}(100)$ or greater, bins, and a fractional asymmetry is calculated directly in each bin. Usually bins are *efficiency-corrected* and *background-subtracted* before the fractional asymmetry is calculated. *Efficiency-corrected* means that the reconstruction efficiency across the Dalitz plot has been calculated (using MC) and efficiency effects are removed from (divided out of) the data. *Background-subtracted* means the number of events in each bin has been corrected for background contamination. One way to do this is by looking at a sideband region in the plot of the decaying particle's mass, where only background events are expected to occur, and subtracting the number of events in the sideband from the number of events in the signal region. (This is also aptly called "sideband-subtraction.") Note that this method of correcting for background contamination only works if the background is linear and there is no peaking background underneath the signal that does not occur in the mass sidebands.

An amplitude analyses is not as straightforward as a binned analysis, but it is required if the goal is to measure the amplitude and phase of resonances in the decay. A Dalitz decay typically involves several intermediate resonances. For example, the $D^0 \rightarrow \pi^+\pi^-\pi^0$ decay discussed in this thesis is dominated by three intermediate ρ resonances, where we observe $D^0 \rightarrow \rho(\pi\pi)\pi$ (all possible $\rho\pi$ charge combinations occur) instead of $D^0 \rightarrow \pi^+\pi^-\pi^0$ directly. Other resonances are also present. Rather than calculating an asymmetry in each bin, an amplitude analysis involves performing a 2-D fit to the Dalitz plot and calculating the amplitude and relative phase of each resonance plus a non-resonant component. *Fit*

fractions for the resonances can then be determined. A *fit fraction* is similar to a branching fraction but for a resonance in a Dalitz plot. The fit fraction FF_i for amplitude i is given by

$$FF_i = \frac{\int |c_i A_i|^2 \, d\Phi}{\int |\sum_j c_j A_j|^2 \, d\Phi} \tag{A.3}$$

where c_i is a complex amplitude, A_i contains kinematic dependencies for amplitude i (e.g., Breit-Wigner resonance shapes and spin factors), and $d\Phi$ indicates phase space. Importantly, the sum of the FF_i is not one due to the presence of interference cross-terms in the denominator.

Amplitude analyses have also been used to search for CP asymmetries: a 2-D fit is performed separately to the Dalitz plot of a particle and its charge-conjugate, and the amplitudes and phases of the resonances are compared. A significant difference in either for any individual resonance would be evidence for CPV. However, the unavoidable *model dependence* of this method has been criticized as making it a poor choice for CP asymmetry studies, since using an amplitude model forces the analyst to make particular assumptions about the resonances present in a decay [41]. In addition to model dependence, the commonly-used "sum-of-Breit-Wigners" treatment, often referred to as an *isobar analysis*, violates unitarity. The K-matrix formalism restores two-body unitarity, but model dependence remains [33].

Recently, new types of Dalitz analyses, including a variation on a binned analysis referred to as the "Miranda method" and a novel approach called the "energy test" have been used in several studies [41, 62]. These new methods are discussed in detail in Appendix B.

Appendix B

Miranda Method and Energy Test Discussion

Analyses of $D^0 \to \pi^+\pi^-\pi^0$ in the last fifteen years have made use of two novel techniques for Dalitz analyses: the Miranda method and the energy test. This appendix discusses each of these techniques in more detail, including their advantages and limitations. In section B.3, these techniques are compared to the analysis method used in this thesis.

B.1 Miranda method

The "Miranda method" was introduced by Bediaga *et al.* in a 2009 paper [62]. Taking a lesson from astronomers, the authors proposed analyzing the significance rather than the fractional asymmetry to determine whether a CP asymmetry is present in a Dalitz plot and, if it is, where it occurs. For a bin i with $N(i)$ events of one type and $\bar{N}(i)$ charge-conjugate events, the fractional asymmetry is defined as:

$$\Delta(i) \equiv \frac{N(i) - \bar{N}(i)}{N(i) + \bar{N}(i)} \tag{B.1}$$

and the significance is defined as:

$$S_{CP}(i) \equiv \frac{N(i) - \bar{N}(i)}{\sqrt{N(i) + \bar{N}(i)}} \tag{B.2}$$

where the denominator in Equation B.2 is the standard deviation of the numerator. Equation B.1 is the quantity most traditional binned analyses calculate, and what we use in this thesis. Dividing the Dalitz plot into a large number of bins, limited by the (reasonable but arbitrary) requirement that each bin contain at least 20 entries, $S_{CP}(i)$ is calculated for each bin. The distribution of this variable is plotted for all bins and fit to a Gaussian; if an acceptable fit to a Gaussian cannot be achieved, this is evidence for the presence of CPV. This procedure can be repeated for subdomains of the Dalitz plot to isolate the region in which CPV occurs. Bediaga *et al.* test this procedure on toy MC samples of $B^\pm \to K^\pm\pi^+\pi^-$ and $D^\pm \to \pi^\pm\pi^+\pi^-$ decays, with and without CPV. They demonstrate that this procedure

does not give false positives yet is sensitive to CPV as small as a 3.6° phase difference in the ρ^0 resonance in the decay $D^\pm \to \pi^\pm\pi^+\pi^-$ [62]. Figure B.1 shows these findings. The shape of the $S_{CP}(i)$ distributions is clearly not well-represented by a Gaussian (overlaid on each $S_{CP}(i)$ plot), indicating the presence of CPV.
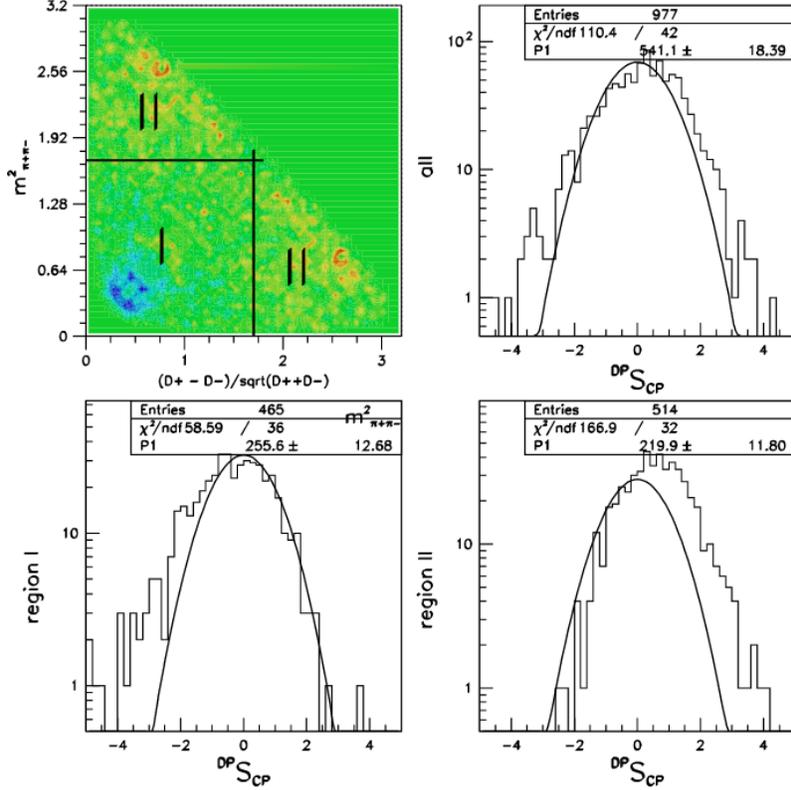
Figure B.1: Dalitz plot of $D^\pm \to \pi^\pm\pi^+\pi^-$ (top left) in a model with a 1% (3.6°) phase difference in the ρ^0 amplitude and $S_{CP}(i)$ distributions for the whole plot (top right), region I (bottom left), and region II (bottom right). This figure is taken from Bediaga *et al.* (2008), and the misplaced x-axis label on the top-left plot is present in the original [62].

