Technische Universität München
Fakultät für Physik

Abschlussarbeit im Masterstudiengang Physik

Application and Verification of Model-Selection Techniques for Diffractively Produced Three-Pion Final States

Florian Markus Kaspar

02. Mai 2017

Erstgutachter (Themensteller): Prof. Dr. Stephan Paul
Zweitgutachter: Prof. Dr. Allen Caldwell

Abstract

Quantum chromodynamics (QCD) successfully describes the strong interaction in the high-energy limit by means of perturbation theory. Unfortunately, these calculations break down in the low-energy regime. To study QCD at low energies one has to rely on numerical lattice QCD simulations, effective theories, or models to explain the experimental findings. The light-meson spectrum is an important probe for low-energy QCD. Light-meson resonances can for example be produced as intermediate states in diffractive dissociation reactions.

The COMPASS experiment at CERN acquired a large dataset for the diffractive scattering reaction $\pi^- + p \to \pi^-\pi^-\pi^+ + p_{\text{recoil}}$.
To extract information about the produced intermediate three-pion resonances, a technique called partial-wave decomposition is applied, which decomposes the measured intensity distribution into the contributions from intermediate resonances with different quantum numbers. The method relies on a model with partial waves as model components that need to be selected carefully to ensure sensible inference. Up to now the relevant partial waves have been selected by hand. The selection procedure is complicated due to strong interplay of the model components and therefore hand-selected models may not be optimal.
With the large dataset available from COMPASS, one would like to improve the analysis to study smaller effects in detail. These advances in data analysis require a more systematic way of selecting the relevant partial waves. Recently, two model-selection methods based on regularized maximum likelihood estimation have been developed. In this thesis the different model-selection procedures are applied and verified on simulated and measured data. The results lead to the development of a new method that combines the advantages of the previously introduced approaches. This new model-selection procedure has then successfully been applied to measured data, solving issues of the other selection procedures. The method is able to systematically select sets of waves for the partial-wave decomposition in a fast and practical manner.

Kurzfassung

Die starke Wechselwirkung im Bereich hoher Energien lässt sich erfolgreich durch die Quantenchromodynamik (QCD) beschreiben. Für niedrige Energien brechen diese Berechnungen leider zusammen. Um QCD bei niedrigen Energien zu untersuchen, muss man auf numerische Berechnungen mittels Gitter-QCD und effektive Feldtheorien zurückgreifen, welche die experimentellen Ergebnisse modellieren. Das Spektrum leichter Mesonen stellt einen Zugang für Niederenergie-QCD-Tests dar. Leichte Mesonresonanzen können beispielsweise als Zwischenzustände in diffraktiven Streuprozessen erzeugt werden. Das COMPASS Experiment am CERN hat einen großen Datensatz zum diffraktiven Streuprozess $\pi^- + p \to \pi^-\pi^-\pi^+ + p_{\text{recoil}}$ aufgezeichnet. Um Informationen über die im Zwischenzustand erzeugten Drei-Pion-Resonanzen zu extrahieren, wird die sogenannte Partialwellenzerlegung angewandt, welche die gemessene Intensitätsverteilung in Beiträge von Zwischenresonanzen, verschiedener Quantenzahlen, zerlegt. Die Methode basiert auf einem Modell, dessen Modellkomponenten, sogenannte Partialwellen, sorgfältig ausgewählt werden müssen, um sinnvolle Inferenz zu garantieren.
Bisher wurden die relevanten Partialwellen von Hand ausgewählt. Die Selektion der Komponenten ist aufgrund starker wechselseitiger Beeinflussung kompliziert, weswegen die von Hand selektierten Modelle suboptimal sein können.
Mit der Verfügbarkeit der großen Datensätze von COMPASS möchte man die Analyse verbessern, um kleinere Effekte im Detail zu untersuchen. Diese Verbesserungen der Datenanalyse benötigen eine systematische Methode die relevanten Partialwellen zu selektieren. Vor Kurzem wurden zwei Methoden entwickelt, welche auf regularisierter Maximum-Likelihood-Schätzungen basieren. In dieser Arbeit werden die verschiedenen Modellselektionsmethoden auf simulierten und gemessenen Daten angewandt und verifiziert. Die Ergebnisse führten zu der Entwicklung einer neuen Methode, welche die Vorteile der vorher vorgestellten Ansätze vereint.
Diese neue Modellselektionsmethode wurde erfolgreich auf gemessenen Daten angewandt und war in der Lage Probleme der anderen Prozeduren zu lösen. Die Methode ermöglicht auf schnelle und praktikable Weise die systematische Selektion von Wellensets für die Partialwellenzerlegung.

Contents

Chapter 1

Introduction

The theory of quantum chromodynamics successfully describes the strong interaction in the high-energy limit, by the means of perturbative expansion in the strong coupling constant α_S. In the low-energy regime, however, these calculations break down due to the running of α_S. In order to study the properties of the strong interaction one has to rely on numerical lattice QCD calculations and effective theories to model experimental findings. One interesting probe for low-energy QCD, that is accessible via experiments, is the spectrum of mesons built from the light up, down and strange quarks. The light-meson spectrum consists of many heavily-overlapping resonances of the strong interaction. With effective theories it is possible to extract the properties of these resonances from measured data. This sector is of special interest, because it is expected that, with increasing computational power, lattice QCD will be able to make predictions that can be compared to the experimental results in the near future.

The Common Muon and Proton Apparatus for Structure and Spectroscopy (COMPASS) is a fixed-target experiment at the CERN Super Proton Synchrotron (SPS), that measures the light-meson spectrum.
At COMPASS, resonances are produced via diffractive dissociation reactions. For this a negatively charged π beam with a momentum of $190\,\mathrm{GeV}/c$ was shot on a liquid hydrogen target during the 2008 data-taking campaign. The scattering process excites the beam particle to short-lived resonances that decay via the strong interaction into hadronic final states. The decay products are measured in the COMPASS spectrometer.
In this thesis the focus will be on the analysis of the reaction $\pi^- + p \to \pi^-\pi^-\pi^+ + p_{\mathrm{recoil}}$. By measuring the momenta of the final-state particles, it is possible to separate the contribution from different intermediate three-pion resonances to the measured kinematic distributions. The analysis technique for extraction of individual contributions of resonances with different quantum numbers is called partial-wave analysis. Intermediate resonances with different quantum numbers and their decay channels are associated with so-called partial waves.
The analysis is separated in two steps. In the first step, the so-called partial-wave

decomposition, complex transition amplitudes of the partial waves are extracted from the data. The second step is called resonance extraction. In it, the results of the partial-wave decomposition are fitted to extract the resonance parameters. The partial-wave analysis makes it possible to look for resonances with specific quantum numbers.

This thesis is concerned with the first of the two steps. It is a priori unknown which partial waves contribute significantly to the observed data. To arrive at sensible results for the transition amplitudes it is therefore crucial to make an appropriate selection of partial waves that are taken into account. This collection of partial waves is known as the wave set. The selection procedure is complicated due to interplay of the partial waves and will be the main topic of this work.

The three-pion final-state data measured by COMPASS has been analyzed in detail in Ref. [Haa13]. In the work the selection of relevant partial waves was performed manually and a wave set of 88 partial waves has been used.
For the analysis of the five-pion final state the construction of a wave set is more challenging and therefore a systematic way of selecting the relevant waves is required. In Ref. [Bic16] an approach for automatic selection of a wave set was presented. This method, called Biggest Conceivable Model (BCM) method, has then been applied to the three-pion final state in Ref. [Dro15], to introduce a more formal way of deducing a wave set. The results looked promising, but further manual intervention was required to fix issues with the model stability.
Recently another model-selection method has been introduced in Ref. [Gue+15], which is based on a well-known statistical method, called the LASSO [Tib96].

These previous analyses are the starting point for this thesis. The goal of this work is to develop a method that is applicable in practice for the selection of a model for partial-wave decomposition. The three-pion final-state data is used to test the selection method, but the method itself can be applied to other final states as well.
In chapter 2 a short introduction to the COMPASS experiment and the underlying physical processes will be given. Furthermore the event selection, used to produce the dataset for this analysis, will be summarized. Afterwards, in chapter 3 the theoretical background of the partial-wave decomposition analysis method will be presented. Chapter 4 continues with the introduction of the model selection methods. These methods are first applied and tested in chapter 5 on simulated datasets, restricted to certain mass-regions. A new model selection technique is introduced. The same mass-regions are then analyzed for measured data in chapter 6. Chapter 7 extends the analysis on the complete available mass-region, while still being restricted to a subset of the available data, and serves as a case-study of the different methods. A

modification of the newly introduced model selection procedure is presented. This modification takes the available phase space of the decay channels into account to improve the inference. In chapter 8 this new approach is applied to the full available dataset of the three-pion final state. The findings are then discussed and compared to the previous model.

The conclusions of this analysis are presented in chapter 9 and an outlook on future applications and developments is given.

Chapter 2

Meson Production in Diffractive Reactions at COMPASS

The data for the analysis, presented in this thesis, were measured by the COMPASS experiment in 2008. In this chapter, a brief introduction to the experimental setup, the underlying physical processes and the selection of the dataset is given. This introduction is based on the following material [Ado+17; Bic16; Haa13; Dro15]

2.1 Experimental Setup

In fig. 2.1, a rendering of the experimental setup, as used during the 2008 data-taking campaign, is shown. The following section summarizes the most important aspects of the particle beam and the detector hardware. A detailed description of the setup and its individual components can be found in Ref. [Abb+15].

Beam and Target COMPASS can be supplied with a variety of particle beams to study different physics. For this analysis, the dataset consists of reactions induced by a 190 GeV/c secondary hadron beam that is composed predominantly out of negatively charged pions. In fig. 2.1 it is marked by an arrow on the left. The secondary beam is produced by a 400 GeV/c primary proton beam from the Super Proton Synchrotron (SPS) that is converted in a Beryllium production target. After the conversion, the beam is directed to the COMPASS experiment via the 1.1 km long M2 beam line [Abb+15]. At the liquid hydrogen target, the hadronic component consists of 96.8% π^-, 2.4% K^- and 0.8% $\bar{p}$ with a momentum of 190 GeV/c [Ado+17]. The 400 mm long liquid hydrogen target is placed within a barrel-shaped recoil-proton detector (RPD), which consists of two layers of scintillator material [Abb+15]. During the reaction with the target, protons are kicked out of the target. These slow recoil protons are then detected by the RPD.

Detector Setup The final-state particles produced in the reaction are measured by a variety of detectors used for the spectrometer and as triggers. High-resolution silicon-microstrip detectors, placed directly behind the target, detect the interaction

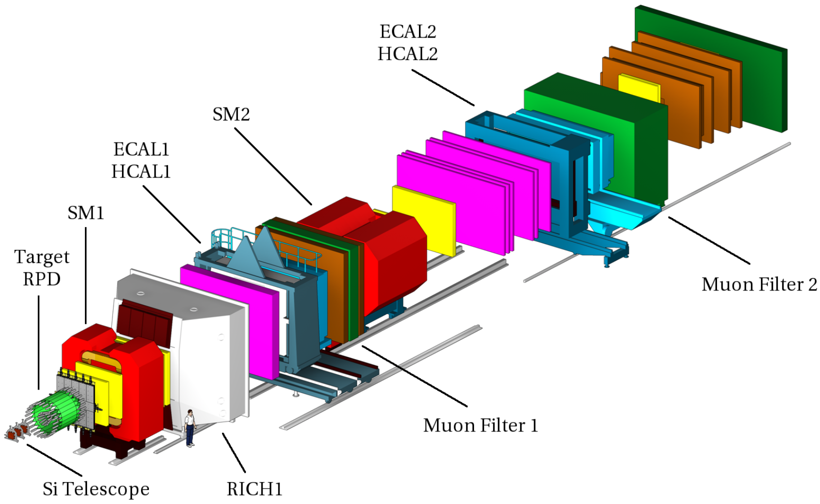


Figure 2.1: Rendering of the COMPASS setup as used during the 2008 data-taking campaign. The 190 GeV/c π^- beam is hitting the target from the left. It interacts with the liquid hydrogen target inside the RPD. The produced final-state particles are detected in the forward spectrometer behind the target, which consists of detectors and two dipole magnets. [Abb+15]

vertices. The forward spectrometer consists of two stages that are optimized for high and low momenta of the outgoing particles. Each stage has a bending magnet to enable the measurement of charged-particle momenta. The tracking of the final-state particles is performed by gas detectors. Different types of these tracking detectors are used, for example GEM or Micromega detectors. The first stage is additionally equipped with a Ring Imaging Cherenkov Counter (RICH) that allows particle identification.

2.2 Meson Production in Diffractive Reactions

To study the light-meson spectrum, excited meson resonances have to be produced. This is, for example, possible by a process known as diffractive dissociation. As already mentioned before, this thesis focuses on the special case of $\pi^- + p \to \pi^-\pi^-\pi^+ + p_{\text{recoil}}$. The Feynman diagram of the reaction is shown in fig. 2.2. The negatively charged beam pion scatters off a target proton. At the target vertex, the scattering is elastic, so that the proton stays intact, while the beam pion is excited to an intermediate resonance X, which then decays into three charged pions. Regge theory forms the theoretical basis for the description of such processes [Haa13; Ado+17]. In this framework, the scattering process is described by exchange of a so-called Reggeon, which is the effective mediator of the strong force. For the 190 GeV/c beam the process is expected to be dominated by the exchange of a Reggeon called the Pomeron ($\mathbb{P}$).

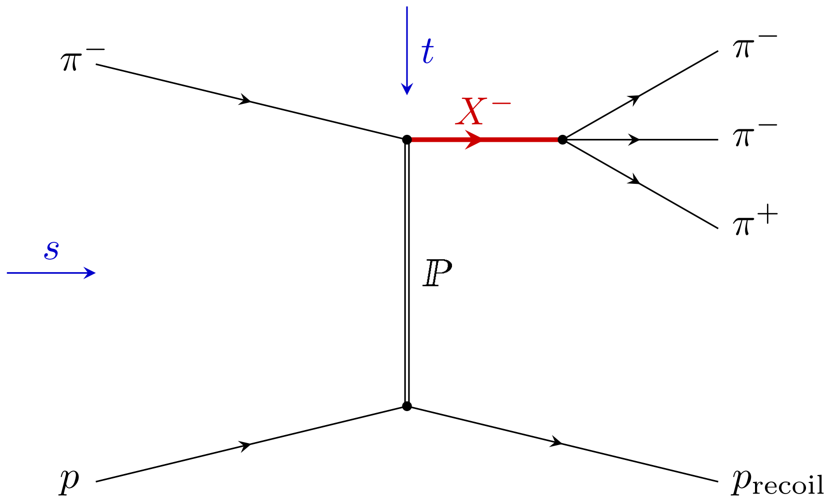


Figure 2.2: Diffractive dissociation reaction: $\pi^- + p \to \pi^-\pi^-\pi^+ + p_{\text{recoil}}$. The beam π^- gets excited to an intermediate resonance X^- via soft scattering off a proton. The X^- then decays into three charged pions. [Ado+17]

The intermediate resonances X are characterized by their quantum numbers $I^G J^{PC} M^\epsilon$. Where I is the isospin, G is the G-parity, C the C-parity, J the spin and P the parity. The quantum numbers are given in the reflectivity basis in which $M \geq 0$ is the spin projection and $\epsilon = \pm 1$ the reflectivity. The quantum numbers of the Reggeon fix the reflectivity in this process. Because the Pomeron is the dominant Regge trajectory at these energies, the reflectivity is expected to be mostly positive [Ado+17]. It is possible that other trajectories contribute, for which resonances of negative reflectivity may be produced. The G-parity is an extension of the C-parity to charged light mesons, as the C-parity is only defined for neutral particles. It is obtained from the charged conjugated state by a 180° rotation around the second component in isospin-space.

$$G = Ce^{i\pi I_2} = C(-1)^I \tag{2.1}$$

For charged states, the C-parity of the neutral partner state is typically assigned.

The reaction in fig. 2.2 imposes certain restrictions on the possible quantum numbers of X [Ado+17]. Due to conservation of charge, the intermediate resonance must be negatively charged. This implies for the isospin $I > 0$. Since no mesons with $I \geq 2$ are known, $I = 1$ is assumed. Furthermore the strong interaction is known to preserve G-parity. The pseudoscalar π^- have $G = -1$. Because the Pomeron has vacuum quantum numbers [Ado+17], the G-parity of X is fixed to $G = -1$ by the G-parity of the beam pion. By the definition of the G-Parity, $C = +1$ can be assigned to X^-. This restricts the resonances, that can be observed, to mesons of the a_J type [1], with $J^{PC} = J^{++}$, and the π_J type, with $J^{PC} = J^{-+}$, where J is the spin.

