



Fakultät für Physik

Development of a track-fitting toolkit  
and analysis of  $D^0 \rightarrow \pi^+ \pi^- \pi^+ \pi^-$  at  
Belle

Johannes Rauch

Vollständiger Abdruck der von der Fakultät für Physik der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigten Dissertation.

Vorsitzender: apl. Prof. Dr. Norbert Kaiser

Prüfende der Dissertation:

1. Prof. Dr. Stephan Paul
2. Prof. Dr. Thomas Kuhr

Die Dissertation wurde am 20.11.2020 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 18.02.2021 angenommen.



## Abstract

Modern particle accelerators and detectors like Belle II are complex machines that produce large amounts of data.

To enable the study of scattering events, particle tracks have to be composed and reconstructed from many individual measurements. The first part of this thesis covers the development of a toolkit for track fitting, GENFIT. This open-source software can be easily adapted to various experiments, and is, among others, being used in the Belle II and  $\bar{\text{P}}\text{ANDA}$  experiments.

Reconstructed track parameters form the starting material for further analyses. In the second part of this thesis, the decay of neutral  $D$  mesons into four charged pions is studied. The decay occurs via many possible intermediary states, which interfere quantum-mechanically. With a partial wave analysis, the relative intensities and phases of the intermediary states can be determined. For that purpose, eligible events are extracted from the Belle data set. The decay with its possible intermediary states is modeled in the isobar formalism. The model parameters are determined with a fit to the data.

## Zusammenfassung

Moderne Teilchenbeschleuniger und Detektoren wie Belle II sind komplexe Maschinen, die eine große Menge an Daten produzieren.

Um die Streuereignisse untersuchen zu können, müssen zunächst die Teilchenspuren aus vielen Einzelmessungen zusammengesetzt und rekonstruiert werden. Der erste Teil dieser Arbeit behandelt die Entwicklung einer Software zu Spurrekonstruktion, GENFIT. Diese quelloffene Software kann einfach für verschiedene Experimente adaptiert werden, und wird u.a. bei Belle II und  $\bar{\text{P}}\text{ANDA}$  eingesetzt.

Rekonstruierte Spurparameter bilden das Ausgangsmaterial für weitere Analysen. Im zweiten Teil dieser Arbeit wird der Zerfall von neutralen  $D$  Mesonen in vier geladene Pionen untersucht. Der Zerfall läuft über viele mögliche Zwischenzustände ab, die quantenmechanisch miteinander interferieren. Mit einer Partialwellenanalyse kann man die relativen Intensitäten und Phasen der Zwischenzustände bestimmen. Dazu werden zunächst geeignete Ereignisse aus dem Datensatz des Belle Experiments extrahiert. Der Zerfall mit seinen möglichen Zwischenzuständen wird im Isobar-Formalismus modelliert. Die Modellparameter werden schließlich mit einem Fit des Modells an die Daten bestimmt.



# Contents

Acronyms and Abbreviations	ix
Motivation	1
<b>I Development of the track-fitting toolkit Genfit</b>	<b>3</b>
1 Motivation for an experiment-independent track-fitting toolkit	5
<b>2 The basics of track fitting</b>	<b>9</b>
2.1 Track parameterization . . . . .	9
2.2 Kalman filter . . . . .	10
2.3 Measurements and observation model . . . . .	12
<b>3 Genfit</b>	<b>13</b>
3.1 The track data structure . . . . .	13
3.1.1 Constructing tracks from track candidates . . . . .	15
3.1.2 Pruning . . . . .	15
3.2 Measurements . . . . .	15
3.2.1 (Virtual) detector planes . . . . .	15
3.3 Track representations . . . . .	16
3.3.1 Runge Kutta track representation . . . . .	17
3.4 Fitting algorithms . . . . .	17
3.4.1 Smoothing . . . . .	18
3.4.2 Handling of multiple measurements in one TrackPoint . . . . .	18
3.4.3 Iterations and convergence criteria . . . . .	19
3.4.4 Kalman fitter . . . . .	19
3.4.5 Kalman fitter with reference track . . . . .	20
3.4.6 Deterministic annealing filter . . . . .	21

<b>4</b>	<b>Performance and visualisation</b>	<b>23</b>
4.1	Performance study . . . . .	23
4.2	Event display . . . . .	24
<b>5</b>	<b>Conclusion and outlook</b>	<b>27</b>
<b>II</b>	<b>Analysis of <math>D^0 \rightarrow \pi^+ \pi^- \pi^+ \pi^-</math></b>	<b>29</b>
<b>6</b>	<b>Introduction</b>	<b>31</b>
<b>7</b>	<b>The Belle Experiment</b>	<b>33</b>
7.1	The KEKB accelerator . . . . .	33
7.2	The Belle detector . . . . .	35
7.3	Particle identification . . . . .	35
7.4	Data sample . . . . .	36
7.5	Monte Carlo samples . . . . .	37
<b>8</b>	<b>Event selection and classification</b>	<b>39</b>
8.1	Flavour tagging . . . . .	39
8.2	Cut-based event selection . . . . .	40
8.2.1	PID and impact parameter cuts . . . . .	40
8.2.2	Mass and vertex cuts . . . . .	40
8.3	Fits of signal and background shapes and yields . . . . .	41
8.3.1	Signal- and background-yield results . . . . .	44
8.4	Boosted Decision Trees . . . . .	47
8.4.1	Input variables . . . . .	49
8.4.2	Results . . . . .	49
8.5	Kaon cut . . . . .	52
<b>9</b>	<b>Partial Wave Analysis</b>	<b>53</b>
9.1	YAP – Yet Another Partial Wave Analysis Toolkit . . . . .	53
9.2	Parameterization of the phase space . . . . .	54
9.3	Isobar formalism . . . . .	54
9.3.1	Admixtures . . . . .	55
9.3.2	Bose-Einstein symmetrization . . . . .	55
9.4	Spin amplitudes . . . . .	55
9.4.1	Clebsch-Gordan coefficients . . . . .	56
9.4.2	Helicity angles . . . . .	56
9.4.3	Wigner $D$ functions . . . . .	58
9.4.4	Helicity spin amplitude . . . . .	58
9.5	Blatt-Weisskopf barrier factors . . . . .	59

## CONTENTS

---

9.6	Mass Shapes . . . . .	59
9.6.1	Breit-Wigner . . . . .	60
9.6.2	Flatté . . . . .	60
9.6.3	Gounaris-Sakurai . . . . .	61
9.6.4	$a_1$ . . . . .	63
9.6.5	$(\pi\pi)_S$ . . . . .	63
9.7	Maximum likelihood fit . . . . .	66
9.7.1	Likelihood . . . . .	66
9.7.2	Integration . . . . .	67
9.7.3	Fit fractions . . . . .	68
9.7.4	Goodness of fit . . . . .	68
9.8	Overlap integrals . . . . .	69
9.9	Model selection . . . . .	69
<b>10</b>	<b>PWA fit results</b>	<b>71</b>
10.1	Phase space plots . . . . .	71
10.2	Model components . . . . .	74
10.3	Fits . . . . .	74
10.3.1	Fits without model selection . . . . .	74
10.3.2	Fits with model selection . . . . .	77
10.4	Fit results . . . . .	77
10.4.1	Fit results without model selection . . . . .	77
10.4.2	Fit results with model selection . . . . .	78
10.5	Uncertainties . . . . .	82
10.5.1	Statistical uncertainties . . . . .	82
10.5.2	Fits with different BDT cuts . . . . .	84
10.6	Discussion . . . . .	84
10.6.1	Background . . . . .	84
10.6.2	$a_1$ . . . . .	86
10.6.3	$(\pi\pi)_S, (\pi\pi)_{S\text{-like}}$ . . . . .	87
10.6.4	$\rho^0$ . . . . .	87
10.6.5	Heavier resonances . . . . .	87
<b>11</b>	<b>Conclusion and outlook</b>	<b>89</b>
	<b>Own contributions</b>	<b>91</b>
	<b>List of figures</b>	<b>95</b>
	<b>List of tables</b>	<b>97</b>

<b>Bibliography</b>	<b>103</b>
<b>Acknowledgments</b>	<b>105</b>

# Acronyms and Abbreviations

In alphabetical order:

ACC	aerogel threshold Cherenkov counter
BAT	Bayesian Analysis Toolkit
basf	Belle analysis framework
basf2	Belle II analysis framework
BCM	biggest conceivable model
BDT	boosted decision tree
BEBC	Big European Bubble Chamber
CDC	central drift chamber
CKM	Cabibbo-Kobayashi-Maskawa
CM	center of mass
CMS	center-of-mass system
DAF	deterministic annealing filter
DCB	double Crystal Ball
DoF	degrees of freedom
ECL	electromagnetic calorimeter
FWHM	full width at half maximum
GBL	general broken lines

HER	high-energy ring
IP	interaction point
KEK	High Energy Accelerator Research Organisation
KLM	$K_L^0$ and $\mu$ detector
LER	low-energy ring
LINAC	linear accelerator
LMU	Ludwig-Maximilians-Universität München
MC	Monte Carlo
MCMC	Markov chain Monte Carlo
MWPC	multi-wire proportional chamber
NDF	number of degrees of freedom
PDF	probability density function
PID	particle identification
POCA	point of closest approach
PWA	partial wave analysis
RAVE	Reconstruction (of vertices) in Abstract, Versatile Environments
RKTrackRep	Runge Kutta track representation
SVD	silicon vertex detector
TMVA	Toolkit for Multivariate Data Analysis with ROOT
TOF	time-of-flight
TPC	time projection chamber
TrackCand	track candidate
TrackRep	track representation
TUM	Technische Universität München

## ACRONYMS AND ABBREVIATIONS

---

UML	unified modeling language
YAP	Yet Another Partial Wave Analysis Toolkit



# Motivation

The discovery of the electron by J. J. Thomson in 1897 marks the beginning of a new era in physics. He realized that atoms are not at all indivisible, but instead are built from more fundamental building blocks.

From 1908 to 1913, H. Geiger, E. Marsden and E. Rutherford conducted their famous scattering experiments [30]. They involved a particle beam (from an alpha emitter in a lead collimator), a target (a thin gold foil) and a particle detector (a fluorescent screen). By measuring the deflection angles of the alpha particles, they did a basic form of track reconstruction, and by analyzing the scattering pattern of many events, they concluded that the positive charge of the atom must be concentrated in a tiny volume in its center.



Figure 1: Geiger and Rutherford [1].

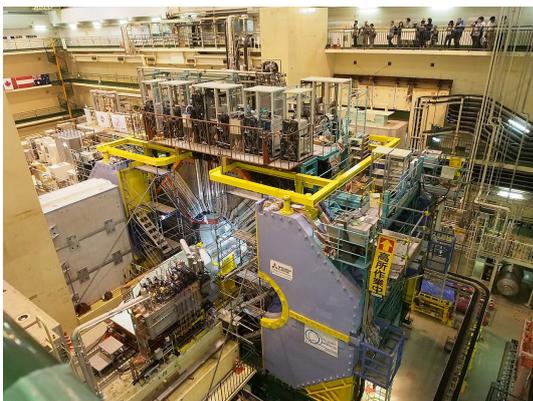


Figure 2: The Belle II detector [2].

Since then, nuclear and particle physics have made an incredible progress. Today, particle accelerators are kilometer-sized machines, detectors are many stories high and produce petabytes of data, which are stored and analyzed in large computing centers. These experiments can only be realized by joint efforts of international collaborations. Our theoretical understanding of the building blocks of matter and their interactions has also come a long way, and is condensed into the Standard Model of particle physics.

Nevertheless, today's experiments still follow the very same basic principles as Rutherford's scattering experiments: Particles (from radioactive decays, cosmic

radiation or accelerators) interact with other particles (in a static target or from another particle beam), and the reaction products are recorded or visualized with particle detectors. From their output, tracks and vertices are reconstructed. The tracks and vertices of many interactions are in turn the basic material for further analysis (e.g. discovering new particles; measuring form factors, cross sections, or branching fractions; partial wave analysis, etc.).

Until the 1970s, track reconstruction was a laborious manual task. For example, ten thousands of photographs of the Big European Bubble Chamber had to be analyzed by dozens of scanners every day. These women measured the curvature of charged tracks, calculated their momenta and reconstructed the decay chains of each event.

Progress in detector technology, signal processing and computing allowed to build a new generation of particle detectors, like multi-wire proportional chambers and solid state detectors. They record electrical signals, which are digitized. Algorithms like the Hough-transform, conformal mapping, and Kalman filters enable a completely automated event reconstruction by computers.

The Kalman filter was first applied for track fitting in the late 80s. It is an algorithm to fit track parameters like position and momentum from a series of noisy measurements. Unlike with a helix fit, tracks in arbitrary magnetic fields and in the presence of material can be fitted precisely.

My thesis is divided into two distinct parts. The first part covers the development of a software framework for track fitting with Kalman filters and related algorithms. This framework, GENFIT, is now being used, among others, in the Belle II and  $\bar{P}$ ANDA experiments.

The second part covers the partial wave analysis of the decay of  $D^0$  mesons (which were first discovered with the Big European Bubble Chamber) into four charged pions, with data from the Belle experiment.

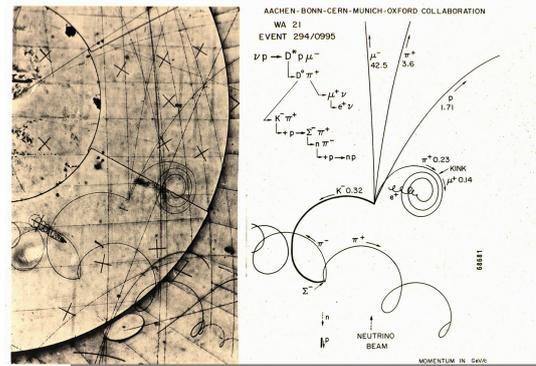


Figure 3: Particle tracks from the Big European Bubble Chamber [3].

# Part I

## Development of the track-fitting toolkit Genfit



# Chapter 1

## Motivation for an experiment-independent track-fitting toolkit

Track reconstruction, or tracking, is the process of extracting track and vertex parameters, like positions and momenta, from detector measurements.

Modern particle detectors like Belle II [16] record voltage signals that are generated by particles passing through them. E.g. an ionizing particle passing a semiconductor detector produces electron-hole pairs. These pairs are separated by an electric field and travel towards electrodes, where they induce a voltage. In a similar manner, signals are produced in a drift chamber, where gas is ionized, producing electrons that accelerate towards a wire and create an avalanche that induces a voltage pulse in the wire.

These signals are amplified, digitized and further processed into measurements of position with uncertainty. A measurement can be one-dimensional, e.g. a line measured by a semiconductor strip detector; two-dimensional, e.g. a position on a semiconductor pixel detector or a drift isochrone of a wire detector;<sup>1</sup> or three-dimensional, e.g. a point in space measured by a time projection chamber (TPC). It can also contain a measurement of the particle's momentum.

Because there are many particles, measurements that are likely to originate from one particle have to be grouped together. This step is called *track finding*. *Track fitting* then reconstructs track parameters and their covariances at any given point along the track from such a group of measurements.

---

<sup>1</sup>A drift isochrone is a surface from where electrons have an equal drift time towards the wire. For a pure radial electric field, the drift isochrone is cylindrical, with a radius proportional to the drift time.

Most high-energy-physics experiments implement their own track fitting programs, using the same standard Kalman filter algorithms. This is laborious, redundant, and error-prone. The high-energy particle physics community greatly benefits from an open-source, modular, and extensible framework for track-fitting and other related tasks, that can easily be adapted to various experimental setups. It is especially useful for smaller experiments, that do not have the manpower to develop their own track fitter, or new experiments, that need a working tool to do research and development.

## Genfit—a track-fitting toolkit

GENFIT [39] is an experiment-independent track-fitting toolkit. It was originally developed for the GEM-TPC project [27] and the  $\bar{\text{P}}\text{ANDA}$  experiment [49]. GENFIT is open-source software and is hosted on GitHub [40].

It is now used in Belle II,  $\bar{\text{P}}\text{ANDA}$ , and other experiments, and is maintained by the Belle II group of the Karlsruhe Institute of Technology. An example of its usage in Belle II is shown in Fig. 1.1.

GENFIT natively supports many detector types like silicon strip and pixel detectors, drift chambers, wire detectors, and TPCs. Tracks can be fitted with different Kalman filter implementations, and a deterministic annealing filter (DAF). The DAF can reject outlier measurements and resolve detector ambiguities, e.g. from multiple hits in strip detectors, or the left-right ambiguities of wire detectors. The track extrapolation can use the full detector geometry to calculate material effects, and propagate through arbitrary magnetic fields. Multiple particle hypotheses can be fitted in parallel. The fitted tracks can easily be stored in ROOT [20] files, and can be visualized with a 3D display. GENFIT provides an interface for detector alignment with the general broken lines (GBL) method [43]; and GFRave, an interface to RAVE.<sup>2</sup> RAVE is a detector-independent toolkit for vertex reconstruction, originally developed for CMS [24]. GFRave takes full advantage of the GENFIT material model, as well as of the sophisticated algorithms of RAVE, allowing for precise and fast vertex reconstruction.

Part of the work shown here has been published in [51]. In the first part of this thesis, I present a revised version of GENFIT. We improved the design and implementation, based on experience gathered in the development of software for the Belle II and  $\bar{\text{P}}\text{ANDA}$  experiments.

---

<sup>2</sup>Reconstruction (of vertices) in Abstract, Versatile Environments[58]

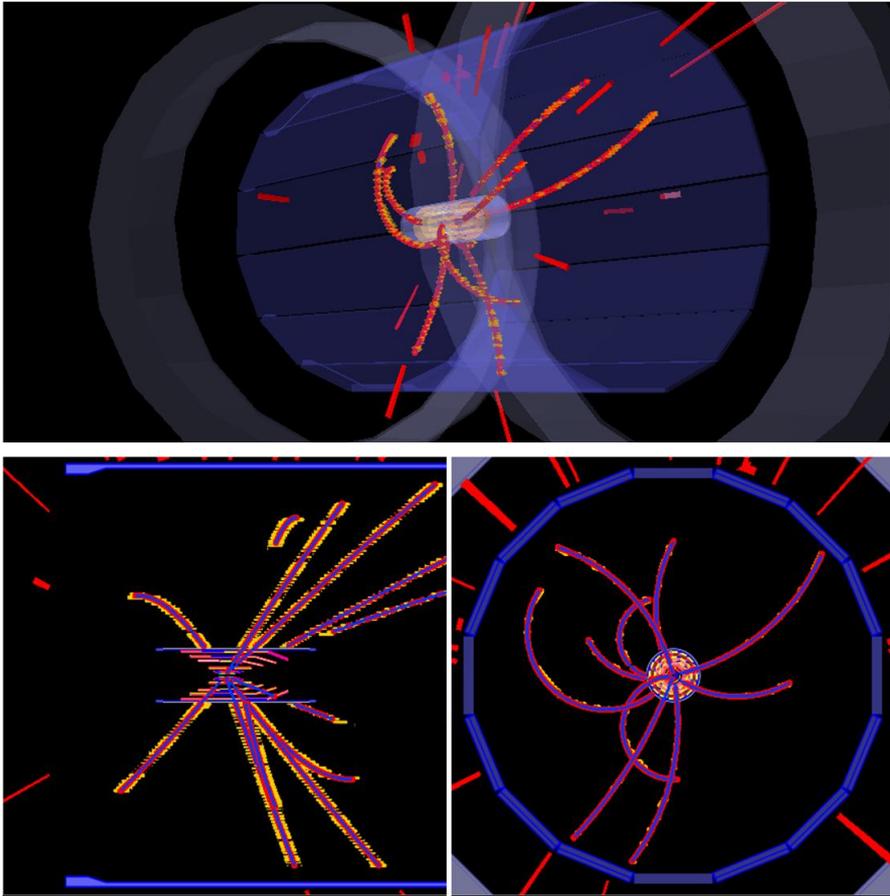


Figure 1.1: Tracking in the Belle II experiment [42]. Measurements (red, with covariance in yellow) are combined into track candidates and fitted with GENFIT (blue).



# Chapter 2

## The basics of track fitting

Before going into the implementation details of GENFIT, I want to outline the basic mathematics of track fitting with a Kalman filter, and how the track states and measurements are parameterized.

### 2.1 Track parameterization

The state of a particle track,  $\mathbf{s}$ , and its corresponding covariance matrix  $\mathbf{C}$ , can be parametrized with 5 coordinates in a local plane coordinate system, as shown in Fig. 2.1. The plane with origin  $\mathbf{o}$  is spanned by two orthonormal vectors  $\mathbf{u}$  and  $\mathbf{v}$ . The track intersects the plane at the Cartesian position  $\mathbf{x}$ , with direction  $\mathbf{a}$ . The state  $\mathbf{s}$  is defined as

$$\mathbf{s} = (q/p, u', v', u, v)^\top \quad (2.1)$$

with

$$\begin{aligned} u' &= \frac{\mathbf{a} \cdot \mathbf{u}}{\mathbf{a} \cdot \mathbf{n}} \\ v' &= \frac{\mathbf{a} \cdot \mathbf{v}}{\mathbf{a} \cdot \mathbf{n}} \\ u &= (\mathbf{x} - \mathbf{o}) \cdot \mathbf{u} \\ v &= (\mathbf{x} - \mathbf{o}) \cdot \mathbf{v}, \end{aligned} \quad (2.2)$$

where  $q/p$  is the particle's charge over its momentum,  $u$  and  $v$  are the Cartesian coordinates of the point  $\mathbf{x}$  in the planar coordinate system, and  $u'$  and  $v'$  are the direction tangens.

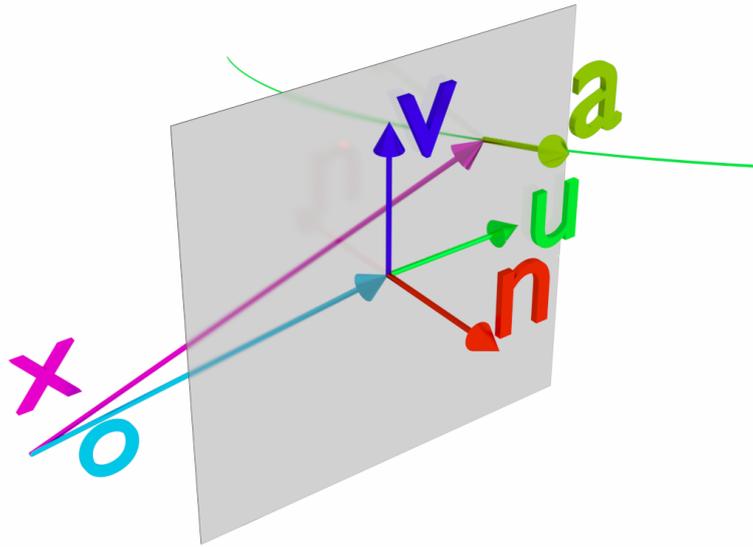


Figure 2.1: Track parametrization. The particle track (lime) with direction  $\mathbf{a}$  intersects the plane (gray) at point  $\mathbf{x}$ . The plane has origin  $\mathbf{o}$  and is spanned by the orthonormal vectors  $\mathbf{u}$  and  $\mathbf{v}$ .

## 2.2 Kalman filter

The Kalman filter [41] is an iterative algorithm that can be used to estimate the state  $\mathbf{s}$  and covariance  $\mathbf{C}$  of a system from a series of measurements  $\mathbf{m}_k$  with covariances  $\mathbf{V}_k$ . The Kalman filter can be applied in real time, and the measurements do not need to describe the full state. The system needs to be modeled, so that its state and covariance can be extrapolated through space or time.

An example application from every-day life is navigation with a car or smartphone. The state is the 2D position- and speed-vector on the map. The device has several sensors, e.g. an accelerometer and a GPS receiver. None of these sensors measure the full state. Integrating over the accelerometer's output would yield a speed vector that drifts over time, while the GPS position would jump around and be updated infrequently.

The Kalman filter solves these issues. Each time we get a new measurement from one of the sensors, we can improve the estimate of the state. First, we have to extrapolate the last known state to the time of the measurement: We multiply the speed vector with the time interval from the last measurement and add it to the position vector. This is the *prediction* step. Then the predicted state is corrected or *updated*. The Kalman filter calculates the optimal state estimation from the predicted state and the measurement (and their covariances). If we could measure the full state of the system, the updated state would simply be the weighted

mean of the state prediction and the measurement. This is generally not the case, though, but the effect is similar.

In particle tracking, the Kalman filter is not applied in real time. Rather, we get a collection of  $n$  measurements from the track finding stage. For each measurement  $\mathbf{m}_k$ , a *prediction* of the state ( $\mathbf{s}_{k|k-1}$ ) at the same position has to be calculated. A new *updated* state  $\mathbf{s}_{k|k}$  is then calculated from measurement and prediction. We have to start with estimates for  $\mathbf{s}_{0|0}$  and  $\mathbf{C}_{0|0}$ , and after repeating this procedure for every measurement, we get the final state  $\mathbf{s}_{n|n}$  with covariance  $\mathbf{C}_{n|n}$ .

Single indices (e.g.  $\mathbf{m}_k$ ) indicate an entity defined for or at measurement  $k$ . Double indices (e.g.  $\mathbf{s}_{k|l}$ ) indicate an entity, that is defined at measurement  $k$ , and takes into account the information of measurements 0 to  $l$ .  $\mathbf{s}_{k|k-1}$  is the state prediction at measurement  $k$ . It takes into account the information of all previous measurements up to  $k - 1$ .  $\mathbf{s}_{k|k}$  is the state update at measurement  $k$ . It takes into account the information of all measurements up to  $k$ .

State predictions and updates are calculated with the following formulae, which are adapted from [32]:

Prediction:

$$\mathbf{s}_{k|k-1} = \mathbf{F}_k \mathbf{s}_{k-1|k-1} + \mathbf{c}_k \quad (2.3)$$

$$\mathbf{C}_{k|k-1} = \mathbf{F}_k \mathbf{C}_{k-1|k-1} \mathbf{F}_k^\top + \mathbf{N}_k.$$

$\mathbf{F}_k$  and  $\mathbf{c}_k$  are the Jacobian and the  $\Delta$  of the state (due to the material effects) for an extrapolation from measurement  $k - 1$  to measurement  $k$ . In our application, they are numerically calculated by extrapolating the state  $\mathbf{s}_{k-1|k-1}$  to the position of the measurement  $\mathbf{m}_k$ , through the magnetic field and detector material. The noise covariance matrix  $\mathbf{N}_k$  describes the added uncertainty due to the material effects (cf. Sec. 3.3).