The first practical application of the Miranda method was an analysis of $D^0 \to \pi^+\pi^-\pi^0$ and $D^0 \to K^-K^+\pi^0$ done by BaBar in 2008 [38]. The authors used four different methods, including a modified version of the Miranda method, to search for CPV. (The three additional methods are an analysis of angular moments, a traditional amplitude analysis, and a Dalitz-integrated analysis that looks for global CPV.) The Miranda method is modified in two ways. First, a slightly different version of the significance is calculated in order to account for different reconstruction efficiencies across the Dalitz plot as well as for global asymmetries, such as the production asymmetry in the detector. The new formula for the significance is:

$$S_{CP,\text{new}}(i) \equiv \frac{N(i) - \alpha\bar{N}(i)}{\sqrt{N(i) + \alpha^2\bar{N}(i)}} \tag{B.3}$$

where $\alpha = N/\bar{N}$, or the ratio of the *total* number of events type N to the *total* number of charge-conjugate events $\bar{N}$, across the whole Dalitz plot. Second, rather than compare the

distribution of the significances to a Gaussian, the authors calculate a reduced χ^2:

$$\chi^2/\nu = \frac{\sum_{i=1}^{\nu}(S_{CP,\text{new}}(i))^2}{\nu} \tag{B.4}$$

where ν = the number of degrees of freedom, i.e., the number of Dalitz plot bins. (The authors use 1,429 bins for $D^0 \to \pi^+\pi^-\pi^0$ and 726 bins for $D^0 \to K^-K^+\pi^0$.) They then compare the result of Equation B.4 in data to the mean of the distribution of this variable for a large number of MC simulations with no CPV. From testing on MC simulations with CPV, if the data contained a 1% difference in amplitude or a 1° phase change for any of the main Dalitz resonances, they would expect to find a χ^2/ν value about 2σ away from the MC mean. They find no evidence for CPV with this method (or with any of the other three methods they use).

In two papers from LHCb, the Miranda method is applied to $D^+ \to K^-K^+\pi^+$ [47] and $D^+ \to \pi^-\pi^+\pi^+$ [48]. As in the BABAR paper, the authors of both of these papers use MC samples with and without CPV to assess the sensitivity of the method, and they perform the analysis on Cabibbo-favored (CF) control modes and sideband regions in the signal mode to determine whether asymmetries from non-CPV sources (such as a forward-backward production asymmetry) would cause false-positive results. The "mirandized" analysis of $D^+ \to K^-K^+\pi^+$ is performed with four different binning schemes, each expected to be sensitive to CPV occurring in either amplitude or phase differences between the resonances. The analysis of $D^+ \to \pi^-\pi^+\pi^+$ uses the Miranda method with a single binning scheme and compares the results to an unbinned analysis using the kth nearest neighbor technique (this is similar to the energy test method, discussed below). Both the $D^+ \to K^-K^+\pi^+$ and $D^+ \to \pi^-\pi^+\pi^+$ analyses use the same modified $S_{CP,\text{new}}(i)$ value defined in Equation B.3, and for each binning scheme they fit the significances to a Gaussian with zero mean and unit width. All distributions are consistent with a normal Gaussian. They also perform the same reduced χ^2 test used by BABAR and calculate a p-value for each distribution. All p-values are consistent with the no-CPV hypothesis. In the unbinned analysis of $D^+ \to \pi^-\pi^+\pi^+$, results are also consistent with the no-CPV hypothesis.

B.1.1 Advantages

Unlike an amplitude analysis, the Miranda method is model-independent, which is an advantage when the goal is to look for CPV. Additionally, the Miranda method is expected to be more sensitive to CPV than either an amplitude analysis or a traditional binned analysis. The modified version of the Miranda method used by BABAR and LHCb also has the advantage that global nuisance asymmetries can be taken into account without having to calculate them explicitly. Testing the sensitivity of the method and its susceptibility to false positives is also straightforward using MC simulations.

B.1.2 Limitations

Perhaps the most significant limitation of the Miranda method is that although it can identify if CPV is present in a Dalitz decay, it cannot give a number for the CP asymmetry. (One could calculate an asymmetry for a selected region based on where outlier values of

$S_{CP,\text{new}}(i)$ occur, but the choice of region would be biased by the previous results.) The best it can do is, using MC simulations, determine the lower bound for the CPV sensitivity. A separate analysis using a different method would need to be performed if evidence for CPV in a particular region was found. Compounding this problem, MC simulations are not model-independent. Since one of the main advantages of the Miranda method is its model-independence, it is unfortunate that model-dependence is required in order to quantify the level of CPV the analysis could reasonably detect.

Another disadvantage is the lack of detailed consideration for background contamination. If the background levels in the Dalitz plot are negligible or if they are flat and able to be accounted for via sideband-subtraction, then the Miranda method is able to account for them, but there is no provision for non-negligible peaking background that necessitates a fit to determine the signal yield. There is similarly no way to account explicitly for fake asymmetries within the background; the best that can be achieved is to test the method's sensitivity to potential false positives using CF control modes and sideband regions in the signal mode.

Although the modified version of the Miranda method used by BaBar and LHCb can account for fake global asymmetries via the parameter α (see Equation B.3), another limitation of this method is that it cannot account for fake asymmetries that may vary bin-to-bin (such as an internal D^0 asymmetry coming from the decay products' differing momentum spectra). Again, control modes must be used to demonstrate a lack of sensitivity to these potential asymmetries, rather than accounting for them explicitly.

B.2 Energy test

The energy test was proposed by Williams in 2011 as an alternative to traditional Dalitz analyses, including the Miranda method version of a binned analysis [41]. Williams demonstrates that the energy test is expected to be more sensitive to CPV than either a binned analysis or an amplitude model analysis. Table B.1 shows Williams' observation of what fraction of CP-violating datasets (from toy, Dalitz-model MC) are found to exhibit CPV at the one, two, and three σ significance levels using the energy test as compared to a binned χ^2 test. The energy test is demonstrated to be more reliable than a binned χ^2 test at detecting CPV .

test	$1\sigma(\%)$	$2\sigma(\%)$	$3\sigma(\%)$
χ^2	38 ± 5	3 ± 2	0 ± 1
energy	87 ± 3	52 ± 5	13 ± 3

Table B.1: Percentage of CP-violating datasets in which CPV is detected at the one, two, and three σ significance levels with a binned χ^2 test as compared to the energy test [41].

Figure B.2 shows the results from the Miranda method; the distribution of $S_{CP,\text{new}}(i)$ values is consistent with a Gaussian even for Williams' toy models that contain CPV. Only the energy test is sensitive enough to reliably detect CPV at the level present in the toy model ($\sim 2\%$ in terms of the integrated direct CP asymmetry).
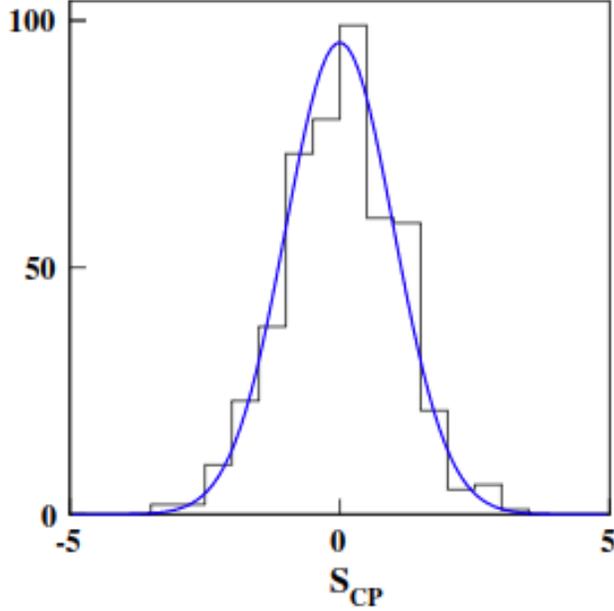
Figure B.2: $S_{CP,\text{new}}(i)$ distribution for the same datasets analyzed with a binned χ^2 and the energy test in Table B.1 [41].