In the simple constituent quark model, mesons are described as quark-antiquark states ($q\bar{q}$). Under this assumption, certain combinations of J^{PC} are forbidden. A short summary of the rules implied by the constituent quark model is given in Ref. [Haa13]. The following states are forbidden by these rules:

$$J^{PC} = 0^{--}, (\text{odd})^{-+}, (\text{even})^{+-} \tag{2.2}$$

In QCD, in principle more complicated objects can be formed than in the constituent quark model. For example, hybrid states, in which gluon fields contribute to the quantum numbers. Such states can bear quantum numbers that are forbidden in the constituent quark model. These states are referred to as exotic states. Because $C = +1$, it is evident from eq. (2.2) that only exotic states with odd spin are accessible in this reaction.

Apart from their quantum numbers, resonances are also characterized by m_X. The invariant mass m_X of the resonance X^- is equivalent to the invariant mass $m_{3\pi}$ of the three-pion final state. It is calculated from the four-momenta p_i of the final state particles, which themselves are given by the three-momenta and the charged pion mass.

$$m_{\mathrm{X}}^2 = m_{3\pi}^2 = \left(\sum_{i=1}^{3} p_i \right)^2 / c^2 \tag{2.3}$$

The production of resonances also depends on the squared four-momentum transfer $t = (p_{\text{beam}} - p_X)^2$. The excitation of the beam pion to a resonance with mass $m_{3\pi}$ requires a minimal squared four-momentum transfer $|t|_{\min}$ [2]. Therefore, for practical reasons, the reduced squared four-momentum transfer is defined:

$$t' \equiv |t| - |t|_{\min} \tag{2.4}$$

[1] $J^{PC} = 0^{++}$ or a_0 resonances cannot be produced during the diffractive scattering reaction.

[2] more details can be found in [Ado+17]

2.3 Background Processes

Different processes contribute to the scattering reaction. Of all these reactions, only diffractive dissociation with $\pi^-\pi^-\pi^+$ in the final state is of interest in this analysis. An overview of the event selection that separates the different final states is given in section 2.4. For the analysis, one would like to have a sample of diffractively produced $\pi^-\pi^-\pi^+$, that is as clean as possible. Unfortunately different processes that lead to the same three-pion final state, but do not have an intermediate three-pion resonance, contribute as background to the analysis. Two of these nonresonant processes are presented here.

2.3.1 Central Production

One possible nonresonant process is known as central production, for which the Feynman diagram is shown in fig. 2.3. Instead of the exchange of a single Pomeron, a double exchange appears. The beam pion scatters elastically via one exchange and the proton via the second one. The Pomerons form an intermediate resonance that decays into a neutral system of two charged pions. As remarked in [Haa13], the kinematics of the final state pions is in general different for central production and diffractive dissociation. The centrally produced subsystem is typically slow, while the scattered beam pion remains fast. In the event selection, cuts on the rapidity of the subsystems are therefore applied in order to suppress the contribution of central production to the background. This will be discussed in section 2.4.

2.3.2 The Deck Effect

A second process known to be part of the nonresonant background is the so-called Deck effect [Dec64]. In this process, the beam is excited to an intermediate $\pi^+\pi^-$ resonance and the third pion is scattered off the proton via Pomeron exchange. One possible Feynman diagram for this process is shown in fig. 2.4. A suppression of this background contribution during the event selection is not possible because the final state kinematics are practically indistinguishable. It is therefore the main source of background for this analysis.

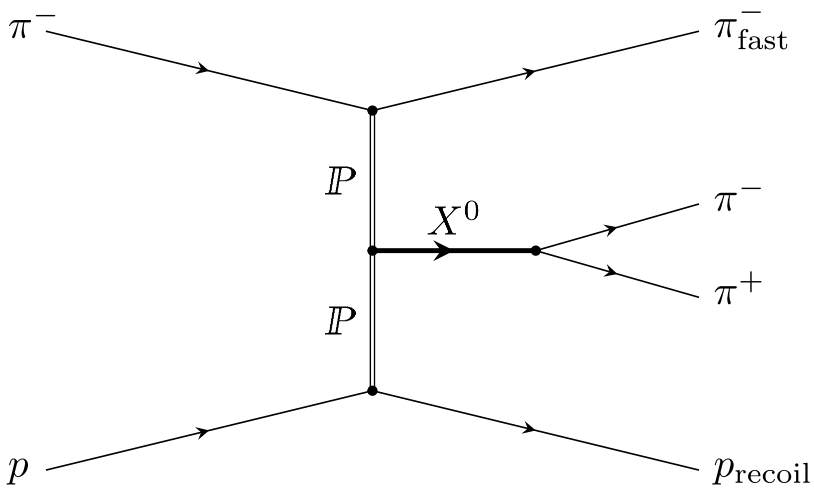


Figure 2.3: Example for a central production reaction. Both beam pion and target proton scatter elastically. The double pomeron exchange produces an intermediate state that decays into a $\pi^-\pi^+$ subsystem. [Haa13]

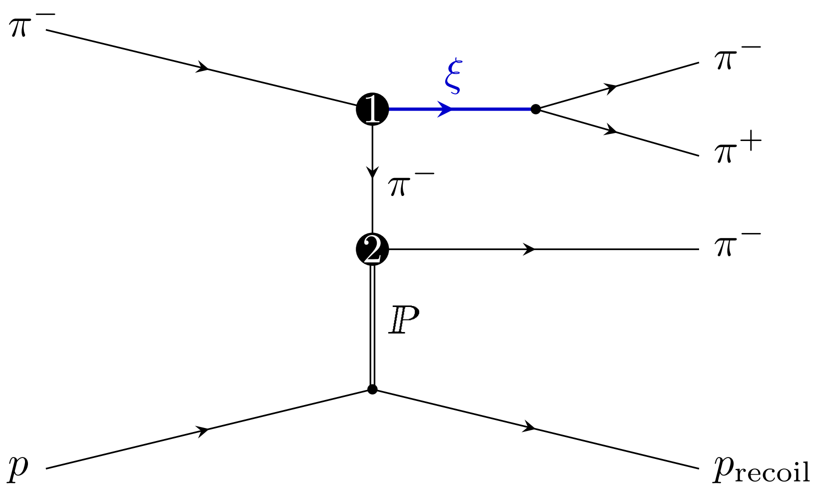


Figure 2.4: Example of a possible Feynman diagram for the Deck effect for the three-pion final state. [Ado+17]

2.4 Event Selection

As already mentioned before, many different reactions and final states may occur in the high-energy scattering. The raw dataset, that is used for the event selection, has been recorded by using the so-called diffractive trigger (DT0) of COMPASS. This minimum-bias hardware trigger [Ado+17] is intended for preselection of relevant events. It requires the interaction of the beam with the target and the detection of a recoil proton in the RPD. A more detailed summary can be found in Ref. [Ado+17; Abb+15]. This preselected dataset consists of $6.4 \cdot 10^9$ events [Ado+17].
In order to obtain a clean sample of diffractively produced $\pi^-\pi^-\pi^+$, the recorded events have to be filtered first. The event selection has already been performed by [Haa13] and the resulting datasets are the basis for this analysis. The complete description of the selection can be found in Ref. [Haa13]. In the following, the most important cuts for the event selection are summarized:

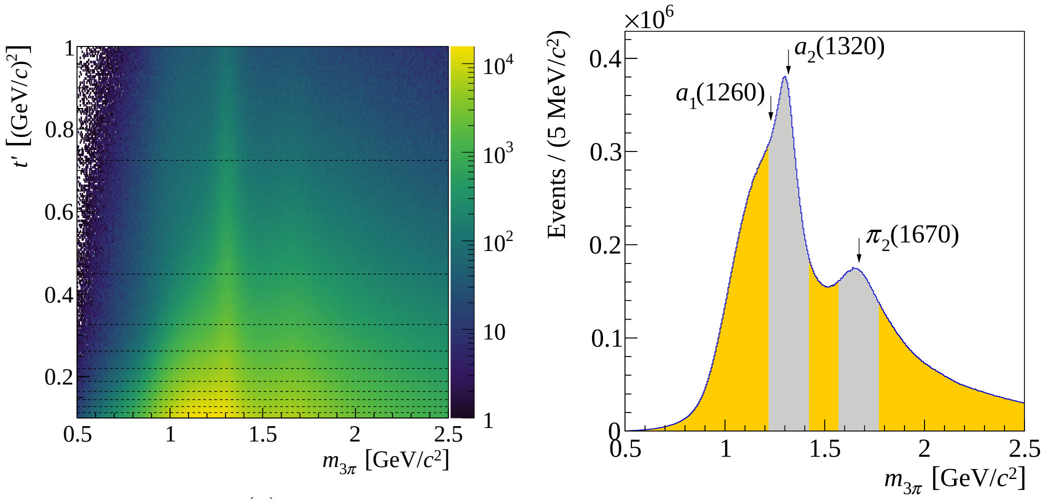
1. A vertex in the target region with three outgoing particles is required. The total charge of the final state particles must equal -1.

2. To suppress processes, in which the target does not stay intact, transverse momentum conservation is required.

3. The beam energy is not directly measured at COMPASS, but calculated from the four-momenta of the outgoing particles. The calculated energy was required to deviate less than two standard deviations from the nominal beam energy.

4. Additional cuts were applied that were intended to suppress other background processes. More detail can be found in Refs. [Ado+17; Haa13].

The event selection reduces the number of measured events from $6.4 \cdot 10^9$ to $46 \cdot 10^6$. The selected dataset consists of mostly three-pion diffractive dissociation events in a kinematic region of $0.5 \leq m_X \leq 2.5\,\mathrm{GeV}/c^2$ and $0.1 \leq t' \leq 1.0\,(\mathrm{GeV}/c)^2$. The two-dimensional histogram of the selected events is shown in fig. 2.5a and the respective projection on the mass axis in fig. 2.5b. Enhancements, originating from resonances, are visible in both representations of the data. The analysis method, used to disentangle the contributions of individual resonances, will be introduced in the next chapter.



(a)

(b)

Figure 2.5: Histograms of the data after event selection for the $\pi^-\pi^-\pi^+$ final state. (a) $m_{3\pi}$ and t' resolved histogram of the event selected data. The dominant resonances are visible as band-like structures. The dashed-lines mark the non-equidistant bins in t'. (b) Invariant mass spectrum of the $\pi^-\pi^-\pi^+$ final state. Only the dominant resonances are visible as peak-like enhancements in the spectrum. [Ado+17]

Chapter 3

Partial-Wave Analysis Method

The overall goal of the analysis of diffractive-dissociation data is the extraction of short-lived resonances of the strong interaction. Because in this type of reaction resonances are typically overlapping, only the dominant ones can be directly discovered as enhancements of the total cross section as demonstrated in fig. 2.5. More information can be extracted by taking into account the four-momenta of the final state particles and measuring the differential cross section. With this additional information contributions from resonances of different quantum numbers can be disentangled by the means of so-called partial-wave decomposition.

Due to the nature of the reaction, neither the mass $m_{3\pi}$ of the 3π final-state nor the squared four-momentum transfer t' are fixed. On one hand, this allows the extraction of the resonance spectrum with a fixed beam energy. On the other hand, the dependency of the differential cross section, in the kinematic variables, on $m_{3\pi}$ and t' is unknown prior to the experiment.

With large datasets, like the ones available from the COMPASS experiment, this problem can be circumvented by subdividing the data into bins of final-state mass $m_{3\pi}$ and reduced squared four-momentum transfer t'. For sufficiently small widths of the bins, the dependency on these variables is negligible and any derived quantity can be assumed to be constant in the binning variables. This motivates a partial-wave analysis method consisting of two steps. In the first step the amplitudes of the partial waves are extracted for each bin in $m_{3\pi}$ and t'. This step is referred to as partial-wave decomposition or sometimes as mass-independent fit. In the second step the m_X and t' dependence is parametrized for the extraction of the resonance parameters. It is therefore called the resonance extraction or mass-dependent fit.

This thesis is concerned with the first analysis step and the construction of a physical model for the partial-wave decomposition. For the extraction of the partial-wave amplitudes a statistical model is required. The following sections, on the derivation of this statistical model, summarize previous works and introduce simplifications that can be used for this specific final-state. More details on the method can be found in the original works [Ado+17; Bic16; Haa13].

3.1 Statistical Model for Partial-Wave Decomposition

The quantities measured by the experiment are the three-momenta of the final-state particles. For known final-state particle mass and fixed beam energy, the four-momenta can be constructed. The kinematics of the three-body final state is fully defined by m_X and five phase-space variables [1] combined in τ. This allows COMPASS to measures the differential intensity distribution, which is proportional to the differential cross-section.

$$\mathcal{I}(\tau, m_X, t', s) = \frac{\mathrm{d}N}{\mathrm{d}m_X \mathrm{d}t' \mathrm{d}\tau} \propto \frac{\mathrm{d}\sigma}{\mathrm{d}m_X \mathrm{d}t' \mathrm{d}\tau} = \sigma_0 \left(m_X, t', s\right) \cdot \left|\mathcal{M}_{fi}\left(\tau; m_X, t', s\right)\right|^2 \tag{3.1}$$

Where $\mathrm{d}\tau$ is the five-dimensional differential phase-space element and $\mathcal{M}_{fi}$ the transition matrix element. The beam energy is fixed in the experiment, therefore s is dropped from now on to simplify the notation. By subdividing the data in narrow bins of m_X and t' these variables can be assumed constant within one bin. Writing down eq. (3.1) for one of these two-dimensional bins, with fixed m_X and t', the notation simplifies further.

$$\mathcal{I}(\tau) = \frac{\mathrm{d}N}{\mathrm{d}\tau} \propto \frac{\mathrm{d}\sigma}{\mathrm{d}\tau}(\tau) = \sigma_0 \cdot \left|\mathcal{M}(\tau)\right|^2 \tag{3.2}$$

By assuming that the observed intensity is dominated by the decay of intermediate 3π resonances, the contributing amplitudes factorize and can be written as a product of the production amplitudes ${}^r\tilde{\mathcal{T}}_i^\epsilon$ of the resonances and the sum over the decay amplitudes ${}^r\tilde{\Psi}_{i,j}^\epsilon$. The index i labels the quantum numbers of the resonance $I^G J^{PC} M$ and the index j its decay modes. The ${}^r\tilde{\mathcal{T}}_i^\epsilon$ describe the strength and phase with which resonances are produced, while the ${}^r\tilde{\Psi}_{i,j}^\epsilon$ describe the decay. The intensity as a function of the kinematic variable τ can now be expressed by taking into account coherent and incoherent processes. The coherent processes interfere and are therefore written according the superposition principle as a sum over the probability amplitudes. The incoherent processes are non-interfering and correspond to sums over the probabilities.

$$\mathcal{I}(\tau) = \sum_{\epsilon=\pm1} \sum_r \left| \sum_i {}^r\tilde{\mathcal{T}}_i^\epsilon \sum_j \tilde{\Psi}_{i,j}^\epsilon(\tau) \right|^2 \tag{3.3}$$

The decay amplitudes $\tilde{\Psi}_{i,j}^\epsilon(\tau)$ can be calculated by assuming successive two-body decays. The Feynman diagram for the decay, in this so-called isobar model, is shown

[1] The three four-momenta of the outgoing particles have in principle 12 degrees of freedom (DOF). Because the particle type and therefore the masses are known, three DOF are fixed. Additionally, four-momentum conservation constraints another four DOF, leaving five remaining DOF that are required to describe the final state.

in fig. 3.1. First the decay $X^- \to \pi^- + \xi$ appears, where the neutral isobar ξ is assumed to be an intermediate two-pion resonance. This assumption can be justified by looking at resonant behavior of the two-pion subsystem. So-called Dalitz plots make the resonance content of the subsystem visible for fixed $m_{3\pi}$ mass range. In fig. 3.2 three of the isobars, used in this analysis are visible as enhancements. In this analysis six possible two-pion resonances are taken into account as isobars. They are described in section 3.2. Between the isobar ξ and the bachelor pion a relative angular momentum L appears. The ξ then further decays into two pions.