Update:

$$\begin{aligned} \mathbf{s}_{k|k} &= \mathbf{s}_{k|k-1} + \mathbf{K}_k (\mathbf{m}_k - \mathbf{H}_k \mathbf{s}_{k|k-1}) \\ \mathbf{C}_{k|k} &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{C}_{k|k-1} \end{aligned} \quad (2.4)$$

with

$$\mathbf{K}_k = \mathbf{C}_{k|k-1} \mathbf{H}_k^\top (\mathbf{V}_k + \mathbf{H}_k \mathbf{C}_{k|k-1} \mathbf{H}_k^\top)^{-1}.$$

The prediction is updated with the product of the Kalman gain  $\mathbf{K}_k$  and the residual of the projected state prediction from the measurement  $\mathbf{m}_k$ . The covariance matrix is calculated in a similar way, where  $\mathbf{I}$  is the unit matrix. Since the measurement generally has a lower dimensionality than the state, we need an observation model  $\mathbf{H}_k$ . It is a matrix that projects the track state and covariance into the coordinate system of the measurement. The optimal Kalman gain  $\mathbf{K}_k$  is calculated from the covariance matrices of the state and the measurement.

## 2.3 Measurements and observation model

In the update step of the Kalman filter,  $\mathbf{s}$  and  $\mathbf{C}$  have to be projected into the measurement coordinate system of  $\mathbf{m}$  and  $\mathbf{V}$ . This projection is described by the observation model  $\mathbf{H}$ .

For example, let's take a silicon strip detector, whose strips are parallel to  $\mathbf{u}$ . It measures the coordinate  $v$  in the local plane system:

$$\mathbf{m} = (v). \tag{2.5}$$

The  $\mathbf{H}$  matrix now simply has to project out the  $v$  component of  $\mathbf{s}$ :

$$\mathbf{H} = (0 \ 0 \ 0 \ 0 \ 1). \tag{2.6}$$

Now,  $\mathbf{H}\mathbf{s}$  is the projected state, and  $\mathbf{H}\mathbf{C}\mathbf{H}^T$  is the projected covariance.

For a detector which measures  $u$  and  $v$ , the following applies:

$$\mathbf{m} = (u \ v)^T \tag{2.7}$$

$$\mathbf{H} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.8}$$

# Chapter 3

## Genfit

GENFIT centers around the track data structure, where everything from raw measurements to fit results is stored.

The measurements provide a detector plane, their measurement coordinates and covariance in that plane, and the  $\mathbf{H}$  matrices necessary for the Kalman update step.

Track representations define a particle hypothesis and handle the extrapolation of states and covariances through the detector material and magnetic field.

The fitting algorithms use the measurements and track representations to calculate state predictions and updates, and store them in the track points.

### 3.1 The track data structure

The track object (`Track`, cf. Fig. 3.1) holds a sequence of track points (`TrackPoint`), which can contain any type of data relevant to the track fit. Currently, this can be measurements and thin scatterers.<sup>1</sup> Each track point can also have one fitter information (`FitterInfo`) per track representation, which are filled by the fitter.

A track contains one or more track representations (`TrackRep`), representing the particle hypotheses which should be fitted. One of them has to be selected as the ‘cardinal’ track representation. This can either be done by the user, or by GENFIT, where the track representation is selected, which fits the measurements best.

Furthermore, the track contains general information, like the type of fitter used, the number of fit iterations, if the fit succeed and converged;  $\chi^2$ , number of degrees of freedom (NDF), and  $p$ -value of the fit; track length, etc. (`FitStatus`).

---

<sup>1</sup>The interaction with the material between two measurements can effectively be described by an interaction with a thin layer of material at the position of each measurement. These thin scatterers are used in the GBL fitter.

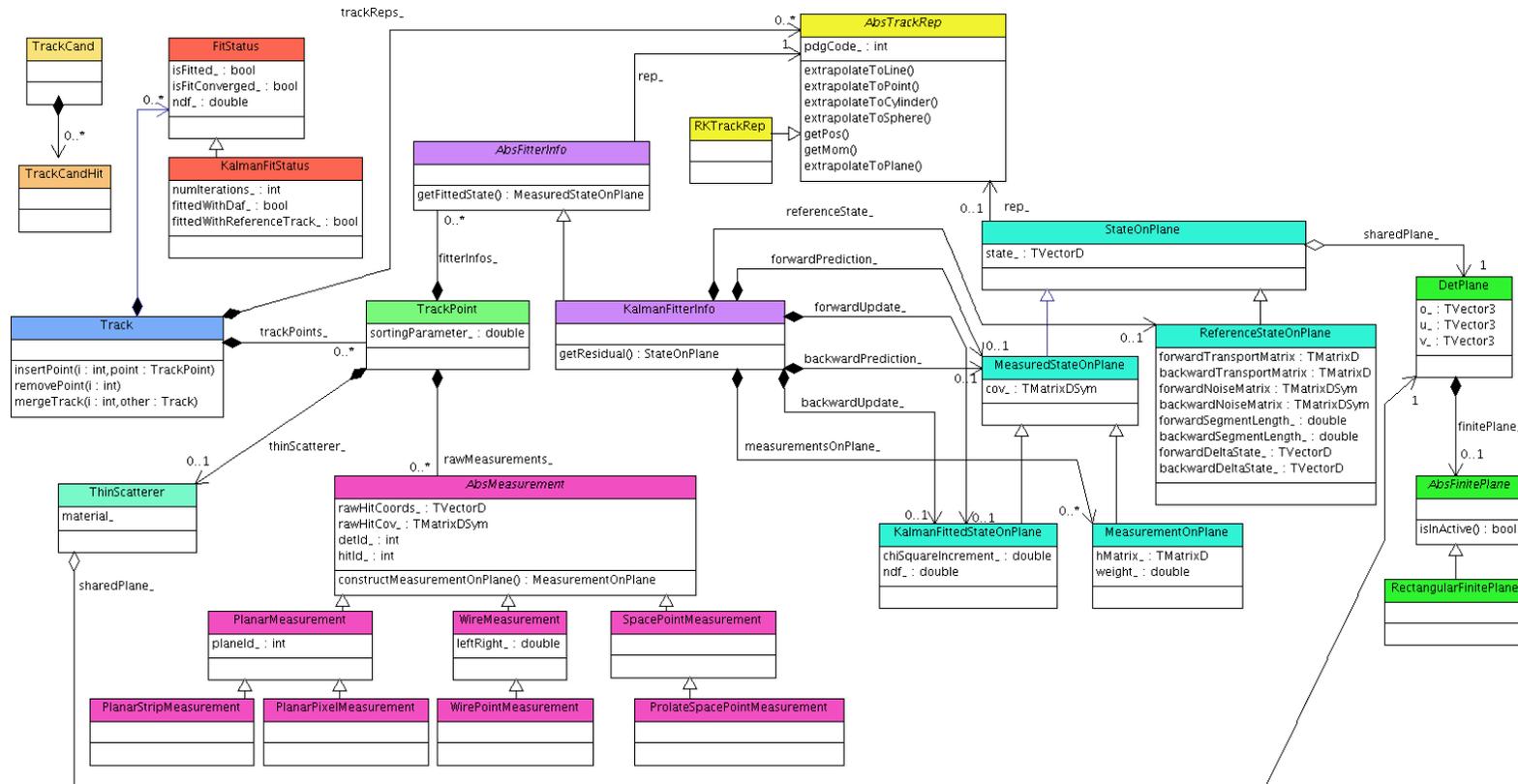


Figure 3.1: GENFIT UML class diagram. This diagram is not complete, it illustrates the inheritance structure and dependencies of the classes related to the Track.

### 3.1.1 Constructing tracks from track candidates

The track candidate (`TrackCand`) is a lightweight helper class for the track finding stage. Indices of raw detector hits are stored in a track candidate hit (`TrackCandHit`). Derived classes can store additional information. E.g. track candidate hits for wire measurements (`WireTrackCandHit`) can store how the left-right ambiguity should be resolved.

Once a track candidate has been composed, it can also be sorted after assigning a sorting parameter to each hit.

A factory class (`MeasurementFactory`) can then be used to build `Track` objects from track candidates.

Alternatively, `Track` objects can be created and filled with measurements directly without the intermediary step of a track candidate.

### 3.1.2 Pruning

After fitting, `Track` objects contain a lot of data: Track representations, track points with measurements, fitter infos, etc. Usually, not all of this information is needed for further processing, or to store fitted tracks in ROOT files. To save memory, the user can decide what data to keep; e.g. only the fitted state of the first track point, and only for the cardinal track representation.

## 3.2 Measurements

A measurement contains measured coordinates from a detector. It provides functions to construct a detector plane and to provide the measurement coordinates  $\mathbf{m}$  and covariance matrix  $\mathbf{V}$  in that plane.

The measurement base class (`AbsMeasurement`) defines the interface. GENFIT comes with predefined measurement classes for various detector types, e.g. planar detectors, drift-chambers, or TPCs.

### 3.2.1 (Virtual) detector planes

In GENFIT, state vectors  $\mathbf{s}$  and measurements  $\mathbf{m}$  are given in local coordinate systems, which are defined by detector planes (implemented in `DetPlane`, cf. Fig. 2.1).

For planar detectors, the detector plane is determined by the detector geometry. For spacepoint- and wire-measurements, so called virtual detector planes have to be constructed by extrapolating the track to the point of closest approach (POCA) to the spacepoint or wire:

- *Extrapolate to point* finds the POCA of the track to a given spacepoint. The virtual detector plane contains the spacepoint and the POCA and is perpendicular to the track.

A *weight matrix*  $\mathbf{G}$  can be used as a metric, defining the space in which the POCA will be calculated. Thereby the weighted distance  $\sqrt{\boldsymbol{\delta}^T \mathbf{G} \boldsymbol{\delta}}$ , where  $\boldsymbol{\delta}$  is the three vector pointing from the spacepoint to the POCA, will be minimized. By default, the inverted 3D-covariance of a spacepoint-measurement is used as a metric, which gives correct fitting results also for spacepoints with arbitrary covariance shapes. The virtual detector plane contains the spacepoint and the POCA, and is perpendicular to the track in the space defined by the metric.

- *Extrapolate to line* finds the POCA of the track to a given line or wire, respectively. The virtual detector plane contains the line and the POCA. This routine is used for fitting wire measurements.

Further arbitrary information can be used, allowing e.g. to correct detector deformations like plane bending, wire sag, and misalignments. Also drift-time corrections are possible.

The intersection of the virtual detector plane with the 3D-covariance gives the 2D-covariance.<sup>2</sup> For wire-measurements, the intersection of the drift isochrone with the detector plane gives two lines, reflecting the left-right ambiguity of such detectors.

### 3.3 Track representations

Track representations combine track parameterization with extrapolation code. The state  $\mathbf{s}$  and the covariance matrix  $\mathbf{C}$  can be extrapolated through material and magnetic field, to detector planes; to the POCA to points, lines, cylinders, and cones; or by a certain distance. The state and covariance can be converted to and from a state defined in the global Cartesian coordinate system,  $(\mathbf{x} \quad \mathbf{p})$ , where  $\mathbf{x}$  is the position, and  $\mathbf{p}$  is the momentum vector.

The track-representation base class (`AbsTrackRep`) defines the interface, and GENFIT comes with the Runge Kutta track representation (`RKTrackRep`) as a concrete implementation.

---

<sup>2</sup>2D measurement positions and covariances are stored in `MeasurementOnPlane` objects.

### 3.3.1 Runge Kutta track representation

The Runge Kutta track representation (`RKTrackRep`) parameterizes the track state as described in Sec. 2.1, and can extrapolate the state through arbitrary magnetic fields and materials. It is based on a Runge Kutta extrapolator from `GEANT 3` [23], which has been ported to `C` by I. Gavrilenko. It is also used in the `PHAST` analysis package [7] of the `COMPASS` experiment [6], which served as a basis for our development.

#### Material handling

The Runge Kutta track representation uses an abstract interface class to interact with the detector geometry. An implementation using `ROOT`'s `TGeoManager` is available. During fitting, material properties are used to calculate the following material effects: Energy loss and energy-loss straggling for charged particles according to the Bethe Bloch formula, multiple scattering (according to [29] or using the Highland formula<sup>3</sup>), where the full noise matrix is calculated, and soft Bremsstrahlung energy loss and energy-loss straggling for electrons and positrons.

#### Adaptive step size calculation

The step sizes used for the Runge Kutta extrapolation should be as large as possible, to save unnecessary computations, while still being small enough to keep errors reasonably small. Therefore, an adaptive step size calculation is done in the Runge Kutta track representation, which takes magnetic-field inhomogeneities and track curvature into account. To calculate material effects correctly, extrapolation stops at material boundaries, and steps can only be so large that a maximum relative momentum loss in the material is not exceeded.

## 3.4 Fitting algorithms

A fitting algorithm uses the measurements and track representations to calculate fit results.<sup>4</sup> Usually a start value is needed for the fit, which has to be provided by the user and can come e.g. from the track finding stage.

Currently, four different track-fitting algorithms are implemented in `GENFIT`: Two extended Kalman fitter implementations, one which linearizes the transport

---

<sup>3</sup>The Highland formula [37] uses a logarithmic term to correct for the non-Gaussian distribution in thin scatterers. Detailed information about the implementation in `GENFIT` can be found in [47]

<sup>4</sup>Fit results are stored in `FitterInfo` objects in the `TrackPoint` objects, and in the `FitStatus` object in the `Track`.

around the state predictions (cf. Sec. 3.4.4) and one which linearizes around a reference track (cf. Sec. 3.4.5). Moreover, a deterministic annealing filter (DAF) (cf. Sec. 3.4.6) and a general broken lines (GBL) fitter [43] are implemented.

Each fitter inherits from an abstract base class (`AbsFitter`) and implements a function (`processTrackWithRep`), which fits the track with the given track representation (i.e. particle hypothesis). The base class can then process all track representations consecutively, starting with the cardinal track representation.

### 3.4.1 Smoothing

Smoothing [31] is a standard technique: The track is fitted in forward and backward direction, where predictions and updates are saved at each track point. With both of these, smoothed track states can be calculated: The weighted average between forward and backward prediction gives the unbiased state, whereas the average between prediction in one direction and update of the opposite direction results in the biased smoothed state. This smoothed states give a better estimate of the true state than either forward or backward updates alone.

Calculating the difference of the measurement and the corresponding smoothed states gives biased and unbiased residuals.

### 3.4.2 Handling of multiple measurements in one Track-Point

Wire detectors measure a drift isochrone (a cylinder in the simplest case). Intersected with the detector plane, this gives two (one dimensional) measurements, since the particle could have passed on either side of the wire. GENFIT provides the possibility to store several measurements of the same type in one track point, mainly for using the DAF to assign weights to them. These tracks can also be fitted with the Kalman fitter, therefore GENFIT provides several options how to handle multiple measurements:

- Average: The average of the individual measurements is calculated. This option is primarily used for the DAF.
- Closest to prediction: The measurement which is closest to the state prediction is selected.
- Closest to reference: The measurement which is closest to the reference track is selected. This can only be used with the Kalman fitter with reference track (cf. Sec. 3.4.5).

- Closest to prediction/reference for wire measurements: If the `TrackPoint` has one `WireMeasurement`, the side that is closest to the prediction or reference is selected; otherwise the average is selected.

The average weighted measurement  $\mathbf{m}$  and its covariance  $\mathbf{V}$  are calculated from the individual measurements  $i$  with weights  $w_i$  as follows:

$$\begin{aligned}\mathbf{m} &= \mathbf{V} \sum_i (w_i \mathbf{V}_i^{-1} \mathbf{m}_i) \\ \mathbf{V} &= \left( \sum_i (w_i \mathbf{V}_i^{-1}) \right)^{-1}.\end{aligned}\tag{3.1}$$

### 3.4.3 Iterations and convergence criteria

In order to reduce bias of the final fit result due to inaccurate start values, usually several fitting passes are made until the fit converges. The last backward update of one iteration is usually taken as start value for the next iteration. In order to minimize bias of the start values, their covariance has to be very large. However, if it is too large, numerical problems can arise. It is practical to multiply the covariance of the start value with a factor of 500 to 1000 between the iterations.

The fitter has a minimum and maximum number of iterations, which are 2 and 4 by default. As soon as the minimum number of iterations have been done, it is checked if the  $p$ -value has changed less than  $1 \times 10^{-3}$  with respect to the previous iteration. However, tracks with a  $p$ -value close to zero are often considered as “converged” with this criterion, while the  $\chi^2$ , albeit big, is still changing significantly, indicating that the fit is still improving. This occurs often for tracks which are given bad start values. To cure this issue, a non-convergence criterion has been introduced: If the relative change in  $\chi^2$  from one iteration to the next is bigger than 20%, the fit will continue.

The number of iterations and convergence criteria values can be adjusted by the user.

### 3.4.4 Kalman fitter

The Kalman fitter implements the formulae described in Sec. 2.2.  $\mathbf{F}_k$  and  $\mathbf{c}_k$  are not explicitly calculated. The predictions are rather obtained by extrapolating the previous update to the current measurement with the track representation.

Our Kalman fitter also supports an implementation based on a square-root formalism, which is adapted from [12]. It provides greater numerical stability at the expense of execution time.

### 3.4.5 Kalman fitter with reference track

Propagating a track state through arbitrary detector material and an inhomogeneous magnetic field is a non-linear problem. However, for the Kalman filter prediction step, the transport has to be linearized, resulting in the Jacobian  $\mathbf{F}$ .

The choice of the linearization point can significantly affect the performance of the track fit. In particular, if the first few measurements lead to a large misestimate of the curvature, the prediction may stray very far from the measurements. As a consequence, the linearization is not optimal, and material- and magnetic-field-lookup will be done at the wrong place. It is therefore common to use an alternative linearization strategy, where the trajectory estimated from the track-finding code is used as initial choice of linearization point for the track fit [32]. At later iterations, the smoothed trajectory from the previous iteration is used as point of linearization. The formulae for calculating the predictions are:

Prediction:

$$\mathbf{s}_{k|k-1} \approx \mathbf{F}_k \mathbf{s}_{k-1|k-1} + \mathbf{c}_k \quad (3.2)$$

$$\mathbf{C}_{k|k-1} \approx \mathbf{F}_k \mathbf{C}_{k-1|k-1} \mathbf{F}_k^T + \mathbf{N}_k$$

Since the predicted state generally does not lie exactly on the reference track,  $\mathbf{s}_{k|k-1}$  and  $\mathbf{C}_{k|k-1}$  can only be approximated. In contrast to the Kalman fitter,  $\mathbf{F}_k$  and  $\mathbf{c}_k$  are stored for each track point (in `ReferenceStateOnPlane`) and describe the reference track:

$$\mathbf{s}_{k,r} = \mathbf{F}_k \mathbf{s}_{k-1,r} + \mathbf{c}_k. \quad (3.3)$$

For the update step, equations 2.4 are used.

In practice, the track is prepared first, i.e. the initial state is extrapolated from measurement to measurement with the track representation, and all  $\mathbf{F}_k$  and  $\mathbf{c}_k$  are calculated. At this stage it is also possible to sort the measurements along the reference track, which can improve fitting accuracy.

For the following iterations, the smoothed states of the previous iteration are taken as new reference states. A  $\chi^2$  value is calculated according to

$$\begin{aligned} \mathbf{r}_k &= \mathbf{s}_{k|n} - \mathbf{s}_{k,r} \\ \chi_k^2 &= \sum_i \frac{\mathbf{r}_{k|n,i}}{\mathbf{V}_{k|n,i,i}}, \end{aligned} \quad (3.4)$$

where  $\mathbf{s}_{k|n}$  is the smoothed state, and  $i$  are the indices of the elements of the state vector. If the  $\chi^2$  is below a certain value (default is 1), the corresponding reference state will not be updated to reduce computing time.

After preparing the track, predictions and updates are calculated with formulae 3.2. The same convergence criteria as for the Kalman fitter are used. Additionally, the fit is converged if the reference states do not change anymore, according to Eq. 3.4.

### 3.4.6 Deterministic annealing filter

The deterministic annealing filter (DAF) is a powerful tool for rejecting outliers and resolving ambiguities of wire measurements or double-sided strip detectors. It is a Kalman filter which uses a weighting procedure between iterations, based on the measurement residuals, to determine the proper weights.

The Kalman fitters weigh all measurements according to their covariances, but otherwise treat them equally. Measurements from noise signals or other particles that were added to the track by the track finder mistakenly can bias the fit. A DAF [33] can lower the weight of these measurements by introducing an annealing scheme: The track is fitted with a Kalman fitter several times, beginning with a high *temperature*, i.e. a factor that the covariances of the measurements are increased by. The user can select between GENFIT's two Kalman implementations and specify the annealing scheme, i.e. a series of decreasing temperatures.

After one fitting pass, the weights of all measurements are updated. The weight depends on the distance to the track (in terms of the measurement's covariance) and looks like a Fermi-Dirac distribution. For small distances, it has a value close to 1, and for large distances the value goes to 0. For high temperatures, the transition from 1 to 0 is smooth, and becomes more step-like for lower temperatures.

The temperature is lowered, and the track is fitted again. In the end, a low temperature of e.g. 0.1 is reached, and the weight function has almost become a step function. Therefore, each measurement has now a weight either close to 1 or 0.

Competing measurements can be stored in one track point, e.g. two or more planar measurements in the same detector.

After the iteration with the last temperature of the annealing scheme, convergence is checked: If the absolute change of all weights is less than  $1 \times 10^{-3}$  (user configurable), the fit is regarded as converged. Otherwise more iterations with the last temperature are done, until the fit converges or a maximum number of iterations is reached.

#### Weight initialization for wire measurements

The DAF is also perfectly suitable to resolve the left-right ambiguities of wire measurements. However, the following problem can occur: The weights of all wire measurements have to be initialized, the basic solution is to initialize both left and

right measurements with a weight of 0.5. Effectively the wire positions are taken as measurements in the first iteration, and their covariance is twice the mean of the individual covariances (cf. Eq. 3.1). This means that all the wire positions now have the same covariance, no matter how far from the actual trajectory they are. This systematic false estimate of the covariances can bias the fit. GENFIT implements a technique to initialize the weight  $w$ , which helps to improve fitting efficiency when using the DAF to resolve left-right ambiguities of wire measurements:

$$w = \frac{1}{2} \left( 1 - \frac{r_{\text{drift}}}{r_{\text{drift, max}}} \right)^2, \quad (3.5)$$

where  $r_{\text{drift, max}}$  is the maximum possible drift radius. Measurements with larger drift radii  $r_{\text{drift}}$  get assigned smaller weights, which in turn lead to bigger covariances, since the wire position is expected to be farther away from the trajectory. In contrast, measurements with small drift radii, which will be closer to the trajectory, get higher weights.

# Chapter 4

## Performance and visualisation

### 4.1 Performance study

GENFIT's performance will depend on many parameters like hardware and compiler, number and types of measurements, momentum, geometry, magnetic field, number of iterations, convergence settings, and so on. Nevertheless, a small study was conducted in the Belle II analysis framework (basf2), to give an idea of the performance one can expect from GENFIT.

The execution time of basf2's `GenFitter` module was measured on an early 2010s office PC with a quad-core 3.4 GHz CPU in single threaded operation. All code was compiled with `-O3` optimization settings.

The module does the fitting and a few more things, like producing the GENFIT track from a track candidate and storing the fitted track in an output array, resulting in an overhead of a little less than 1 ms. Tracks are generated with an azimuth angle of  $100^\circ$  and a momentum of 0.9 GeV in a constant magnetic field. The resulting tracks have 72 track points on average. GENFIT is configured to do three to ten iterations with default convergence criteria.

From the results shown in Tab. 4.1, one can see that the Kalman fitter is always converged after three iterations. One iteration without material effect calculation takes only 0.8 ms. Material lookup takes around 2.2 ms per iteration.

Without material effects, the Kalman fitter with reference track is slower than the Kalman fitter, since the reference track has to be prepared and stored first. However, the material effect calculation takes less time, around 2.0 ms per iteration, since the reference states are not recalculated if they are close to the smoothed states of the previous iteration. This is also the reason why the Kalman fitter with reference track often converges after two iterations. The reference track is already so close to the smoothed track, that it would not change in subsequent iterations (cf. Sec. 3.4.5).

Fitter	w/o matFX	w/ matFX	∅ iterations
Kalman	3.4 ms	10 ms	3
Reference Kalman	4.0 ms	8.2 ms	2.13
DAF	7.0 ms	11 ms	4.03

Table 4.1: Execution time of the GenFitter module in basf2, without and with material effect calculation, and average number of iterations.

If no material effects need to be calculated, the Kalman fitter is the fastest option. To make it even faster, the minimum number of iterations can be lowered to two or even one.

With a complex detector geometry, the material lookup can be the bottleneck. Since the Kalman fitter with reference track needs less material lookup, it can be faster than the Kalman filter, and also give more accurate fit results.

The DAF also uses a reference track. An annealing scheme with four temperatures was used. The calculation of the weights adds a small overhead. Nevertheless, it is not much slower than the other fitters.

## 4.2 Event display

GENFIT features a sophisticated 3D event display, which visualizes fitted tracks (cf. Fig. 4.1). Detector geometry, measurements, detector planes, reference track, forward- and backward-fits (predictions and updates), smoothed track and covariances of measurements and track can be drawn. Tracks can be refitted with the different algorithms and settings, and fit results can be viewed instantly.