The actual procedure for performing the energy test is similar to the modified Miranda method used by BABAR [38]. In both cases, a test statistic is calculated in data and compared to a distribution of that statistic in an ensemble of MC datasets that do not contain CPV. In the modified Miranda method, that statistic is the reduced χ^2/ν from Equation B.4. In the energy test, the test statistic is:

$$T \approx \frac{1}{n(n-1)} \sum_{i,j>i}^{n} \psi(\Delta \vec{x}_{ij}) + \frac{1}{\bar{n}(\bar{n}-1)} \sum_{i,j>i}^{\bar{n}} \psi(\Delta \vec{x}_{ij}) - \frac{1}{n\bar{n}} \sum_{i,j}^{n,\bar{n}} \psi(\Delta \vec{x}_{ij}) \qquad \text{(B.5)}$$

where $\Delta \vec{x}_{ij} = |\vec{x}_i - \vec{x}_j|$ is the difference between the position of two points in the Dalitz plot[1] and n ($\bar{n}$) is the number of events of one type (the charge conjugate type) over the whole Dalitz plot. This is also similar to a kth nearest neighbor technique, except instead of considering only nearest neighbors, every point in the Dalitz plot is compared to every other point. The $\approx$ in Equation B.5 (as opposed to an $=$) indicates the method is developed for a continuous distribution; summing over a finite number of points, albeit a very large number of finite points, is an approximation. Williams suggests multiple ways to define the weighting function $\psi(\Delta \vec{x}_{ij})$. The one that is later used by LHCb is:

$$\psi(\Delta \vec{x}_{ij}) = \exp(-(\Delta \vec{x}_{ij})^2/2\sigma^2) \qquad \text{(B.6)}$$

where σ is a tunable parameter of the order of the mean distance to the kth nearest neighbor. Equation B.5 compares every point in the Dalitz plot to every other point in the Dalitz plot

---

[1]For a decay into three particles of masses m_1, m_2, and m_3, the two Dalitz plot axes are chosen from m_{12}^2, m_{23}^2, and m_{13}^2. To avoid an arbitrary choice of axis, one may define $\vec{x} = (m_{12}^2, m_{23}^2, m_{13}^2)$ [40].

and to every point in the Dalitz plot of the charge-conjugate decay. It requires no binning and is not model-dependent.

In 2015, LHCb applied the energy test to a Dalitz analysis of $D^0 \to \pi^+\pi^-\pi^0$ [40]. Their results are consistent with the no-CPV hypothesis with a p-value of $(2.6 \pm 0.5)\%$. Their p-value is defined as the fraction of MC simulations that exhibit a T value larger than the one observed in data; a very low p-value would be evidence for CPV. As part of their analysis, the authors compare the sensitivity of the energy test to that of the Miranda method. They find that the relative improvement in sensitivity depends on the selection efficiency across phase space. In regions where the efficiency is especially low, the energy test is comparable to the Miranda method, but in regions where the selection efficiency is higher, the energy test outperforms the Miranda method.

B.2.1 Advantages

The energy test has several advantages over both an amplitude analysis and a binned analysis. Notably, Williams demonstrates that the energy test is expected to be more sensitive to CPV than either of these two more common methods. LHCb confirms this in an actual analysis that used this method, although they stipulate that the degree of improvement depends on the reconstruction efficiency. Like the Miranda method, the energy test is model-independent, which is an advantage when trying to detect CPV. The test statistic given in Equation B.5 does not account for background asymmetries and efficiency differences, but the method can be adapted to account for known asymmetries and differences in detector efficiencies using weighting functions.

B.2.2 Limitations

The energy test has some of the same limitations as the Miranda method, including that while it can predict if CPV is present in a decay, it cannot give a number for the asymmetry and must rely on other methods to fully quantify any observed CPV or to put a limit on the method's sensitivity. However, the major limitation of the energy test is not in the results but in the method itself: since every point in the Dalitz plot must be compared to every other point both in its own Dalitz plot and in the charge-conjugate plot, the energy test is enormously computationally expensive. Additionally, the computing cost scales quadratically with the amount of data. This means that the available computing power plays a larger role in determining whether the energy test is feasible than it does for a binned or amplitude analysis, and it means it may be difficult to scale up this method as an experiment gathers more data. In addition to the computing power required to calculate T, this method requires many, $\mathcal{O}(1000)$, MC simulations be generated in order to get a precise p-value. Generating these simulations and calculating the distribution against which to compare the test statistic in data requires even more computing power. LHCb cites the available computing time as a limiting factor for future, higher-statistics analyses [40].

Another limitation is that the modified method to account for asymmetries and differences in detector efficiencies depends on such differences being known. There may be cases where it is known that there is a fake asymmetry that must be accounted for, but an exact number for this asymmetry is more difficult to determine.

B.3 Discussion

The method used for the analysis described in this thesis, described in detail in chapter 5, is partly inspired by the modified versions of the Miranda method that use the factor α to remove global asymmetries from consideration. Our method also relies on differentiating between local CPV and CPV that is constant across the Dalitz plot (i.e., average or global CPV), rather than between "real" and "fake" CPV explicitly; the fact that most of the "fake" CPV is global makes this distinction very useful. However, there are some key differences between our dataset and the data used by BaBar and LHCb that meant we could not rely on the Miranda method entirely. For one, we observe peaking backgrounds in our signal region which could not be accounted for with sideband subtraction. This means we need to perform a fit to separate the signal from the background. Performing a fit requires more events per bin than simply counting the signal events would, so we cannot use the more than 100 bins that BaBar and LHCb use.

As for the energy test, the limiting factor of available computing power is a serious downside for our purposes. Since the analysis in this thesis is meant to be just the first step in a more complete analysis on this mode with a much larger dataset, an analysis technique that is difficult to scale up is not as useful as one that scales more easily. Additionally, LHCb analyzed the mode we are looking at, $D^0 \to \pi^+\pi^-\pi^0$, with the energy test already, so there is little value added by *Belle II* repeating the same analysis. Instead, we use an entirely different technique to complement LHCb and other studies, rather than duplicate them.

It is worth noting that both the Miranda method and the energy test rely on comparisons to MC to detect CPV. While this technique has some advantages, it may be limited by how well the MC mimics the data. Thus, it is useful to have a method of analysis that does not rely on MC at all to use as a crosscheck. Our analysis uses the MC for background studies and for some systematic uncertainties, but all of the calculations for the central values of the asymmetries and the statistical uncertainties come from the data itself.