Each decay is parametrized in a suitable coordinate system, which are both described in detail in Ref. [Ado+17]. For this part of the analysis the exact choice of parametrization is unimportant and the reference frames are merely introduced to explain the formalism.
The decay of the X^- into the isobar ξ and the bachelor pion is parametrized in the Gottfried-Jackson (GJ) reference frame, the decay of the isobar ξ into the two-pion subsystem in the helicity reference frame (HF). In the GF the decay is characterized by two angles, ϑ_{GF} and ϕ_{TY}. The decay in the HF is also characterized by a pair of angles, ϑ_{HF} and ϕ_{HF}. Together with the mass of the isobar m_ξ and the mass of the intermediate resonance $m_X = m_{3\pi}$ the kinematics of the final state is described. The four angles and the mass of the isobar are grouped into the five-dimensional variable τ:

$$\tau \equiv (\vartheta_{\mathrm{HF}}, \phi_{\mathrm{HF}}, m_\xi, \vartheta_{\mathrm{GJ}}, \phi_{\mathrm{TY}}) \tag{3.4}$$

τ and m_ξ can be calculated from the three momenta of the three outgoing pions. For this analysis it is only important, that these values are calculable. The exact description of these calculations can be found in Ref. [Ado+17].

To obtain the decay amplitudes $\tilde{\Psi}^{\epsilon}_{i,j}(\tau)$, one has to calculate the amplitudes of the two-body processes first. The two-body amplitudes factorize into an angular and a dynamic part [Ado+17]. The decay of the isobar ξ is described by the amplitude $\mathcal{A}^{\xi}_{\lambda}$.

$$\mathcal{A}^{\xi}_{\lambda}(\vartheta_{\mathrm{HF}}, \phi_{\mathrm{HF}}, m_\xi) = \underbrace{D^{J_\xi *}_{\lambda 0}(\vartheta_{\mathrm{HF}}, \phi_{\mathrm{HF}}, 0)}_{\text{angular part}} \underbrace{f^{J_\xi}_{00}(m_\xi; m_\pi, m_\pi)}_{\text{dynamic part}} \tag{3.5}$$

The angular part is expressed by the Wigner D-function $D^{J_\xi *}_{\lambda 0}(\vartheta_{\mathrm{HF}}, \phi_{\mathrm{HF}}, 0)$. The dynamic part describes the mass dependence and is a product of a normalization

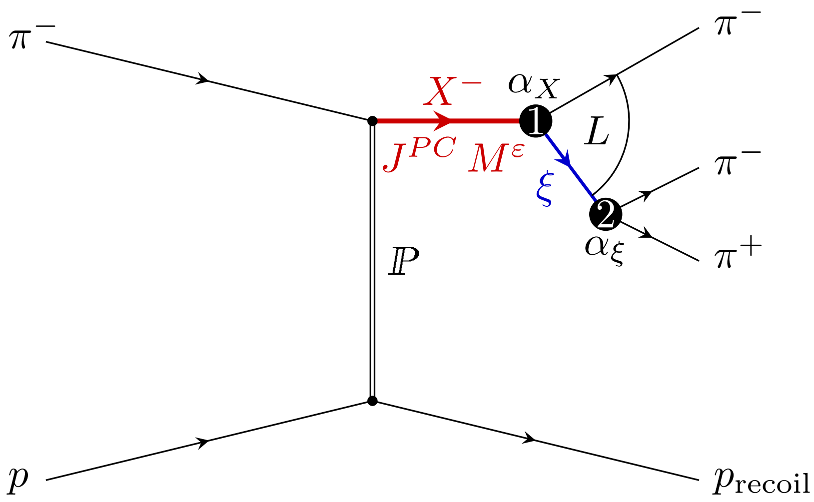


Figure 3.1: Diffractive dissociation in the isobar model. The intermediate resonance X^- first decays in a negatively charged bachelor π^- and a neutral isobar ξ. The isobar then decays into two charged pions. [Ado+17]

factor, the coupling constant α_ξ, the barrier factor F_{J_ξ} and the isobar line shape Δ_ξ, explained in section 3.2.

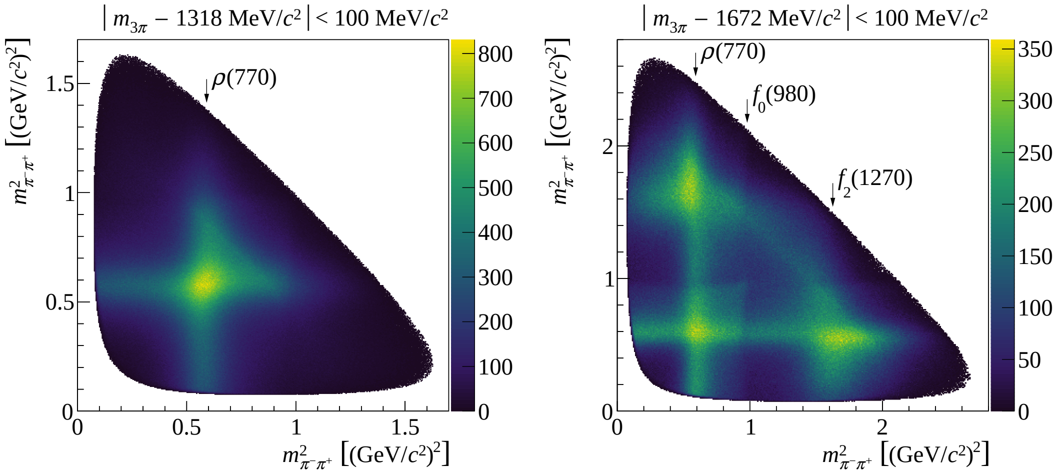
$$f_{00}^{J_\xi}(m_\xi; m_\pi, m_\pi) = \underbrace{\sqrt{2J_\xi + 1}}_{\text{normalization}} \underbrace{\alpha_\xi F_{J_\xi}(m_\xi; m_\pi, m_\pi)\Delta_\xi(m_\xi; m_\pi, m_\pi)}_{\text{dynamics}} \tag{3.6}$$

The decay of the resonance X into the isobar and the bachelor pion is described by the two-body amplitude $\mathcal{A}_M^X$. It is expressed as a sum over the helicity λ of the isobar.

$$\mathcal{A}_M^X(\vartheta_{\text{GF}}, \phi_{\text{TY}}, m_{3\pi}) = \sum_\lambda D_{M\lambda}^{J_*}(\phi_{\text{GF}}, \vartheta_{\text{TY}}, 0) f_{\lambda 0}^J(m_{3\pi}; m_\xi, m_\pi) \tag{3.7}$$

The dynamic part for the decay into the isobar and the bachelor pion looks similar to the one of the isobar decay eq. (3.6). It consists of a normalization term and a dynamic term, consisting of a barrier factor and the coupling constant α_X. Additonally, a Clebsch-Gordan coefficients from the coupling of the spins is required.

$$f_{\lambda 0}^J(m_{3\pi}; m_\xi, m_\pi) = \underbrace{\sqrt{2J_\xi + 1}}_{\text{normalization}} \underbrace{(L0J_\xi|J\lambda)}_{\substack{\text{L-S coupling} \\ \text{Clebsch-Gordon}}} \underbrace{\alpha_X F_L(m_{3\pi}; m_\xi, m_\pi)}_{\text{dynamics}} \tag{3.8}$$



(a) (b)

Figure 3.2: Dalitz plots for 3π masses around the $a_2(1320)$ and $\pi_2(1670)$ resonances. In (a) the $a_2(1320)$ and $a_1(1260)$ resonances are visible in their decay to $\rho(770)$ π. In (b) decays of the $\pi_2(1670)$ resonance to $\rho(770)$ π, $f_2(1270)$ π and $f_0(980)$ π are visible. [Ado+17]

The decay amplitudes can now be written as the product of the two two-body decay amplitdes.

$$\psi_{i,j}(\underbrace{\vartheta_{\mathrm{HF}}, \phi_{\mathrm{HF}}, m_\xi, \vartheta_{\mathrm{GJ}}, \phi_{\mathrm{TY}}}_{\tau}; m_{3\pi}) = \sum_\lambda D^{J*}_{M\lambda}(\phi_{\mathrm{TY}}, \vartheta_{\mathrm{GJ}}, 0) f^J_{\lambda 0}(m_{3\pi}; m_\xi, m_\pi) \times \mathcal{A}^\xi_\lambda(\vartheta_{\mathrm{HF}}, \phi_{\mathrm{HF}}, m_\xi) \tag{3.9}$$

One additional step is necessary to arrive at the expression of the decay amplitudes in the case of the three-pion final state. Because the two negatively charged pions are indistinguishable, there are two possibilities that can make up the neutral subsystem in which the isobar decays. The decay amplitudes must therefore be bose symmetrized.

$$\tilde{\Psi}^\epsilon_{i,j}(\tau_{1,3}, \tau_{3,1}) = \frac{1}{\sqrt{2}}\left[\psi_{i,j}(\tau_{1,3}) + \psi_{i,j}(\tau_{2,3})\right] \tag{3.10}$$

The couplings α_ξ and α_X are unknown. They are assumed to be independent of the kinematics. By combining the couplings with the production amplitudes, one obtains the transition amplitudes.

$${}^r\bar{\mathcal{T}}^\epsilon_{i,j} \equiv \alpha_\xi \alpha_X {}^r\tilde{\mathcal{T}}^\epsilon_{i,j} \tag{3.11}$$

$$\bar{\Psi}^{\epsilon}_{i,j} \equiv \frac{\tilde{\Psi}^{\epsilon}_{i,j}(\tau)}{\alpha_{\xi}\alpha_{X}} \tag{3.12}$$

The index a is now a combination of the resonance quantum numbers $i \equiv I^G J^{PC} M$, ϵ and the decay channel $j \equiv \xi L$, which is described by the isobar and the relative angular momentum. Each partial wave is then labeled by a.

$$a \equiv (i, j, \epsilon) \equiv I^G J^{PC} M^{\epsilon} \xi \pi L \tag{3.13}$$

Because $I^G = 1^-$ holds for resonances produced for this type of reaction, it is often omitted for the sake of shorter notation.

One last transformation of the transition and decay amplitudes is performed.

$$\Psi^{\epsilon}_{a}(\tau) \equiv \frac{\bar{\Psi}^{\epsilon}_{a}(\tau)}{\sqrt{\int \mathrm{d}\varphi_3(\tau')|\bar{\Psi}^{\epsilon}_{a}(\tau')|^2}} \tag{3.14}$$

$${}^{r}T^{\epsilon}_{a} \equiv {}^{r}\bar{\mathcal{T}}^{\epsilon}_{a}\sqrt{\int \mathrm{d}\varphi_3(\tau')|\bar{\Psi}^{\epsilon}_{a}(\tau')|^2} \tag{3.15}$$

This phase-space normalization of the decay amplitudes leads to transition amplitudes ${}^{r}T^{\epsilon}_{a}$ for which the intensity of a single wave $|{}^{r}T^{\epsilon}_{a}|^2$ is given in number of events.

The intensity of a single event then simplifies to the from given in eq. (3.3):

$$\mathcal{I}(\tau) = \sum_{\epsilon=\pm1}\sum_{r}\left|\sum_{a}{}^{r}T^{\epsilon}_{a}\,{}^{r}\Psi^{\epsilon}_{a}(\tau)\right|^2 \tag{3.16}$$

Because no non-interfering processes except for the different reflectivities and an additional background term described in fig. 5.17b are used in this analysis, the incoherent sum over r is dropped from now on. These fits are then called rank 1 fits in contrast to rank n fits for n non-interfering processes. A discussion on rank 2 fits can be found in Ref. [Haa13].
To further simplify notation the coherent sum over the amplitudes is replaced by the scalar product of the two complex vectors $\boldsymbol{T}^{\epsilon}(m_X, t', s)$, $\boldsymbol{\Psi}^{\epsilon}(\tau; m_X, t', s) \in \mathbb{C}^n$. This results in the simple expression eq. (3.17) for the intensity of a single event.

$$\mathcal{I}(\tau) = \sum_{\epsilon=\pm1}\left|\boldsymbol{T}^{\epsilon}(m_X)\cdot\boldsymbol{\Psi}^{\epsilon}(\tau; m_X)\right|^2 \tag{3.17}$$

However the data is most-certainly not free from contributions of background processes. To account for this effect an additional isotropic amplitude is added to the intensity, the so-called flat wave. This leads to a modified formula eq. (3.18) with an additional third incoherent term.

$$\mathcal{I}(\tau) = \sum_{\epsilon=\pm 1} |\boldsymbol{T}^{\epsilon} \cdot \boldsymbol{\Psi}^{\epsilon}(\tau)|^2 + |T_{\text{flat}}|^2 \tag{3.18}$$

The probability of a single measured event τ_k, given in eq. (3.20), is just the ratio of the intensity of the single event and the total expected intensity given by eq. (3.19), where $0 \leq \eta(\tau) \leq 1$ describes effects due to the detector-acceptance and $\mathrm{d}\varphi_3(\tau)$ is the three-body phase-space volume element.

$$\bar{N} = \int \mathrm{d}\varphi_3(\tau') I(\tau') \eta(\tau') \tag{3.19}$$

$$P(\tau) = \frac{\mathcal{I}(\tau)}{\int \mathrm{d}\varphi_3(\tau') \mathcal{I}(\tau')} \tag{3.20}$$

The number of measured events N follows a Poisson distribution with the rate being the total expected intensity $\bar{N}$. The extended likelihood for a dataset $D = \{\tau_k\}$ eq. (3.21) can now be written as the product of the Poisson probability for the measured number of events and the product of the probabilities for every individual event.

$$\mathcal{L}(\boldsymbol{T}; D) = \frac{e^{-\bar{N}} \bar{N}^N}{N!} \prod_k^N P(\tau_k) \tag{3.21}$$

This can be further simplified as the denominator of the probability for a single event is the total expected intensity. The product over all the events results in a term of $\bar{N}^N$ in the total denominator. This term cancels with the same part in the Poisson probability and the new formula is given by eq. (3.22).

$$\mathcal{L}(\boldsymbol{T}; D) = \frac{e^{-\bar{N}}}{N!} \prod_k^N \mathcal{I}(\tau_k) \tag{3.22}$$

In the usual Frequentist approach one would like to find the parameters, in this case the complex transition amplitudes, that maximize the likelihood. This corresponds to the idea that because the data has been observed the observation should be very probable. The parameters are then chosen in such a way that the probability of the data given the set transition amplitudes becomes maximal.
The maximization is performed with numerical methods that require stable calculations. Because products are usually numerically problematic the logarithm of the likelihood can be used to transform any products into sums. Taking the logarithm does not change the position of the maximum as it is a monotonically increasing

function so the optimization can be conducted on the log-likelihood instead of the likelihood.

$$\log \mathcal{L}\left(\boldsymbol{T}; D\right) = -\bar{N} - \log N! + \sum_{k}^{N} \log \mathcal{I}\left(\tau_k\right) \tag{3.23}$$

Any constant terms in the log-likelihood correspond to a constant scaling of the likelihood and can thus be dropped because they are irrelevant for the position of the maximum.

$$\log \mathcal{L}\left(\boldsymbol{T}; D\right) = -\bar{N} + \sum_{k}^{N} \log \mathcal{I}\left(\tau_k\right) \tag{3.24}$$

3.2 Isobar Parametrization

In the previous section the statistical model, under the assumption of the isobar model, was introduced. For the calculation of the decay amplitudes the isobar amplitudes Δ_ξ are required, the parametrization of which are introduced in this section.

Six isobars have been used in this analysis. The four isobars, $\rho(770)$, $f_2(1270)$, $f_0(1500)$ and $\rho_3(1690)$, are described by a Breit-Wigner amplitude, given in eq. (3.25), with different parametrization of the width $\Gamma(m)$. More details on the exact parametrization of the widths can be found in Refs. [Ado+17; Dro15]. As an example the parametrization of the $f_0(1500)$ isobar with a constant-width Breit-Wigner is shown in fig. 3.3.