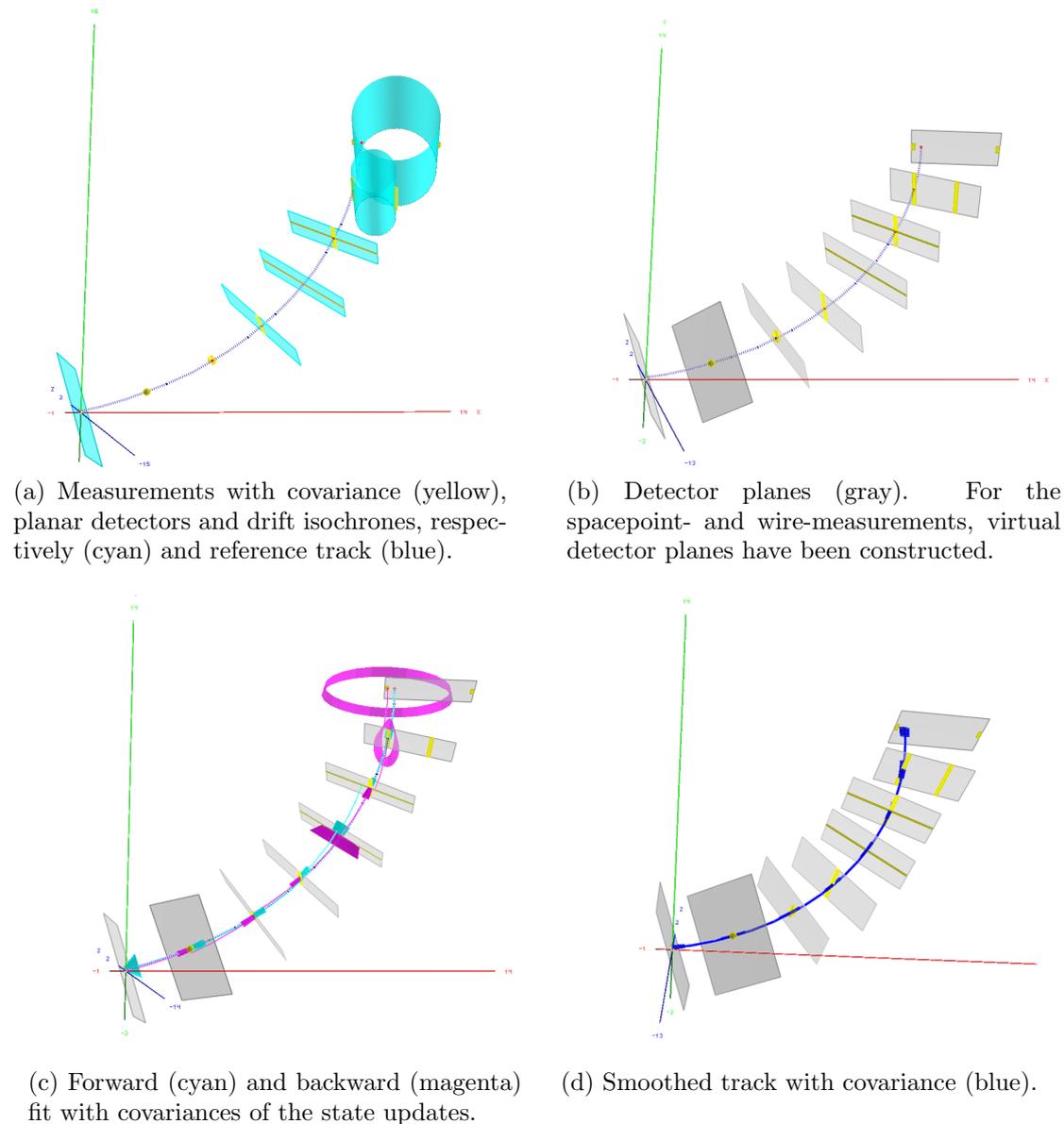


Figure 4.1: GENFIT event display screenshots. The fit of a set of measurements with the Kalman fitter with reference track is shown. For demonstration purposes, the different measurement types supported by GENFIT are used (starting from the origin): Planar pixel measurement, spacepoint measurement, prolate spacepoint measurement, two perpendicular planar strip measurements, double sided planar strip measurement, wire measurement, and wire measurement with second coordinate measurement.



# Chapter 5

## Conclusion and outlook

From the inception of GENFIT, the goal was to develop a modular, experiment-independent framework for track fitting. With the revised version of GENFIT I worked on, these ambitious objectives are accomplished better than ever.

GENFIT is a well tested and proven open-source framework, with an active user and developer community. Our publication [51] has over 30 citations.

GENFIT is the standard track fitter in two large particle physics experiments, Belle II and PANDA. The SπRIT TPC [53] uses GENFIT and RAVE to study heavy nuclear collisions. Two future neutrino experiments, Baby-MIND [13] and DUNE [34], integrated GENFIT into their reconstruction software. GENFIT is even used in the medical field, in the FOOT experiment [11] which aims to improve cancer treatment with hadron therapy.

GENFIT has a firm place in many experiments of all scales and is a vital part to enable new measurements, discoveries and achievements for many years to come.



## Part II

Analysis of  $D^0 \rightarrow \pi^+ \pi^- \pi^+ \pi^-$



# Chapter 6

## Introduction

In this part of my thesis, I analyze the decays of neutral  $D$  mesons into four charged pions. The basic material are the parameters of reconstructed charged pion tracks. They were recorded with the Belle detector, which is described in Chap. 7, and reconstructed with Kalman filters and other techniques similar to those described in Part I.

Particle decays are quantum-mechanical, statistical processes. They can proceed via various intermediate short-lived resonances, until eventually only long-lived particles remain, which are recorded by the tracking detectors. Neutral  $D$  mesons can decay in numerous modes, and in about 0.76% of all decays [48], the final state particles are four charged pions.<sup>1</sup> Considering the conservation laws of the weak and strong interactions, there are many possible intermediary resonances.

These resonances cause interference effects, which materialize in the phase-space distribution of the final state particles. In order to reveal this pattern, we need to consider many decay events. In Chap. 8, I describe how events are selected from the Belle data set.

From the interference pattern we can learn about the intermediary resonances. We need a mathematical description of all possible decay chains and their interference. Then we can fit this model to the data and determine the relative strengths and phases. This technique is called partial wave analysis (PWA) and is described in Chap. 9.

The results of the PWA are presented and discussed in Chap. 10.

---

<sup>1</sup>Pions are not stable particles, but they live long enough to pass the tracking detectors, even at low momenta.



# Chapter 7

## The Belle Experiment

The Belle experiment was located at the High Energy Accelerator Research Organisation (KEK) in Tsukuba, Japan. Between 1999 and 2010, it collected more than  $1 \text{ ab}^{-1}$  of data. Among the most notable results of the experiment is the observation of CP violation in the neutral  $B$  system, which confirmed the Cabibbo-Kobayashi-Maskawa (CKM) theory and yielded the 2008 nobel price for Makoto Kobayashi and Toshihide Maskawa. The experiment consists of the KEKB accelerator and the Belle detector, which is described briefly in the following sections. More detailed descriptions can be found in [9] and [8].

### 7.1 The KEKB accelerator

KEKB is an asymmetric-energy collider (cf. Fig. 7.1). Electrons and positrons are accelerated in a linear accelerator (LINAC) to maximum energies of 8 GeV and 3.5 GeV, respectively. The electrons are inserted into the high-energy ring (HER), the positrons into the low-energy ring (LER). The two storage rings have circumferences of 3016 m and cross at the interaction point in Tsukuba hall. The center-of-mass energy of the electron and positron can be adjusted to different  $\Upsilon$  resonances.

The energy asymmetry leads to a Lorentz boost of the  $B\bar{B}$  system. Due to the good spacial resolution of the silicon vertex detector (SVD), the displacement  $\Delta z$  of the decay vertices of the two  $B$  mesons can be measured and converted into a decay time measurement.

KEKB reached a peak luminosity of  $2.1 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$  in 2009, making it the highest luminosity machine up to date.

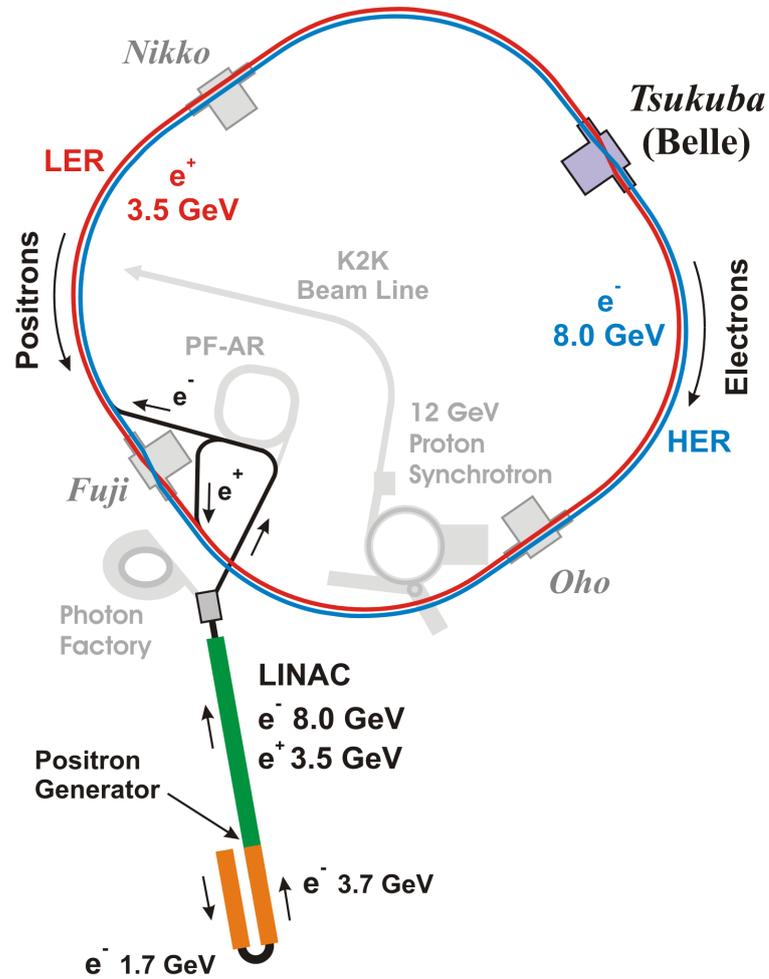


Figure 7.1: The KEKB  $B$  factory [57]. Electrons and positrons are accelerated in a LINAC and stored in the HER and LER. The Belle detector covers the interaction point in Tsukuba hall.

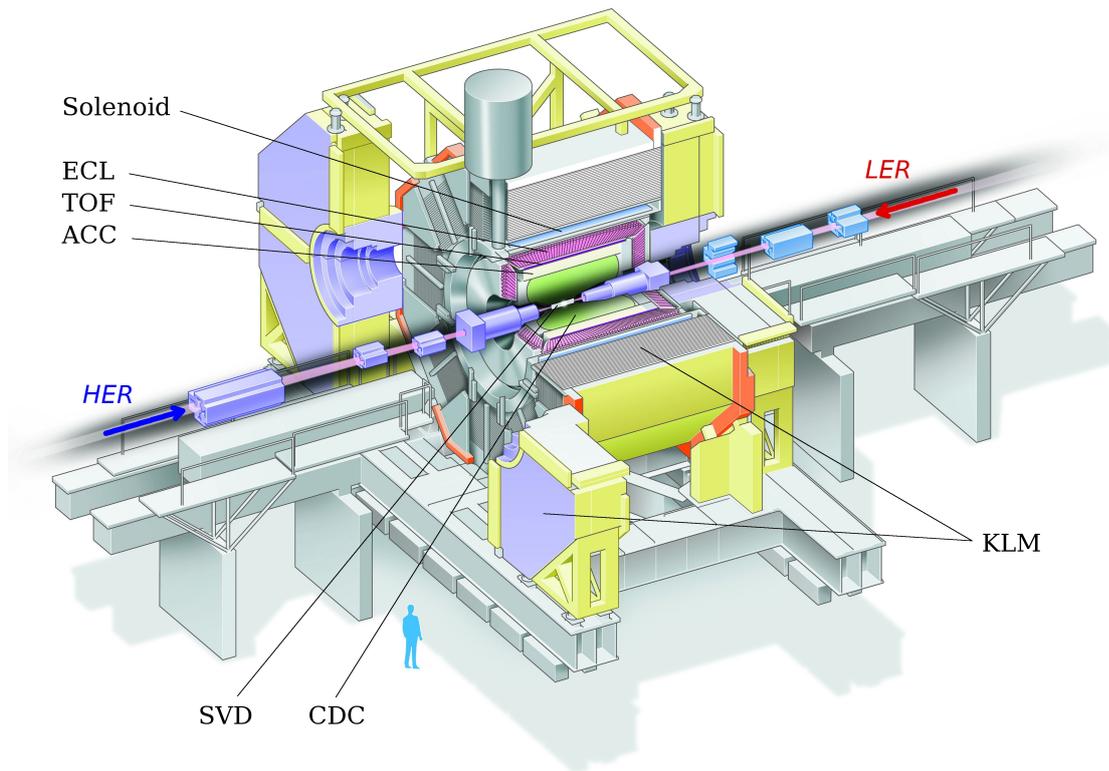


Figure 7.2: Cutaway view of the Belle detector and its subdetectors [59].

## 7.2 The Belle detector

The Belle detector (cf. Fig. 7.2) was located around the collision point in Tsukuba hall. It is a large-solid-angle magnetic spectrometer, covering an azimuthal angle from  $17^\circ$  to  $150^\circ$ . Charged particles follow helical paths in the 1.5 T magnetic field generated by the superconducting solenoid coil. The SVD and a 50-layer central drift chamber (CDC) deliver tracking information. The CDC also measures the energy loss per distance,  $dE/dx$ . An array of aerogel threshold Cherenkov counters (ACCs), time-of-flight (TOF) scintillation counters, and the electromagnetic calorimeter (ECL) are used for particle identification (PID). The  $K_L^0$  and  $\mu$  detector (KLM) is located outside of the solenoid coil.

## 7.3 Particle identification

For the analysis of  $D^0 \rightarrow \pi^+\pi^-\pi^+\pi^-$ , charged pions have to be identified. They have to be distinguished from other long-lived charged particles, i.e. electrons, muons, protons and kaons.

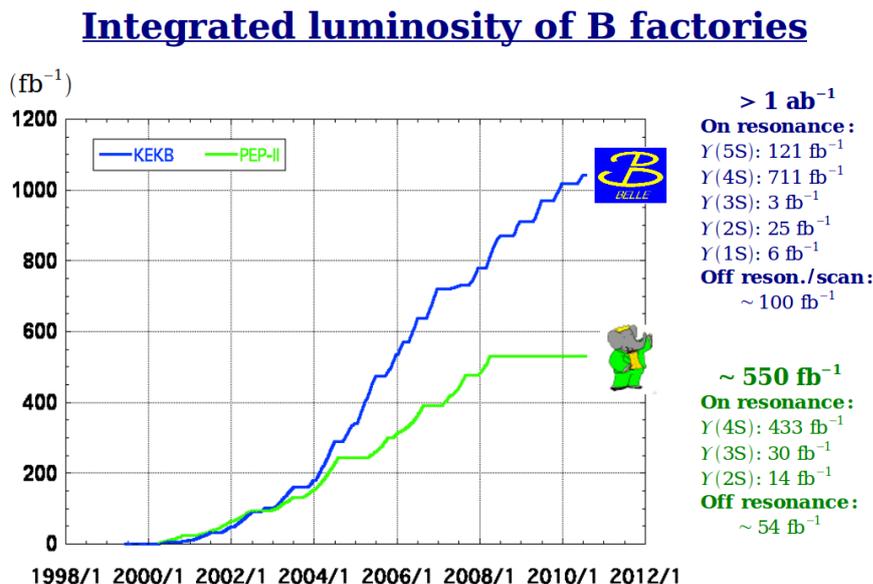


Figure 7.3: Integrated luminosities over time of the  $B$  factories KEKB (Belle), and PEP-II (BaBar) [17].

Electrons are identified by their energy deposit and shower shape in the ECL, a match with a track in the CDC, momentum and  $dE/dx$  measurement in the CDC, light yield in the ACC, and time-of-flight measurement in the TOF.

Muons are identified in a similar way. Due to their larger mass, they do not produce a shower in the ECL, and are tracked by the KLM detectors.

Protons and pions are distinguished by their pulse height in the ACC,  $dE/dx$ , and time of flight.

Charged kaons and pions are distinguished by  $dE/dx$ , time of flight, and the number of photoelectrons produced in the ACC.

## 7.4 Data sample

The integrated luminosity of the data sample recorded by the Belle detector is around  $1 \text{ ab}^{-1}$ .

The major part ( $711 \text{ fb}^{-1}$ ) of the data was recorded at a center of mass (CM) energy of  $10.58 \text{ GeV}$ . This corresponds to the mass of the  $\Upsilon(4S)$  resonance, a bound  $b\bar{b}$  state that decays dominantly into neutral or charged  $B$  meson pairs. Belle and its competitor experiment BaBar are therefore often referred to as the  $B$  factories. However, the cross section for non-resonant  $c\bar{c}$  production at a CM

energy of 10.58 GeV is 1.3 nb, larger than the resonant  $b\bar{b}$  cross section of 1.1 nb. The  $c\bar{c}$  states decay into various  $D$  meson pairs, also including charged  $D^*$  mesons, which are the basis for this analysis.

Another  $155\text{ fb}^{-1}$  of data were recorded at other  $\Upsilon$  resonances, and about  $100\text{ fb}^{-1}$  were recorded off resonance (cf. Fig. 7.3).

## 7.5 Monte Carlo samples

Several streams of simulated data are provided by the Belle collaboration. One stream of Monte Carlo (MC) data resembles the Belle data set and has the same integrated luminosity. A total of six streams are used for this analysis.

Events are generated with EvtGen [44] and Pythia [54], based on known branching fractions and cross sections. Final state radiation is simulated with PHOTOS [15]. The generated particles are propagated through the detector material and magnetic field with Geant3 [23], and the detector response is simulated.

The decay  $D^0 \rightarrow \pi^+\pi^-\pi^+\pi^-$  is simulated as a 4-particle decay with even distribution in the available phase space. The same final state can also be reached indirectly via other decay chains. e.g. via  $K_S^0 K_S^0$  or  $\rho^0 K_S^0$ . However, these decays are not summed coherently, since only one decay channel at a time is randomly picked, based on the given branching fractions.



# Chapter 8

## Event selection and classification

The Belle experiment recorded about one billion events. This corresponds to more than one PB of data, which are stored at the KEK computing center. For the analysis, events with  $D^0$  candidates have to be selected from this data set.

The event selection is done in two stages. The goal of the first stage is to retain as many of the  $D^0 \rightarrow \pi^+\pi^-\pi^+\pi^-$  events as possible, while bringing the amount of data down to a more manageable amount. Since the complete Belle data set has to be skimmed, the selection procedure has to be fast. Therefore, a cut-based event selection is performed at the KEK computing center. It yields 8.6 million events, containing about 0.5 million signal events, and is described in Sec. 8.2. This data sample, about 4 GB in size, is then transferred to Technische Universität München (TUM) for further analysis.

The goal of the second stage is to extract a large and pure sample of signal events. For this purpose, the events are classified with boosted decision trees (BDTs), which are trained with the MC data. A sample of 150 thousand events with more than 90% purity is obtained for PWA. This is described in Sec. 8.4.

### 8.1 Flavour tagging

It was originally planned to search for CP violation in the  $D^0$  decay, which however turned out to be beyond of the scope of this thesis. For such an analysis, it is necessary to determine the flavour of the  $D^0$ , i.e. whether it is a  $D^0$  or  $\bar{D}^0$ . This can be achieved by selecting  $D^0$  candidates from  $D^*$  decays. The charged  $D^*$  decays into a  $D^0$  and a charged pion.

The  $D^*$  has a rest mass of  $2010.26 \pm 0.05$  MeV, the  $D^0$  has a rest mass of  $1864.83 \pm 0.05$  MeV, and the pion has a rest mass of 139.57 MeV [48]. This means only 5.86 MeV of kinetic energy are available in this decay. The pion is therefore called ‘soft pion’. Its charge determines the flavour of the  $D^0$  at the moment of

the  $D^*$  decay:

$$D^{*+} \rightarrow D^0 + \pi^+$$

$$D^{*-} \rightarrow \bar{D}^0 + \pi^-.$$

The selection scheme was kept, and selecting  $D^0$  candidates from  $D^*$  decays also yields more possibilities to suppress background.

## 8.2 Cut-based event selection

The cut-based event selection is developed with the Belle MC sample (cf. Sec. 7.5). The distribution of various variables can be analyzed separately for signal and background events. This allows to choose appropriate variables and cut values, which are described in the following sections.

### 8.2.1 PID and impact parameter cuts

Since the  $D^*$  decays into 5 charged pions, at least 5 charged particles which are likely to be pions are required. The PID system (cf. Sec. 7.3) gives likelihoods of the particle being an electron, a muon, a proton vs. a kaon, and a pion vs. a kaon. Electrons, muons and protons have likelihoods that sharply peak at 100%, so a conservative cut at 99% discards them, and leaves mostly pions and kaons. They are harder to distinguish: In the pion vs. kaon likelihood, the kaons peak at 0%, the pions peak at 100%, but they have a smooth overlap. Particles with a pion likelihood greater than 60% are kept.

The  $D^*$  decays promptly via the strong interaction, and the  $D^0$  has a decay length of about 0.2 mm, so all pions should originate from close to the interaction point (IP). This is achieved by cutting on the impact parameters, which are the distance of the point of closest approach (POCA) of the pion to the IP in radial ( $\Delta r$ ) and in axial direction ( $\Delta z$ ) of the  $z$ -axis, which is in the opposite direction of the  $e^+$  beam. Cut values of 2 cm for  $\Delta r$  and 4 cm for  $\Delta z$  are used to cut away pions from cosmic rays and secondary vertices from long lived resonances.

### 8.2.2 Mass and vertex cuts

The invariant mass is calculated for every possible  $\pi^+\pi^-\pi^+\pi^-$  combination. If it is within a 100 MeV window around the nominal  $D^0$  mass [48]—the full width at half maximum (FWHM) of the  $D^0$  peak is about 27 MeV—an unconstrained (uc) and a mass-constrained (c) vertex fit are performed. If the pions do not originate

from the same point, the  $p$ -value of the vertex fits will be very small. Therefore, a minimal  $p$ -value is required to reject combinatoric background.

$D^*$  candidates are formed by combining the  $D^0$  candidates with the remaining pions, and again the invariant mass is calculated. A cut is performed on  $\Delta m = m_{D^*} - m_{D^0}$ . A vertex fit with a constraint to the IP is done. The  $D^*$  has a very short lifetime and decays promptly. The IP-constraint therefore reduces background and contribution from  $D^*$  which are not directly produced from  $c\bar{c}$  pairs. For further analysis, the quantity  $\sqrt{q} = \sqrt{\Delta m - m_{\pi^\pm}}$ , where  $m_{\pi^\pm}$  is the nominal mass of a charged pion, is defined.

$$x_{D^*} = \frac{|p_{D^*,\text{CMS}}|}{\sqrt{\frac{s}{4} - m_{D^*}^2}} \quad (8.1)$$

is the momentum of the  $D^*$ , in terms of the fraction of its maximum possible momentum in the center-of-mass system (CMS). A cut on  $x_{D^*}$  reduces combinatoric background and also helps to suppress  $D^*$  which do not originate from an initial  $c\bar{c}$  pair, but rather from decays of  $B$  mesons. Preferably selecting  $D^*$  mesons that are directly produced from  $c\bar{c}$  pairs would reduce the bias from possible CP asymmetries in intermediate resonances for the intended CP studies.

The cuts used for the event selection are summarized in Tab. 8.1. The cuts are rather loose in order to keep as much signal events as possible. We also need a window around  $m_{D^0}$  and  $\sqrt{q}$  for the fits of signal and background shapes and yields, which are described in the next section. The training of the BDT, which is performed with the MC data subjected to the same cuts, also needs a significant amount of background events.

A total of 8 592 981 events remain after the event selection.

### 8.3 Fits of signal and background shapes and yields

We now have a selection of 8.6 million events, consisting of the signal and several background components. In order to quantify each component, their distributions are modeled based on the MC sample, and then fitted to the data sample. The signal and background contributions of the MC sample after the cut-based event selection are shown in Fig. 8.1. The one-dimensional distributions can also be seen in Fig. 8.3. The signal shows up as a very clear peak in both  $m_{D^0}^2$  and  $\sqrt{q}$ . There is a slight tail in  $m_{D^0}^2$  towards lower masses, which is caused by additional

PID cuts	
electron	< 0.99
muon	< 0.99
proton-pion	< 0.99
pion-kaon	> 0.6
Impact parameter cuts	
$\Delta z$	< 4 cm
$\Delta r$	< 2 cm
Mass and vertex cuts	
$m_{\pi^+\pi^+\pi^-\pi^-}$	[1815, 1915] MeV
$m_{D^0}$ after vertex fit	[1815, 1915] MeV
$\Delta m = m_{D^*} - m_{D^0}$	[142, 149] MeV
$x_{D^*}$	[0.25, 1]
$p$ -value of $D^0$ and $D^*$ vertex fits	[0.001, 1]

Table 8.1: Event selection cuts.

photons produced in the  $D^0$  decay. The combinatoric background has a very smooth distribution. The  $m$ -peaking background consists of correctly identified  $D^0$  particles, which however have no  $D^*$  as a mother, or have a random pion assigned as the soft pion. The  $\sqrt{q}$ -peaking background consists of correctly identified  $D^*$  particles, which either did not decay via a  $D^0$ , or had the wrong pions assigned to the  $D^0$ . Its amount is negligible.

The binned distribution of signal candidates in  $m_{D^0}^2$  and  $\sqrt{q}$  is fitted with the Bayesian Analysis Toolkit (BAT) [21]. The shapes are selected to match the distributions obtained from the MC sample.

The signal shape is modeled as the product of a double Crystal Ball (DCB) function in each dimension. The DCB function has a Gaussian core and power-law tails. When normalized, it has 6 free parameters: The mode (i.e.  $m_{D^0}^2$  and  $\sqrt{q}$ ), the width of the Gaussian, and for both sides the position where the power law functions take over, and their exponents. A bivariate Gaussian is added to this 2D-DCB function. It has 5 parameters: Two modes and widths and the correlation coefficient. Also the weight with respect to the 2D-DCB function is fitted.

The sum of these two functions is normalized and multiplied with the signal yield, which is also fitted.