Appendix C

Dalitz-Integrated Analysis

The binned Dalitz analysis described in this thesis determines the local CP asymmetry in each bin ($\mathcal{A}^i_{CP,\mathrm{loc}}$), defined as the deviation of the asymmetry present in the bin ($\mathcal{A}^{*i}_{\mathrm{bin}}$) from the average asymmetry across the Dalitz plot ($\mathcal{A}_{\mathrm{avg}}$). This definition removes nuisance asymmetries that are constant across the Dalitz plot, specifically $\mathcal{A}_{\mathrm{FB}}$ and $\mathcal{A}_{\pi_s}$, as well as the bin-specific nuisance asymmetry $\mathcal{A}^i_{D^0}$. However, as a result, this method is insensitive to real CPV that either is antisymmetric across the $m(\pi^\pm\pi^0)^2$ symmetry axis or occurs across the entire phase space of the decay ($\mathcal{A}_{CP,\mathrm{avg}}$). The former issue can be greatly reduced by using a larger number of bins, which would be possible with more data. The latter can be addressed by supplementing the results of the binned analysis with an analysis that *is* able to detect $\mathcal{A}_{CP,\mathrm{avg}}$, and hence $\mathcal{A}^i_{CP}$. It is interesting to explore how a future analysis might accomplish this, and this appendix discusses two such methods. In section C.1, we discuss some preliminary work on these methods using different decay modes. In section C.2, we discuss how to determine $\mathcal{A}_{\pi_s}$ and $\mathcal{A}_{\mathrm{FB}}$ explicitly, which would provide the missing information we need to determine $\mathcal{A}^i_{CP}$ and $\mathcal{A}_{CP,\mathrm{avg}}$. In section C.3, we discuss a modified analysis method that removes the need to calculate $\mathcal{A}_{\mathrm{FB}}$ explicitly.

C.1 Preliminary work

We performed a preliminary analysis on D^*-tagged and untagged $D^0 \to K^-\pi^+$ as part of our investigation into a possible Dalitz-integrated analysis. The dataset used for this analysis was the same $400\,\mathrm{fb}^{-1}$ of MC described in section 4.1. The `basf2` release used to analyze it was `light-2205-abys`, and `ROOT v6.24/06` was used to analyze the final, selected events. Table C.1 (Table C.2) summarizes the cuts used to select events in the tagged (untagged) modes; they are very similar to the cuts described in Table 4.1 with some changes made to account for 1) the presence of a kaon in the final state, 2) the absence of a π^0 in the final state, 3) the absence of a D^{*+} in the untagged decay, and 4) a much lower percentage of events containing multiple candidates. Specifically, considering point 3, in order to maintain consistency between the tagged and untagged modes, we move the center-of-mass (CM) momentum cut from the D^{*+} to the D^0 candidate. Considering point 4, since both the tagged and untagged $K\pi$ decays have a lower percentage of events with multiple candidates than the $\pi^+\pi^-\pi^0$ mode ($\sim$1%, rather than $\sim$18%), we remove all events containing multiple

candidates instead of using a ranking algorithm to determine the best candidate in these events. After the cuts in Table C.1 and Table C.2 are applied, the signal region is defined as $1.852\,\text{GeV}/c^2 < M(K^-\pi^+) < 1.878\,\text{GeV}/c^2$ (for both the tagged and untagged modes) and $144.8\,\text{MeV}/c^2 < \Delta M < 146.2\,\text{MeV}/c^2$ (for the tagged mode only), and only events from the signal regions are retained. Plots in this appendix use truth-tagging and show only true signal events within the signal region.

Particle	Selection criteria
$\pi_s^+,\ \pi^+$	$dr < 0.5\,\text{cm}$ and $|dz| < 2.0\,\text{cm}$
	in CDC acceptance
K^-	$dr < 0.5\,\text{cm}$ and $|dz| < 2.0\,\text{cm}$
	in CDC acceptance
	binaryPID$(K, \pi) > 0.6$
D^0	$1.65\,\text{GeV}/c^2 < M(\pi^+\pi^-\pi^0) < 2.08\,\text{GeV}/c^2$
	$p^* > 2.3\,\text{GeV}/c$
	flightSig > 0 (applied after vertex fit)
D^{*+}	$\Delta M < 0.155\,\text{GeV}/c^2$
Vertex	Perform one fit with TreeFitter on full decay chain:
fitting	with ipConstraint, require chiProb > 0.001
Other	Remove events with multiple candidates

Table C.1: Cuts for a preliminary analysis of $D^{*+} \to D^0(K^-\pi^+)\pi_s^+$ (i.e., D^*-tagged $D^0 \to K^-\pi^+$) in MC only.

Particle	Selection criteria
π^+	$dr < 0.5\,\text{cm}$ and $|dz| < 2.0\,\text{cm}$
	in CDC acceptance
K^-	$dr < 0.5\,\text{cm}$ and $|dz| < 2.0\,\text{cm}$
	in CDC acceptance
	binaryPID$(K, \pi) > 0.6$
D^0	$1.65\,\text{GeV}/c^2 < M(\pi^+\pi^-\pi^0) < 2.08\,\text{GeV}/c^2$
	$p^* > 2.3\,\text{GeV}/c$
	flightSig > 0 (applied after vertex fit)
Vertex	Perform one fit with TreeFitter on full decay chain:
fitting	with ipConstraint, require chiProb > 0.001
Other	Remove events with multiple candidates

Table C.2: Cuts for a preliminary analysis of untagged $D^0 \to K^-\pi^+$ in MC only.

C.2 Determining $\mathcal{A}_{\pi_s}$ and $\mathcal{A}_{\text{FB}}$

The key part of the binned analysis that prevents us from determining $\mathcal{A}_{CP,\text{avg}}$ is that, per Equation 5.4, $\mathcal{A}_{\pi_s}$ and $\mathcal{A}_{\text{FB}}$ are never explicitly determined. Rather, they are demonstrated

to be constant across the Dalitz plot (plus or minus some systematic uncertainty) and then cancelled in the final subtraction. To determine $\mathcal{A}_{CP,\text{avg}}$ with our current analysis, we need to determine both of these nuisance asymmetries explicitly. There are a few possible ways to accomplish this.

C.2.1 Determining $\mathcal{A}_{\pi_s}$ with $D^0 \to K^- \pi^+$

One way to determine $\mathcal{A}_{\pi_s}$ would be to analyze a non-self-conjugate, Cabibbo-favored (CF) D^0 decay where we expect no real CPV to be present. The only asymmetries we expect to see in decays of this type are the nuisance asymmetries due to detector effects. One possible decay we could examine is $D^0 \to K^- \pi^+$. Since this decay is not self-conjugate, we do not need to rely on D^*-tagging to determine the flavor of the D meson. (Because the doubly Cabibbo-suppressed (DCSD) decay $D^0 \to K^+ \pi^-$ also occurs, we do not know for sure 100% of the time that the flavor of of the D^0 meson is correct. Since we are trying to measure an asymmetry rather than a branching fraction, and since the DCSD decay occurs the same amount of time for D^0 as it does for $\overline{D}^0$, this ultimately is irrelevant for this analysis.) We can therefore calculate an asymmetry in untagged $D^0 \to K^- \pi^+$ ($\mathcal{A}_{\text{untag}}^{K\pi}$) and an asymmetry in tagged $D^{*+} \to D^0(K^- \pi^+)\pi_s$ ($\mathcal{A}_{\text{tag}}^{K\pi}$). Since any internal D^0 asymmetry present in this decay will mostly cancel in the subtraction, the difference between these two asymmetries is the slow pion asymmetry:

$$\mathcal{A}_{\pi_s} = \mathcal{A}_{\text{tag}}^{K\pi} - \mathcal{A}_{\text{untag}}^{K\pi}. \tag{C.1}$$

Complicating factors

There are a few complications to this approach which at first glance seems straightforward. For one, asymmetries due to detector effects are dependent on the kinematics of the decay. The spectrum and angular distribution of the D^0 and its decay products in the tagged and untagged modes are not exactly the same. This means that, in particular, $\mathcal{A}_{\text{FB}}$ and the internal D^0 asymmetry may not exactly cancel in the subtraction in Equation C.1. Figure C.1 shows the $\cos \theta_{D^0}^*$ distributions (the asterisk on θ indicates this variable is measure in the CM frame) in the tagged and untagged decays. Clearly these distributions do not agree. Figure C.2 (Figure C.3) shows that the same problem occurs in the angular distribution of the kaon (pion). It is to be expected that these distributions are not the same, since the full signal decay chains for tagged vs. untagged decays are different. However, it is important they be comparable in order for $\mathcal{A}_{\text{FB}}$ and any internal D^0 asymmetry to cancel out in Equation C.1. In theory, one way to address the former would be to account for $\mathcal{A}_{\text{FB}}$ separately in the tagged and untagged modes, rather than relying on the subtraction to cancel it. This is common in many analyses, and it is done by binning the data according to $\cos \theta_{D^0}^*$ and taking an unweighted average of the asymmetries in opposing bins. However, it turns out this method presents some additional complications in this case and is not appropriate for our analysis, as we now discuss.