For the two isobars $[\pi\pi]_S$ and $f_0(980)$ different parameterizations are required. As closely described in Ref. [Ado+17], the so-called M solution separates the $J^{PC} = 0^{++}$ isobar sector. As mentioned above, the $f_0(1500)$ isobar is described by a Breit-Wigner amplitude. The other components are described by the broad $[\pi\pi]_S$ isobar with a slowly rising phase, shown in fig. 3.5, and the narrow $f_0(980)$ isobar, parametrized by a Flatté amplitude [Ado+17], shown in fig. 3.4.

$$\Delta_{\mathrm{BW}}(m; m_0, \Gamma_0) = \frac{m_0 \Gamma_0}{m_0^2 - m^2 - i m_0 \Gamma(m)} \tag{3.25}$$

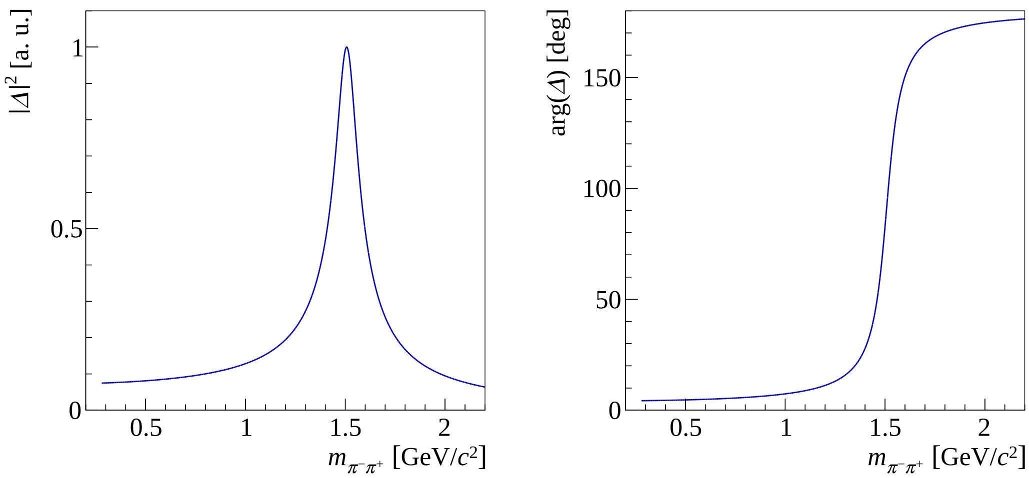


Figure 3.3: Fixed-width Breit-Wigner amplitudes of the $f_0(1500)$ isobar. (a) intensity and (b) phase of the Breit-Wigner.

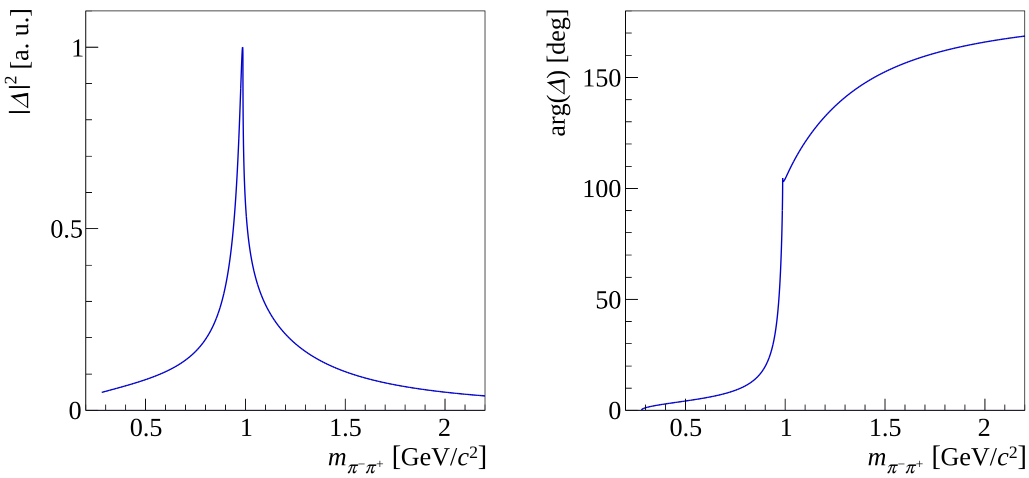


Figure 3.4: Flatté amplitude [Ado+17] of the $f_0(980)$ isobar. (a) intensity and (b) phase.

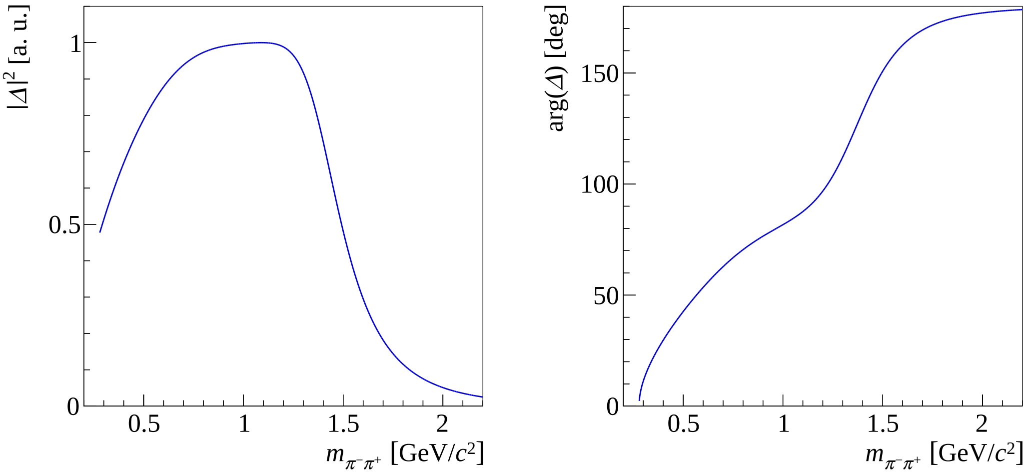
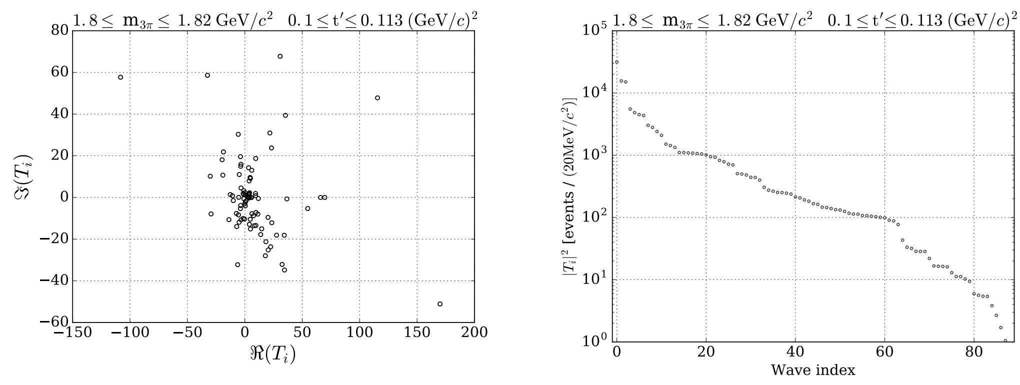


Figure 3.5: Parametrization of the $[\pi\pi]_S$ isobar [Ado+17]. (a) intensity and (b) phase.

3.3 Partial-Wave Decomposition

In the previous section the statistical model has been introduced. The analysis is performed independently in two-dimensional bins in mass m_X and squared four-momentum transfer t'. The mass range between $0.5\,\mathrm{GeV}/c^2$ and $2.5\,\mathrm{GeV}/c^2$ is split into 100 equidistant bins of $20\,\mathrm{MeV}/c^2$ width and the t' range between $0.1\,\mathrm{GeV}/c$ and $1.0\,\mathrm{GeV}/c$ is split in 11 non-equidistant bins. The binning in t' has been chosen such that each bin contains a similar number of events [2]. The details of the t' binning can be found in table 3.1.

For each of these 1100 bins a set of complex transition amplitudes is obtained from maximizing the likelihood. The result can be compactly written as the complex vector of transition amplitudes $\boldsymbol{T}$. The components of the vector can then be represented in the complex plane as shown in fig. 3.6a as an example, where each point represents the transition amplitude T_i of one wave. Neglecting the phase information the result can also be represented by plotting the magnitudes square of the transition amplitudes $|T_i|^2$. The intensities are then plotted in descending order on logarithmic scale as shown in fig. 3.6b.



(a) (b)

Figure 3.6: Example of a fit result obtained in a single (m_X, t') bin. (a) Every transition amplitude T_i is a point in the complex plane. (b) The intensity of the each wave is the magnitude squared of the transition amplitudes $|T_i|^2$ sorted according to the intensity.

By inferring the amplitudes in each bin independently, the m_X and t' dependence

[2] For the last two bins the number of events is smaller

can be recovered. As an example, the $a_2(1320)$ resonance is shown in fig. 3.7. The intensity in fig. 3.7a shows a peak around the resonance mass. The phase information, shown in fig. 3.7b, shows a rising phase motion with respect to the $1^{++}0^{+}\rho(770)\pi S$ reference wave. Resonant behavior occurs also in the reference wave, therefore usually no ideal 180° phase motion is visible.

The example shows the strength of the analysis method. Without any assumptions about the resonance shape, the fit is able to recover the mass dependence of the transition amplitudes. To extract the resonance parameters these results are supplied to a second fit.

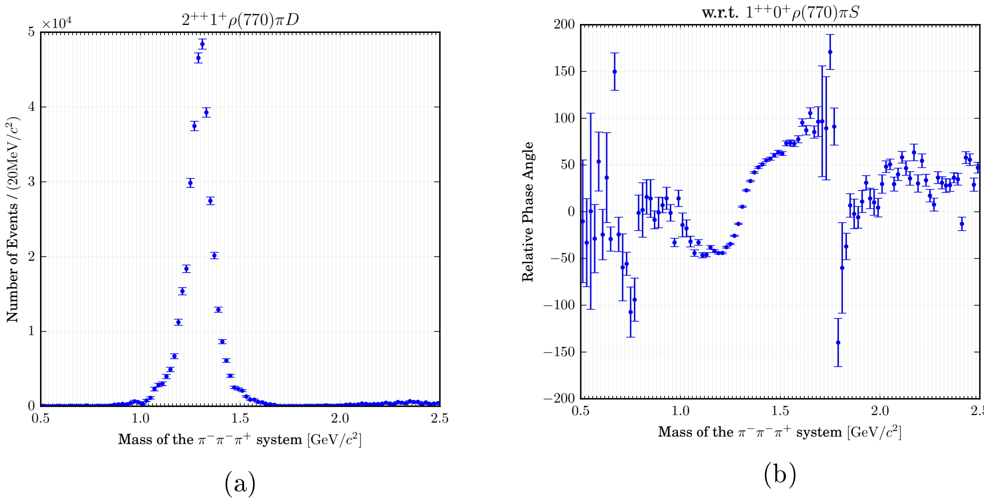


Figure 3.7: The result of the partial-wave decomposition for the $2^{++}1^{+}\rho(770)\pi D$ wave. The fits have been performed in the lowest t' bin over all m_X bins. In (a) the intensity spectrum is shown. A pronounced peak is visible that originates from the $a_2(1320)$ resonance. In (b) the phase relative to the $1^{++}0^{+}\rho(770)\pi S$ wave is shown. The rising phase motion coincides with the peak, indicating true resonant behavior.

t' bin [(GeV/c)^2]	Number of Events
0.100000-0.112853	5018351
0.112853-0.127471	4951799
0.127471-0.144385	4864385
0.144385-0.164401	4774514
0.164401-0.188816	4681303
0.188816-0.219907	4586075
0.219907-0.262177	4482491
0.262177-0.326380	4359759
0.326380-0.448588	4217365
0.448588-0.724294	3115895
0.724294-1.000000	873990

Table 3.1: List of the non-equidistant t' bins from lowest to highest. The two highest bins contain less events than the rest.

Chapter 4

Selection of Model Components

In the previous chapter, the partial-wave analysis formalism was introduced, which is based on modeling the observed intensity distribution. However the model in principle allows for an infinite number of partial waves. In practice, only a finite subset of partial waves can be used for the inference. Because of computational limitations and finite datasets, the relevant partial waves have to be included in this finite wave set, while irrelevant ones should be excluded from the model.

Selecting the significant components of a model is a challenging problem. The inclusion of too many waves leads to cross-talk between the waves that starts to describe statistical fluctuations and deviations from the model. This overadaption or overfitting of the model breaks the interpretability of the individual waves. On the contrary, models that leave out waves, that are relevant for the description of the data, suffer from so-called model leakage. The waves in the model try to describe structure in the data, that otherwise would have been described by additional waves. This effect also breaks the interpretability. To obtain sensible results, the wave set needs to minimize both of these undesired effects, consisting only of waves that are required to describe the data.

The wave set, used for all the analyses of the 3π final state [Ado+17; Haa13; Uhl16; Wal15; Sch14], has been selected by hand. It consists of 87 partial waves and the isotropic flat wave. For simplicity it is often referred to as the 88-wave set. Its waves are listed in appendix A. A detailed description can be found in Ref. [Haa13]. Selecting the model by hand may be problematic due to the large number of waves that need to be considered. This model relies of course on the experimentalist's intuition. The selection procedure is very time consuming and hard to reproduce. It may therefore suffer from an unknown selection bias.

The problem of selecting an appropriate wave set becomes significantly harder in less well-studied decay-channels like the $\pi^-\pi^-\pi^-\pi^+\pi^+$ final state. Analyses of this channel initiated the search for a systematic approach for the selection procedure. In Refs. [Bic16; Neu12] a genetic algorithm, mimicking evolution by combining waves from more successful models to effectively 'breed' a suitable wave set, has been developed. Because the algorithm experienced slow convergence and lacked a clear convergence criterion, this approach has been discarded.

An alternative method, called the Biggest Conceivable Model (BCM) method, was developed in Ref. [Bic16]. It is based on regularizing a Maximum Likelihood Estimation (MLE) fit of a large and systematically constructed set of waves via a penalty term and has produced promising results. The BCM method has been applied to $\pi^-\pi^-\pi^+$ data in the three-pion final state in Ref. [Dro15] and has been tested against the hand-selected 88-wave model.

Yet another method has been developed by the authors of Ref. [Gue+15]. This approach is based on a statistical method called the Least Absolute Shrinkage and Selection Operator (LASSO), developed by the author of Ref. [Tib96]. The LASSO has been successfully applied in the context of variable selection in many different contexts.

This chapter is concerned with introducing the different methods that have been tested on simulated and measured data. The properties of the methods will be discussed and possible advantages and disadvantages will be stressed.
First the selection by hand will be summarized in section 4.1. Then the notion of regularized maximum likelihood will be introduced in section 4.2. This forms the basis for the BCM and LASSO methods, which will be introduced in section 4.2.3.1 and section 4.2.3.2. Finally a new approach, called the MBCM method, that is aiming to combine the advantages of both methods while getting rid of undesired properties, is introduced in section 4.2.3.3.

4.1 Selection by Hand

The traditional approach for building a wave set is the selection of waves, that are considered relevant, by hand. Certain waves obviously need to be included because the corresponding resonances are directly visible in the measured intensity spectra of the $\pi^-\pi^-\pi^+$ final state and its $\pi^-\pi^+$ subsystem. However, this works only for dominant waves that usually have low spin. Smaller signals are much harder to detect as interference effects distort the intensities.
Combinatorially, this problem is impossible to resolve, because the number of possible wave sets quickly becomes large. One has to fall back to trial and error and one's physical intuition when constructing the wave set.
Additionally, some waves become very large, especially for lower 3π masses. These large intensities are the result of destructive interference between the waves. This

behavior is usually resolved by excluding problematic waves below a certain 3π mass threshold.

In Ref. [Haa13] a set of 88 waves, which consists of 80 waves with positive reflectivity, 7 with negative reflectivity and the isotropic background wave, was derived by removing waves with relative intensities below 10^{-3} from a larger set of 128 waves [Ado+17]. The building procedure of the model includes prior knowledge and therefore does not include waves that are not considered sensible to be included in the first place. The inclusion of this knowledge is, however, not performed in a systematic way. The resulting model may lack relevant waves because they were not considered or may miss combinations of waves that need to be included simultaneously to produce sensible results.

Overall, the procedure is very time consuming and hard to reproduce. The largest source of uncertainties for the partial-wave analysis in this final state is of a systematic nature. It is therefore important to use a systematic and reproducible way to obtain a wave set, which forms the very basis of this analysis. In the next section a type of methods will be introduced that try to tackle this problem.

4.2 Systematic Selection via Induction of Sparsity

As already mentioned in the introduction of this chapter, it would be advantageous to have a systematic approach for building a wave set. This section discusses different approaches that can be reduced to a common origin. All of these methods rely on the idea of a regularized maximum likelihood fit of a systematically constructed large wave pool of possible waves. The regularization is applied in form of an additional term that is added to the log-likelihood function. The methods differ in the type of regularization used and in the way the result of the fit is interpreted. These methods can be interpreted in a Frequentist or in a Bayesian framework.