The combinatoric background is fitted with a linear function in  $m_{D^0}^2$  and a quadratic function in  $\sqrt{q}$ . Again, they are normalized and multiplied with the

### 8.3 FITS OF SIGNAL AND BACKGROUND SHAPES AND YIELDS

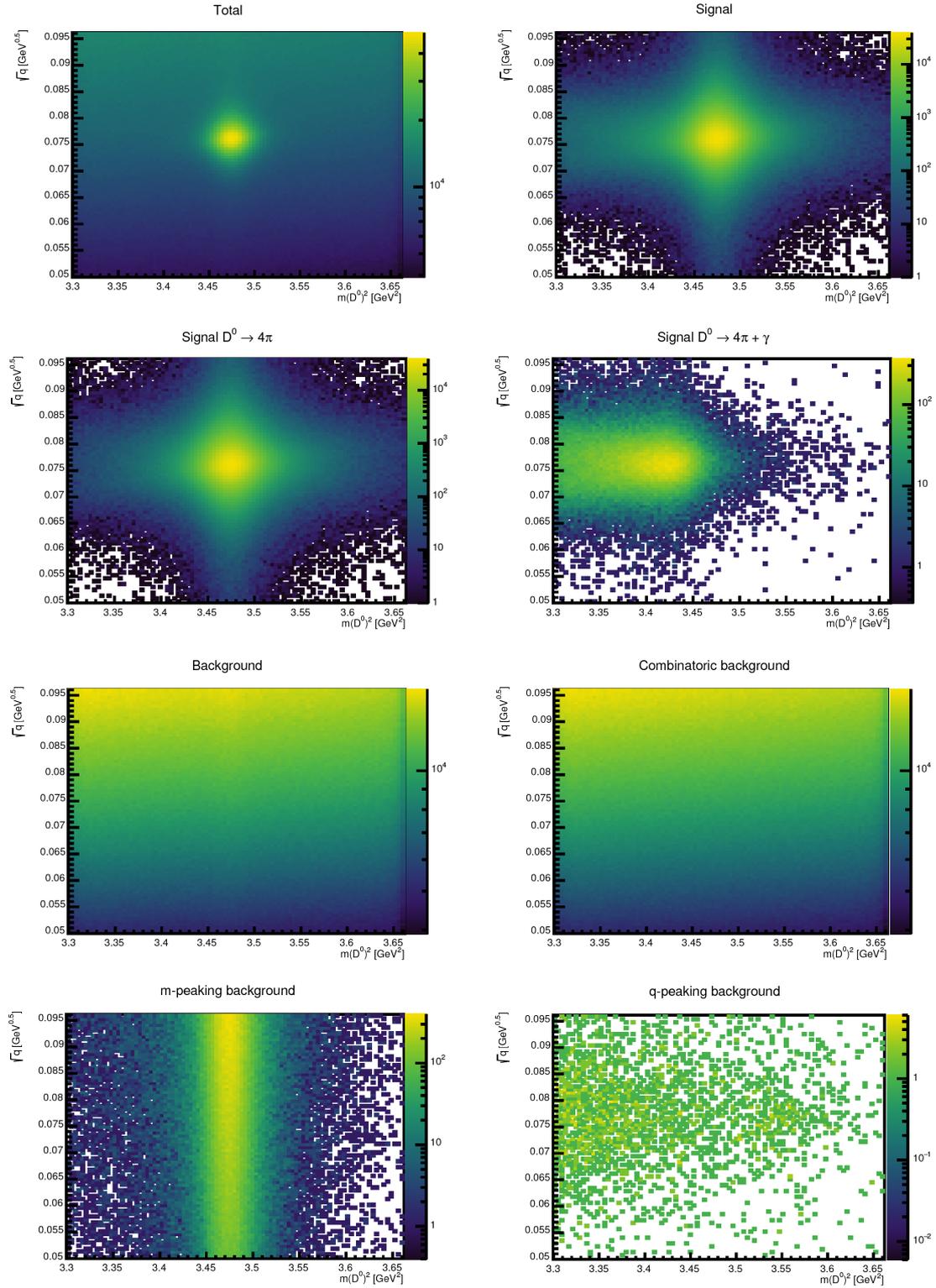


Figure 8.1: MC event distributions of signal and background contributions in  $\sqrt{q}$  over  $m_{D^0}^2$ .

	$m_{D^0}^2$	$\sqrt{q}$	scaling
Signal	DCB (6) bivariate Gaussian (5)	DCB (6) bivariate Gaussian (5)	signal yield (1) relative weight (1)
Combinatoric bg.	linear (1)	quadratic (2)	yield (1)
Mass-peaking bg.	$\equiv$ signal (0)	$\equiv$ combinatoric bg. (0)	yield (1)

Table 8.2: Number of parameters for the signal and background components and scaling.

	$m_{D^0}^2$		$\sqrt{q}$	
	Range/GeV <sup>2</sup>	# bins	Range/GeV	# bins
Fit region	3.317 to 3.630	37	0.050 to 0.096	36
Signal window	3.432 to 3.531	24	0.070 to 0.086	24

Table 8.3: Ranges and number of bins of the fit and signal region.

combinatoric background yield, which is also fitted.

The mass-peaking background is described by a projection of the signal shape in  $m_{D^0}^2$ , and the quadratic shape of the combinatoric background in  $\sqrt{q}$ . Only the yield is fitted. Since both signal and mass-peaking contributions are formed by  $D$  decays, the distributions in  $m_{D^0}^2$  look very similar. Leaving the parameters free does not give satisfactory fit results, so this approach is chosen.

The  $\sqrt{q}$ -peaking background has only very few events and is therefore omitted in the fit.

In total, the fit has 24 free parameters, an overview is given in Tab. 8.2. The probability-density for each parameter, as well as their correlations, are obtained by a Markov-chain. The data are binned, with a higher density of bins in the signal window (cf. Tab. 8.3).

### 8.3.1 Signal- and background-yield results

The results of the fit to the MC data can be seen in Figs 8.2 and 8.3. The plots show the normalized number of events for data and fit over  $m_{D^0}^2$  (left) and  $\sqrt{q}$  (right). The bin contents are always normalized to the bin area. The fit describes the data well. The tails of the signal are slightly underestimated (cf. Fig. 8.3a). The mass-peaking background is overestimated a bit (cf. Fig. 8.3b), due to slight differences between the projected shape of the signal component and the actual shape of the mass-peaking background.

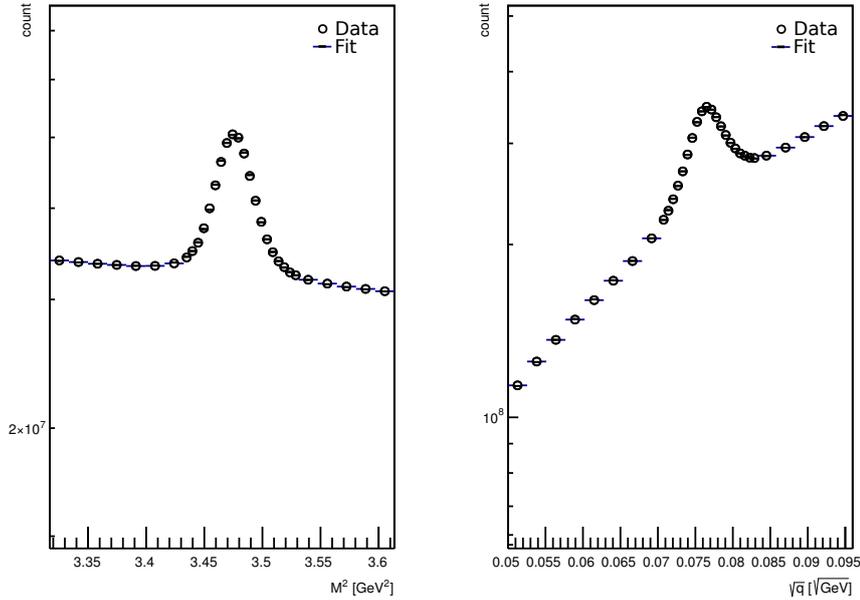


Figure 8.2: Fit of the signal and background shapes to the MC data set. Normalized number of events for data and fit over  $m_{D^0}^2$  (left) and  $\sqrt{q}$  (right).

The fit to the real data (cf. Fig. 8.4) is also very good. The results of the fit are summarized in Tab. 8.4. A total of  $500\,071 \pm 4335$  signal events could be reconstructed. The number of detectable  $D^*$  events per MC stream is around 3.2 million (cf. Tab. 8.5). Assuming an equal amount of detectable  $D^*$  events in the real data, the reconstruction efficiency is around 16%. The signal to background ratio in the signal window is 67%.

	Total	Signal window
Total events	8 592 981	1 353 720
Non-peaking background	$8\,060\,876 \pm 71\,604$	$939\,627 \pm 8346$
M-peaking background	$31\,409 \pm 3873$	$10\,658 \pm 1314$
Signal events	$500\,071 \pm 4335$	$426\,232 \pm 1343$
S/B		$0.6671 \pm 0.0032$

Table 8.4: Fitted signal and background yields of the data set.

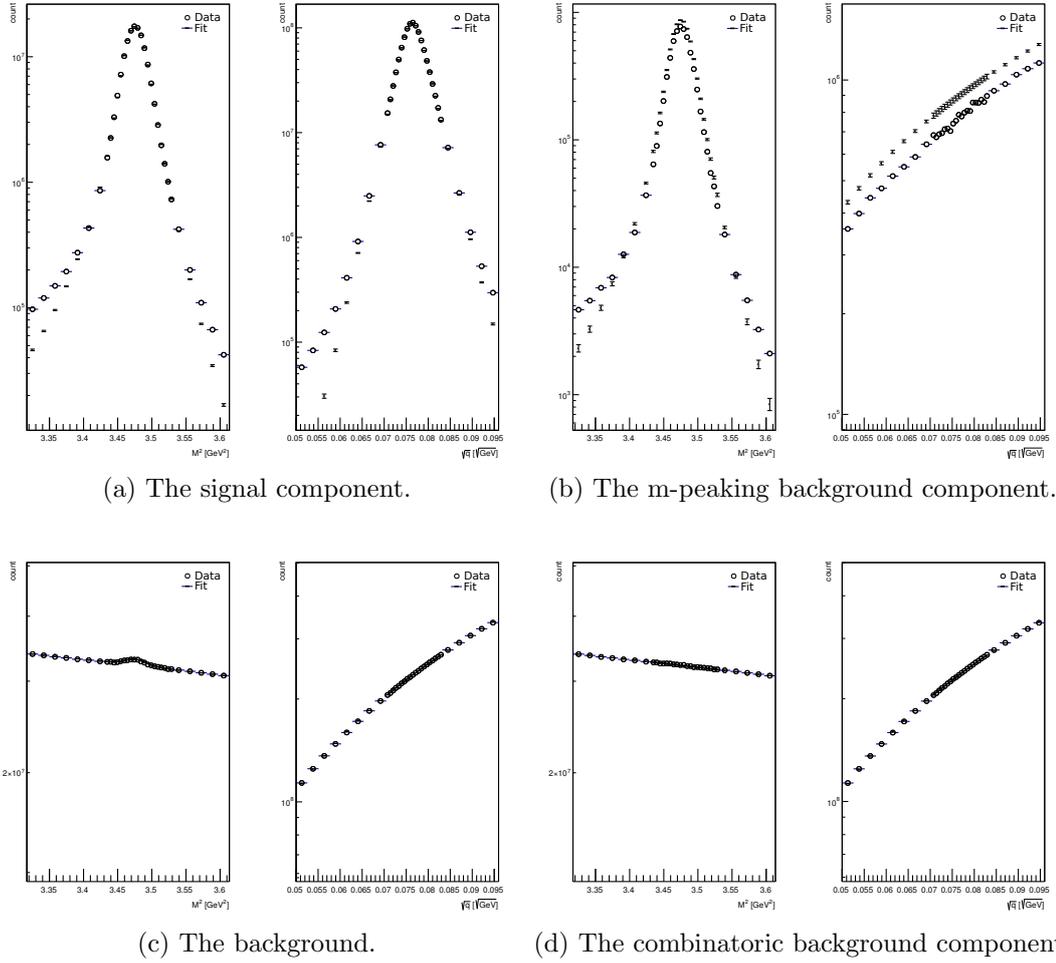


Figure 8.3: Signal and background components of the fit to the MC data set. Normalized number of events for data and fit over  $m_{D^*}^2$  (left in each sub-figure) and  $\sqrt{q}$  (right in each sub-figure).

	Generated		Detectable	
	$D^{*+}$	$D^{*-}$	$D^{*+}$	$D^{*-}$
All	10 380 589	10 382 087	1 626 896	1 625 509
From $c\bar{c}$	5 806 349	5 808 270	1 029 516	1 023 787

Table 8.5: Average number of generated and detectable  $D^*$  per stream from MC simulations.

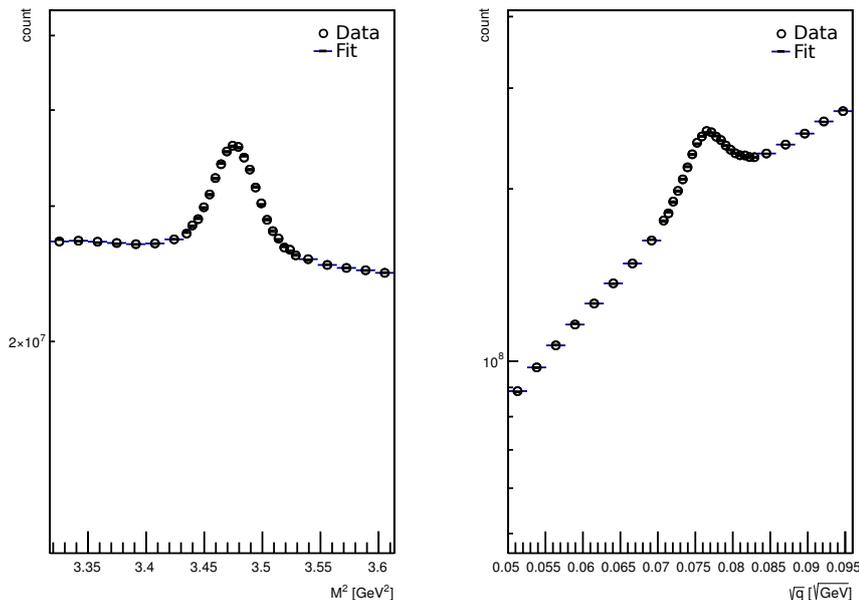


Figure 8.4: Fit of the signal and background shapes to the real data set. Normalized number of events for data and fit over  $m_{D^0}^2$  (left) and  $\sqrt{q}$  (right).

## 8.4 Boosted Decision Trees

The simple cut-based event selection works well for making a preselection, but will not give an optimal signal-to-background ratio needed for the PWA. Consider the signal distribution in Fig. 8.1: With cuts on  $\sqrt{q}$  and  $m_{D^0}^2$ , a rectangle around the signal peak would be selected, whereas e.g. an elliptical selection—taking into account the correlation of the signal distribution in  $\sqrt{q}$  and  $m_{D^0}^2$ —would cut away more background events, and therefore result in a better signal-to-background ratio.

BDTs provide a way to consider the correlations of the signal distribution in *all* event-selection variables. In other words: With simple cuts, events in an  $N$ -dimensional hypercube (where  $N$  is the number of event-selection variables) will be selected. With BDTs, events in a much smaller hypervolume, which is determined to maximize the separation between signal and background events, will be selected.

The Toolkit for Multivariate Data Analysis with ROOT (TMVA) [38] provides numerous classification algorithms. I used decision trees with adaptive boosting, which turned out to have superior classification power, and modest computational demands.

A BDT must be trained on events for which it is known whether they are signal

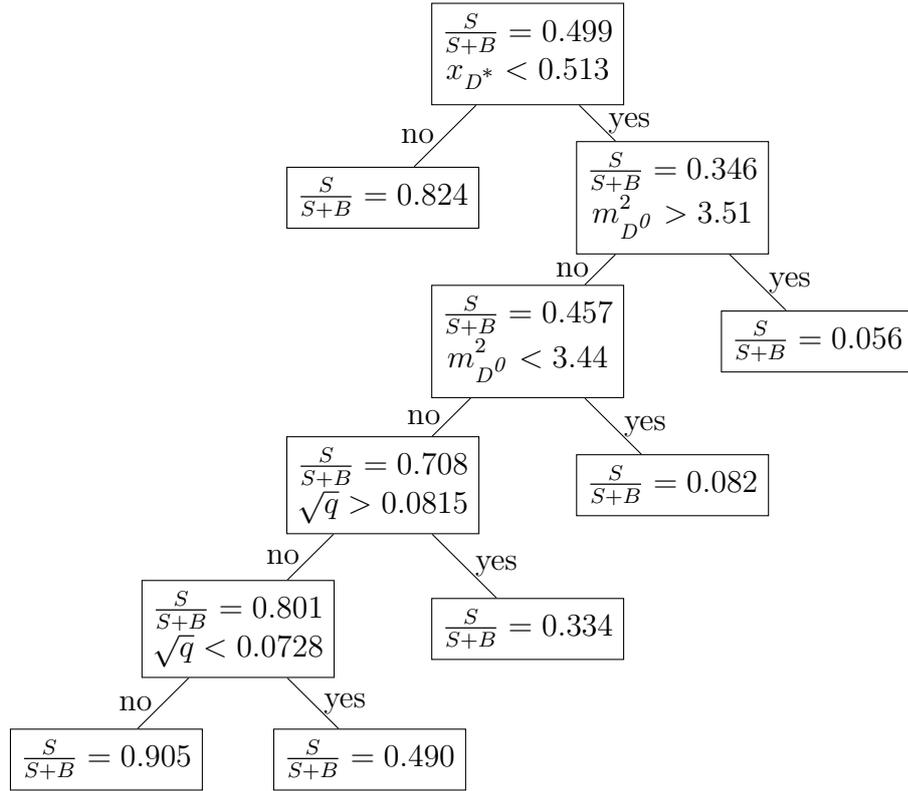


Figure 8.5: Example of a decision tree. To classify an event, its parameters like  $m_{D^0}^2$  and  $\sqrt{q}$  need to be calculated. Starting from the top, they are compared against the given values, until a leaf node is encountered.

or background. An example of a decision tree is depicted in Fig. 8.5. Nodes are created so that they give the best separation power between signal and background, by dividing the sample along one input variable. This is done in an iterative way, until the remaining fraction of events in a node falls below a minimum percentage, or a maximum depth of the tree is reached. This procedure yields one decision tree.

Before creating the next tree of the decision forest, the weights of all events are adjusted. The weighting, or *boosting*, can be done in different ways, the most popular being the adaptive boost: The new weights for misclassified events,  $\alpha$ , are derived from the misclassification rate,  $\text{err}$ , of the previous tree:

$$\alpha = \frac{1 - \text{err}}{\text{err}} \quad (8.2)$$

The boosting gives misclassified events a higher weight for the next iteration. After normalization, the procedure is repeated.

BDTs are prone to overtraining. During the training process, the misclassification rate will always decrease, but at some point the algorithm will start to learn the statistical fluctuations of the training sample. This means that the misclassification rate for a statistically independent sample will then start to rise again. Therefore, the algorithm splits the dataset in halves. Only the first half (the training sample) is used for the actual training. The second half (the test sample) is used to monitor the training process. As soon as the misclassification rate for the test sample rises again, the training is stopped.

The trees can then be used to classify events. Given the same number of signal and background events, a classifier value of 0 means that it is equally likely that the event is signal or background. Higher values indicate that the event is more likely to be a signal event, and vice versa.

### 8.4.1 Input variables

We use every signal event, and every 12<sup>th</sup> background event (in order to get a similar number of signal and background events) of MC stream 1 for training. The following BDT input variables are used:

- $m_{D^0}^2$ ,  $\sqrt{q}$ , and the logarithms of the  $p$ -values of the  $D^0$  and  $D^*$  vertex fits, each for the unconstrained and the constrained vertex fits.
- $x_{D^*}$ .
- The transverse momenta  $p_T$  of the 5 pions.
- The impact parameters  $\Delta z$  and  $\Delta r$ ,  $\mu$  likelihood,  $K$ - $\pi$  separation, and  $p$ - $\pi$  separation. For each of those variables, the maximum of the values of the four pions from the  $D^0$  and the value of the  $\pi_s$  are used.

The distributions for some of these variables, which offer good discrimination between signal and background, are shown in Fig. 8.6.

### 8.4.2 Results

The BDT output distributions are shown in Fig. 8.7. The MC signal (green) and background (red) distributions have been scaled to the fitted yields from Tab. 8.4. Thus, we can compare the MC distribution (pink) to the BDT distribution for the data (blue). Below a BDT value of 0.06, the data is slightly shifted towards higher BDT values. Above 0.08, MC and data align well.

For further analysis, a BDT cut is performed: Events with a lower BDT value (which are more likely to be background events) are discarded, and events with a higher BDT value (which are more likely to be signal events) are kept.

A low BDT cut yields a high signal efficiency, but a low background rejection. With an increasing BDT cut, more background is cut away, so the background

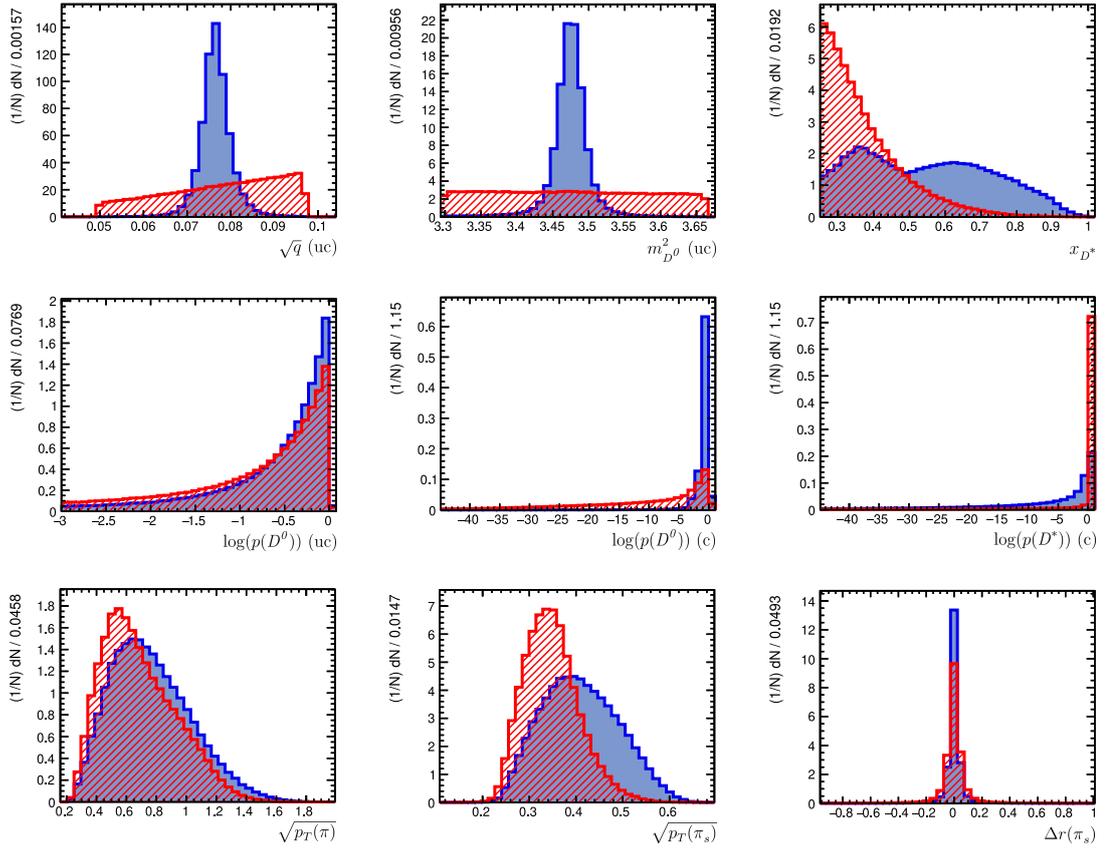


Figure 8.6: Input variable distributions of the MC training set, which is used to train the BDTs with TMVA. The signal is drawn in blue, the background in hatched red. Statistically independent input variables with preferably small overlap between signal and background are key for a good classification power of the BDTs.

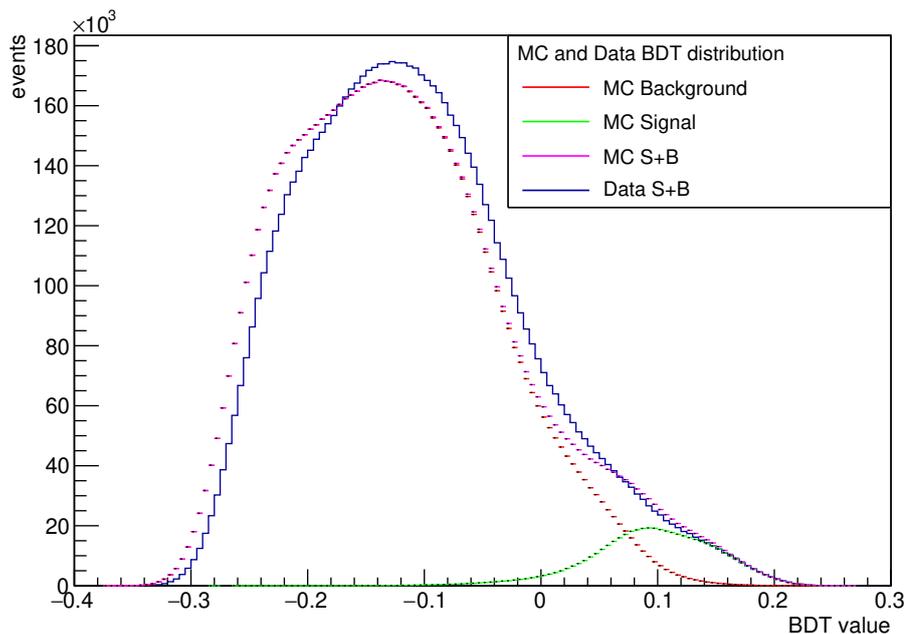


Figure 8.7: Distribution of the BDT response for MC and data. MC signal and background are well separated, and MC and data align well.

BDT cut	# events after BDT cut	# events after $K$ cut	expected purity
0.10	259 647	225 632	0.88
0.12	174 486	152 490	0.92
0.15	79 817	70 405	0.96

Table 8.6: Number of events and expected purity of the data for different BDT cuts.

rejection goes up. Eventually, more and more of the signal is cut away, so the signal efficiency goes down. Thus, a reasonable tradeoff between signal efficiency and background rejection has to be made.

The yield after the BDT cut and  $K$  cut (cf. Sec. 8.5) for different BDT cuts, as well as the expected purity from scaled MC, are summarized in Tab. 8.6. A BDT cut of 0.12 was used for most fits, it yields more than 150 000 events with a purity over 90 %.

## 8.5 Kaon cut

The  $D^0$  can also decay into one or two  $K$  mesons, which can decay into two charged pions. However, the line width of the  $K$  decays is much smaller than the detector resolution, and the fit can therefore not treat these decays correctly without further complication. We decided to reject all events where an invariant mass of a  $\pi^+ \pi^-$  pair is within a range of 11.92 MeV around the  $K^0$  mass.