$\mathcal{A}_{\text{FB}}$ is assumed to be antisymmetric in $\cos \theta_{D^0}^*$, which is why binning in this variable effectively cancels the asymmetry. However, it is important to note that this approach also cancels the antisymmetric part of *any other asymmetry that is present*, including $\mathcal{A}_{\pi_s}$. This presents a problem if we want our results to complement the binned analysis we have

already performed. The asymmetries determined from the binned analysis include both the symmetric and antisymmetric parts, since we have not done any binning in $\cos\theta^*_{D^0}$. This means that simply cancelling $\mathcal{A}_{\mathrm{FB}}$ in the analysis used to calculate $\mathcal{A}_{\pi_s}$ is not sufficient, as this would leave us with only the symmetric part of $\mathcal{A}_{\pi_s}$. We need both the symmetric and anytisymmetric parts to plug into Equation 5.4.
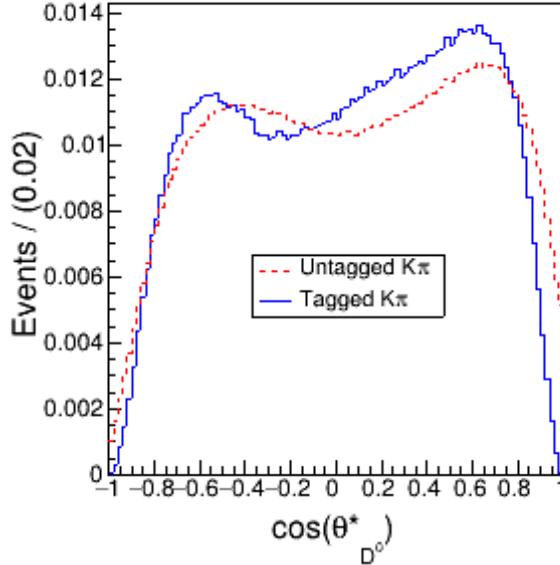

Figure C.1: Comparison of the $\cos\theta^*_{D^0}$ distribution in tagged and untagged $D^0 \to K^-\pi^+$ in truth-tagged signal MC. Both histograms are normalized to unit area, so this is a comparison of the shape of the distributions rather than the efficiency.

In addition to agreement between the tagged and untagged modes, we also have to consider how well the slow pion spectrum in tagged $D^0 \to K^-\pi^+$ agrees with the slow pion spectrum in tagged $D^0 \to \pi^+\pi^-\pi^0$. Since we ultimately want to determine the slow pion asymmetry in the $D^0 \to \pi^+\pi^-\pi^0$ decay, we need to be confident that the slow pion in tagged $D^0 \to K^-\pi^+$ behaves the same way, as the particle kinematics will affect this detector-induced asymmetry. Figure C.4 shows a comparison of the slow pion kinematics in tagged $D^0 \to K^-\pi^+$ and $D^0 \to \pi^+\pi^-\pi^0$. Due to the different momentum cuts and the different number of D^0 decay products in these two decays, it is expected that these distributions will differ. However, it is important they be comparable in order for the value of $\mathcal{A}_{\pi_s}$ determined by analyzing $D^0 \to K^-\pi^+$ to be applicable to our analysis of $D^0 \to \pi^+\pi^-\pi^0$. While the momentum spectra agree fairly well, the angular distribution of the slow pion is noticeably different in the $K^-\pi^+$ and $\pi^+\pi^-\pi^0$ modes. This is likely partly due to the different momentum cuts used in these analyses. In our analysis of $D^0 \to \pi^+\pi^-\pi^0$, there is a cut on $p^*(D^{*+}) > 2.5\,\mathrm{GeV}/c$ (see Table 4.1); however, in tagged $D^0 \to K^-\pi^+$, in order to be consistent with the untagged analysis, we instead make a cut on $p^*(D^0) > 2.3\,\mathrm{GeV}/c$ (see Table C.1 and Table C.2). Since the momenta of the D^{*+} and the D^0 are highly correlated, these cuts are similar, but they are not exactly the same, so we expect to observe some difference in the kinematic distributions for the D^{*+}, D^0, and π^+_s particles.
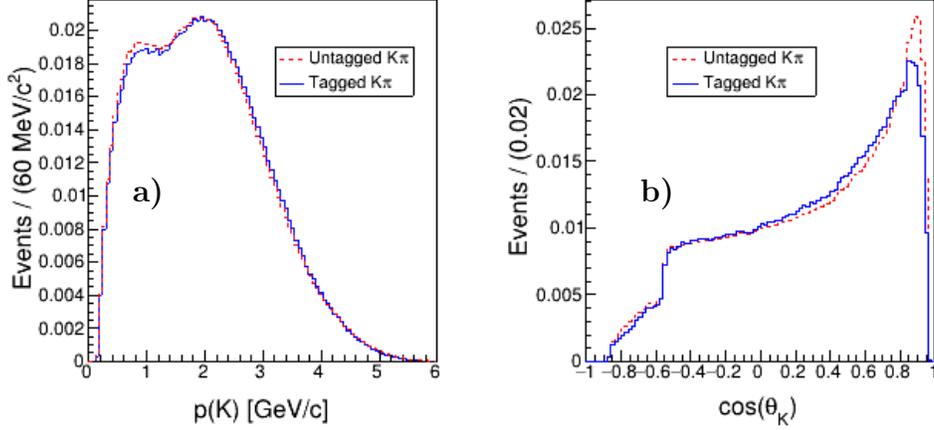
Figure C.2: Comparison of **a)** $p(K)$ and **b)** $\cos\theta_K$ in tagged and untagged $D^0 \to K^-\pi^+$ in truth-tagged signal MC. All histograms are normalized to unit area, so these are comparisons of the shapes of the distributions rather than the efficiencies.
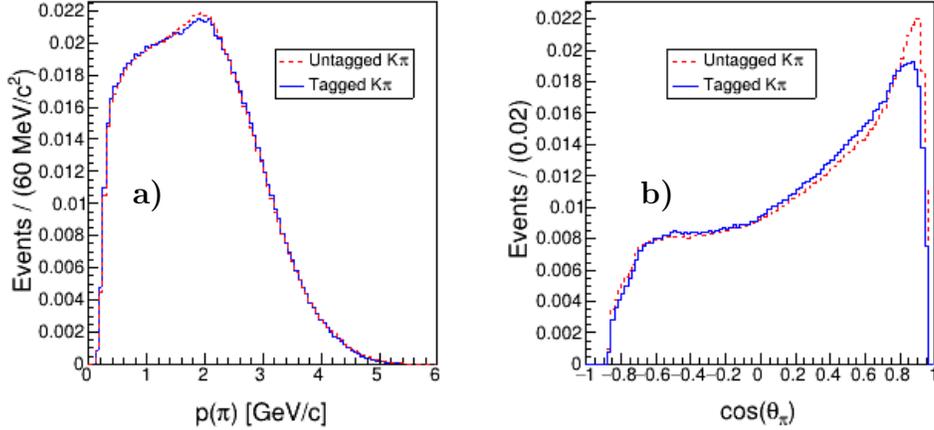
Figure C.3: Comparison of **a)** $p(\pi)$ and **b)** $\cos\theta_\pi$ in tagged and untagged $D^0 \to K^-\pi^+$ in truth-tagged signal MC. All histograms are normalized to unit area, so these are comparisons of the shapes of the distributions rather than the efficiencies.