4.2.1 Construction of the Wave Pool

Before any model-selection fit can be conducted, the infinite number of possible partial waves has to be reduced to a finite number, in a systematic way. Luckily the quantum numbers, that define the partial waves, impose a natural ordering according to which one would expect them to contribute: That is, waves with higher spin quantum number J and higher orbital angular momenta L require more energy to be produced and should therefore be suppressed compared to waves with smaller J and L. For the analysis presented here, the total spin has been limited to $J \leq 6$ and the maximum orbital angular momentum was also limited to $L \leq 6$. Additionally, higher spin projections are expected to be suppressed for the analyzed t' range and

should therefore contribute less. Here, the maximum spin projection is limited to $M \leq 2$.
Together with the restriction to the six possible isobars, presented in section 3.2, this reduces the infinite number of partial waves to a finite wave pool of 432 waves.
The construction of the wave pool can of course be changed if one would like to investigate different isobars or higher spins, angular momenta or spin projections. Because this wave pool has been used in the previous analysis of Ref. [Dro15] it is reused here in order to provide easier comparison.

4.2.2 Regularization

As already mentioned before, the analysis relies on regularization of the MLE fit. The idea is to add an additional term, a so-called penalty or regularization term, to the log-likelihood, that disfavours solutions that are considered problematic for some reason. These terms become necessary in different situations. For example to circumvent overfitting when the number of data points is very small compared to the number of parameters to estimate or when the correlation between parameters becomes too strong for the data to resolve them.

There are special kinds of penalties that are capable of selecting parameters, meaning that they strongly push parameters, that are only weakly supported by the data, towards zero and thereby 'deselecting' them. These are of course the type of penalties interesting for this analysis and several will be introduced in the following subsection.

While the idea of adding penalty terms may sound like an arbitrary idea at first it can be very well motivated in a Bayesian picture of the analysis. The usual MLE fits used in many previous analyses like Ref. [Haa13] can be regarded in the context of Bayes' theorem eq. (4.1). In Bayesian statistics the only sensible probability density (PDF) function for parameter inference is the posterior PDF. For a model M_i, that consists of a specific set of partial waves, and a dataset D, this is the probability of the model given the data $P(M_i \mid D)$. It is expressed as the product of the likelihood $P(D \mid M_i)$, which is the PDF of the data given the model, and the prior PDF $P(M_i)$ of the model, divided by the probability of the data $P(D) = \sum_i^m P(D \mid M) \cdot P(M_i)$ or evidence. The prior PDF contains knowledge of the model, that is independent of the data.

$$\mathcal{P}(M_i \mid D) = \frac{P(D \mid M_i) \cdot P(M_i)}{\sum_i^m P(D \mid M) \cdot P(M_i)} \tag{4.1}$$

For a fixed dataset D, the evidence is a constant with respect to the model M_i. The posterior PDF is therefore proportional to the product of the likelihood and the prior. For this analysis, the model M_i is equivalent to a set of partial waves, that is labeled by the vector of complex transition amplitudes $\boldsymbol{T}$. Writing the likelihood as a function of $\boldsymbol{T}$ [1], this leads to the relation:

$$\mathcal{P}\left(\boldsymbol{T} \mid D\right) \propto \mathcal{L}\left(\boldsymbol{T}; D\right) \cdot P\left(\boldsymbol{T}\right) \tag{4.2}$$

For inference of the transition amplitudes the posterior can be used. In Bayesian statistics, samples can be drawn from the posterior and transition amplitudes can be estimated by the mean. Sampling is however not feasible in high-dimensional settings, like this analysis, where hundreds of parameters need to be considered. Instead, similar to the maximization of the likelihood, the posterior PDF can be maximized. To obtain this maximum a posteriori (MAP) estimate, the transition to the log-posterior is made:

$$\log \mathcal{P}\left(\boldsymbol{T} \mid D\right) = \log \mathcal{L}\left(\boldsymbol{T}; D\right) + \log P\left(\boldsymbol{T}\right) + C \tag{4.3}$$

Where C is an unimportant constant, that does not change the estimate of the maximum. It corresponds to a scaling factor of the posterior. In the log-picture, the prior transforms to a penalty or regularization term that is added to the log-likelihood. In the next subsection, the details of the penalty choice will be elaborated.

4.2.3 Sparsity Inducing Regularization

In this section three different regularization terms or priors will be introduced. They all have in common the idea of treating the different waves as exchangeable and thereby equal. This means that the prior on the complex parameter vector $\boldsymbol{T}$ factorizes in a product of individual and identical priors on the intensity $|T_i|^2$ of the waves.

$$P\left(\boldsymbol{T}\right) = \prod_i P(|T_i|^2) \tag{4.4}$$

In the log-likelihood picture this transforms to a sum of penalty terms.

$$\log P\left(\boldsymbol{T}\right) = \sum_i \log P(|T_i|^2) \tag{4.5}$$

In the case $P(\boldsymbol{T}) = 1$ the posterior PDF is equal to the likelihood, the penalty term vanishes and the MAP estimate is equal to a simple MLE fit. This would then be considered a non-informative ansatz, as any value of $\boldsymbol{T}$ has the same penalty, namely none.

[1] $P\left(D \mid \boldsymbol{T}\right) \equiv \mathcal{L}\left(\boldsymbol{T}; D\right)$

The special type of penalties used for model selection have the common property of concentrating the probability around zero. This forces the MAP estimate to favor waves with small intensities. The different terms, introduced in the following, differ in the details of the high-intensity limit and the region around zero.

More information on sparsity can be found in Refs. [OS09; PC08; Tib96].

4.2.3.1 BCM Regularization

In Ref. [Bic16] the so-called biggest conceivable model method (BCM) has been developed and used for inference of the transition amplitudes for five-pion data. In Ref. [Dro15] it has been applied to measured three-pion data. The motivation for this type of regularization was to favor small intensities while not introducing a very strong bias on big intensities. This resulted in a prior of the from eq. (4.6).

$$P_{\mathrm{BCM}}\left(|T_i|^2;\Gamma\right) = \frac{1}{1+\frac{|T_i|^2}{\Gamma^2}} \tag{4.6}$$

Before, this has been considered a half-cauchy prior PDF on the magnitude of the transition amplitude $|T_i|$. Because the physically meaningful quantities are intensities and relative phases between the waves, in this thesis the term will be interpreted as an improper prior on the intensity $|T_i|^2$ rather than on the magnitude. Where improper denotes a non-normalizable function, which can by definition not be a PDF. Improper priors are no problem for meaningful inference as long as the posterior remains a proper or normalizable PDF. This is the case, as the exponential of the poisson term in the likelihood is dominating the posterior in the large parameter limit and guarantees a finite normalization integral. For simplicity, the improper prior will still be referred to as prior in the following. Regardless of a proper or improper prior the motivation of the regularization remains the same.
The BCM prior has one adjustable parameter or hyper parameter Γ that controls the scale of the prior. It is interesting to point out the limiting behavior. In the limit of a small scale parameter or large intensity, the 1 in the denominator is negligible. This results in the limit eq. (4.7). Because the scale parameter is constant with respect to the model parameters it is unimportant for the optimization procedure.

$$P_{\mathrm{BCM}}\left(|T_i|^2;\Gamma\right) \approx \frac{\Gamma^2}{|T_i|^2} \propto \frac{1}{|T_i|^2} \tag{4.7}$$

From eq. (4.7) the BCM penalty follows:

$$\log P_{\mathrm{BCM}}\left(|T_i|^2;\Gamma\right) = -\log\left(1+\frac{|T_i|^2}{\Gamma^2}\right) \tag{4.8}$$

It is plotted in the complex plane in fig. 4.1 and a cut along the real axis is shown in fig. 4.2. The logarithm causes this type of penalty to not be significantly stronger for a wave that has an already large intensity and a wave that has a larger intensity. In this case one can speak of a weakly-informative prior or penalty.

The penalty has been applied to all waves but the isotropic flat wave, which is untented to absorb background events.

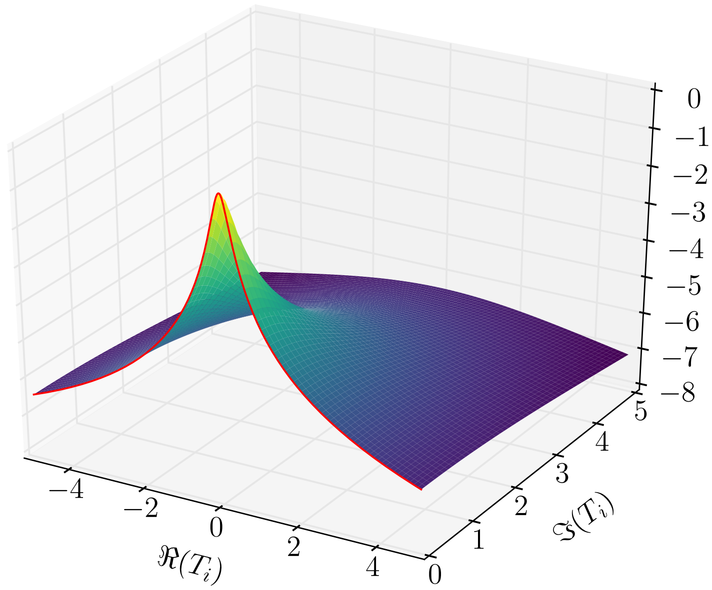


Figure 4.1: Two-dimensional plot of the BCM penalty in the complex plane. For easier visualization of the shape the function has been sliced at the zero axis. A clear spike with a smooth top around zero is visible. The slice is also plotted separately in fig. 4.2

4.2.3.2 LASSO regularization

This regularization technique has been introduced to amplitude analysis in the work of Ref. [Gue+15]. The authors apply the so-called Least Absolute Shrinkage and Selection Operator (LASSO) regularization, that was originally introduced in the

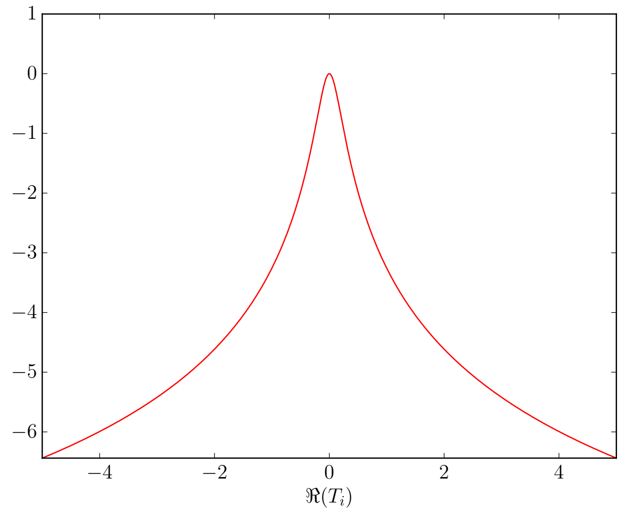


Figure 4.2: Slice of the two-dimensional and symmetric BCM penalty.

context of ordinary least square fits by R. Tibshirani in Ref. [Tib96], to an extended likelihood fit. The idea is originating from a frequentist's perspective and was originally formulated as a penalty term of the following form:

$$\log P(\boldsymbol{\beta}) = -\lambda \sum_i |\beta_i| \tag{4.9}$$

Where $\boldsymbol{\beta}$ is the vector of parameters, λ is a free hyper-parameter that adjusts the strength of the penalty and $|\beta_i|$ is the absolute value or l_1 norm of the individual parameter. This type of penalty is also sometimes called the l_1 penalty or regularization. For $\lambda = 0$ no regularization is applied and the fit is equal to a MLE fit. The LASSO can also be interpreted as individual double exponential or laplace priors on the parameters, as has already been remarked by the original author in Ref. [Tib96]. For comparison with the BCM prior it is advantageous to also make the transition $\lambda = \frac{1}{\Gamma}$ and write the prior corresponding to the LASSO with a width parameter Γ as:

$$P(\boldsymbol{\beta}) = \prod_i \exp\left(\frac{-|\beta_i|}{\Gamma}\right) \tag{4.10}$$

Starting from this, a similar penalty as proposed by the authors of Ref. [Gue+15] is used for the LASSO method. As a prior it can be written as:

$$P_{\mathrm{LASSO}}\left(|T_i|^2; \Gamma\right) = \exp\left(-|T_i|/\Gamma\right) \tag{4.11}$$

Instead of the absolute values of the parameters the magnitude of the transition amplitudes is used. This ensures that no information about the relative phase of a wave is used. The authors of Ref. [Gue+15] apply this penalty also to the isotropic flat wave. The same procedure has been applied in this analysis at first, but was then replace by a variant without penalized flat wave for measured data. The differences between the variants were small. The penalty, which is the logarithm of the prior, is plotted in the complex plane in fig. 4.3 and the corresponding slice along the real axis in fig. 4.4.
The absolute value of the complex transition amplitudes can be written as:

$$|T_i| = \sqrt{(\operatorname{Re} T_i)^2 + (\operatorname{Im} T_i)^2} \tag{4.12}$$

It is worth noting that the absolute value is not smooth at 0 and therefore is problematic for gradient descent algorithms used to find the mode. For the LASSO problem several specialized solvers exist. Tests with some of this fitters however did seem to have problems with the complex likelihood landscape of this problem and had trouble to converge. Therefore, similar to the authors of Ref. [Gue+15], ignoring the non-differentiability at 0 still seems to produce reasonable results. The time for the algorithm to converge, on the other hand suffers massively by this, so a smoothed version of the absolute value is used.

$$|T_i| \approx \sqrt{T_i T_i^* + \epsilon} \quad \text{for } \epsilon \to 0 \tag{4.13}$$

Where $\epsilon = 10^{-5}$ has been chosen as a reasonable small value. This improved the convergence time and made the approach applicable in practice while keeping the convexity of the penalty. The difference between the smoothed and non-smoothed version is shown in fig. 5.21.
The strong concentration around 0 and the non-differentiability of this penalty are capable of producing actually sparse solutions. This is the reason why the LASSO is so popular for model selection.
In contrast to the BCM penalty the exponential tails of the LASSO are not heavy. This results in the property that the penalty grows strongly also for already large intensity waves and therefore significant bias is introduced to the bigger intensity waves as well.

The LASSO is strongly dependent on the choice of its tuning parameter. In Ref. [Gue+15] the usage of two, so-called information criteria, is suggested by the authors. These criteria are a measure for the trade-off between better likelihood and the number of parameters in the model. The Akaike Information Criterion (AIC) was introduced in Ref. [Aka74]. It is defined as follows:

$$\mathrm{AIC} = -2 \log \mathcal{L} + 2r \tag{4.14}$$

Where r are the degrees of freedom of the model. In Ref. [Gue+15] the number of waves over a certain intensity threshold has been used to estimate r. Here the number of parameters is used, which is twice the number of waves minus one parameter in each incoherent sector, because within one sector the global phase in undefined.
The second suggested information criterion is the Bayesian Information Criterion (BIC), which was first introduced in Ref. [Sch78]. It is given by the following formula:

$$\mathrm{BIC} = -2\log\mathcal{L} + r\log n \tag{4.15}$$

Where n is the size of the sample under study.
Both criteria of the $-2\log\mathcal{L}$ term, which decreased for better values of the likelihood, and a term that penalizes more parameters of the model. The criteria can be used to find an appropriate parameter of the LASSO, by 'scanning' different values of Γ and calculating the information criteria. The parameter value that minimizes the information criterion is then selected.
The criteria have a different origin and are derived under different assumptions, so one should decide on one of the two. The authors of [Gue+15] suggest the usage of both and including the different results in the systematics of the fit.
It is not clear, whether the assumptions used to derive the criteria hold for this analysis. Therefore, they are applied to serve as a guide for the parameter choice.

4.2.3.3 MBCM Regularization

Finally a method based on a modification of the BCM penalty is introduced in this section, which will be called modified BCM or MCBM for short in the following.
In the course of this thesis it will become evident that the LASSO penalty has some desirable regularization properties due to its stronger bias on larger intensities while the BCM method shows good properties required for selection of the model. These observations will be explained in detail in chapters 5 to 7.

A modification of the BCM method can be made that combines the two properties into one prior term with two free parameters Γ and λ. The LASSO can be expressed as a so-called scale-mixture of Gaussians, as has for example been remarked in Ref. [PC08]. Where a scale-mixture denotes an integral over the variance of the Gaussian of a product of the Gaussian and a PDF of the variance. So in a mathematical sense the scale-mixture is just a convolution.