The Kaons have strangeness and therefore decay weakly into two pions. The  $K^0$  is a mixture of the flavour eigenstates  $K_S^0$  and  $K_L^0$ . The  $K_S^0$  decays primarily into  $\pi^+ \pi^-$ , with a branching fraction of  $(69.20 \pm 0.05)\%$ . The  $\pi^+ \pi^-$  branching fraction for the  $K_L^0$  is much smaller, only  $(1.967 \pm 0.010) \times 10^{-3}$  [48].

The  $K_S^0$  has a mean life of around  $10^{-10}$  s, which translates into a line width in the order of  $10^{-5}$  eV. This is many orders of magnitude smaller than the detector resolution: A gaussian fit of the  $K$  peak in the  $\pi^+ \pi^-$  invariant mass spectrum gives a width of 3.97 MeV.

In reality, the smearing due to the detector resolution happens after the interference of the partial waves. In the fit, the unsmearred model is compared to the smeared data and the detector resolution is neglected. This can only work well if the detector resolution is small compared to the widths of the partial waves. This is indeed the case for all resonances in our model, which decay strongly, and therefore have much smaller lifetimes and line widths in the order of 100 MeV. Only the  $f_0(980)$  has a sharp peak where the detector resolution has a somewhat noticeable effect.

In order to treat the Kaons correctly, the model would have to be smeared after calculating the intensity. This would be complicated and computationally very demanding.

# Chapter 9

## Partial Wave Analysis

We now have a selection of events, where an initial state, the  $D^0$ , decays into a final state,  $\pi^+ \pi^- \pi^+ \pi^-$ . This decay has many possible intermediary states, which quantum-mechanically occur at the same time, and interfere with each other. The final state, i.e. the kinematics of the four pions, is described unambiguously by five parameters (cf. Sec. 9.2), and one event corresponds to a seemingly random point in this 5-dimensional phase space. However, when we look at a large number of events, their distribution in said phase space clearly exhibits a non-random structure. This is the interference pattern of the intermediary states, illustrated in Fig. 10.1.

With a partial wave analysis (PWA), we try to find out about the intermediary states, and their relative intensities and phases. PWA, in a strict sense, is a technique to decompose a total scattering amplitude into its constituent angular momentum components. In our sense, the decay amplitude of the  $D^0$  is decomposed into two-body decay amplitudes via the isobar formalism (cf. Sec. 9.3). Once a model (a set of decay amplitudes) has been composed, the best fit of the free amplitudes and model parameters to the data can be obtained with a maximum likelihood fit (cf. Sec. 9.7).

### 9.1 YAP – Yet Another Partial Wave Analysis Toolkit

We developed a toolkit for PWA, YAP [25]. It implements all the formalisms and amplitude components described in this chapter, and provides a framework to compose models in the isobar formalism, handle data and interface to fitting software, e.g. the Bayesian Analysis Toolkit (BAT).

## 9.2 Parameterization of the phase space

We need five parameters to describe the kinematics of a four body decay with a spin-0 initial state. Each final state particle has a four-momentum, so we have  $4 \times 4 = 16$  parameters. The masses of the final state particles are known (determining 4 parameters) and the total four-momentum has to be conserved (determining another 4 parameters). Since the spin of the initial state particle is 0, the whole system is rotation-invariant (discarding 3 parameters).

In Dalitz plot analyses with three final state particles  $a$ ,  $b$ , and  $c$ , there are two degrees of freedom (DoF). The invariant-mass squares  $m_{ab}^2$  and  $m_{bc}^2$  are chosen to fill 2-dimensional histograms, so called Dalitz plots. Analogously, we choose five invariant-mass squares,  $m_{12}^2$ ,  $m_{14}^2$ ,  $m_{23}^2$ ,  $m_{34}^2$ , and  $m_{14}^2$  as parameters to describe the kinematics of the decay. All other kinematic variables, like the helicity angles (cf. Sec. 9.4.2), can be derived from these five invariant-mass squares.

## 9.3 Isobar formalism

In the isobar formalism, a decay into a final state is looked at as subsequent two-body decays. It was originally used to model pion-nucleon, nucleon-nucleon, and antinucleon-nucleon interactions [55], where the intermediate resonances are isobars of a particular nuclear state. It was later generalized to three-body final states [36], and can be expanded for N-body final states. The complex amplitude of a specific decay chain is

$$\mathcal{A}_r = a_r F_r T_r W_r \prod_d \mathcal{A}_d. \quad (9.1)$$

The free complex amplitude of the decay,  $a_r$ , defines the magnitude and phase of the decay chain, and can be determined by a fit.

$F_r$  is the Blatt-Weisskopf barrier factor, which describes the centrifugal-barrier effect caused by the relative angular momentum  $L$  in the isobar decay. It is described in Sec. 9.5.

The dynamical function  $T_r$  describes the mass shape. The mass shapes used in this analysis are described in Sec. 9.6.

$W_r$  is the spin amplitude resulting from the spin coupling of the decay. It is described in Sec. 9.4.

$\mathcal{A}_d$  are the amplitudes of the daughter particles  $d$ , which are calculated recursively with the same formula. Final-state particles have an amplitude of 1.

Amplitudes of all decay chains are summed coherently:

$$\mathcal{A} = \sum_r \mathcal{A}_r. \quad (9.2)$$

### 9.3.1 Admixtures

The data contain some amount of background events. Most of them are uncorrelated pions which can be described by a flat wave. But also other contributions, where two or three pions have a common mother, and the other is uncorrelated, can occur. The signal and background components have to be added incoherently, i.e. their intensities have to be summed up:

$$\mathcal{I} = \sum_{\text{adm}} \mathcal{I}_{\text{adm}} \quad (9.3)$$

$$\mathcal{I}_{\text{adm}} = \mathcal{A}_{\text{adm}}^* \mathcal{A}_{\text{adm}}. \quad (9.4)$$

### 9.3.2 Bose-Einstein symmetrization

If two or more final state particles are indistinguishable, the amplitude of a decay is the sum over the amplitudes of all indistinguishable permutations of the final state particles. Let's take the decay  $D^0 \rightarrow \rho^0 \rho^0$  as an example. Each  $\rho^0$  decays into a  $\pi^+ \pi^-$  pair. We have recorded the four-momenta of four charged pions, which are labeled as

$$\pi^+ \pi^- \pi^+ \pi^- \equiv \pi_1 \pi_2 \pi_3 \pi_4. \quad (9.5)$$

There are four possible ways for the pions to form the  $\rho$  mesons:

$$D^0 \rightarrow (\rho^0 \rightarrow \pi_1 \pi_2)(\rho^0 \rightarrow \pi_3 \pi_4) \quad (9.6)$$

$$D^0 \rightarrow (\rho^0 \rightarrow \pi_3 \pi_2)(\rho^0 \rightarrow \pi_1 \pi_4) \quad (9.7)$$

$$D^0 \rightarrow (\rho^0 \rightarrow \pi_1 \pi_4)(\rho^0 \rightarrow \pi_3 \pi_2) \quad (9.8)$$

$$D^0 \rightarrow (\rho^0 \rightarrow \pi_3 \pi_4)(\rho^0 \rightarrow \pi_1 \pi_2). \quad (9.9)$$

The Bose-Einstein symmetrization is performed automatically in YAP.

## 9.4 Spin amplitudes

The spin amplitudes are described in the nonrelativistic helicity formalism. The helicity spin amplitude  $W_r$  depends on the helicity angles and angular momenta of the particles in an isobar decay and is calculated from Clebsch-Gordan coefficients and Wigner  $D$  functions.

### 9.4.1 Clebsch-Gordan coefficients

Clebsch-Gordan coefficients are expansion coefficients used in angular momentum coupling.

Consider a system with angular momentum  $J$  and wave function  $\psi_{jm}$ . The eigenvalue of the squared total-angular-momentum vector operator  $\mathbf{J}^2$  is  $j(j+1)$ , and the eigenvalue of  $J_z$ , the  $z$  component of  $\mathbf{J}$ , is  $m$ :

$$\mathbf{J}^2\psi_{jm} = j(j+1)\psi_{jm} \quad (9.10)$$

$$J_z\psi_{jm} = m\psi_{jm}. \quad (9.11)$$

The direct product of the eigenfunctions of two systems with angular momenta  $j_1$  and  $j_2$  is called the uncoupled representation:

$$\psi_{j_1 m_1} \psi_{j_2 m_2}. \quad (9.12)$$

It can be shown that the commutation rule is also valid for the sum of angular momenta:

$$[J_x, J_y] = iJ_z. \quad (9.13)$$

Therefore a coupled representation  $\psi_{JM}$  has to exist, where  $\mathbf{J}^2$  and  $J_z$  (with eigenvalues  $J(J+1)$  and  $M$ ), as well as  $\mathbf{J}_1^2$  and  $\mathbf{J}_2^2$  are diagonal. It is connected with the uncoupled representation by a unitary transformation

$$\psi_{JM} = \sum_{m_1 m_2} C(j_1, m_1, j_2, m_2, J, M) \psi_{j_1 m_1} \psi_{j_2 m_2}, \quad (9.14)$$

with the Clebsch-Gordan coefficients  $C(j_1, m_1, j_2, m_2, J, M)$ . Since  $M$  is the sum of  $m_1$  and  $m_2$ , this argument can be omitted.

A more detailed derivation and the calculation of the Clebsch-Gordan coefficients can be found in [52]. Tabulated values can be found in [48].

### 9.4.2 Helicity angles

The helicity angles are needed to calculate the Wigner  $D$  function of the helicity spin amplitude.

The helicity angles of a decay are the polar and azimuth angles  $\theta$  and  $\phi$  of a daughters' momentum vector of a two-particle decay, in the rest frame of the parent particle, with respect to an orthonormal helicity frame. The daughter particles are emitted back-to-back in the parent's rest frame, so a convention has to be made which of the daughters' momenta to chose to define the helicity angles.

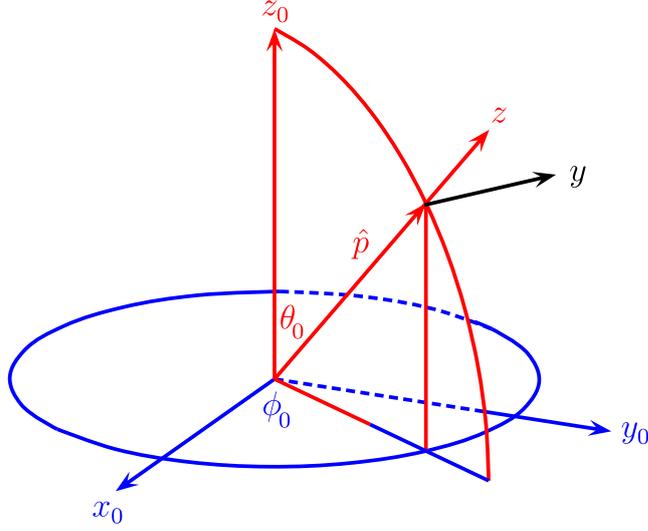


Figure 9.1: In the helicity frame  $(x_0, y_0, z_0)$ , the helicity angles of a decay are the polar and azimuth angles  $\theta_0$  and  $\phi_0$  of a daughters' momentum vector  $\hat{p}$ . The helicity frame  $(x, y, z)$  for a subsequent decay is also defined by  $\hat{p}$

The orientation of the helicity frames is defined by the decay kinematics. The helicity frame for the initial  $D^0$  decay can be chosen arbitrarily, since it has spin 0 (in this case the spin amplitude is always 1). For a subsequent decay of a particle with momentum  $\vec{p}$  in the  $D^0$  rest frame, and the (arbitrary)  $z$ -axis of the  $D^0$  rest frame,  $z_0$ , the helicity frame is defined as follows (cf. Fig. 9.1):

$$z \equiv \hat{p} \quad (9.15)$$

$$y \equiv z_0 \times z \quad (9.16)$$

$$x \equiv y \times z. \quad (9.17)$$

After a Lorentz boost of the helicity frame into the rest frame of the particle, its daughters have momenta  $\vec{q}$  and  $-\vec{q}$ , and the helicity angles are

$$\cos \theta \equiv \hat{q} \cdot z \quad (9.18)$$

$$\cos \phi \equiv \hat{q} \cdot x / \sin \theta. \quad (9.19)$$

A sign convention is chosen so that  $\theta$  is in the range from 0 to  $\pi$ , and  $\phi$  is in the range from  $-\pi$  to  $\pi$ .

An example for the decay  $D^0 \rightarrow \rho^0 \rho^0$  is depicted in Fig. 9.2. If the  $x$ -axis of the initial helicity frame,  $x_0$ , is chosen as shown,  $\phi_1$  is 0.

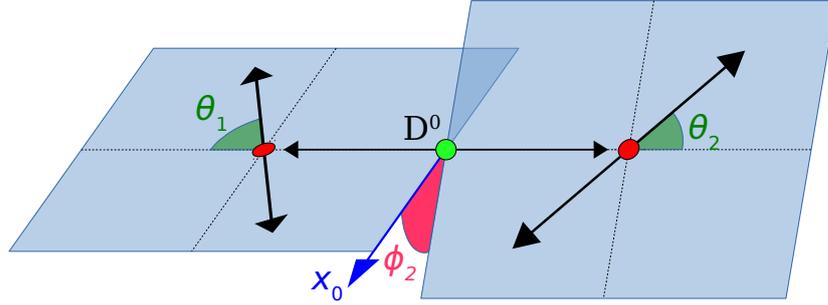


Figure 9.2: Helicity angles of a decay into two isobars. Every decay is drawn in its own helicity frame, which has already been boosted into the rest frame of the mother particle. The daughter momenta are therefore back-to-back.

### 9.4.3 Wigner $D$ functions

The Wigner  $D$  function gives the matrix elements of the rotation operator with Euler angles  $\alpha, \beta, \gamma$  in the  $jm$ -representation [56]:

$$D(j, m, n, \alpha, \beta, \gamma) = e^{-im\alpha} d(j, m, n, \beta) e^{-in\gamma} \quad (9.20)$$

$$d(j, m, n, \beta) = \sum_k (-1)^k \frac{\sqrt{(j+m)!(j-m)!(j+n)!(j-n)!}}{(j-n-k)!(j+m-k)!k!(k+n-m)!} \times \cos^{2j+m-m'-2k} \frac{\beta}{2} \sin^{2k+m'-m} \frac{\beta}{2}. \quad (9.21)$$

### 9.4.4 Helicity spin amplitude

Now we can calculate the helicity spin amplitude  $W_r$ . Consider a parent particle with spin  $J$  and spin projection  $M$  decaying into two daughter particles with spins  $j_1$  and  $j_2$ , and spin projections  $m_1$  and  $m_2$ . The daughters have a total spin  $S$  and relative angular momentum  $L$  which satisfy

$$|j_1 - j_2| \leq S \leq j_1 + j_2 \quad (9.22)$$

$$|J - S| \leq L \leq J + S. \quad (9.23)$$

The spin amplitude for a given spin configuration is

$$\begin{aligned}
 W_r = & \overline{D}(J, M, m_1 - m_2, \phi, \theta, 0) \sqrt{\frac{2L+1}{2J+1}} \\
 & \times C(L, 0, S, m_1 - m_2, J) C(j_1, m_1, j_2, -m_2, S).
 \end{aligned}
 \tag{9.24}$$

$\overline{D}$  is the complex conjugate of the Wigner  $D$  function, the first Clebsch-Gordan coefficient describes the spin-orbit coupling, the second describes the spin-spin coupling, and  $\phi$  and  $\theta$  are the helicity angles of the decay.

## 9.5 Blatt-Weisskopf barrier factors

The Blatt-Weisskopf barrier factor  $F_r$  describes the centrifugal-barrier effect caused by the relative angular momentum  $L$  in the isobar decay. It depends on the relative angular momentum  $L$  between the daughters:

$$L = 0 : F(z) = 1 \tag{9.25}$$

$$L = 1 : F(z) = \sqrt{\frac{2z}{z+1}} \tag{9.26}$$

$$L = 2 : F(z) = \sqrt{\frac{13z^2}{z(z+3)+9}} \tag{9.27}$$

$$L = 3 : F(z) = \sqrt{\frac{277z^3}{z(z(z+6)+45)+225}} \tag{9.28}$$

with

$$z = R^2 q^2, \tag{9.29}$$

where  $R$  is the radial size of the resonance, and  $q$  is the breakup momentum of the daughters (cf. Eq. 9.32).

## 9.6 Mass Shapes

The mass shape  $T_r$  is the mass-dependent part of the decay amplitude. The intensity  $T_r^2$  can be considered as probability density of the invariant mass of the daughter particles of the decay. It peaks around the nominal mass of the resonance  $m_r$ , and its width  $\Gamma$  is related to the lifetime of the resonance via the uncertainty principle.

A resonance also involves a phase motion of  $180^\circ$  over the mass range. The trajectory of the mass shape in the complex plane is called Argand plot. For a Breit-Wigner resonance, it is a circle with center  $0.5i$  and radius 0.5. It starts at the origin and goes around one turn anti-clockwise.

The mass shapes (except for the  $a_1$ ) are normalized so that  $T_r(m_r) = i$ .

### 9.6.1 Breit-Wigner

The dynamical function  $T_r$  is often described by a Breit-Wigner shape [19]. This parameterization resembles the formula of a driven harmonic oscillator and can work well for narrow resonances. It is calculated as

$$T_r = \frac{\Gamma_{ab}}{m_r^2 - m_{ab}^2 - im_r\Gamma_{ab}} \quad (9.30)$$

$$\Gamma_{ab} = \Gamma_r \left( \frac{q_{ab}}{q_r} \right)^{2J+1} \left( \frac{m_r}{m_{ab}} \right) \frac{F^2(R^2 q_{ab}^2)}{F^2(R^2 q_r^2)}. \quad (9.31)$$

Here,  $m_r$  is the nominal mass of the resonance,  $m_{ab}$  is the invariant mass of the daughter particles,  $\Gamma_{ab}$  is the mass-dependent width, and  $\Gamma_r$  is the nominal width.  $F$  are Blatt-Weisskopf factors.  $q_{ab}$  and  $q_r$  are breakup momenta of the daughter particles. They are calculated as

$$q = \sqrt{\frac{(m_p^2 - (m_a + m_b)^2)(m_p^2 - (m_a - m_b)^2)}{4m_p^2}}. \quad (9.32)$$

For  $q_{ab}$ , the parent mass  $m_p$  is taken as the measured invariant mass of the daughters,  $m_{ab}$ ; for  $q_r$ , the parent mass  $m_p$  is taken as the nominal mass of the resonance,  $m_r$ .

### 9.6.2 Flatté

The Flatté coupled-channel form [28] can describe resonances that lie close to the threshold of one or more other channels  $c$ :

$$T_r = \frac{2\Gamma_R}{m_r \left( m_r^2 - m_{ab}^2 - \frac{2i}{m_r} \sum_c g_c q_c \right)}. \quad (9.33)$$

$g_c$  and  $q_c$  are the coupling constants and breakup momenta of the channels  $c$ . This form is used for the  $f_0(980)$ , which decays to  $\pi\pi$  and  $KK$ .

The resulting mass shape has a very sharp and narrow peak, due to the opening of the  $KK$  channel. In the data, the peak is slightly smeared, due to the limited

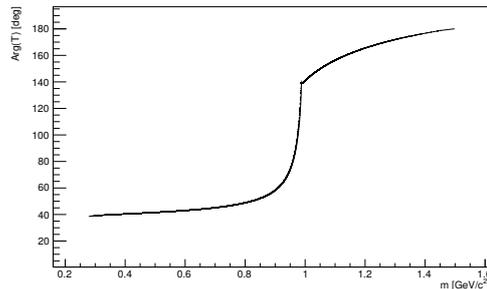


Figure 9.3: Phase of the  $f_0(980)$  mass shape ( $\text{Arg}(T_r)$  in degrees over mass).

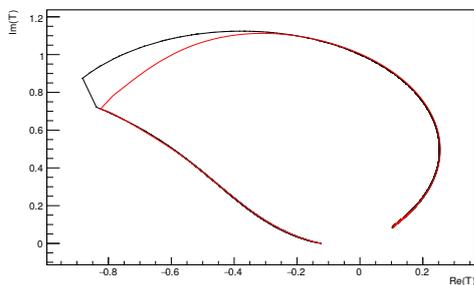


Figure 9.4: Argand plot of the  $f_0(980)$  mass shape ( $\text{Im}(T_r)$  over  $\text{Re}(T_r)$ ). Nominal (black), and smeared (red).

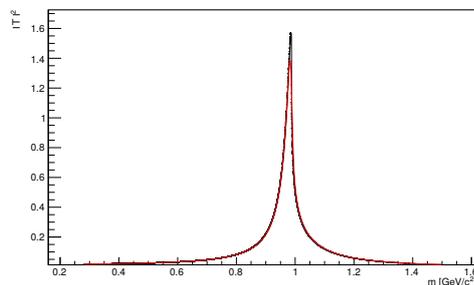


Figure 9.5: Intensity of the  $f_0(980)$  mass shape ( $|T_r|^2$  over mass). Nominal (black), and smeared (red).

detector resolution. To take this effect into account, we numerically convolute the intensity with a Gaussian with a width of 3.97 MeV (cf. Fig. 9.5). This value is obtained from a Gaussian fit to the  $K^0$  peak in the  $\pi^+ \pi^-$  invariant-mass distribution (cf. Sec. 8.5).

This approach is not quite correct, since in reality the smearing due to the detector resolution happens after the interference of the partial waves, and not before. However, smearing the total amplitude  $\mathcal{A}$  would be disproportionately laborious and computationally demanding. Since the difference of the unsmeared and the smeared mass shape is small—essentially only the peak is smoothed a bit—smearing only the mass shape is better than not taking the detector resolution into account.

### 9.6.3 Gounaris-Sakurai

Broad vector resonances like the  $\rho^0(770)$  can be described with a parameterization proposed by Gounaris and Sakurai [35]. Our parameterization is derived from [45].

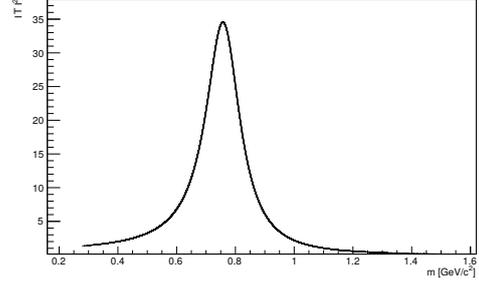
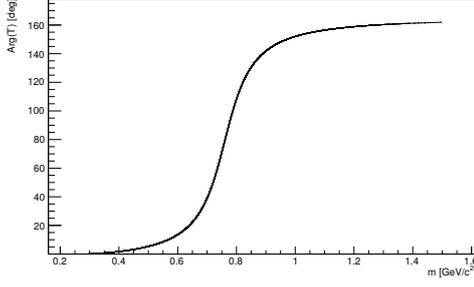


Figure 9.6: Phase of the  $\rho^0$  mass shape. Figure 9.7: Intensity of the  $\rho^0$  mass shape.

The resulting mass shape of the  $\rho^0$  is shown in Figures 9.6 and 9.7. It is calculated as

$$T_r = \frac{m_r^2 + d_0 m_r \Gamma_r}{M^2 - m_{ab}^2 - i M_0 \Gamma} \quad (9.34)$$

with

$$\Gamma = \Gamma_r \frac{m_r}{m_{ab}} \left( \frac{m_{ab}^2 - s_0}{m_r^2 - s_0} \right)^{3/2} \quad (9.35)$$

$$d_0 = \frac{1}{\pi} \frac{m_r}{\sqrt{m_r^2 - s_0}} \left( 1 - \frac{s_0}{m_r^2 - s_0} (2 - 3h(m_r^2, s_0)) \right) \quad (9.36)$$

$$M^2 = m_r^2 + \frac{\Gamma_0 m_r^2}{\pi} \frac{m_{ab}^2 - s_0}{(m_r^2 - s_0)^{3/2}} \left( 2h(m_{ab}^2, s_0) - \left( 2 + \frac{s_0}{m_r^2} \frac{m_{ab}^2 - m_r^2}{m_{ab}^2 - s_0} \right) h(m_r^2, s_0) - \frac{m_r^2 - s_0}{m_r^2} \frac{m_{ab}^2 - m_r^2}{m_{ab}^2 - s_0} \right) \quad (9.37)$$

where

$$s_0 = (m_a + m_b)^2 \quad (9.38)$$

$$h(m^2, s_0) = \frac{\sqrt{m^2 - s_0}}{m} \ln \left( \frac{m + \sqrt{m^2 - s_0}}{\sqrt{s_0}} \right). \quad (9.39)$$

### 9.6.4 $a_1$

The  $a_1$  is described as a Breit-Wigner resonance with a mass-dependent width:

$$\Gamma_{\text{tot}}^{a_1}(s) = \Gamma_{2\pi^0\pi^+}^{a_1}(s) + \Gamma_{2\pi^+\pi^-}^{a_1}(s) + g_{K^*K}^2 \Gamma_{K^*K}^{a_1}(s). \quad (9.40)$$

The total width is the sum of the partial widths of the decays into  $2\pi^0\pi^+$ ,  $2\pi^+\pi^-$ , and  $K^*K$ . The partial widths into  $2\pi^0\pi^+$  and  $2\pi^+\pi^-$  are assumed to be equal and proportional to the integral over the  $a_1$  Dalitz plot at a given invariant mass  $s$ :

$$\Gamma_{2\pi^0\pi^+}^{a_1}(s) = \Gamma_{2\pi^+\pi^-}^{a_1}(s) \quad (9.41)$$

$$\Gamma_{2\pi^+\pi^-}^{a_1}(s) \propto \frac{1}{s^{3/2}} \int ds_1 ds_2 |\mathcal{A}_{a_1}(s_1, s_2)|^2. \quad (9.42)$$

In the so-called Bowler parametrization, only the dominant decay  $a_1 \rightarrow \rho^0\pi^+$  ( $S$ -wave) is taken into account. In our model, we also integrate over the  $a_1 \rightarrow \rho^0\pi^+$  ( $D$ -wave) and  $a_1 \rightarrow (\pi\pi)_S\pi^+$  components. If any free amplitude changes, the integral is re-evaluated. The mass dependent width, phase motion, argand plot, and intensity distribution of the  $a_1$  mass shape for a given set of free amplitudes are shown in Figs 9.8 to 9.11.