It is also possible that the presence of a π^0 in the $\pi^+\pi^-\pi^0$ mode plays a role in the discrepancy in Figure C.4. While at the level of production, how the D^0 decays cannot affect the kinematics of $D^{*+} \to D^0\pi_s^+$, in the detection and reconstruction, it is possible for the number and charge of the D^0 decay products to influence the efficiency as a function of the D^0 polar angle (and by extension the π_s^+ polar angle as well).

If key distributions are not in agreement between tagged and untagged $D^0 \to K^-\pi^+$ or between tagged $D^0 \to K^-\pi^+$ and $D^0 \to \pi^+\pi^-\pi^0$, we can weight events to bring the distributions into agreement. The degree to which this is necessary and the accompanying

systematic uncertainties depend on the level of deviation observed and the expected precision of the result.

Other CF decays could also be considered for this analysis. In particular, $D^0 \to K^-\pi^+\pi^0$ is an option if the presence of a π^0 affects relevant kinematic distributions in the tagging decay.
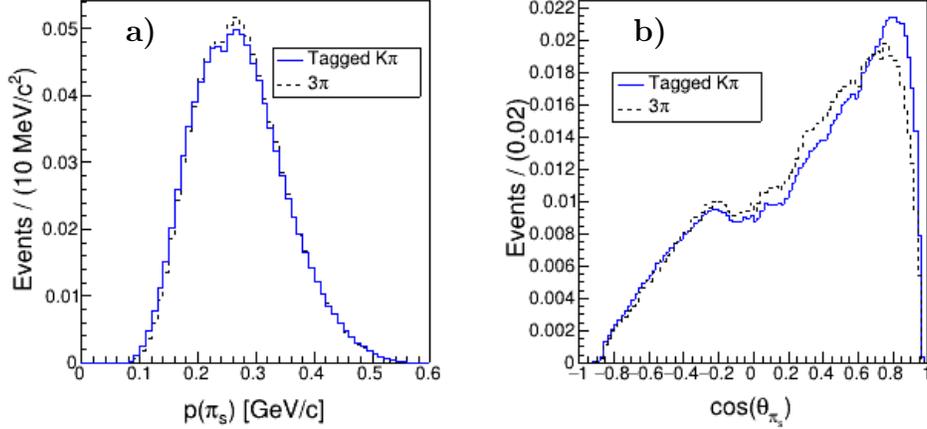

Figure C.4: Comparison of **a)** $p(\pi_s)$ and **b)** $\cos\theta_{\pi_s}$ in tagged $D^0 \to K^-\pi^+$ and $D^0 \to \pi^+\pi^-\pi^0$ in truth-tagged signal MC. All histograms are normalized to unit area, so these are comparisons of the shapes of the distributions rather than the efficiencies.

C.2.2 Determining $\mathcal{A}_{\pi_s}$ with $D^0 \to K^-K^+$

Another approach to determining $\mathcal{A}_{\pi_s}$ is, rather than studying a mode where we expect no real CPV to be present, to study a mode where we do expect CPV to be present but the CPV from real sources is already well-measured. This removes the need to compare a tagged and untagged version of the decay; instead we can determine the raw asymmetry in the tagged decay, subtract off the known CP asymmetry, and what remains will be the slow pion effect. A two-body, self-conjugate decay is best for this, since this type of decay has no internal D^0 asymmetry due to the symmetry of the final state. One possible mode is $D^0 \to K^-K^+$. The most recent and precise measurement of the CP asymmetry in this mode is $\mathcal{A}_{CP}^{KK} = (0.068 \pm 0.054 \pm 0.016)\%$, where the first error is statistical and the second is systematic [63]. This value has an associated error that is much smaller than our expected precision of $\mathcal{O}(1\%)$, so it would be sufficient for our analysis. Defining $\mathcal{A}_{\mathrm{raw}}^{KK}$ as the raw asymmetry in $D^{*+} \to D^0(K^-K^+)\pi_s^+$, we have:

$$\mathcal{A}_{\pi_s} + \mathcal{A}_{\mathrm{FB}} = \mathcal{A}_{\mathrm{raw}}^{KK} - \mathcal{A}_{CP}^{KK}. \tag{C.2}$$

Complicating factors

While using $D^0 \to K^-K^+$ has the advantage that we do not need to compare the kinematics of tagged and untagged samples, it has a similar issue as the $K^-\pi^+$ mode regarding

102

accounting for $\mathcal{A}_{\mathrm{FB}}$. $\mathcal{A}_{\mathrm{FB}}$ does not cancel out in Equation C.2, so it must be either de-termined explicitly or cancelled by binning in $\cos\theta^*_{D^*}$ (we use D^{*+} rather than D^0, since there is no untagged analysis with this method) and averaging. As already mentioned in subsection C.2.1, the latter is incompatible with our existing binned analysis, so this means we must determine $\mathcal{A}_{\mathrm{FB}}$ explicitly if we used $D^0 \to K^-K^+$ to calculate $\mathcal{A}_{\pi_s}$:

$$\mathcal{A}_{\pi_s} = \mathcal{A}^{KK}_{\mathrm{raw}} - \mathcal{A}^{KK}_{CP} - \mathcal{A}_{\mathrm{FB}}. \tag{C.3}$$

In theory this is fine, but since these two measurements now cannot be decoupled, relevant kinematic distributions must agree between tagged $D^0 \to K^-K^+$ and whatever mode is used to determine $\mathcal{A}_{\mathrm{FB}}$.

Additionally, we need to consider imperfect cancellation due to disagreement between key kinematic distributions in $D^0 \to K^-K^+$ compared to $D^0 \to \pi^+\pi^-\pi^0$. Based on the analysis of tagged $D^0 \to K^-\pi^+$, it is likely the distribution of $\cos\theta_{\pi_s}$ for $D^0 \to K^-K^+$ will also not be in agreement with that distribution in the $\pi^+\pi^-\pi^0$ mode. Substantial disagreement between these distributions would again possibly require event weighting so that we can be confident the value of $\mathcal{A}_{\pi_s}$ we determine with Equation C.3 is applicable to our $D^0 \to \pi^+\pi^-\pi^0$ analysis.

C.2.3   Determining $\mathcal{A}_{\mathrm{FB}}$

As alluded to in subsection C.2.1, most analyses at *Belle II* with the goal of measuring an asymmetry never measure $\mathcal{A}_{\mathrm{FB}}$ directly. Instead, they take advantage of the fact that $\mathcal{A}_{\mathrm{FB}}$ is antisymmetric in $\cos\theta^*$ and calculate the raw asymmetry in bins of $\cos\theta^*$. In principle though, it is possible to determine $\mathcal{A}_{\mathrm{FB}}$ explicitly. To do so, one would need to plot the raw asymmetry for a particular mode in bins of $\cos\theta^*$ and fit a line to the resulting distribution. The slope of the line is partially $\mathcal{A}_{\mathrm{FB}}$. However, in order for this to work, one would need non-trivial information about $\mathcal{A}_{\pi_s}$ because it also has a slope with respect to $\cos\theta^*$ (along with higher order contributions) and thus contaminates the measurement of $\mathcal{A}_{\mathrm{FB}}$. Figure C.5 shows the raw asymmetry in truth-tagged signal MC for the decay $D^{*+} \to D^0(K^-\pi^+)\pi^+_s$ in bins of $\cos\theta^*_{D^*}$. Clearly, the shape of this distribution is not well-described by a linear function, and even if it were, the slope of that line would be a combination of $\mathcal{A}_{\pi_s}$ and $\mathcal{A}_{\mathrm{FB}}$. (We expect to see a similar shape in data.) This makes extracting $\mathcal{A}_{\mathrm{FB}}$ by itself more complicated.