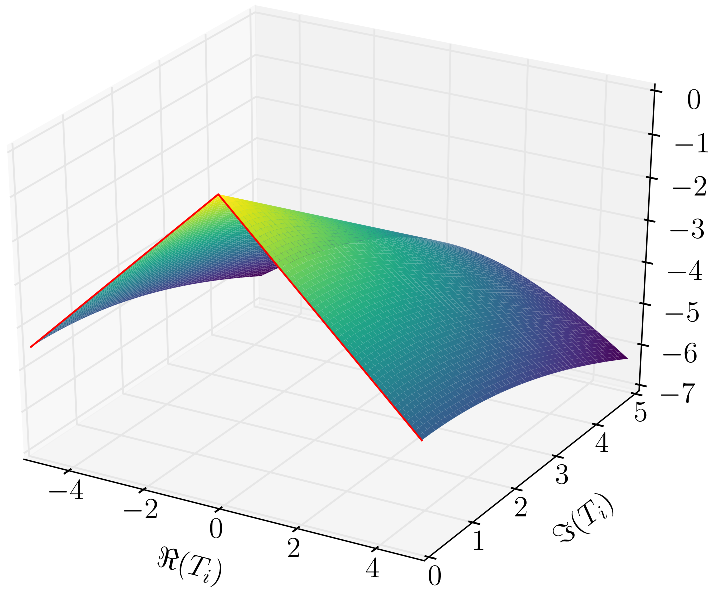


Figure 4.3: Two-dimensional plot of the LASSO penalty in the complex plane. For easier visualization of the shape the function has been sliced at the zero axis. The penalty is a symmetric cone around zero, where it is non-differentiable.

The LASSO prior is therefore expressed as:

$$P_{\text{LASSO}}\left(|T_i|^2;\Gamma\right) = \mathcal{N}_0 \int_0^\infty \underbrace{\frac{1}{\sigma}\exp\left(-\frac{|T_i|^2}{\sigma^2}\right)}_{\text{Gaussian in } |T_i|} \underbrace{\exp\left(-\sigma^2/(4\Gamma^2)\right)}_{\text{mixture PDF of } \sigma^2} d\sigma^2 = \bar{\mathcal{N}_0}\exp\left(-|x|/\Gamma\right) \tag{4.16}$$

Where $\mathcal{N}_0$ and $\bar{\mathcal{N}_0}$ are unimportant normalization constants.
The BCM prior can be expressed in a similar fashion:

$$P_{\text{BCM}}\left(|T_i|^2;\Gamma\right) = \mathcal{N}_0 \int_0^\infty \underbrace{\frac{1}{\sigma}\exp\left(-\frac{|T_i|^2}{\sigma^2}\right)}_{\text{Gaussian in } |T_i|} \underbrace{\frac{1}{\sigma^3}\exp\left(-\frac{\Gamma^2}{\sigma^2}\right)}_{\text{mixture PDF of } \sigma^2} d\sigma^2 = \frac{\mathcal{N}_0}{\Gamma^2 + |T_i|^2} = \frac{\bar{\mathcal{N}_0}}{1 + \frac{|T_i|^2}{\Gamma^2}} \tag{4.17}$$

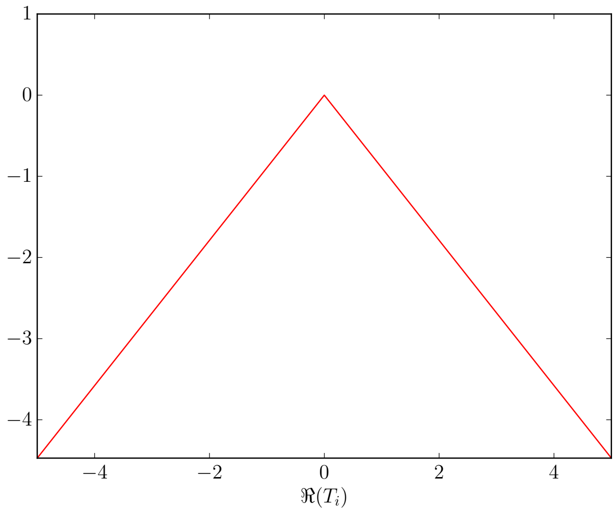


Figure 4.4: Slice of the two-dimensional and symmetric LASSO penalty.

Both priors concentrate their probability at zero, but differ in their tail behavior and around zero. From eq. (4.16) and eq. (4.17) this can be understood by their different mixing PDFs. During the study of suitable penalties for this analysis, the following modification of the BCM mixing PDF was motivated empirically:

$$\begin{aligned} P_{\text{MBCM}}\left(|T_i|^2;\Gamma,\lambda\right) &= \mathcal{N}_0 \int_0^\infty \underbrace{\frac{1}{\sigma}\exp\left(-\frac{|T_i|^2}{\sigma^2}\right)}_{\text{Gaussian in } |T_i|} \underbrace{\frac{1}{\sigma^3}\exp\left(-\frac{\Gamma^2}{\sigma^2}\right)\exp\left(-\frac{\sigma^2}{\lambda}\right)}_{\text{mixture PDF of } \sigma^2} d\sigma^2 \\ &= \mathcal{N}_0 \frac{2\mathcal{K}_1\left(\frac{2\sqrt{|T_i|^2+\Gamma^2}}{\sqrt{\lambda}}\right)}{\sqrt{\lambda}\sqrt{|T_i|^2+\Gamma^2}} \end{aligned} \tag{4.18}$$

Where $\mathcal{K}_i(\cdot)$ is the modified Bessel function of the second kind. The additional exponential term $\exp\left(-\frac{\sigma^2}{\lambda}\right)$ in the mixing PDF modifies the tails of the BCM prior for large intensities $|T_i|^2$. This can be understood by looking at eq. (4.18). In the limit of large λ the BCM prior is recovered, because the additional exponential becomes unity:

$$P_{\text{MBCM}}\left(|T_i|^2;\Gamma,\lambda\right) \approx \frac{\mathcal{N}_0}{\Gamma^2+|T_i|^2} \mathrel{\hat{=}} P_{\text{BCM}}\left(|T_i|^2;\Gamma\right) \quad \text{for } \lambda\to\infty \tag{4.19}$$

So the scale parameters of BCM and MBCM priors correspond to each other $\Gamma_{\mathrm{BMC}} \hat{=} \Gamma_{\mathrm{MBCM}}$. For finite λ the MBCM prior is dominated by an exponential in the high intensity limit:

$$P_{\mathrm{MBCM}}\left(|T_i|^2;\Gamma,\lambda\right) \approx \mathcal{N}_0 \frac{\sqrt{\pi}}{\sqrt[4]{\lambda}} \left(\frac{1}{|T_i|^2}\right)^{3/4} \exp\left(-\frac{2\sqrt{|T_i|^2+\Gamma^2}}{\sqrt{\lambda}}\right) \quad \text{for } |T_i|^2 \to \infty \tag{4.20}$$

This limiting behavior should introduce a bias on large intensities that is similar to the LASSO. The scale of the LASSO Γ_{LASSO} can loosely be related to the decay parameter λ of the MBCM prior by comparing the mixing PDF of the LASSO with the exponential modification term of MBCM prior. The following relation can be made:

$$\Gamma_{\mathrm{LASSO}} \hat{=} \frac{\sqrt{\lambda}}{2} \tag{4.21}$$

The MBCM prior is able to combine the properties of the BCM and LASSO priors and should therefore provide a more flexible approach. It was discovered that this type of prior was introduced, in a different context, before by the authors of Ref. [CD08].

In fig. 4.5 the MBCM penalty, which is the logarithm of the MBCM prior, is plotted for $\lambda = 5$. The corresponding slice along the real axis is shown in fig. 4.6. The penalty looks similar to the BCM penalty shown in fig. 4.1, but falls off steeper. This behavior is better visible in fig. 4.7, where it is compared with the BCM penalty for the same scale parameter $\Gamma = 0.2$, but different λ. The BCM penalty corresponds to the case $\lambda = \infty$. The plot is a slice through the symmetric two-dimensional penalties. For smaller values of λ, the penalty decreases faster away from zero and introduces more bias on waves with large intensity. For larger values the MBCM approaches the BCM penalty, shown in red.

Like for the BCM penalty, the flat waves has not been penalized by the MBCM method.

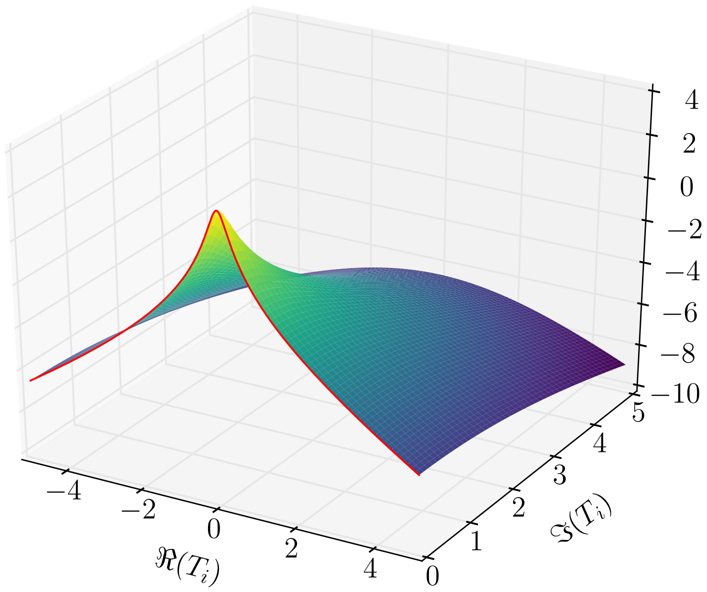


Figure 4.5: Two-dimensional plot of the MBCM penalty in the complex plane. For easier visualization of the shape the function has been sliced at the zero axis. The penalty is similar to the BCM penalty (compare fig. 4.1), but steeper for values that are further away from zero.

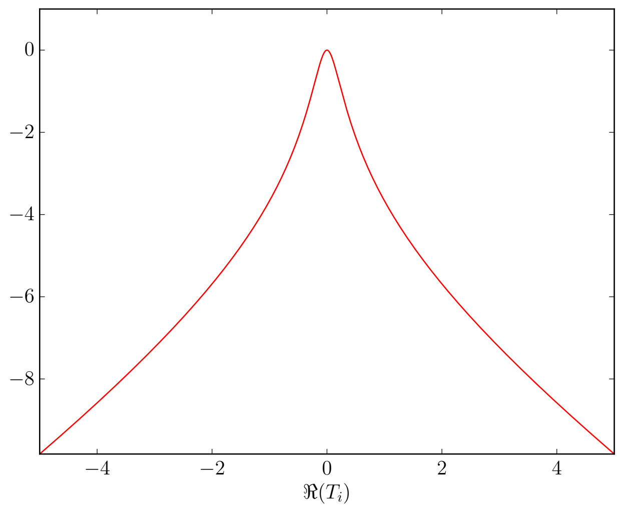


Figure 4.6: Slice of the two-dimensional and symmetric MBCM penalty.

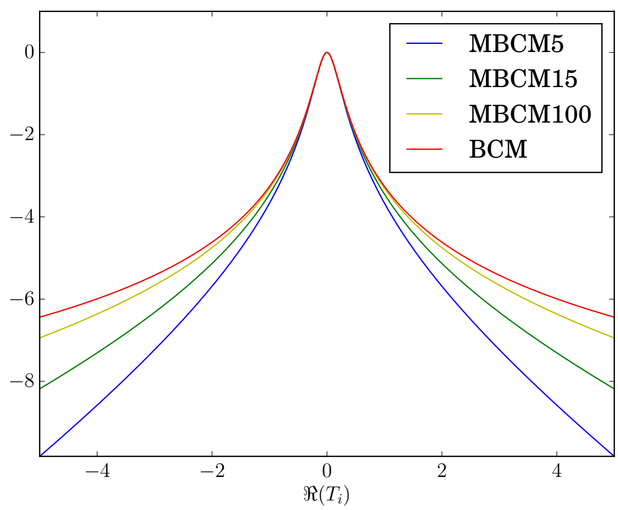


Figure 4.7: Comparison of the MBCM penalty for different parameters λ with the BCM penalty. The MBCM is plotted for three different values of λ (blue: $\lambda = 5$, green: $\lambda = 15$, yellow: $\lambda = 100$). For larger values the penalty approaches the BCM penalty, plotted in red.

Chapter 5

Model-Component Selection on Simulated Data

A method for selecting the relevant model components is expected to perform best in the case when the true model is among all possible models considered. On measured data, it is in general highly unlikely that the statistical model used for inference is the true model that is describing the data. In fact, it has been remarked that true models do not exists and statistical models can merely be a useful approximation [Gel08; Box76]. Still, if a model selection procedure fails in the idealized case of simulated data generated according to the model later used for inference, then it is expected to fail also in the case of measured data.

In this chapter such an idealized test, that can be conducted on simulated data, is discussed. Monte Carlo methods are used to generate datasets according to the model used for inference. On these datasets the different model selection procedures have been tested. In section 5.1 their important properties are presented and a connection to the problems observed on measured data will be made.
For comparison, fits without any of the model selection penalties have been performed in section 5.2. The first model-selection technique that is discussed in section 5.3 is the BCM method. For the three-pion final state, this method has so far only been applied to measured data and a verification on simulated data has only been conducted for the five-pion final state, where it showed deviations in the behavior compared to the application on measured data. This issue will be discussed and a resolution is presented. It has also been observed in previous studies that the fit with the BCM method does not always produce unique solutions and therefore the reasons for this instability of the inference will be discussed in section 5.4.
In section 5.5 the LASSO method will be tested and the results will also be discussed in the context of inference stability.
Finally the new MBCM method is applied and its advantages over the two other methods are discussed in section 5.6.

5.1 Details of the Simulated Data

The analysis method relies on independent inference in bins of m_X and t'. Past analyses revealed that optimization of the likelihood may not always result in unique solutions for the inferred amplitudes, depending on the mass-region of the bin. For this study, data in two mass-bins has been generated that should represent regions for which the inference works fine and for which it is more problematic.

One bin has been generated at a mass of $1.81\,\mathrm{GeV}/c^2$. This high-mass region is known to be quite unproblematic for inference on measured data [Dro15; Haa13]. It is expected to lead to the most well-behaved results.

A second bin has been generated at $0.99\,\mathrm{GeV}/c^2$. The low-mass region is more problematic. There fit attempts with random start values may lead to results that strongly differ. The reason for this are many local maxima of the likelihood and no global maximum can be easily identified. Additionally the largest maximum, which was found, did not necessarily produce physically reasonable results. This required thresholding of certain waves on measured data in the low-mass region.
This low-mass bin should serve exemplary for this problematic behavior.

The simulated data serves as a test in order to understand to which extend problems of the model selection are intrinsic to the analysis method and appear even in the case when the true model is contained as a subset of the wave pool.
To account for local minima several hundred fit attempts with random starting positions have been performed. In some studies the number of attempts has been increased to thousand, but this was also not enough to reliably find a global optimum.

The data has been generated according to the results obtained by fits with the hand-selected model on measured data. The contribution of the isotropic background wave has not been generated.
In fig. 5.1a and fig. 5.1b the results of a MLE fit with the exact input model, used for generation of the set, are shown. The complex transition amplitudes are plotted in the complex plane. These fits will from now on be called reference fits.
The phase of the amplitudes will be unimportant for the model selection methods as the prior or penalty term does not depend on it. A simpler representation of the fit result can be achieved by looking at the intensity of the waves. This is simply the magnitude squared of the transition amplitude. In fig. 5.1c and fig. 5.1d the intensities of the reference fits are plotted in descending order on a logarithmic scale. The waves span an intensity range over several orders of magnitude.
These types of plots will also be used for the fits with different model selection techniques, with the reference fits, shown in this section, superimposed in red.

name	mass [GeV/c^2]	number of events	number of waves
Set A1	1.81	38675	87
Set A2	1.81	10000	87
Set B1	0.99	115700	65
Set B2	0.99	30000	65

Table 5.1: Details of the simulated datasets used to test the model selection methods. Most of the studies have been performed on set A1 and B2. Set A2 was used to test properties of the model selection for smaller sample sizes in the high-mass bin. Set B1 was used to test properties of the model selection for larger sample sizes in the low-mass bin. To cope with long convergence times in the low-mass region, set B2 was created as a subset of B1 and the detailed studies have been performed on this subset.