### 9.6.5 $(\pi\pi)_S$

The decay into two pions in a relative  $S$  state is described according to [10]. It is based on a parametrization of the  $(\pi\pi)_S$ -wave from [14], which they extracted from  $\pi\pi$  elastic scattering data. The  $f_0(980)$  is removed from the amplitude (and treated as a separate resonance, cf. Sec. 9.6.2). The phase motion, argand plot, and intensity distribution of the  $(\pi\pi)_S$  mass shape are shown in Fig. 9.12, 9.13, and 9.14.

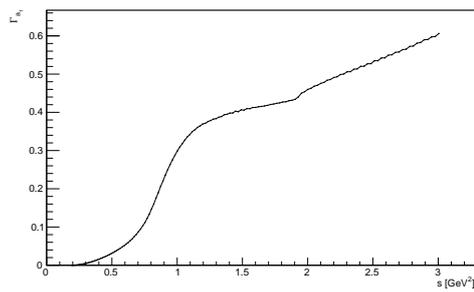


Figure 9.8: Mass dependent width of the  $a_1$  mass shape ( $\Gamma_{\text{tot}}^{a_1}(s)$  over  $s$ ).

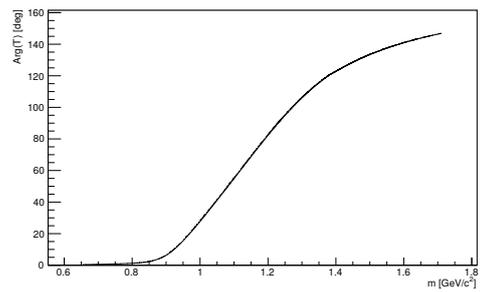


Figure 9.9: Phase of the  $a_1$  mass shape.

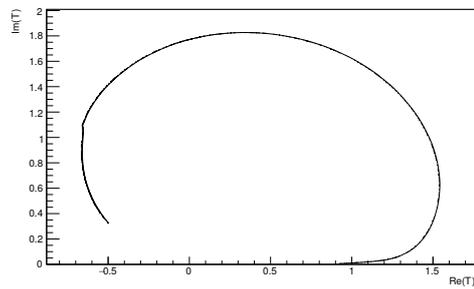


Figure 9.10: Argand plot of the  $a_1$  mass shape.

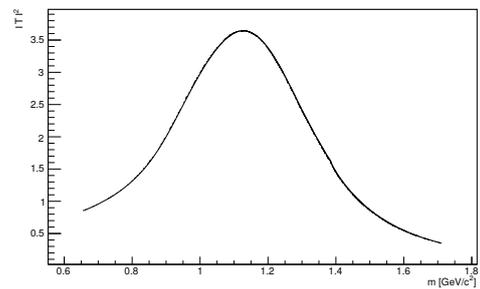


Figure 9.11: Intensity of the  $a_1$  mass shape.

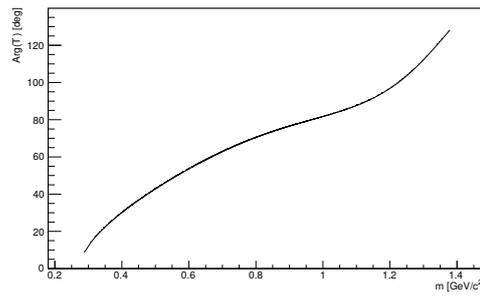


Figure 9.12: Phase of the  $(\pi\pi)_S$  mass shape.

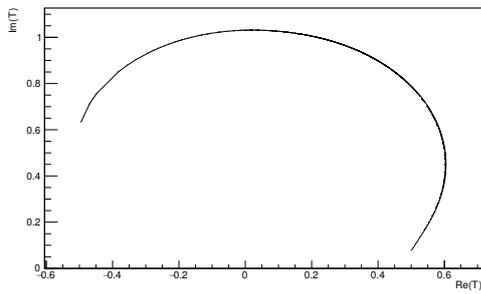


Figure 9.13: Argand plot of the  $(\pi\pi)_S$  mass shape.

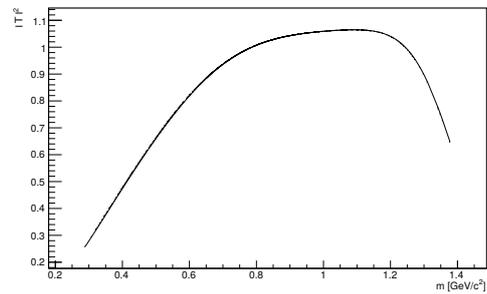


Figure 9.14: Intensity of the  $(\pi\pi)_S$  mass shape.

## 9.7 Maximum likelihood fit

The previous sections detailed how a decay amplitude and its components are described mathematically. After a model has been compiled from several signal and background components, we want to determine the parameters which yield the best fit to the data. The parameters are the complex free amplitudes of each component, the admixtures of signal and background, and possibly others like masses and widths.

The figure of merit how well the model fits the data is the likelihood  $\mathcal{L}$  (cf. Sec. 9.7.1). We use ROOT's MINUIT to maximize it. MINUIT is basically a gradient follower. At a given point in the parameter space, it numerically calculates the gradient of the likelihood by varying each parameter slightly. It then performs a line search along the steepest gradient for the maximum likelihood. From this new point it starts over, until it converges towards a maximum. Depending on the starting point, it might be a local maximum. A simple approach to find the global maximum is to do many fits with random starting points.

Admittedly, there are more sophisticated methods to find the global maximum, like Markov chain Monte Carlo (MCMC) methods. With BAT, the user can easily switch from MINUIT to various MCMC methods, but they turned out to be too computationally demanding.

### 9.7.1 Likelihood

The likelihood  $\mathcal{L}$  is the product of the intensities  $\mathcal{I}$  (cf. Eq. 9.3) of all data points  $i$  divided by the integral  $I$  (cf. Sec. 9.7.2). With a high number of data points, the likelihood can easily exceed the range of a double-precision floating point number. Maximizing the natural logarithm of the likelihood yields the same fit results, since it is a strictly monotonic function, but makes its numerical calculation manageable. The natural logarithms of the intensities have to be summed up. The numerical accuracy of this summation is increased by using the Kahan summation algorithm.

$$\mathcal{L} = \prod_i \frac{\mathcal{I}_i}{I} \quad (9.43)$$

$$\ln \mathcal{L} = \sum_i (\ln \mathcal{I}_i - \ln I). \quad (9.44)$$

### 9.7.2 Integration

The integral  $I$  is the normalized sum of intensities  $\mathcal{I}$  (cf. Eq. 9.3) of the events  $i$  of an integration sample (cf. following section) with  $N$  points:

$$I = \frac{1}{N} \sum_{i=1}^N \mathcal{I}_i. \quad (9.45)$$

Calculating this integral during the PWA fit takes about a second. However, if only the values of the free amplitudes  $a_r$  are changed, the integral can be evaluated fast without summing over the events  $i$ . With Eq.s 9.3, 9.4 and 9.2 we get

$$\sum_i \mathcal{I}_i = \sum_i \sum_{\text{adm}} \mathcal{I}_{\text{adm}}(i, \vec{p}) = \sum_{\text{adm}} \sum_i \mathcal{I}_{\text{adm}}(i, \vec{p}) \quad (9.46)$$

$$\sum_i \mathcal{I}_{\text{adm}}(i, \vec{p}) = \sum_i \mathcal{A}_{\text{adm}}^*(i, \vec{p}) \mathcal{A}_{\text{adm}}(i, \vec{p}) \quad (9.47)$$

$$= \sum_i \left( \sum_r \mathcal{A}_r^*(i, \vec{p}_r) \sum_r \mathcal{A}_r(i, \vec{p}_r) \right) \quad (9.48)$$

$$= \sum_i \sum_{r, s \geq r} 2\mathcal{A}_r^*(i, \vec{p}_r) \mathcal{A}_s(i, \vec{p}_s). \quad (9.49)$$

$\vec{p}$  and  $\vec{p}_r$  are the parameters of the model and a single resonance  $r$ , respectively, like masses and widths. With  $\mathcal{A}_r(i, \vec{p}_r) \equiv a_r A_r(i, \vec{p}_r)$ , we can pull the free amplitudes  $a$  out of the sum over the events  $i$ .

$$\sum_i \mathcal{I}_{\text{adm}}(i, \vec{p}) = \sum_i \sum_{r, s \geq r} 2a_r^* A_r^*(i, \vec{p}_r) a_s A_s(i, \vec{p}_s) \quad (9.50)$$

$$= \sum_{r, s \geq r} a_r^* a_s \sum_i 2A_r^*(i, \vec{p}_r) A_s(i, \vec{p}_s) \quad (9.51)$$

$$\equiv \sum_{r, s \geq r} a_r^* a_s I_{rs}(\vec{p}). \quad (9.52)$$

The values  $I_{rs}$  are stored in a matrix, and only have to be reevaluated if  $\vec{p}$  is changed. We can now rewrite the integral  $I$  in terms of free amplitudes  $a$  and matrix elements  $I_{rs}$ :

$$I = \frac{1}{N} \sum_{\text{adm}} \sum_{r, s \geq r} a_r^* a_s I_{rs}(\vec{p}). \quad (9.53)$$

### Integration sample

The integration is performed with phase-space-distributed  $D^0 \rightarrow \pi^+\pi^-\pi^+\pi^-$  events from the Belle MC data. The same event-selection cuts, BDT cut and  $K$  cut as for the data sample are performed (cf. Chap. 8). From the remaining events,  $D^0 \rightarrow \pi^+\pi^-\pi^+\pi^-$  decays are selected based on MC information. This yields  $1.52 \times 10^6$  events.

Since the integration sample is processed in the same way as the data sample, acceptance effects of the detector and event selection are incorporated.

### 9.7.3 Fit fractions

The complex amplitudes yielded by the fit depend on the choice of normalizations, phase conventions, and choice of mass shapes. Therefore, they are not necessarily comparable between different experiments and analyses. Fit fractions, on the other hand, are straightforward to calculate and typically quoted in PWAs. They allow to compare different analyses and estimate the branching fractions of the decay modes. A total fit fraction of more than one means there is destructive interference between the components of the model, and a total fit fraction of less than one means there is constructive interference between the components of the model.

The fit fraction FF of a component  $j$  is defined as

$$\text{FF}_j = \frac{I_{jj}}{I}, \quad (9.54)$$

where  $I_{jj}$  is the integral of the component  $j$ , and  $I$  is the total integral, according to Sec. 9.7.2, both calculated with the fit fraction sample.

### Fit fraction sample

Unlike the integration sample, the fit fraction sample must not include any acceptance effects. Therefore,  $10^7$  events with an even distribution in the phase space were generated with ROOT's `TGenPhaseSpace`.

### 9.7.4 Goodness of fit

The quality of the fit is assessed by a  $\chi^2$  test. The data are binned with an adaptive binning scheme: The phase space is repeatedly divided along five mass axes, such that each bin  $i$  contains the same number of events  $N_i$ . The number of expected events  $\langle N_i \rangle$  is the weighted number of events  $j$  from the integration sample in the bin. The events are weighted with their intensities  $w_j$ . The weighted total number

of events from the integration sample is normalized to the total number of events in the data sample,  $N$ . The total  $\chi^2$  is the sum of the  $\chi^2$  of each bin.

$$\chi^2 = \sum_i \frac{(N_i - \langle N_i \rangle)^2}{\langle N_i \rangle} \quad (9.55)$$

$$\langle N_i \rangle = \frac{N}{\sum_i w_i} w_i \quad (9.56)$$

$$w_i = \sum_{j \in i} w_j. \quad (9.57)$$

The statistical uncertainty of  $\langle N_i \rangle$  is taken into account by adding the normalized sum of the squared intensities in the denominator.

$$\chi^2 = \sum_i \frac{(N_i - \langle N_i \rangle)^2}{\langle N_i \rangle + \bar{\sigma}_i^2} \quad (9.58)$$

$$\bar{\sigma}_i^2 = \frac{N}{\sum_i w_i} w_i^2. \quad (9.59)$$

In order to get a comparable measure of the goodness of fit, the  $\chi^2$  has to be divided by the number of degrees of freedom (NDF). The NDF is the number of bins  $i$  minus the number of free parameters of the fit.

A perfect fit has a reduced  $\chi^2$  of 1. Values lower than 1 indicate over-fitting, and values much larger than 1 indicate a poor fit.

## 9.8 Overlap integrals

Normalized overlap integrals are a way to quantify how similar two components are. The overlap integral  $O$  of two components  $r$  and  $s$  is defined as

$$O_{rs} = \frac{I_{rs}}{\sqrt{I_{rr}I_{ss}}}. \quad (9.60)$$

Components with a large overlap integral are hard to distinguish by the fit, which can create large destructive interferences if no countermeasures are taken.

## 9.9 Model selection

A good model should achieve two goals: It should describe the data well; and it should do so with few parameters, which means that the correlations between the parameters should be minimal.

The maximum-likelihood fit, however, only tries to achieve the first goal. It can create large destructive interference between similar components as a side effect. The amplitudes of the interfering components will have large correlations, and the total fit fraction can become quite large.

There are many components which can possibly contribute to the model. Adding more components can lead to a higher likelihood value, at the cost of more parameters.

To achieve the second goal as well, the likelihood can be modified, so that solutions with smaller total fit fractions (and less components) will be favored. Two common approaches are LASSO [50] and biggest conceivable model (BCM) [18].

$$\text{LASSO: } \Delta\mathcal{L} = - \sum_i f_i/\Gamma \tag{9.61}$$

$$\text{BCM: } \Delta\mathcal{L} = - \sum_i \ln(1 + f_i/\Gamma^2) \tag{9.62}$$

LASSO penalizes the likelihood with the sum over all fit fractions  $f_i$  divided by a scaling factor  $\Gamma$ . The BCM approach is similar, but the penalty is logarithmic, so that single large contributing components are not penalized as much as in the LASSO approach.

I use BCM for the fits with model selection, with  $\Gamma = 0.75$ .

For components where the two previous analyses by FOCUS [46] and CLEO-c [26] give comparable results, Gaussian priors are used for the fit fractions:

Components	Mean	Width
$a_1^+\pi^-, a_1^+ \rightarrow \rho^0\pi^+$ (S-wave)	40	5
$a_1^+\pi^-, a_1^+ \rightarrow (\pi\pi)_S\pi^+$	9.6	3
$\rho^0\rho^0$	24	5

# Chapter 10

## PWA fit results

I performed several maximum likelihood fits, with different models, with and without model selection, and with different BDT cuts. The PWA has been performed with YAP, BAT, and MINUIT. The results are presented in the following sections. I will discuss and compare the results with previous measurements in Sec. 10.6.

### 10.1 Phase space plots

Before going into the quantitative details of the fits, I want to picture the data and fit results. Fig. 10.1 shows the event distribution of the data in all possible combinations of two- and three-pion invariant masses. The  $\rho^0$  peak is clearly visible in all two-pion invariant masses, and also the cusp of the  $f_0(980)$  is evident. The Kaon cut leaves a gap in the invariant masses of oppositely charged pion pairs (e.g. in  $m_{12}$ , but not in  $m_{13}$ ). The plots also illustrate the phase space boundaries. E.g. in  $m_{12}$  vs.  $m_{23}$ , the populated area is a triangle, since the sum of the invariant masses of the two pion pairs cannot exceed the invariant mass of the  $D^0$ . Due to the Bose-Einstein symmetrization (cf. Sec. 9.3.2), there are several sets of plots which are identical, disregarding statistical fluctuations.

Similar plots can be obtained for the fit results. Events from the integration sample (cf. Sec. 9.7.2) are weighted with their intensity in the fitted model. The result of the baseline fit is shown in Fig. 10.2, and it looks very similar for all other fits.

The fit has more events towards the phase space boundary, because the integration sample has more events overall. The  $\rho^0$  looks slightly attenuated and shifted towards higher masses, but overall, data and fit look very similar.

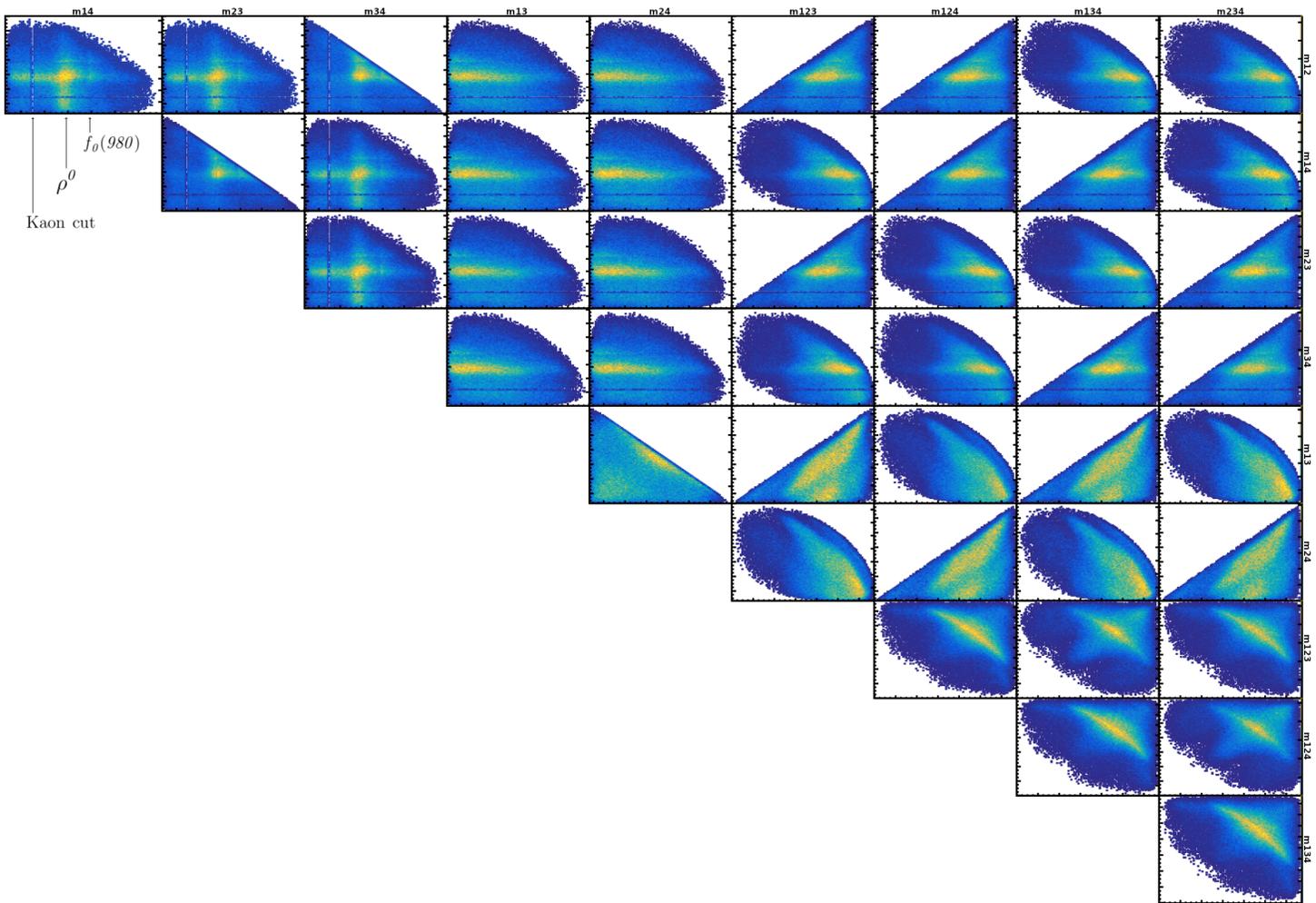


Figure 10.1: The *data*. Event distribution of the data in all possible combinations of invariant masses of two- and three-particle subsystems.

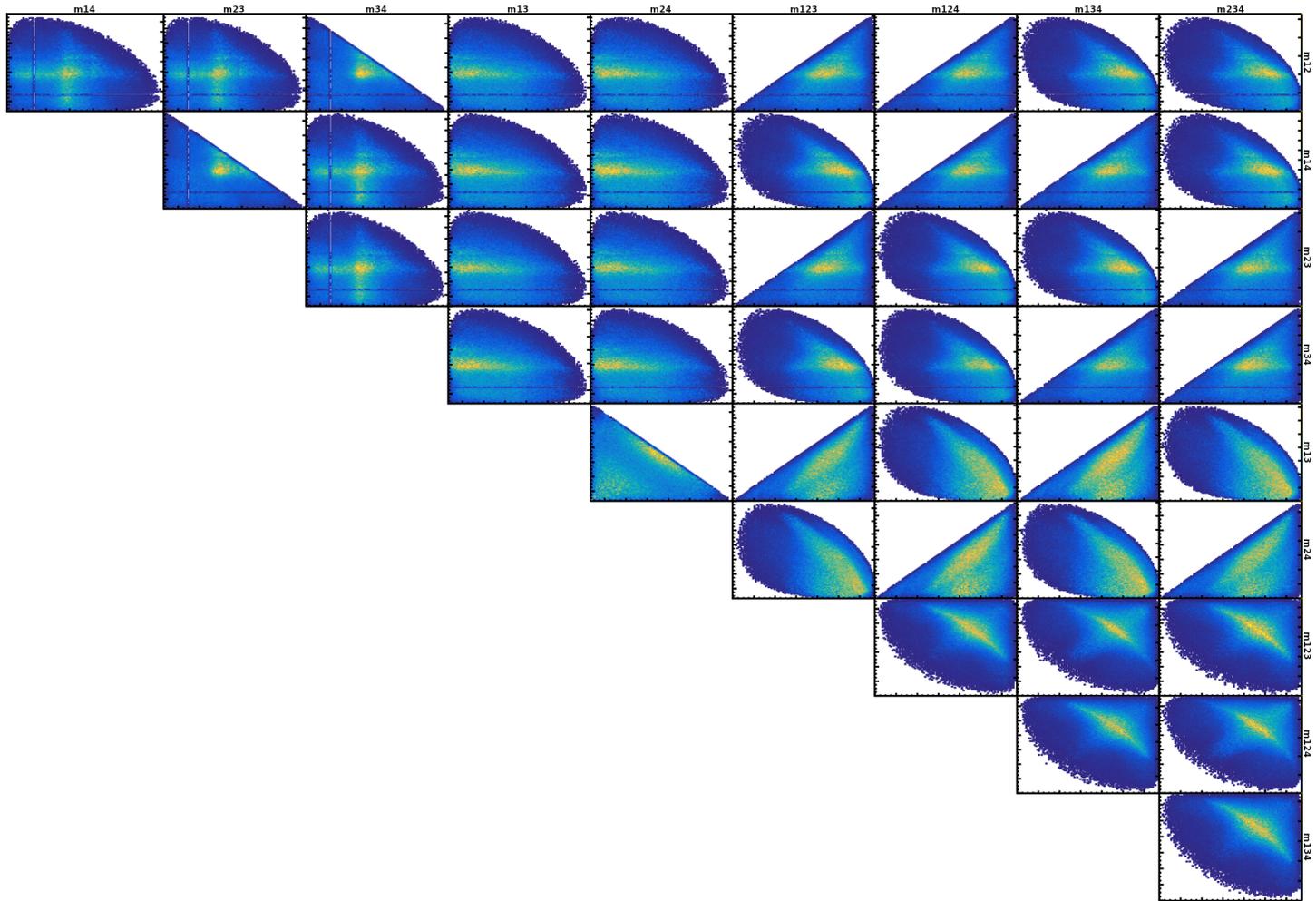


Figure 10.2: The *fit*. Phasespace-distributed events weighted with the baseline fit results (A), in all possible combinations of invariant masses of two- and three-particle subsystems.

## 10.2 Model components

The  $D^0$  decays weakly, so its daughter particles must be strange or unflavored. Since the Kaons are cut out (cf. Sec. 8.5), we are left with the unflavored daughter particles. The  $D^0$  can either decay into a charged resonance and a pion, where the charged resonance then decays into a neutral resonance and a pion; or into two neutral resonances. The neutral resonances always decay into two pions. Due to energy conservation, we must consider charged resonances in a mass range from three pion masses (419 MeV) to the  $D^0$  mass minus one pion mass (1725 MeV), and neutral resonances in a mass range from two pion masses (280 MeV) to the  $D^0$  mass minus two pion masses (1585 MeV). We can also observe the low-mass tails of resonances above these thresholds. Charge, spin, isospin and parity must also be conserved, and of course the resonances must have the required decay mode into charged pions.

The model components, together with the relative angular momenta  $J$  of the decay daughters are summarized in Tab. 10.1. For the remainder of this chapter, the model components are denoted with their short-hand notation. For the decays into a charged resonance and a pion, the charge conjugate decays are also added. The neutral unflavored resonances  $\rho^0$ ,  $f_0$  and  $f_2$  all decay into  $\pi^+\pi^-$  pairs. The  $(\pi\pi)_S$  wave is a parameterization of a  $\pi^+\pi^-$  pair in a relative  $S$  wave (cf. Sec. 9.6.5).  $\pi(1800)^+$  and  $f_2$  are above threshold.

Every model component is a decay chain from  $D^0 \rightarrow \pi^+\pi^-\pi^+\pi^-$  with defined resonances and relative angular momenta. The  $a_1^+$ ,  $P$ ,  $\rho^0$ ,  $S$  component always has a fixed amplitude of 1, and the other free amplitudes are fitted relative to it.

The masses, widths, and parameters of the resonances that were used in the fit are summarized in Tab. 10.2.

The background is modeled as a direct  $D^0 \rightarrow \pi^+\pi^-\pi^+\pi^-$  decay.

## 10.3 Fits

I performed fits with different models, with and without model selection, and with different BDT cuts. A brief overview of all fits is given in Tab. 10.3.

### 10.3.1 Fits without model selection

The baseline model contains all decays with relative angular momenta up to  $D$ , and omits the tensor mesons  $\pi_2(1670)^+$  and  $a_2(1320)^+$ . Events with a BDT value above 0.12 are selected.

In fit (A), the free amplitudes are fitted.