C.3   Modified analysis

In the previous section, we discuss determining $\mathcal{A}_{\pi_s}$ and $\mathcal{A}_{\mathrm{FB}}$ explicitly to supplement our binned analysis and allow for determination of $\mathcal{A}_{CP,\mathrm{avg}}$. However, there is a way in which we could avoid having to calculate $\mathcal{A}_{\mathrm{FB}}$. It requires some modification to our already existing analysis.

The need to calculate $\mathcal{A}_{\mathrm{FB}}$ explicitly arises because we do not account for $\mathcal{A}_{\mathrm{FB}}$ in our measurements of $\mathcal{A}^{*i}_{\mathrm{bin}}$ and rather use Equation 5.4 to cancel it. But in principle, we could account for $\mathcal{A}_{\mathrm{FB}}$ the same way we account for $\mathcal{A}^i_{D^0}$: in addition to dividing the Dalitz bins along the $m(\pi^\pm\pi^0)^2$ symmetry axis, we could also divide each bin into regions of $\cos\theta^*_{D^*}$ and average over fits in opposing regions of this variable in addition to over the regions A and B
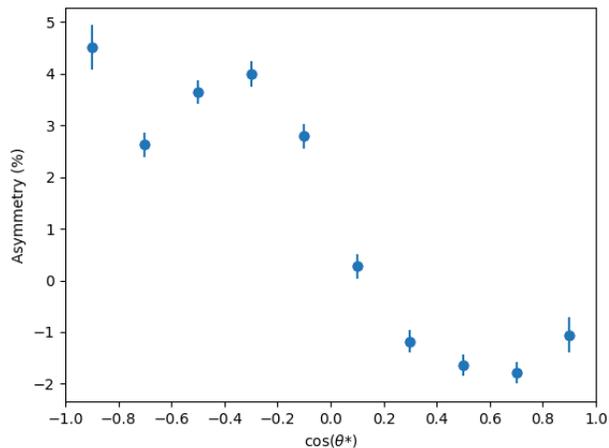
Figure C.5: Asymmetry vs. $\cos\theta^*_{D*}$ bin in tagged $D^0 \to K^-\pi^+$ in signal MC. There are 10 equal-size (width 0.2) bins of $\cos\theta^*_{D*}$.

we already use. Assuming we use 10 $\cos\theta^*_{D*}$ bins, this means that now each Dalitz bin needs to accommodate $10 \times 4 = 40$ separate fits, rather than just the 4 fits we currently perform. This means that substantially more data would need to be collected in order to increase the per-bin statistics before this would be a reasonable modification, but it would be possible in a future version of this analysis.

Another option would be to perform a true Dalitz-integrated analysis that does not rely on dividing the Dalitz plot into bins. In a Dalitz-integrated analysis, dividing the data into regions of $\cos\theta^*_{D*}$ does not overly dilute the statistics. However, the advantage of our binning scheme is to account for $\mathcal{A}^i_{D^0}$; an integrated analysis that did not rely on binning would need another way to account for this nuisance asymmetry, and the result would be independent of (rather than complementary to) the binned analysis.

C.4 Summary

One of the advantages of the binned analysis described in this thesis is that it fully accounts for nuisance asymmetries without having to calculate them directly. However, this feature of the analysis becomes less straightforward when we start asking how to go from $\mathcal{A}^i_{CP,\mathrm{loc}}$ to $\mathcal{A}^i_{CP}$ and $\mathcal{A}_{CP,\mathrm{avg}}$ more generally.

One of the reasons binned Dalitz analyses are attractive is that if we observe CPV in a multi-body decay, we expect the CPV to occur in a particular region of phase space, not across the entire Dalitz plot. Binning the data makes an analysis more sensitive to regional CPV. It is extremely unlikely that CPV is completely independent of location in the Dalitz plot, so an analysis that only captures $\mathcal{A}^i_{CP,\mathrm{loc}}$ is still a valid and useful approach. Determining $\mathcal{A}_{CP,\mathrm{avg}}$ would be nice though, and as presented in this appendix there are various ways future analyses may attempt to do so.

Bibliography

1. Lee, T. D. & Yang, C. N. Question of parity conservation in weak interactions. *Phys. Rev.* **104,** 254 (1956).

2. Wu, C. S., Ambler, E., Hayward, R. W., Hoppes, D. D. & Hudson, R. P. Experimental test of parity conservation in beta decay. *Phys. Rev.* **105,** 1413 (1957).

3. Landau, L. On the conservation laws for weak interactions. *Nucl. Phys.* **3,** 127 (1957).

4. Christenson, J. H., Cronin, J. W., Fitch, V. L. & Turlay, R. Evidence for the 2π decay of the K_2^0 meson. *Phys. Rev. Lett.* **13,** 138 (1964).

5. Bloom, E. D. *et al.* High-energy inelastic $e-p$ scattering at 6° and 10°. *Phys. Rev. Lett.* **23,** 930 (1969).

6. Breidenbach, M. *et al.* Observed behavior of highly inelastic electron-proton scattering. *Phys. Rev. Lett.* **23,** 935 (1969).

7. Cabibbo, N. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.* **10,** 531–533 (12 1963).

8. Kobayashi, M. & Maskawa, T. CP-violation in the renormalizable theory of weak interaction. *Prog. Theor. Phys.* **49,** 652 (1973).

9. Bjørken, B. J. & Glashow, S. L. Elementary particles and SU(4). *Phys. Lett.* **11,** 255 (1964).

10. Glashow, S. L., Iliopoulos, J. & Maiani, L. Weak interactions with lepton-hadron symmetry. *Phys. Rev. D* **2,** 1285 (1970).

11. Aubert, J. J. *et al.* Experimental observation of a heavy particle J. *Phys. Rev. Lett.* **33,** 1404 (1974).

12. Augustin, J. E. *et al.* Discovery of a narrow resonance in e^+e^- annihilation. *Phys. Rev. Lett.* **33,** 1406 (1974).

13. Herb, S. W. *et al.* Observation of a dimuon resonance at 9.5 GeV in 400-GeV proton-nucleus collisions. *Phys. Rev. Lett.* **39,** 252 (1977).

14. Abe, F. *et al.* Observation of top quark production in $\overline{p}p$ collisions with the Collider Detector at Fermilab. *Phys. Rev. Lett.* **74,** 2626 (1995).

15. Aubert, B. *et al.* Observation of CP violation in the B^0 meson system. *Phys. Rev. Lett.* **87,** 091801 (2001).

16. Abe, K. *et al.* Observation of large CP violation in the neutral B meson system. *Phys. Rev. Lett.* **87,** 091802 (2001).

17. Aubert, B. *et al.* Direct CP violating asymmetry in $B^0 \to K^+\pi^-$ decays. *Phys. Rev. Lett.* **93,** 131801 (2004).

18. Chao, Y. *et al.* Evidence for direct CP violation in $B^0 \to K^+\pi^-$ decays. *Phys. Rev. Lett.* **93,** 191802 (2004).