The number of generated events is of the order of several thousand events, which is similar to the typical number of events measured in the experiment. For the low-mass region a dataset size of 30000 events has been used for most of the studies. This sample is a subset of a larger dataset of 115700 events. The subset has been used to cope with the computational effort required for the analysis. The large sample was used to verify certain properties of the model selection. This will be discussed in more detail in the respective sections.
For the high-mass bin the studies have been conducted on a dataset with 38675 events. Similar to the low-mass sample, some effects needed to be verified for a different number of events. Here a smaller dataset has been generated, containing 10000 events.
A summary of the datasets can be found in table 5.1. In the right column the number of waves is given. This number is smaller in the $0.99\,\mathrm{GeV}/c^2$ bin because of the thresholds used in the hand-selected model according to which the data has been generated.

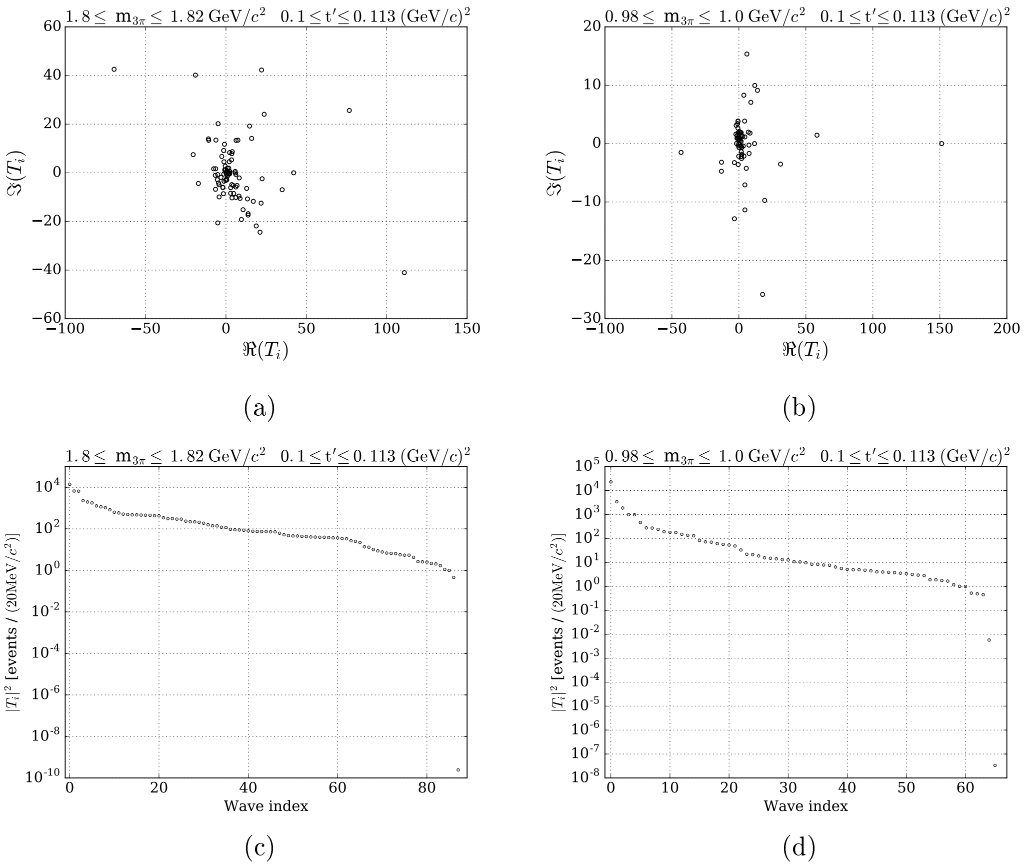


Figure 5.1: Refit of the high- and low-mass bin of simulated data with the respective input model.

5.2 Fit without regularization

A first test is fitting the simulated datasets with the complete wave pool without any regularization. This is then analogue to the MLE approach used for inference with the hand-selected model, but with an increased number of waves in the model. The results serve to visualize the importance of regularization in these types of fits. Different properties of the fit can be observed in the different mass bins.
Multiple local maxima of the likelihood lead to instabilities of the fit, meaning that multiple solutions can be found by the optimization process. A very prominent source of instabilities of the model selection is connected with the inclusion of waves with negative reflectivity. This effect will be discussed in the respective subsections.

First the results of the higher mass bin at $1.81\,\mathrm{GeV}/c^2$ will be discussed. It is expected from the results obtained on measured data that this bin should be well-behaved.
In fig. 5.2 the results of the several hundred fit attempts from random starting positions are presented. The intensities of the waves for the best fit, which is the result with the largest likelihood, are shown in grey. They are plotted in descending order.
The results with a smaller likelihood are sorted with respect to the best fit and drawn underneath. The difference in log-Likelihood between the result and the best fit is color-coded from dark purple to yellow. Where purple corresponds to a smaller difference or better likelihood and yellow to the opposite.
If the scattering of the intensity in the different solutions is small, the wave is found in a stable manner.
To gauge the quality of the results the reference fits are overlaid in red such that they are drawn on top of the corresponding wave of the best fit.
In the logarithmic scale no scattering of the solutions is visible for waves with intensities above several hundreds of events and the reference result matches the best fit. For lower intensities the fit attempts show a clear scattering over the different solutions.
In the lower part of fig. 5.2 the difference in intensity between the reference fit and the best fit is plotted for the waves of the generating model on a linear scale. For the largest wave this difference is of the order of thousand events.
It is interesting to perform the fits with a reduced wave pool from which all waves of negative reflectivity have been removed. The MLE fit for the residual 236 waves shows that the stability over the different fit attempts greatly increases as can be seen in fig. 5.3. No scattering is visible in this case, except for the smallest wave, which is the isotropic background wave. Furthermore the difference in intensity has decreased to several hundred events.
For the high-mass bin one can note the following findings. In both cases waves which were not generated in the dataset do pick up some intensity. The inference for the positive waves that were generated looks reasonable. The inclusion of the negative sector worsens the result. At this point it is not clear if this occurs simply because of the higher dimensionality of the problem when the negative reflectivity waves are included. This will be discussed in more detail for the application of the model selection.

The same fits are performed again in the lower mass bin at $0.99\,\mathrm{GeV}/c^2$. The fit results for the complete wave pool is shown in fig. 5.4 and the one for the reduced wave pool in fig. 5.5. Here in both cases the deviations from the reference fit result

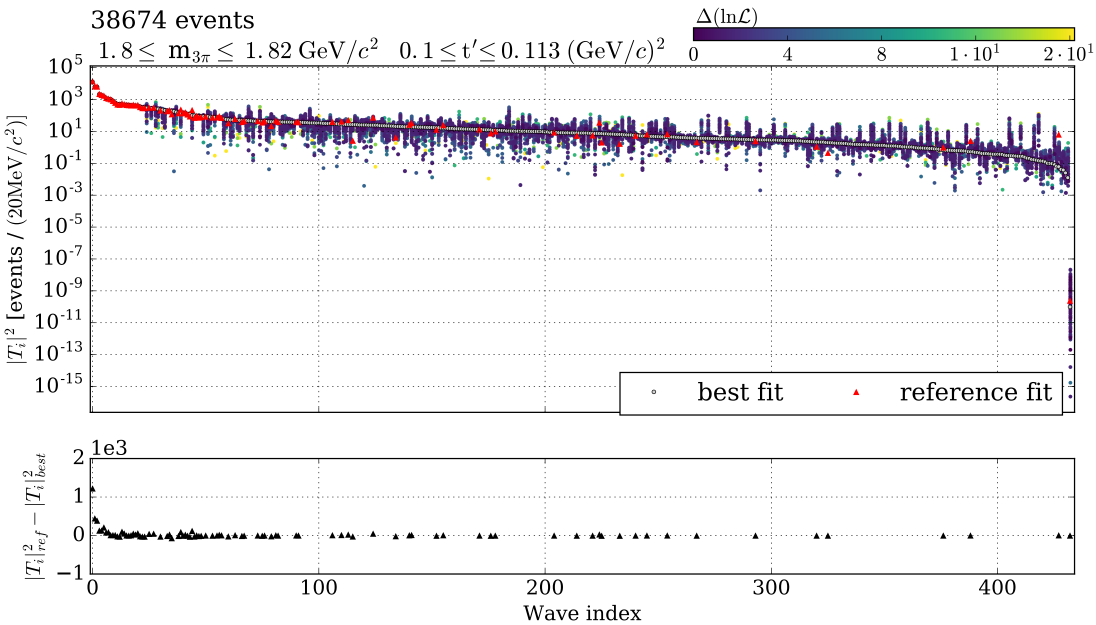


Figure 5.2: Intensities of a fit with the complete wave pool of 432 waves to simulated data at 1.81 GeV/c^2. The result with the largest likelihood is shown in grey. The intensities have been sorted in descending order. Additional results are ordered according to the best result. The result of the reference fit is superimposed in red. The other attempts are color-coded according to their absolute difference in likelihood with respect to the best result.
The lower plot shows the absolute difference in intensity between the reference fit and the best fit.

become very large. It is not unexpected that the fits are more unstable in this bin as the same behavior is found for measured data. The main source of instability in the higher mass region was introduced by waves with negative reflectivity. In contrast to the well-behaved high-mass bin, the fluctuations over the fit attempts persist even with the reduced wave pool.
Furthermore it had to be excluded that the fluctuations are not merely a result of the minimally smaller sample size of $30 \cdot 10^3$ events used in this sample. Therefore the fit has also been repeated on the larger sample of about $116 \cdot 10^3$ events. In fig. 5.6 the result of the fit with the reduced wave pool is shown. The increased size of the sample is not able to resolve the issues.
The application of the unregularized MLE fit on the simulated datasets revealed the following. Like for measured data the fits are better behaved for data at high masses. For the 1.81 GeV/c^2 bin the inference worked quite well, but a selection of

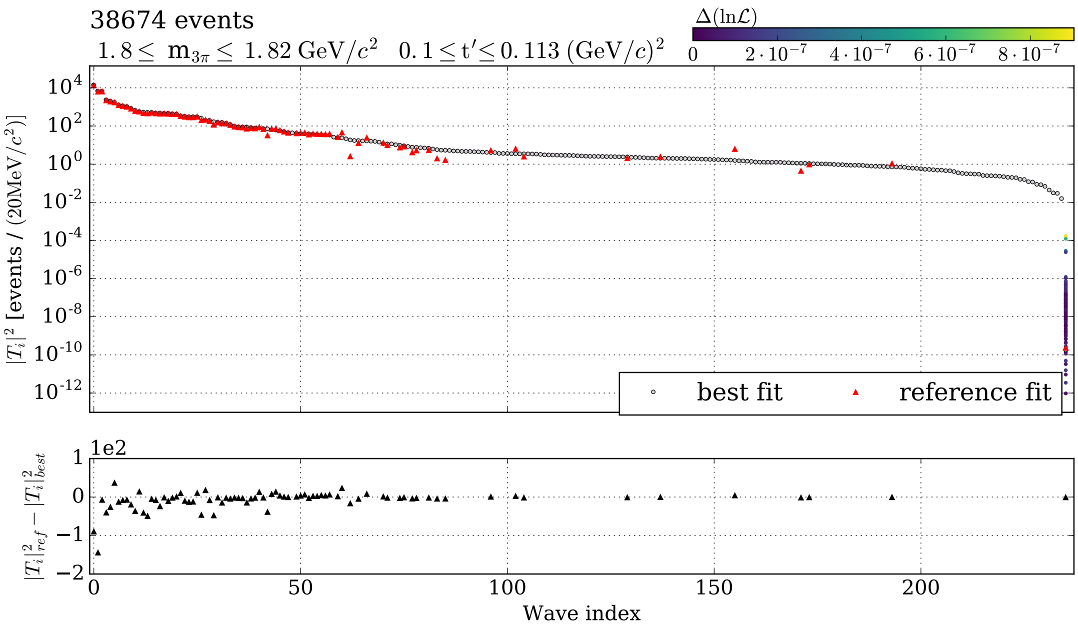


Figure 5.3: Fit of the reduced wave pool of 236 waves to simulated data at 1.81 GeV/c^2. This plot is analogue to fig. 5.2. The fit without negative reflectivity waves shows less scattering over the different fit attempts.

the important waves is not possible as all waves obtain some intensity.
Inference in the low-mass bin at 0.99 GeV/c^2 is practically impossible and the fit is unstable in this region.
For both bins the inclusion of negative reflectivity waves makes the results worse.

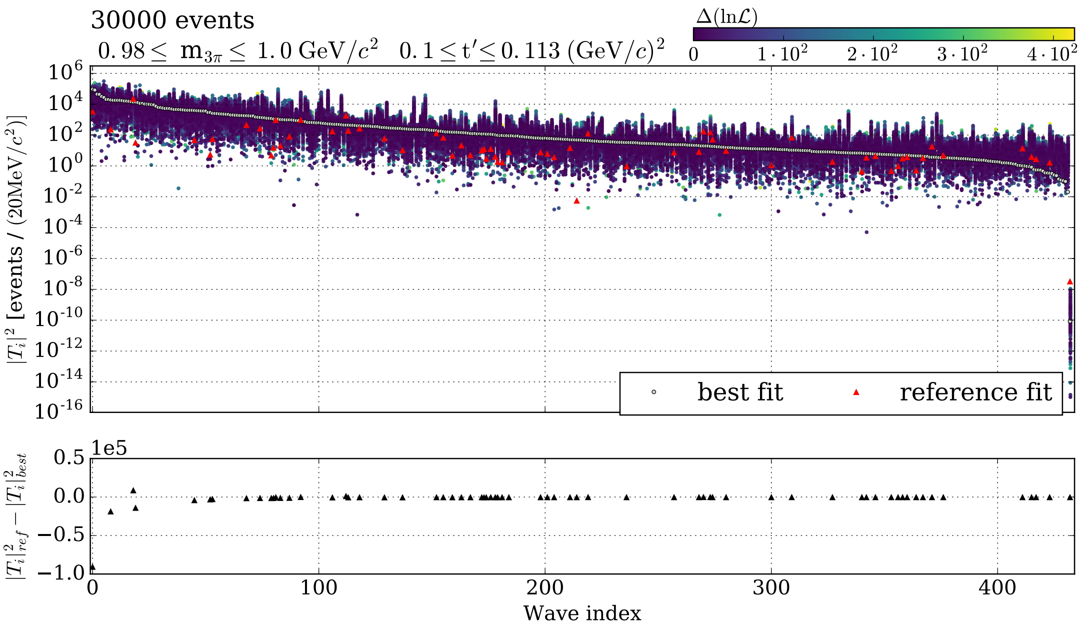


Figure 5.4: Intensities of a fit with the complete wave pool of 432 waves to simulated data at $0.99\,\mathrm{GeV}/c^2$. This plot is analogue to fig. 5.2 but for the low-mass bin. Here many waves show large intensities and the reference fit result cannot be recovered well.

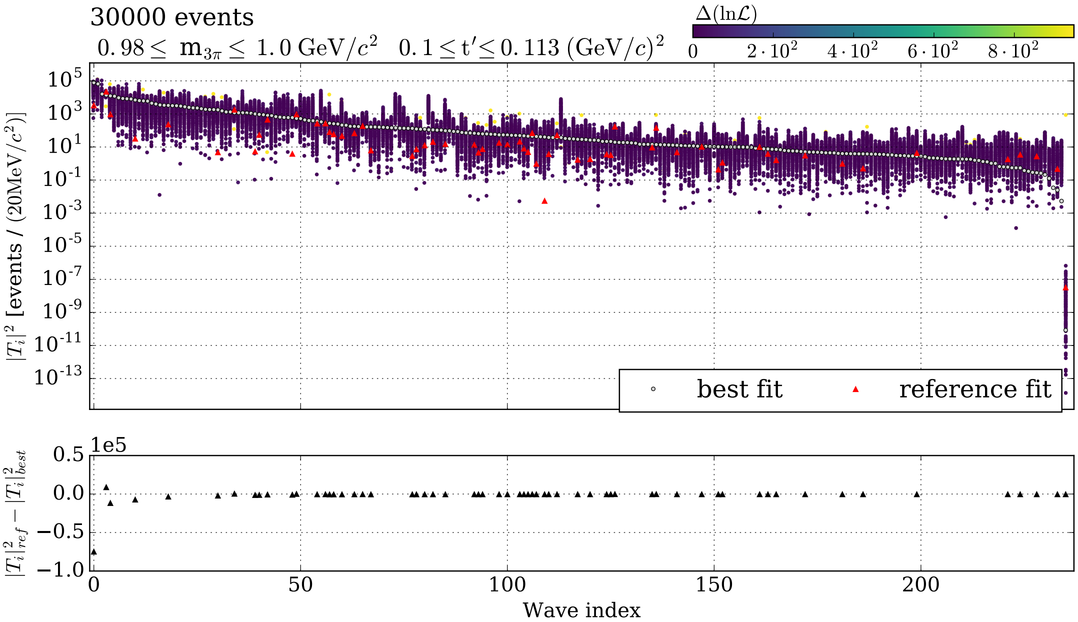


Figure 5.5: Fit of the reduced wave pool of 236 waves to simulated data at $0.99\,\mathrm{GeV}/c^2$. This plot is analogue to fig. 5.2 but for the low-mass bin and with the reduced wave pool without negative reflectivity waves. Even with the reduced wave pool the reference fit result cannot be recovered.