	$J$		$J$	short-hand notation		
$D^0 \rightarrow a_1^+ \pi^-$	$P$	$a_1^+ \rightarrow \rho^0 \pi^+$	$S$	$a_1^+, P, \rho^0, S$		
			$D$	$a_1^+, P, \rho^0, D$		
			$P$	$a_1^+, P, (\pi\pi)_S \pi^+$		
			$P$	$a_1^+, P, f_2, P$		
			$F$	$a_1^+, P, f_2, F$		
	$P$	$a_1(1420)^+ \pi^-$	$P$	$a_1(1420)^+ \rightarrow f_0 \pi^+$	$P$	$a_1(1420)^+, P, f_0, P$
	$P$	$a_1(1640)^+ \pi^-$	$P$	$a_1(1640)^+ \rightarrow \rho^0 \pi^+$	$S$	$a_1(1640)^+, P, \rho^0, S$
					$D$	$a_1(1640)^+, P, \rho^0, D$
					$P$	$a_1(1640)^+, P, f_2, P$
					$F$	$a_1(1640)^+, P, f_2, F$
	$D$	$a_2(1320)^+ \pi^-$	$D$	$a_2(1320)^+ \rightarrow \rho^0 \pi^+$	$D$	$a_2(1320)^+, D, \rho^0, D$
	$S$	$\pi(1300)^+ \pi^-$	$S$	$\pi(1300)^+ \rightarrow (\pi\pi)_S \pi^+$	$S$	$\pi(1300)^+, S, (\pi\pi)_S, S$
					$P$	$\pi(1300)^+, S, \rho^0, P$
	$D$	$\pi_2(1670)^+ \pi^-$	$D$	$\pi_2(1670)^+ \rightarrow (\pi\pi)_S \pi^+$	$D$	$\pi_2(1670)^+, D, (\pi\pi)_S, D$
					$P$	$\pi_2(1670)^+, D, \rho^0, P$
					$F$	$\pi_2(1670)^+, D, \rho^0, F$
					$S$	$\pi_2(1670)^+, D, f_2, S$
					$D$	$\pi_2(1670)^+, D, f_2, D$
	$S$	$\pi(1800)^+ \pi^-$	$S$	$\pi(1800)^+ \rightarrow f_0(1500) \pi^+$	$S$	$\pi(1800)^+, S, f_0(1500), S$
	$D^0 \rightarrow \rho^0 \rho^0$	$S$			$\rho^0, \rho^0, S$	
$P$				$\rho^0, \rho^0, P$		
$D$				$\rho^0, \rho^0, D$		
$S$		$(\pi\pi)_S (\pi\pi)_S$			$(\pi\pi)_S, (\pi\pi)_S, S$	
$S$		$(\pi\pi)_S f_0$			$(\pi\pi)_S, f_0, S$	
$D$		$(\pi\pi)_S f_2$			$(\pi\pi)_S, f_2, D$	
$S$		$f_0 f_0$			$f_0, f_0, S$	
$S$		$f_2 f_2$			$f_2, f_2, S$	
$P$		$f_2 f_2$			$f_2, f_2, P$	
$D$		$f_2 f_2$			$f_2, f_2, D$	
$F$		$f_2 f_2$			$f_2, f_2, F$	

Table 10.1: Model components used in the fits. The neutral mesons decay into  $\pi^+ \pi^-$  pairs, so each row represents one decay chain from  $D^0$  to  $\pi^+ \pi^- \pi^+ \pi^-$ , with defined resonances and relative angular momenta  $J$ . The components are referred to via the short-hand notation throughout this section.

Resonance	mass/GeV	width/GeV	other parameters
$a_1^+$	1.2299	0.459	$g_{K^*K}^2$ : 0.06
$a_1(1420)^+$	1.4105	0.1606	
$a_1(1640)^+$	1.647	0.254	
$a_2(1320)^+$	1.3145	0.1066	
$\pi(1300)^+$	1.3	0.04	
$\pi_2(1670)^+$	1.642	0.3111	
$\pi(1800)^+$	1.812	0.208	
$\rho^0$	0.775 26	0.1478	
$f_0$	0.970 09	0.0055	
$f_2$	1.2755	0.1867	

Table 10.2: Masses, widths and parameter values used in the fits.

Fit	Model	Model sel.	BDT cut	Description
A	Baseline		0.12	
B	Baseline		0.12	Fixed amplitudes, free parameters
C	Baseline		0.12	Free amplitudes, free parameters
D	Baseline	BCM	0.12	Starting with amplitudes from fit (A)
E	Baseline	BCM	0.12	Starting with normalized fit fractions
F	Baseline	BCM	0.12	Starting with amplitudes from fit (E)
G	Baseline+		0.12	Rel. angular momenta $J$ up to F
H	Extended		0.12	
I	Reduced	BCM	0.12	Normal BDT cut
J	Reduced	BCM	0.15	Increased BDT cut
K	Reduced	BCM	0.10	Decreased BDT cut
L	Extended+	BCM	0.12	Free amplitudes, free parameters, $J$ up to F

Table 10.3: Overview of the different fits.

In fit (B), the amplitudes are fixed to the results from fit (A). Instead, some of the parameters of the resonances are fitted. These are the mass, width, and  $K^* K$  coupling of the  $a_1$  mass shape; the mass of the  $f_0(980)$  Flatté mass-shape; and the mass and width of the  $\rho^0(770)$  Gounaris-Sakurai mass-shape (cf. Sec. 9.6).

In fit (C), both the amplitudes and parameters are fitted.

In fit (G), decays with relative angular momenta up to  $F$  are included (except  $f_2 f_2$ ), and the free amplitudes are fitted.

In fit (H), the tensor mesons  $\pi_2(1670)^+$  and  $a_2(1320)^+$  are added to the model, and the free amplitudes are fitted.

### 10.3.2 Fits with model selection

Fits (D), (E), (F), (I), (J), (K), and (L) employ the BCM approach for model selection (cf. Sec. 9.9), and the free amplitudes are fitted.

Fit (D) uses the resulting free amplitudes from fit (A) as start values.

For fit (E), the start values are scaled, so that all components have similar fit fractions.

Fit (F) uses the results from fit (E) as start values.

Fit (I) uses a model with less components, and decays with relative angular momenta up to  $D$ . Fits (I), (J), and (K) differ only in the BDT cuts.

Fit (L) uses the extended model of fit (H), plus additional  $f_2, f_2$  components with relative angular momenta up to  $F$ . An additional  $\rho^0$  background component is added. The free amplitudes and free parameters (like in fit (B)) are fitted.

## 10.4 Fit results

### 10.4.1 Fit results without model selection

The fit results are illustrated in Fig. 10.3 and summarized in Tab. 10.4. The resulting fit fractions for the fits without model selection are very similar. The total fit fraction is around 10 (1000 %), so there is a lot of destructive interference between the model components.

The background admixture is estimated to be 35 %. From the model selection and the given BDT cut, we would expect a background of around 8 % (cf. Tab. 8.6). In fits (C) and (H), with additional free parameters and more model components, respectively, the background admixture lowers to 29 % and 31 %, respectively.

The additional components in fit (H) have fit fractions up to the percent level, and it is the best fit, with a reduced  $\chi^2$  of 2.71.

The  $(\pi\pi)_S, (\pi\pi)_{S\text{-like}}$  components<sup>1</sup> all have large fit fractions above 1. Fig. 10.4 shows that these four components have particularly large overlap integrals between 0.76 and 0.95. These components look almost identical and cannot be clearly distinguished by the fit.

The same holds for the  $a_1^+, P, \rho^0, S$ -like components,<sup>2</sup> which have an overlap integral of 0.74.

These components account for the large total fit fraction, as you will see in the next section.

## 10.4.2 Fit results with model selection

Fit (D) yields results very similar to fit (A). The reason might be that the result of fit (A) still forms a local maximum in the parameter space for fit (D), and MINUIT is stuck in this local maximum.

In fit (E), the total fit fraction is 0.91. The model selection works well, and the fit is now generating more constructive than destructive interference. The fit fractions of components with large mutual overlap integrals ( $(\pi\pi)_S, (\pi\pi)_{S\text{-like}}$ ; and  $a_1^+, P, \rho^0, S$ -like) now all have fit fractions around 0.01 and below.

Fit (F) yields similar results.

For fit (I), some smaller components were added, but the  $\pi(1300)$  was omitted. The reduced  $\chi^2$  is 4.08, indicating a worse fit. Fits (I), (J), and (K) are discussed in Sec. 10.5.2.

In fit (L), the additional  $\rho^0$  background changes the results drastically. The flat background component now has a fit fraction of only 8%, consistent with the prediction from the model selection. The  $\rho^0$  background, however, has 28% fit fraction. The  $(\pi\pi)_S, (\pi\pi)_{S\text{-like}}$  components are suppressed, similar to fits (E) and (F), but  $a_1^+, P, \rho^0, S$  still has 25% fit fraction, and also  $a_1(1640)^+, P, \rho^0, S$  is not suppressed as much. Most other components have results similar to the baseline fit (A). The fit quality is the best of the model-selection fits, but slightly worse than in fit (H), which is expected due to the model selection.

---

<sup>1</sup> $(\pi\pi)_S, f_0, S; (\pi\pi)_S, (\pi\pi)_{S\text{-like}}, S; f_0, f_0, S; \pi(1300)^+, S, (\pi\pi)_S, S$   
<sup>2</sup> $a_1^+, P, \rho^0, S; a_1(1640)^+, P, \rho^0, S$



	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)	(L)
total fit fraction	11	11	9.1	11	0.91	1.7	11	12	1.3	1.5	0.93	1.6
flat background	0.35	0.35	0.29	0.35	0.45	0.54	0.35	0.31	0.40	0.41	0.58	0.080
$\rho^0$ background												0.28
$a_1^+, P, \rho^0, S$	0.34	0.34	0.49	0.34	0.012	0.0046	0.34	0.27	0.19	0.12	0.073	0.25
$a_1^+, P, \rho^0, D$	0.031	0.029	0.042	0.029	0.012	0.093	0.030	0.031	0.040	0.026	0.013	0.0032
$a_1^+, P, (\pi\pi)_S, P$	0.37	0.36	0.46	0.36	0.22	0.31	0.37	0.28	0.44	0.44	0.000 011	0.50
$a_1^+, P, f_2, P$	0.0016	0.0013	0.000 93	0.0013	0.016	0.039	0.000 43	0.000 12	0.012	0.0032	0.0035	0.0036
$a_1^+, P, f_2, F$							0.000 040	0.000 080				0.000 61
$a_1(1420)^+, P, f_0, P$	0.0017	0.0017	0.000 63	0.0017	0.0019	0.0070	0.0016	0.0017				0.0026
$a_1(1640)^+, P, \rho^0, S$	0.30	0.29	0.28	0.29	0.0080	0.015	0.30	0.25	0.19	0.12	0.16	0.054
$a_1(1640)^+, P, \rho^0, D$	0.000 22	0.000 22	0.000 18	0.000 22	0.0030	0.000 87	0.000 19	0.0037	0.019	0.016	0.043	0.029
$a_1(1640)^+, P, f_2, P$	0.000 15	0.000 15	0.000 069	0.000 15	0.0061	0.048	0.000 12	0.000 79	0.013	0.0015	0.000 63	0.0094
$a_1(1640)^+, P, f_2, F$							0.000 82	0.000 63				0.000 44
$a_2(1320)^+, D, \rho^0, D$								0.000 19	0.000 016	0.000 22	0.0021	0.000 054
$\pi(1300)^+, S, (\pi\pi)_S, S$	1.3	1.3	1.0	1.3	0.0018	0.003	1.3	1.2				0.0077
$\pi(1300)^+, S, \rho^0, P$	0.041	0.04	0.029	0.040	0.011	0.017	0.041	0.033				0.045
$\pi_2(1670)^+, D, (\pi\pi)_S, D$								0.044	0.025	0.0061	0.0034	0.053
$\pi_2(1670)^+, D, \rho^0, P$								0.0088	0.0049	0.000 90	0.071	0.048
$\pi_2(1670)^+, D, \rho^0, F$								0.011				0.0063
$\pi_2(1670)^+, D, f_2, S$								0.0016	0.0030	0.000 015	0.0045	0.0025
$\pi_2(1670)^+, D, f_2, D$								0.0021	0.0059	0.020	0.0068	0.0029
$\pi(1800)^+, S, f_0(1500), S$	0.036	0.035	0.021	0.035	0.027	0.019	0.035	0.046	0.14	0.10	0.014	0.053
$\rho^0, \rho^0, S$	0.13	0.13	0.11	0.13	0.45	0.69	0.13	0.14	0.32		0.59	0.15
$\rho^0, \rho^0, P$	0.000 64	0.000 63	0.000 13	0.000 63	0.0032	0.021	0.000 59	0.014	0.026	0.037	0.014	0.000 63
$\rho^0, \rho^0, D$	0.23	0.23	0.16	0.23	0.073	0.20	0.23	0.19	0.24	0.28	0.0040	0.15
$(\pi\pi)_S, (\pi\pi)_S, S$	2.9	2.8	2.2	2.8	0.0011	0.0016	2.9	2.9	0.076	0.069	0.000 76	0.0018
$(\pi\pi)_S, f_0, S$	3.4	3.3	2.7	3.3	0.013	0.048	3.4	3.9	0.023	0.29	0.0033	0.064
$(\pi\pi)_S, f_2, D$	0.002	0.002	0.0027	0.0020	0.020	0.0049	0.0020	0.0096	0.017	0.0099	0.0017	0.014
$f_0, f_0, S$	2	1.9	1.5	1.9	0.014	0.10	2.0	2.1	0.066	0.18	0.019	0.11
$f_2, f_2, S$	0.015	0.015	0.010	0.015	0.0091	0.036	0.015	0.015				0.016
$f_2, f_2, P$												0.000 19
$f_2, f_2, D$												0.000 96
$f_2, f_2, F$												0.000 005
$\ln \mathcal{L}$	50016.82	50135.60	50489.97	50016.82	41078.51	41136.89	50035.58	51540.20	45597.58	22199.46	48505.45	50309.30
$\chi^2$												
NDF	3.13	3.09	2.96	3.13			3.12	2.71	4.08	4.48	5.44	3.07

Table 10.4: Fit fraction results and goodness of fit.

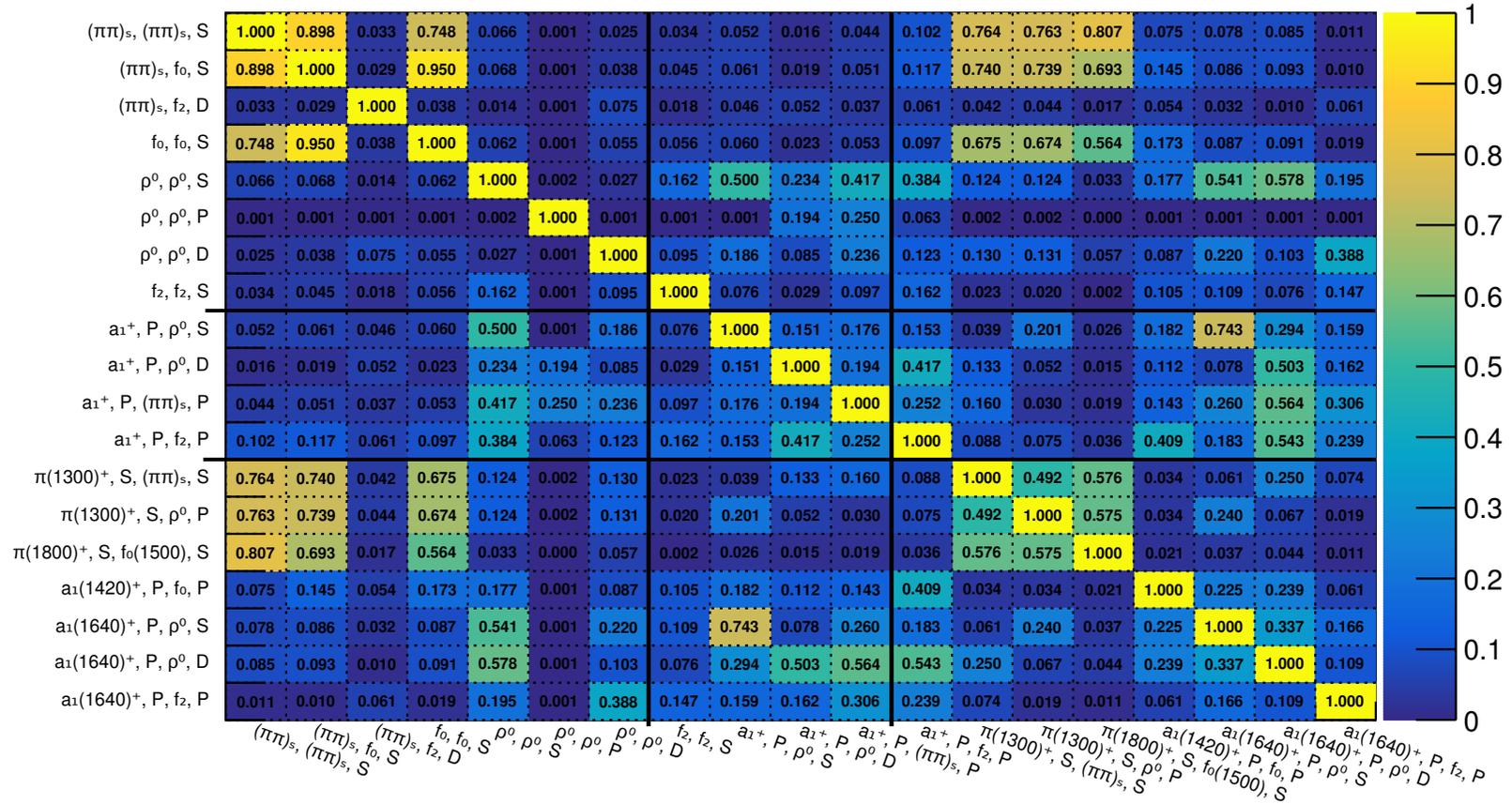


Figure 10.4: Overlap integrals of the components used in the baseline fit (A). Large overlap integrals close to 1 imply that the components look very similar.

## 10.5 Uncertainties

The statistical uncertainties of the fit fractions are evaluated with a MCMC study. The systematic influence of the BDT cut on the fit fractions is studied.

### 10.5.1 Statistical uncertainties

The statistical uncertainties for the fit fractions can be determined as follows: MINUIT gives statistical uncertainties for the real and imaginary parts of the free amplitudes. They are interpreted as Gaussian probability density functions (PDFs) and used as priors for a MCMC. It samples from the priors, i.e. it repeatedly draws the free amplitudes from the Gaussian PDFs, and histograms the resulting fit fractions. These histograms now represent the PDFs of the fit fractions. A Gaussian fit to these PDFs yields the statistical uncertainties.

The correlations between the free amplitudes are not taken into account, so this procedure yields an upper bound, and the mean values can be slightly shifted when compared to the original fit fraction results.

The results are summarized in Tab. 10.5. The uncertainties are in the single digit percent range for the larger fit fractions, and increase to 100% or more for the smallest fit fractions. Keep in mind that a fit fraction of 0.001 corresponds to only 140 events. The statistical uncertainties are small in comparison to the variations of the fit fractions between the different fits.

	(A)	(C)	(D)	(E)
$a_1^+, P, \rho^0, S$	0.3485 ± 0.0095	0.499 ± 0.020	0.3522 ± 0.0063	0.012 17 ± 0.000 08
$a_1^+, P, \rho^0, D$	0.0305 ± 0.0033	0.0403 ± 0.0083	0.0296 ± 0.0017	0.012 18 ± 0.000 33
$a_1^+, P, (\pi\pi)_S, P$	0.369 ± 0.016	0.469 ± 0.017	0.3765 ± 0.0089	0.2298 ± 0.0024
$a_1^+, P, f_2, P$	0.002 46 ± 0.001 40	0.0015 ± 0.0011	0.001 42 ± 0.000 43	0.016 65 ± 0.000 59
$a_1(1420)^+, P, f_0, P$	0.002 16 ± 0.000 62	0.000 68 ± 0.000 23	0.0030 ± 0.0029	0.0019 ± 0.0001
$a_1(1640)^+, P, \rho^0, S$	0.298 ± 0.014	0.276 ± 0.020	0.3012 ± 0.0083	0.008 08 ± 0.000 17
$a_1(1640)^+, P, \rho^0, D$	0.000 36 ± 0.000 36	0.000 27 ± 0.000 33	0.000 24 ± 0.000 15	0.003 07 ± 0.000 14
$a_1(1640)^+, P, f_2, P$	0.000 49 ± 0.000 39	0.000 21 ± 0.000 18	0.000 51 ± 0.000 43	0.006 21 ± 0.000 22
$\pi(1300)^+, S, (\pi\pi)_S, S$	1.325 ± 0.072	1.038 ± 0.052	1.329 ± 0.024	0.001 92 ± 0.000 15
$\pi(1300)^+, S, \rho^0, P$	0.0414 ± 0.0036	0.0323 ± 0.0035	0.0420 ± 0.0016	0.0112 ± 0.0011
$\pi(1800)^+, S, f_0(1500), S$	0.0366 ± 0.0044	0.0209 ± 0.0030	0.0364 ± 0.0014	0.027 20 ± 0.000 39
$\rho^0, \rho^0, S$	0.1312 ± 0.0075	0.1139 ± 0.0079	0.1370 ± 0.0043	0.4613 ± 0.0033
$\rho^0, \rho^0, P$	0.001 00 ± 0.000 63	0.000 24 ± 0.000 25	0.000 82 ± 0.000 45	0.003 06 ± 0.000 37
$\rho^0, \rho^0, D$	0.2338 ± 0.0094	0.1624 ± 0.0083	0.237 ± 0.030	0.0741 ± 0.0012
$(\pi\pi)_S, (\pi\pi)_S, S$	2.89 ± 0.12	2.23 ± 0.13	2.919 ± 0.052	0.001 087 ± 0.000 065
$(\pi\pi)_S, f_0, S$	3.38 ± 0.24	2.61 ± 0.24	3.447 ± 0.063	0.013 41 ± 0.000 59
$(\pi\pi)_S, f_2, D$	0.002 13 ± 0.000 49	0.002 83 ± 0.000 51	0.002 10 ± 0.000 22	0.020 66 ± 0.000 54
$f_0, f_0, S$	1.98 ± 0.10	1.571 ± 0.083	1.991 ± 0.037	0.013 77 ± 0.000 28
$f_2, f_2, S$	0.000 69 ± 0.000 23	0.000 35 ± 0.000 12	0.000 698 ± 0.000 099	0

Table 10.5: Statistical uncertainties of the fit fractions.

### 10.5.2 Fits with different BDT cuts

The systematic uncertainties due to the choice of the BDT cut are evaluated. Fit (I) is repeated with a higher BDT cut of 0.15 (Fit (J)), and with a lower BDT cut of 0.10 (fit (K)). The results are shown in Fig. 10.5.

Fit (I) and (J) yield similar results for the components with fit fractions greater than 0.1, except for the  $(\pi\pi)_S, f_0, S$  component, which is an order of magnitude larger in fit (J). The fit fraction of the background increases slightly in fit (J), even though there should be less background in the sample with the higher BDT cut. Most of the smaller components have smaller fit fractions in fit (J), but they suffer from larger statistical uncertainties, since the higher BDT cut yields only half as many events. Fit (J) has a larger reduced  $\chi^2$  than fit (I).

In fit (K), the background has a higher fit fraction of 0.58. While it should contain more background than fit (I), the value is still much higher than the 12% expected from the event selection (cf. Tab. 8.6). The  $(\pi\pi)_S, (\pi\pi)_S$ -like components are strongly suppressed, and the  $a_1^+, P, (\pi\pi)_S, P$  component seems to vanish almost completely. The goodness of fit is poor, with a reduced  $\chi^2$  of 5.44.

This study shows that a BDT cut of 0.12 is a reasonable choice. A higher BDT cut does not change the results significantly, but yield much less events, while a lower BDT cut seems to add more background which disturbs the fit drastically.

## 10.6 Discussion

I will discuss the results and compare them to previous measurements: Two PWAs by the FOCUS [46] and CLEO-c [26] collaborations, which are based on  $6360 \pm 115$  events, and  $7250 \pm 56$  (stat)  $\pm 46$  (syst) events, respectively; and results from diffractive production of  $\pi^- \pi^- \pi^+$  final states from the COMPASS collaboration [10], which are based on  $6.4 \times 10^9$  events.

### 10.6.1 Background

The fitted background admixture (except in fit (L)) is around 5 times as large as expected from the event selection. This points out that part of the signal is mistaken as background, or the model is missing a flat component. On the other hand, fit (C) and (H) show that the issue can be partly remedied, and the goodness of fit can be improved, by freeing additional parameters of the components, and by adding more components to the model.

Fits (J) and (K) suggest that the description of the background as a single flat component is simplistic, and some components of the background are mistaken as signal components.



Figure 10.5: Fit fractions for fits with different BDT-cut values. Lines connecting the points are guides to the eye.

A MC study shows that the background consists of many components, but often pions have a common mother resonance. E.g. around 5% of the pion pairs originate from a  $\rho^0$  decay.

Adding such an incoherent  $\rho^0$  background (fit (L)), brings down the flat background to a level perfectly consistent with the expectation from the event selection. However, the  $\rho^0$  background contribution is now too large. It has a fit fraction of 28% instead of the expected 0.4%. This suggests that there is now significant leakage of signal components to the  $\rho^0$  background. The total amount of background is still overestimated.

I also tried to deduce an effective background distribution from events with low BDT values. However, the BDT cut affects the background distribution, and different slices of BDT values look different.

### 10.6.2 $a_1$

The  $a_1^+, P, \rho^0, S$  should be the dominant component with around 40% fit fraction, according to previous analyses. In the fits without model selection, the fit fractions are between 27% and 49%, which is more or less compatible with the FOCUS and CLEO-c results ( $(43.3 \pm 2.5 \pm 1.9)\%$  and  $(36.7 \pm 2.4)\%$ ). COMPASS, which measured the relative intensities of three-pion final states, measured a relative intensity of 32.7% for the  $\rho^0 \pi S$  component.

However, the  $a_1(1640)^+, P, \rho^0, S$  has to be included in this discussion. It has a large overlap integral with the  $a_1^+, P, \rho^0, S$ , and therefore these components look very similar. In the fits without model selection, it has a similar fit fraction as the  $a_1^+, P, \rho^0, S$ . In the fits (E) and (F), the fit fraction of both components are strongly suppressed by the model selection, in spite of the prior for the  $a_1^+, P, \rho^0, S$  fit fraction.<sup>3</sup> This suggests that a large part of the fit fraction of the  $a_1(1640)^+, P, \rho^0, S$  in the fits without model selection originates from destructive interference with the  $a_1(1640)^+, P, \rho^0, S$ .

On the other hand, in fit (L), the  $a_1^+, P, \rho^0, S$  fit fraction is 25%, while the  $a_1(1640)^+, P, \rho^0, S$  fit fraction is only 5.4%. Part of the signal in fit (L) might also have leaked to the  $\rho^0$  background component.