19. Aaij, R. *et al.* Observation of CP violation in $B^\pm \to DK^\pm$ decays. *Phys. Lett. B* **712,** 203 (2012).

20. Aaij, R. *et al.* First observation of CP violation in the decays of B_s^0 mesons. *Phys. Rev. Lett.* **110,** 221601 (2013).

21. Aubert, B. *et al.* Evidence for D^0-$\overline{D}^0$ mixing. *Phys. Rev. Lett.* **98,** 211802 (2007).

22. Starič, M. *et al.* Evidence for D^0-$\overline{D}^0$ mixing. *Phys. Rev. Lett.* **98,** 211803 (2007).

23. Aaltonen, T. *et al.* Evidence for D^0-$\overline{D}^0$ mixing using the CDF II detector. *Phys. Rev. Lett.* **100,** 121802 (2008).

24. Aaij, R. *et al.* Observation of D^0-$\overline{D}^0$ oscillations. *Phys. Rev. Lett.* **110,** 101802 (2013).

25. Aaij, R. *et al.* Observation of CP violation in charm decays. *Phys. Rev. Lett.* **122,** 211803 (2019).

26. Sakharov, A. D. Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe. *JETP Lett.* **5,** 24 (1967).

27. *Belle II* Collaboration. *Super KEKB and Belle II* https://www.belle2.org/project/super_kekb_and_belle_ii/. Accessed: 8 August 2022.

28. Wolfenstein, L. Parametrization of the Kobayashi-Maskawa matrix. *Phys. Rev. Lett.* **51,** 1945 (1983).

29. Kou, E. *et al.* The Belle II Physics Book. *Prog. Theor. Exp. Phys.* **2019.** doi:10.1093/ptep/ptz106 (2019).

30. Abe, T. *et al. Belle II Technical Design Report* (eds Doležal, Z. & Uno, S.) 2010. arXiv:1011.0352.

31. Dalitz, R. H. CXII. On the analysis of τ-meson data and the nature of the τ-meson. *Phil. Mag. Ser.* **44,** 1068 (1953).

32. Dalitz, R. H. Decay of τ mesons of known charge. *Phys. Rev.* **94,** 1046 (1954).

33. Workman, R. L. *et al.* Review of Particle Physics. *PTEP* **2022,** 083C01 (2022).

34. Bevan, A. *et al.* The Physics of the B Factories. *Eur. Phys. J. C* **74,** 3026 (2014).

35. Aubert, B. *et al.* Precise branching ratio measurements of the decays $D^0 \to \pi^-\pi^+\pi^0$ and $D^0 \to K^-K^+\pi^0$ relative to the $D^0 \to K^-\pi^+\pi^0$ decay. *Phys. Rev. D* **74,** 091102 (2006).

36. Aubert, B. *et al.* Measurement of CP violation parameters with a Dalitz plot analysis of $B^\pm \to D_{\pi^+\pi^-\pi^0}K^\pm$. *Phys. Rev. Lett.* **99,** 251801 (2007).

37. Gaspero, M., Meadows, B., Mishra, K. & Soffer, A. Isospin analysis of D^0 decay to three pions. *Phys. Rev. D* **78,** 014015 (2008).

38. Aubert, B. *et al.* Search for CP violation in neutral D meson Cabibbo-suppressed three-body decays. *Phys. Rev. D* **78,** 051102 (2008).

39. Arinstein, K. *et al.* Measurement of the ratio $\mathcal{B}(D^0 \to \pi^+\pi^-\pi^0)/\mathcal{B}(D^0 \to K^-\pi^+\pi^0)$ and the time-integrated CP asymmetry in $D^0 \to \pi^+\pi^-\pi^0$. *Phys. Lett. B* **662,** 102 (2008).

40. Aaij, R. *et al.* Search for CP violation in $D^0 \to \pi^+\pi^-\pi^0$ decays with the energy test. *Phys. Lett. B* **740,** 158 (2015).

41. Williams, M. Observing CP violation in many-body decays. *Phys. Rev. D* **84,** 054015 (2011).

42. Kopp, S. *et al.* Dalitz analysis of the decay $D^0 \to K^-\pi^+\pi^0$. *Phys. Rev. D* **63,** 092001 (2001).

43. Link, J. *et al.* Dalitz plot analysis of the $D^+ \to K^-\pi^+\pi^+$ decay in the FOCUS experiment. *Phys. Lett. B* **653,** 1 (2007).

44. Aitala, E. M. *et al.* Dalitz plot analysis of the decay $D^+ \to K^-\pi^+\pi^+$ and indication of a low-mass scalar $K\pi$ resonance. *Phys. Rev. Lett.* **89,** 121801 (2002).

45. Aitala, E. M. *et al.* Model-independent measurement of S-wave $K^-\pi^+$ systems using $D^+ \to K\pi\pi$ decays from Fermilab E791. *Phys. Rev. D* **73,** 032004 (2006).

46. Rubin, P. *et al.* Search for CP violation in the Dalitz-plot analysis of $D^\pm \to K^+K^-\pi^\pm$. *Phys. Rev. D* **78,** 072003 (2008).

47. Aaij, R. *et al.* Search for CP violation in $D^+ \to K^-K^+\pi^+$ decays. *Phys. Rev. D* **84,** 112008 (2011).

48. Aaij, R. *et al.* Search for CP violation in the decay $D^+ \to \pi^-\pi^+\pi^+$. *Phys. Lett. B* **728,** 585 (2014).

49. Akai, K., Furukawa, K. & Koiso, H. SuperKEKB collider. *Nucl. Instrum. Methods A* **907,** 188 (2018).

50. KEK. *Electrons and positrons collide for the first time in the SuperKEKB Accelerator* https://www.kek.jp/en/newsroom/2018/04/26/0700/. Apr. 26, 2018.

51. KEK. *SuperKEKB collider achieves the world's highest luminosity* https://www.kek.jp/en/newsroom/2020/06/26/1400/. June 26, 2020.

52. *KEK Roadmap Open Symposium* (Tsukuba, Japan and remote connection, July 6, 2020).

53. Adachi, I., Browder, T., Križan, P., Tanaka, S. & Ushiroda, Y. Detectors for extreme luminosity: Belle II. *Nucl. Instrum. Methods A* **907,** 46 (2018).

54. Briere, R. A. *Belle II Particle ID: dE/dx from the CDC* http://www-hep.phys.cmu.edu/users/rbriere/B2_dEdx/B2_SummerSchool_dEdx.pdf. July 16, 2021.

55. Tavernier, S. *Experimental techniques in nuclear and particle physics* (Springer, 2010).

56. Kuhr, T., Pulvermacher, C., Ritter, M., Hauth, T. & Braun, N. The Belle II Core Software. *Comput. Softw. Big Sci.* **3,** 1 (2019).

57. Hulsbergen, W. D. Decay chain fitting with a Kalman filter. *Nucl. Instrum. Methods A* **552,** 566 (2005).

58. Krohn, J.-F. *et al.* Global decay chain vertex fitting at Belle II. *Nucl. Instrum. Methods A* **976,** 164269 (2020).

59. Johnson, N. L. Systems of frequency curves generated by methods of translation. *Biometrika* **36,** 149 (1949).

60. Johnson, N. L. Bivariate distributions based on simple translation systems. *Biometrika* **36,** 297 (1949).

61. Fabri, E. A study of tau-meson decay. *Nuovo Cim.* **11,** 479 (1954).

62. Bediaga, I. *et al.* On a CP anisotropy measurement in the Dalitz plot. *Phys. Rev. D* **80,** 096006 (2009).

63. LHCb Collaboration. Measurement of the time-integrated CP asymmetry in $D^0 \to K^- K^+$ decays. `arXiv:2209.03179` (2022).