The  $a_1^+, P, \rho^0, P$  component is suppressed with respect to the  $S$  component by about a factor of 10 in most fits, which is also consistent with the results from FOCUS and COMPASS.

---

<sup>3</sup>cf. Sec. 9.9. The actual  $\ln \mathcal{L}$  penalties for small fit fractions are  $\mathcal{O}(10)$ . This is small in comparison to the total  $\ln \mathcal{L}$ , and the differences in  $\ln \mathcal{L}$  between the fits.

The  $a_1^+, P, (\pi\pi)_S, P$  has a fit fraction of around 37%. This result is quite consistent across all fits, and in the same region as the  $a_1^+, P, \rho^0, S$  fit fraction. In the other measurements, the  $a_1^+, P, (\pi\pi)_S, P$  is always significantly smaller than the  $a_1^+, P, \rho^0, S$ . FOCUS and CLEO-c measured fit fractions of  $(8.3 \pm 0.7 \pm 0.6)\%$  and  $(10.9 \pm 1.5)\%$  for  $a_1^+, P, \sigma, P$ . However, the  $\sigma$  or  $f_0(500)$  is not very well defined, therefore we replaced it with the more general  $(\pi\pi)_S$ . COMPASS, which used the same  $(\pi\pi)_S$  parameterization, also measured a small relative fit fraction of only 8.0%.

All other  $a_1$  components, as well as the heavier  $a_1(1420)^+$  and  $a_1(1640)^+$ , have comparably small fit fractions, which is also consistent with the other measurements.

In conclusion, the  $a_1$  results are comparable—by their order of magnitude—with other analyses, except for the  $a_1^+, P, (\pi\pi)_S, P$ , which seems to be too large.

### 10.6.3 $(\pi\pi)_S, (\pi\pi)_S$ -like

$(\pi\pi)_S, f_0, S; (\pi\pi)_S, (\pi\pi)_S, S; f_0, f_0, S;$  and  $\pi(1300)^+, S, (\pi\pi)_S, S$  have either large fit fractions greater than 100%, which is unphysical, or very small fit fractions when the model selection is enabled. This makes it hard to discuss their fit fractions quantitatively.

### 10.6.4 $\rho^0$

The  $\rho^0, \rho^0$  components have a total fit fraction around 34%, while FOCUS and CLEO-c measured  $(24.5 \pm 1.3 \pm 1.0)\%$  and 22.7%, so the results are comparable in this regard. FOCUS measured in the transversity basis, so we can only compare the perpendicular component (which corresponds to the  $P$  component). FOCUS and CLEO-c measured  $(6.4 \pm 0.6 \pm 0.5)\%$  and  $(7.1 \pm 0.5)\%$ , while I measured below 0.1%. This is a significant discrepancy.

The  $D$  wave is not suppressed with respect to the  $S$  wave, as one would expect due to the higher relative angular momentum. In most fits, it even has the larger fit fraction. This is consistent with the CLEO-c results, where the  $D$  wave is also enhanced.

### 10.6.5 Heavier resonances

The  $\pi_2(1670)^+$  is just below threshold, the  $\pi(1800)^+$  and  $f_2, f_2, S$  are above threshold.  $\pi_2(1670)^+$  and  $f_2$  also have a large spin of 2. Therefore we expect them to

have small fit fractions. The resulting fit fractions are in the percent range and below, are consistent across the fits, and consistent with CLEO-c results.

# Chapter 11

## Conclusion and outlook

I performed two classes of fits, with and without model selection. There are two groups of similar components with large overlap integrals: The  $(\pi\pi)_S$ ,  $(\pi\pi)_S$ -like components, and the  $a_1^+$ ,  $P$ ,  $\rho^0$ ,  $S$ -like components. These components, especially the  $(\pi\pi)_S$ ,  $(\pi\pi)_S$ -like, have large fit fractions in the fits without model selection. This is because the fit cannot distinguish them well enough, and creates large destructive interference between those components, which yields slightly larger likelihood values, but also large total fit fractions  $\mathcal{O}(10)$ .

The model selection penalizes large fit fractions, so mostly the  $(\pi\pi)_S$ ,  $(\pi\pi)_S$ -like and  $a_1^+$ ,  $P$ ,  $\rho^0$ ,  $S$ -like components get much smaller fit fractions, which also brings the total fit fractions much closer to 1.

In this regard, all fit results are reasonably consistent with each other, and the differences between the fits with and without model selection are comprehensible.

The other components are effected much less by the model selection, and get similar fit fractions in both classes of fits. When compared to previous measurements, my results are in the same ballpark, except for the  $a_1^+$ ,  $P$ ,  $(\pi\pi)_S$ ,  $P$ , for which I measure a larger fit fraction, and the  $\rho^0$ ,  $\rho^0$ ,  $S$ , which is strongly suppressed in my fits.

The model of the decay is not perfect. This is most apparent from the fit fraction results of the background. It is inconsistent with the expectations from the model selection. Even after adding another background component, the total background fit fraction is still too high. On the other hand, this implies that a significant portion of the signal is fitted as background. The model could possibly be improved by adding more signal and background components, but also the existing components could be improved. The Breit-Wigner mass shape is suited to describe narrow resonances, but the heavier charged resonances are relatively broad.

PWA is a complex analysis technique, and very hard to master. Consider the two previous analyses of the same channel: They roughly agree on the fit fractions for the  $a_1$  and the  $\rho^0\rho^0$  components, but otherwise their results are incompatible, if only due to their different choice of model components. The COMPASS data show that already the three-pion subsystem has a very rich resonant structure. Even with several billion events, it is not straightforward to analyze.

The model obviously has a big influence on the results. To eliminate any bias from the experimenter, every component which is possible according to the conservation laws should be included, and the fit and the model selection should figure out which components to take into account. But even then, not all possible resonances are well established, and some have large uncertainties of their masses and widths. Moreover, the mathematical description of the mass shapes is not trivial. Possible openings of other channels and varying partial widths in the mass range have to be taken into account. Another approach would be a model-independent fit, where the mass shapes are fully determined by the fit. This, however, comes at the cost of many more model parameters, and ambiguities which would otherwise be resolved by the isobar model.

In the first part of my thesis it became clear that an open-source framework like GENFIT, with a strong user and developer base, and validated and tested code, brings great benefits for the whole community. The same is true for PWA. While YAP certainly is not as successful as GENFIT in this regard, there are other PWA frameworks, like ROOTPWA [4] or ComPWA [5], which are actively developed and maintained. In hindsight, it might have been more profitable—for our own analyses and the PWA community—to use and contribute to an existing PWA framework instead of writing our own from scratch.

The Belle II experiment is already running, and in within a few years it will collect 40 times more data than its predecessor. This makes it even more important to have a set proven tools and accurate models for physics analyses. At the same time, it will enable much more precise measurements, and searches for new physics beyond the standard model.

# Own contributions

## Genfit

My work on GENFIT started as a technical student in 2010, when I developed the Runge Kutta track representation (`RKTrackRep`), based on a Runge Kutta code used in COMPASS, and material effect calculation from GEANE [22]. This work was supervised by Christian Höppner (Technische Universität München (TUM)).

During my Diploma thesis at TUM, I wrote `GFRave`, an interface to the vertex fitting framework `RAVE`.

From the start of my doctoral thesis in 2012 until 2016, I was the maintainer of GENFIT and a member of the Belle II tracking group.

Driven by the requirements of Belle II, I started working on GENFIT 2 together with Tobias Schlüter (Ludwig-Maximilians-Universität München). I did the main work of designing and implementing the new software and many new features, like the Kalman fitter with reference track.

We also adapted the Belle II analysis framework (`basf2`) tracking code to work with the new version, and in 2014 Belle II successfully switched to GENFIT 2.

I presented GENFIT at the following conferences:

- DPG Frühjahrstagung, Frankfurt, Germany, 2014.
- DPG Frühjahrstagung, Mainz, Germany, 2014.
- 16th International workshop on Advanced Computing and Analysis Techniques in physics research (ACAT), Prague, Czech Republic, 2014.

Together with T. Schlüter, I published a conference proceeding, which has over 30 citations:

- J. Rauch and T. Schlüter. GENFIT – a Generic Track-Fitting Toolkit. *J. Phys. Conf. Ser.*, 608(1):012042, 2015.

I am co-author of the following papers:

- T. Bilka, G. Casarosa, R. Frühwirth, P. Kodys, P. Kvasnicka, J. Lettenbichler, E. Paoloni, J. Rauch, and T. Schlüter.  
Demonstrator of the Belle II online tracking and pixel data reduction on the High Level Trigger system.  
*2014 19th IEEE-NPSS Real Time Conference*, pages 1–4, 2014.
- T. Bilka, N. Braun, T. Hauth, T. Kuhr, L. Lavezzi, F. Metzner, S. Paul, E. Prencipe, M. Prim, J. Rauch, J. Ritman, T. Schlüter, and S. Spataro.  
Implementation of GENFIT2 as an experiment independent track-fitting framework. 2019.

## YAP

Together with Daniel Greenwald (TUM), I developed Yet Another Partial Wave Analysis Toolkit (YAP). The bulk of coding was done from 2015 to 2017, and I contributed about half of the code.

We presented YAP at the DPG Frühjahrstagung, Hamburg, Germany, 2016.

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

D. Greenwald developed frameworks for event selection and analysis on top of the Belle analysis framework (basf). Based on them, I developed the event selection, fits of the signal- and background-shapes and the event selection with boosted decision trees (BDTs).

I developed a signal model in YAP, used the Bayesian Analysis Toolkit (BAT) for fitting, and further modified YAP according to my needs during the analysis.

I presented my analysis at the following conferences:

- 53rd International Winter Meeting on Nuclear Physics, Bormio, Italy, 2015.
- DPG Frühjahrstagung, Wuppertal, Germany, 2015.
- 9th Annual Meeting of the Helmholtz Alliance “Physics at the Terascale”, Hamburg, 2015.
- DPG Frühjahrstagung, Hamburg, Germany, 2016.
- DPG Frühjahrstagung, Münster, Germany, 2017.

I am a co-author of this conference proceeding:

- D. Greenwald, A. Hönle, D. Levit, S. Paul, J. Rauch, and A. Tsipenyuk.  
Singly Cabibbo Suppressed Charm Decays: CP Violation and Amplitude Analysis.  
*PoS*, Bormio 2015(054)



# List of Figures

1	Geiger and Rutherford [1]. . . . .	1
2	The Belle II detector [2]. . . . .	1
3	Particle tracks from the Big European Bubble Chamber [3]. . . . .	2
1.1	Tracking in the Belle II experiment. . . . .	7
2.1	Track parametrization. . . . .	10
3.1	GENFIT unified modeling language (UML) class diagram. . . . .	14
4.1	GENFIT event display screenshots. . . . .	25
7.1	The KEKB $B$ factory. . . . .	34
7.2	Cutaway view of the Belle detector and its subdetectors. . . . .	35
7.3	Integrated luminosities of the $B$ factories. . . . .	36
8.1	MC event distributions of signal and background contributions. . . . .	43
8.2	Fit of the signal and background shapes to the MC data set. . . . .	45
8.3	Signal and background components of the fit to the MC data set. . . . .	46
8.4	Fit of the signal and background shapes to the real data set. . . . .	47
8.5	Example of a decision tree. . . . .	48
8.6	Input variable distributions of the MC training set. . . . .	50
8.7	Distribution of the BDT response for MC and data. . . . .	51
9.1	Helicity frame. . . . .	57
9.2	Helicity angles of a decay into two isobars. . . . .	58
9.3	Phase of the $f_0(980)$ mass shape. . . . .	61
9.4	Argand plot of the $f_0(980)$ mass shape. . . . .	61
9.5	Intensity of the $f_0(980)$ mass shape. . . . .	61
9.6	Phase of the $\rho^0$ mass shape. . . . .	62
9.7	Intensity of the $\rho^0$ mass shape. . . . .	62
9.8	Mass dependent width of the $a_1$ mass shape. . . . .	64
9.9	Phase of the $a_1$ mass shape. . . . .	64

---

9.10	Argand plot of the $a_1$ mass shape. . . . .	64
9.11	Intensity of the $a_1$ mass shape. . . . .	64
9.12	Phase of the $(\pi\pi)_S$ mass shape. . . . .	65
9.13	Argand plot of the $(\pi\pi)_S$ mass shape. . . . .	65
9.14	Intensity of the $(\pi\pi)_S$ mass shape. . . . .	65
10.1	The data. . . . .	72
10.2	The fit. . . . .	73
10.3	Fit fractions of the components of all models and fits. . . . .	79
10.4	Overlap integrals of the components used in the baseline fit. . . . .	81
10.5	Fit fractions for fits with different BDT-cut values. . . . .	85

# List of Tables

4.1	Execution time of the GenFitter module in basf2, without and with material effect calculation, and average number of iterations. . . . .	24
8.1	Event selection cuts. . . . .	42
8.2	Number of parameters for the signal and background components and scaling. . . . .	44
8.3	Ranges and number of bins of the fit and signal region. . . . .	44
8.4	Fitted signal and background yields of the data set. . . . .	45
8.5	Average number of generated and detectable $D^*$ per stream from MC simulations. . . . .	46
8.6	Number of events and expected purity of the data for different BDT cuts. . . . .	51
10.1	Model components used in the fits. . . . .	75
10.2	Masses, widths and parameter values used in the fits. . . . .	76
10.3	Overview of the different fits. . . . .	76
10.4	Fit fraction results and goodness of fit. . . . .	80
10.5	Statistical uncertainties of the fit fractions. . . . .	83



# Bibliography

- [1] <http://ecuip.lib.uchicago.edu/multiwavelength-astronomy/images/gamma-ray/science/ernest-rutherford-and-hans-geiger.jpg>.
- [2] [https://upload.wikimedia.org/wikipedia/commons/thumb/6/67/KEK\\_Belle\\_II\\_Detector\\_%281%29.jpg/1280px-KEK\\_Belle\\_II\\_Detector\\_%281%29.jpg](https://upload.wikimedia.org/wikipedia/commons/thumb/6/67/KEK_Belle_II_Detector_%281%29.jpg/1280px-KEK_Belle_II_Detector_%281%29.jpg).
- [3] <https://cds.cern.ch/record/39469>.
- [4] <https://github.com/ROOTPWA-Maintainers/ROOTPWA>.
- [5] <https://github.com/CompPWA/CompPWA>.
- [6] COMPASS, 2019. <http://wwwcompass.cern.ch/>.
- [7] PHAST (COMPASS), 2019. <http://ges.home.cern.ch/ges/phast/>.
- [8] A. Abashian et al. The Belle Detector. *Nucl. Instr. Meth. Phys. Res. A*, 479:117–232, 2002.
- [9] I. Abe et al. KEK-B: The KEK B-factory. *Nucl. Instr. Meth. Phys. Res. A*, 499(1), 2003. All papers included in this volume.
- [10] C. Adolph et al. Resonance production and  $\pi\pi$   $S$ -wave in  $\pi^- + p \rightarrow \pi^- \pi^- \pi^+ + p_{\text{recoil}}$  at 190 GeV/c. *Phys. Rev. D*, 95:032004, Feb 2017.
- [11] Andrey Alexandrov et al. The FOOT Experiment: Fragmentation Measurements In Particle Therapy. 03 2019.
- [12] B. D.O. Anderson and J. B. Moore. *Optimal Filtering*. Dover Publications, 1979.
- [13] M. Antonova et al. Baby MIND: A magnetised spectrometer for the WA-GASCI experiment. 04 2017.

- 
- [14] K. L. Au, D. Morgan, and M. R. Pennington. Meson dynamics beyond the quark model: Study of final-state interactions. *Phys. Rev. D*, 35:1633–1664, Mar 1987.
- [15] Elisabetta Barberio and Zbigniew Wąs. PHOTOS – a universal Monte Carlo for QED radiative corrections: version 2.0. *Computer Physics Communications*, 79(2):291 – 308, 1994.
- [16] Belle II Collaboration. *Belle II Technical Design Report*. KEK, 2010.
- [17] <http://belle.kek.jp>.
- [18] Karl Alexander Bicker. Model Selection for and Partial-Wave Analysis of a Five-Pion Final State at the COMPASS Experiment at CERN. *Dissertation, Technische Universität München*, 2016.
- [19] G. Breit and E. Wigner. Capture of slow neutrons. *Phys. Rev.*, 49:519–531, 1936.
- [20] R. Brun and F. Rademakers. ROOT Framework. <http://root.cern.ch>.
- [21] A. Caldwell, D. Kollar, and K. Kröninger. BAT – The Bayesian Analysis Toolkit. *Comput. Phys. Commun.*, 180:2197–2209, 2009.
- [22] CERN. GEANE – Average Tracking and Error Propagation Package. *CERN Program Library*, W5013-E, 1991.
- [23] CERN. GEANT – Detector description and simulation tool. *CERN Program Library Long Write-up*, W5013, 1993.
- [24] CMS Collaboration. The CMS experiment at the CERN LHC. *J. Instrum.*, 0803:S08004, 2008.
- [25] Daniel Greenwald, Johannes Rauch, and Paolo Di Giglio. YAP – Yet Another Partial Wave Analysis Toolkit. <https://github.com/yap/YAP>.
- [26] P. d’Argent, N. Skidmore, J. Benton, J. Dalseno, E. Gersabeck, S. T. Harnew, P. Naik, C. Prouve, and J. Rademacker. Amplitude analyses of  $D^0 \rightarrow \pi^+\pi^-\pi^+\pi^-$  and  $D^0 \rightarrow K^+K^-\pi^+\pi^-$  decays. *J. High Energy Phys.*, 2017(5):143, 2017.
- [27] L. Fabbietti et al. The PANDA GEM-based TPC prototype. *Nucl. Instr. Meth. Phys. Res. A*, 628:204–208, 2011.
- [28] S. M. Flatté. On the Nature of  $0^+$  Mesons. *Phys. Lett. B*, 63:228–230, 1976.

## BIBLIOGRAPHY

---

- [29] A. Fontana, P. Genova, L. Lavezzi, and A. Rotondi. Track Following in Dense Media and Inhomogeneous Magnetic Fields. *PANDA Report PV/01-07*, 2007.
- [30] Professor E. Rutherford F.R.S. The scattering of  $\alpha$  particles by matter and the structure of the atom. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 21(125):669–688, 1911.
- [31] R. Frühwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instr. Meth. Phys. Res. A*, 262:444–450, 1987.
- [32] R. Frühwirth, M. Regler, R.K. Bock, H. Grote, and D. Notz. *Data Analysis Techniques for High-Energy Physics*. Cambridge University Press, 2000.
- [33] R. Frühwirth and A. Strandlie. Track Fitting with Ambiguities and Noise: A Study of Elastic Tracking and Nonlinear Filters. *Comput. Phys. Commun.*, 120(2-3):197–214, 1999.
- [34] Maury Goodman. The Deep Underground Neutrino Experiment. *Advances in High Energy Physics*, 2015:1–9, 12 2015.
- [35] G. J. Gounaris and J. J. Sakurai. Finite-width corrections to the vector-meson-dominance prediction for  $\rho \rightarrow e^+e^-$ . *Phys. Rev. Lett.*, 21:244–247, 1968.
- [36] David J. Herndon, Paul Söding, and Roger J. Cashmore. Generalized isobar model formalism. *Phys. Rev. D*, 11:3165–3182, Jun 1975.
- [37] Virgil L. Highland. Some practical remarks on multiple scattering. *Nuclear Instruments and Methods*, 129(2):497 – 499, 1975.
- [38] A. Hoecker et al. TMVA: Toolkit for Multivariate Data Analysis. *PoS, ACAT:040*, 2007.
- [39] C. Höppner, S. Neubert, B. Ketzer, and S. Paul. A Novel Generic Framework for Track Fitting in Complex Detector Systems. *Nucl. Instr. Meth. Phys. Res. A*, 620:518–525, 2010.
- [40] C. Höppner, S. Neubert, J. Rauch, and T. Schlüter. <https://github.com/GenFit/GenFit>.
- [41] R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *Trans. ASME*, 82(Series D):35–45, 1960.
- [42] Kim, Doris Yangsoo. The Software Library of the Belle II Experiment. *ICHEP 2014, Valencia, Spain*.

- 
- [43] C. Kleinwort. General Broken Lines as advanced track fitting method. *Nucl. Instr. Meth. Phys. Res. A*, 673:107–110, 2012.
- [44] D. J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth.*, A462:152–155, 2001.
- [45] Peter Lichard and Martin Vojik. An Alternative parametrization of the pion form-factor and the mass and width of rho(770). 2006.
- [46] J. M. Link et al. Study of the  $D^0 \rightarrow \pi^- \pi^+ \pi^- \pi^+$  decay. *Phys. Rev. D*, 75:052003, 2007.
- [47] Moritz Nadler. Material estimation and low momentum tracking in Belle II. *Dissertation, Institut für Hochenergiephysik*, 2013.
- [48] K. A. Olive et al. *Review of Particle Physics*, volume 38. 2014.
- [49] PANDA Collaboration. PANDA: *Technical Progress Report*. GSI FAIR, 2005.
- [50] Trevor Park and George Casella. The Bayesian Lasso. *Journal of the American Statistical Association*, 103(482):681–686, 2008.
- [51] J. Rauch and T. Schlüter. GENFIT – a Generic Track-Fitting Toolkit. *J. Phys. Conf. Ser.*, 608(1):012042, 2015.
- [52] M.E. Rose. *Elementary Theory of Angular Momentum*. Dover books on physics and chemistry. Dover, 1995.
- [53] R. Shane et al. SpiRIT: A time-projection chamber for symmetry-energy studies. *Nucl. Instrum. Meth.*, A784:513–517, 2015.
- [54] Torbjörn Sjöstrand, Patrik Edén, Christer Friberg, Leif Lönnblad, Gabriela Miu, Stephen Mrenna, and Emanuel Norrbin. High-energy-physics event generation with Pythia 6.1. *Computer Physics Communications*, 135(2):238 – 259, 2001.
- [55] R. M. Sternheimer and S. J. Lindenbaum. Extension of the isobaric nucleon model for pion production in pion-nucleon, nucleon-nucleon, and antinucleon-nucleon interactions. *Phys. Rev.*, 123:333–376, Jul 1961.
- [56] <http://www.theoretical-physics.net/dev/math/wigner.html>.
- [57] <https://upload.wikimedia.org/wikipedia/commons/c/c2/KEKB.png>.
- [58] W. Waltenberger. RAVE: A Detector-independent Toolkit to Reconstruct Vertices. *IEEE Trans. Nucl. Sci.*, 58:434–444, 2011.

## BIBLIOGRAPHY

---

[59] <http://www-f9.ijs.si/~krizan/sola/tokyo/niceBelle.jpg>.



# Danksagung

Ich möchte Prof. Stephan Paul danken, der es mir ermöglicht hat, diese Doktorarbeit zu schreiben, und meine Ergebnisse auf Konferenzen zu präsentieren. Ebenfalls möchte ich Daniel Greenwald danken, der diese Arbeit betreute und immer offen für Fragen und Diskussionen war.

Ein besonderer Dank gebührt Florian Haas und der Firma Actindo, die es mir ermöglichten, meine Doktorarbeit neben einer Teilzeitanstellung fertigzustellen.

Alle drei standen mir, besonders zum Ende dieser Arbeit, mit Rat und Tat zur Seite.

Ich möchte Christian Höppner danken, der mich bereits als Werkstudent betreute, mir die Verantwortung für GENFIT übertrug, und mich in die Belle II Tracking-Gruppe brachte. Weiterer Dank gilt den Kollegen, die ebenfalls zu GENFIT beitrugen, allen voran Tobias Schlüter, mit dem ich die zweite Version von GENFIT konzeptionierte und entwickelte. Ebenfalls beigetragen haben Sverre Dorheim, Martin Berger, Karl Bicker und Felix Böhmer, und, von anderen Instituten, Moritz Nadler, Tadeas Bilka, Sergey Yashchenko und Philipp Öhler. Besonderer Dank gilt auch den Kollegen vom KIT, die GENFIT weiter pflegen, allen voran Markus Prim und Nils Braun.

Ich möchte auch den Mitgliedern des Belle II Experiments, und vor allem der Tracking-Gruppe danken, u.a. Thomas Kuhr, Martin Heck, Eugenio Paoloni, Christian Kiesling, Martin Ritter, Jakob Lettenbichler, Giulia Casarosa und Oliver Frost. Wir hatten viele tolle und konstruktive Meetings in schönen Städten wie Bad Aibling, Tsukuba, Hamburg, Leinsweiler, Prag und Pisa.

Weiterer Dank gilt Elisabetta Prencipe, die die Integration der neuen GENFIT-Version in  $\bar{P}$ ANDA vorantrieb, und den Hauptteil der Arbeit an der neusten Veröffentlichung stemmte.

Ich möchte dem Tracking-Experten Rudolf Frühwirth danken, der Tobias und mich nach HEPHY nach Wien einlud, um gemeinsam an GENFIT zu arbeiten.

Mein weiterer Dank gilt den aktiven und ehemaligen Mitgliedern von E18: Boris Grube, der immer für Fragen und Diskussionen zur Verfügung stand, Igor

Konorov, der sich beharrlich nach dem Stand meiner Doktorarbeit erkundigte, Christian Dreisbach, der immer für einen Plausch zu haben war, meine (ehemaligen) Bürokollegen Stefan Wallner, Florian Kaspar und Dominik Steffen. Außerdem unseren Belle II Masteranden, Andreas Hönle, Arseniy Tsipenyuk und Paolo di Giglio, der auch zu YAP beitrug. Mit Arseniy an der Geige und Paolo an der Gitarre durften wir auch einmal für die musikalische Untermalung der E18 Weihnachtsfeier sorgen. An dieser Stelle darf natürlich unsere geschätzte Sekretärin Karin Frank nicht fehlen, die nicht nur bei organisatorischen Fragen immer hilfsbereit war.

Natürlich möchte ich auch den vielen hier nicht namentlich aufgeführten Menschen danken, die ich während meiner langen Zeit am Lehrstuhl und als Physiker kennenlernen durfte, und bei den Steuerzahlerinnen und -zahlern, die Forschungsarbeiten wie diese erst ermöglichen.

Zu guter Letzt möchte ich meiner Familie, meinen Freunden, und meiner Freundin Tabea König danken, die trotz der langen Zeit, die diese Arbeit in Anspruch nahm immer zu mir gestanden und mich unterstützt haben. Danke.