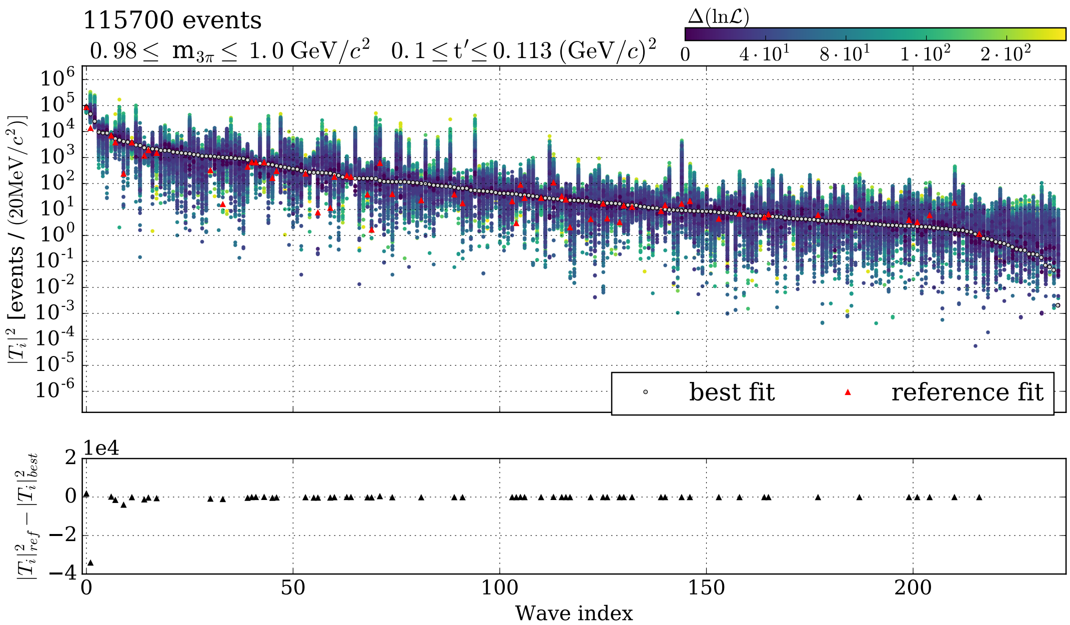


Figure 5.6: Fit of the reduced wave pool of 236 waves to the large sample of simulated data at $0.99\,\mathrm{GeV}/c^2$. This plot is analogue to fig. 5.2 but for the low-mass bin and with the reduced wave pool without negative reflectivity waves. Even the increased sample size does not allow a reliable recovery of the reference fit result.

5.3 Application of the Biggest Conceivable Model Method

In the previous section it has been shown that an unregularized fit with a large wave pool is not always suitable even in the case of simulated data that follows the same model structure. The following section is concerned with the study of the BCM method and its properties on simulated data. These studies should help building an understanding of the validity of this approach. In Ref. [Dro15] problems with the results, in the form of unphysically large intensities, at low masses have been observed. For the unregularized MLE fit the low-mass region is already problematic on simulated data and shows exactly these increases in intensity. This section tries to produce some insight on the effects of the BCM method.

5.3.1 Application of the unaltered BCM method

First the BCM method will be applied in the same manner as it has been applied in the context of the previous analyses. This means that the scale parameter of the penalty is fixed at a value of $\Gamma = 0.5$.

Looking again first at the results of the high-mass bin in fig. 5.7 , it is interesting to see that the fluctuations over the fit attempts have disappeared for most of the large intensity waves and that the difference in intensity has decreased to less than a hundred events.
Two other features are interesting in this plot. First, the two waves with intensities on the order of a few hundred events that show unstable behavior over the fit attempts are of negative reflectivity. Second, about half of the waves exhibit stronger instability. Those are the waves with the smallest intensities on the right half of the plot. Additionally to the continuous spread of the solutions, three discrete levels of intensity are visible for some of them. Most solutions cluster around the best fit. Another set of solutions show extremely small intensities that are smaller than 10^{-12} events. The third type of solutions show intensities of order 10^2 events.
The waves that show these spreads in intensity are later identified to be all negative reflectivity waves.

The most noticeable effect of the method can be seen in the $0.99\,\mathrm{GeV}/c^2$ bin. The simple MLE fit failed to recover the correct intensities of the waves. The BCM method is capable of recovering the intensities of the biggest waves in the best fit. The rest of the solutions show scattering in the intensity. In contrast to the high-mass bin, this is true even for the waves with large intensities of several hundreds of events and above.

In both bins the inference can be improved by the application of the BCM method. The intensity plots are smooth for both bins. Unlike for the application on measured data, no drop in intensity is observed. This is similar to the findings on simulated data for the five-pion final state [Bic16] and makes the selection of a model impossible.

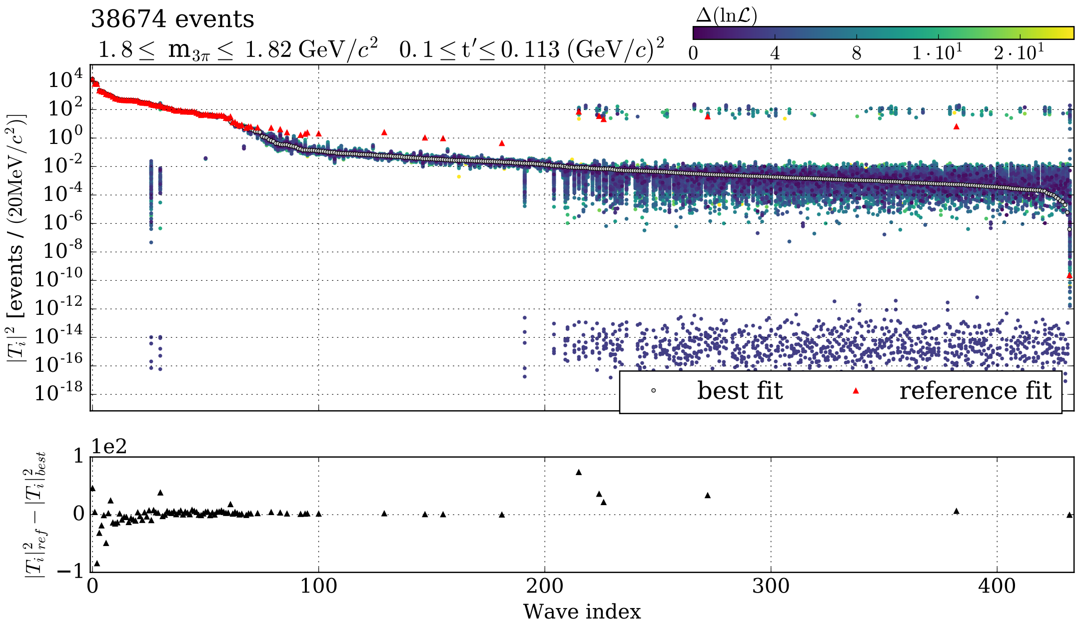


Figure 5.7: Intensities of a fit with the complete wave-pool of 432 waves to simulated data at 1.81 GeV/c^2. The result with the largest likelihood is shown in grey. The intensities have been sorted in descending order. Additional results are ordered according to the best result. The result of the reference fit is superimposed in red. The other attempts are color-coded according to their absolute difference in likelihood with respect to the best result.

5.3.2 Application of the BCM method with decreased scale parameter

Applying the BCM method with the same scale parameter as in the previous works showed no clear intensity drop in the waves. By decreasing the parameter to $\Gamma = 0.2$ the same behavior of the method as on measured data is recovered. An explanation for the appearance of the intensity drop for a different scale parameter will be given in section 6.2.

This result is important as it justifies the interpretation of the intensity drop as the correct criterion to select the relevant waves.

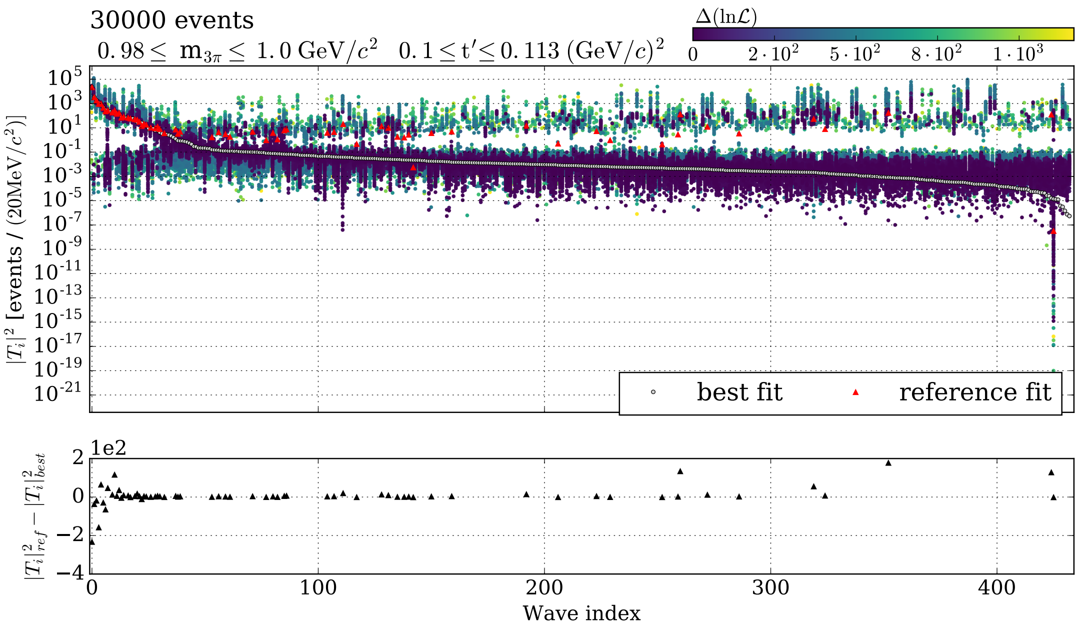


Figure 5.8: Intensities of a fit with the complete wave-pool of 432 waves to simulated data at 0.99 GeV/c^2. The result with the largest likelihood is shown in grey. The intensities have been sorted in descending order. Additional results are ordered according to the best result. The result of the reference fit is superimposed in red. The other attempts are color-coded according to their absolute difference in likelihood with respect to the best result.

The high-mass bin is again discussed first. Several interesting features appear in fig. 5.9. The most important is the drop in intensity the 69th wave. This is exactly the same behavior as observed on measured data for both the three- and the five-pion final state analyses. Most of the waves from the reference fit sit on top of the best fit in this region. Some waves with intensities of order 10 - 100 events are not found. This suggests that the drop intensity can indeed be used to separate the important from the less important waves.

The drop after the 69th wave is not the only one in the plot. Additionally a second intensity drop appears after the 235th wave. This 'double kink'-structure has been observed on measured data as well [Dro15]. The unregularized MLE fit already showed the tendency to suppress waves with negative reflectivity stronger than the one with positive reflectivity. Here the second drop corresponds solely to waves with negative reflectivity. This behavior of the BCM model selection has first been remarked in Ref. [Uhl16]. For most of the following studies a reduced wave pool,

consisting only of waves with positive reflectivity, is used.
Additionally a single wave with large intensity spread over the fit attempts is visible in the selected sector. This wave is the flat wave, which has no been generated. It seems that the BCM model selection prefers to use waves from the dominant positive reflectivity sector and replace the waves with negative reflectivity with the unpenalized flat wave.

For the low mass region shown in Figure 5.10 the same intensity drop is visible. Here not second drop appears in the best fit, but several waves show a similar pattern in the other fit attempts. Also the flat wave obtains intensity in this fit.
All waves show a strong intensity spread over the different fit attempts. For the high-mass bin these spreads were small for the waves with the biggest intensities. For the low-mass bin this is not the case anymore.
From the plots it is not evident which waves are in the selection of the method. In this bin an additional wave is found. The $0^{-+}0^{+}f_0(1500)\pi S$ wave has not been generated in this bin. This wave is especially interest as it is one of the waves know to behave problematically on measured data.

5.3.3 Effect of the Scale Parameter and the Sample Size

The next obvious question is of course the effect a further decrease of the scale parameter has. As already mentioned previously in section 4.2.3.1 the penalty approaches a form that is proportional to the reciprocal intensity of a wave. Decreasing the parameter to $\Gamma = 0.05$ shows similar results for the low- and high-mass region of the fit. The results are shown in fig. 5.13 and fig. 5.14. Most importantly neither the number of selected waves nor the bias is strongly influenced by the scale. The sector of deselected waves get pushed down to smaller intensities while the end of the selected waves remains at an intensity of several events.
Another noticeable effect becomes apparent when looking at the scattering of the different fit solutions. With decreasing scale parameter the spread in the log-likelihood differences increases. The best solutions however remains good especially in the high-mass region.
For the low-mass bin the more additional waves are introduced. This is thought to be connected to the increased instability of the waves over the fit attempts.

The BCM method has also been tried for different event sample sizes, but no dependence on the sample size was found.

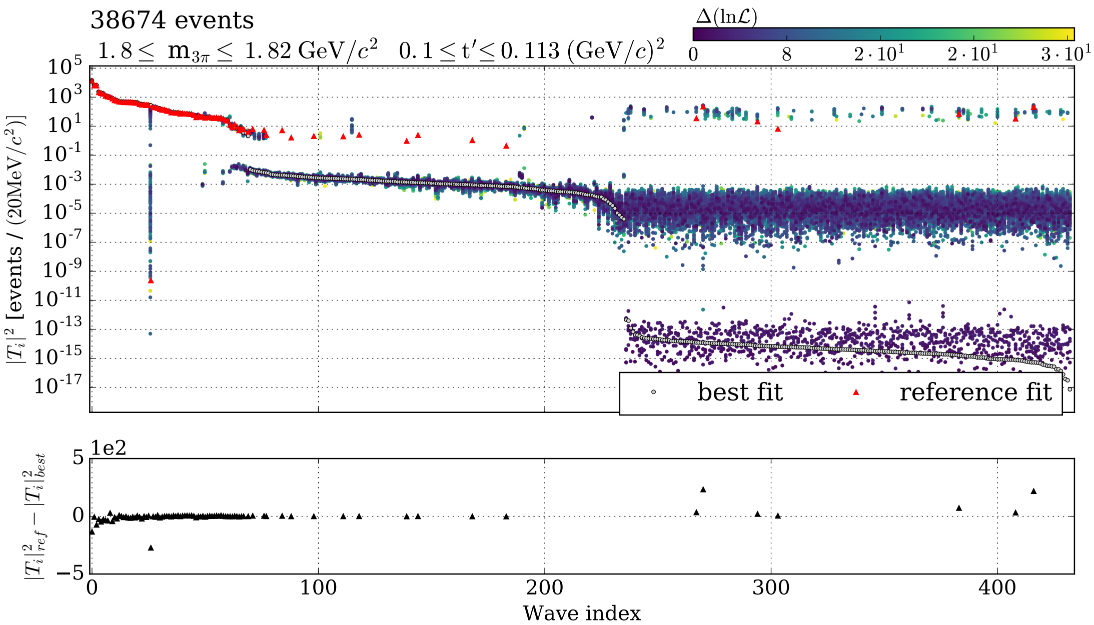


Figure 5.9: Intensities of a fit with the complete wave-pool of 432 waves to simulated data at 1.81 GeV/c^2. The result with the largest likelihood is shown in grey. The intensities have been sorted in descending order. Additional results are ordered according to the best result. The result of the reference fit is superimposed in red. The other attempts are color-coded according to their absolute difference in likelihood with respect to the best result.

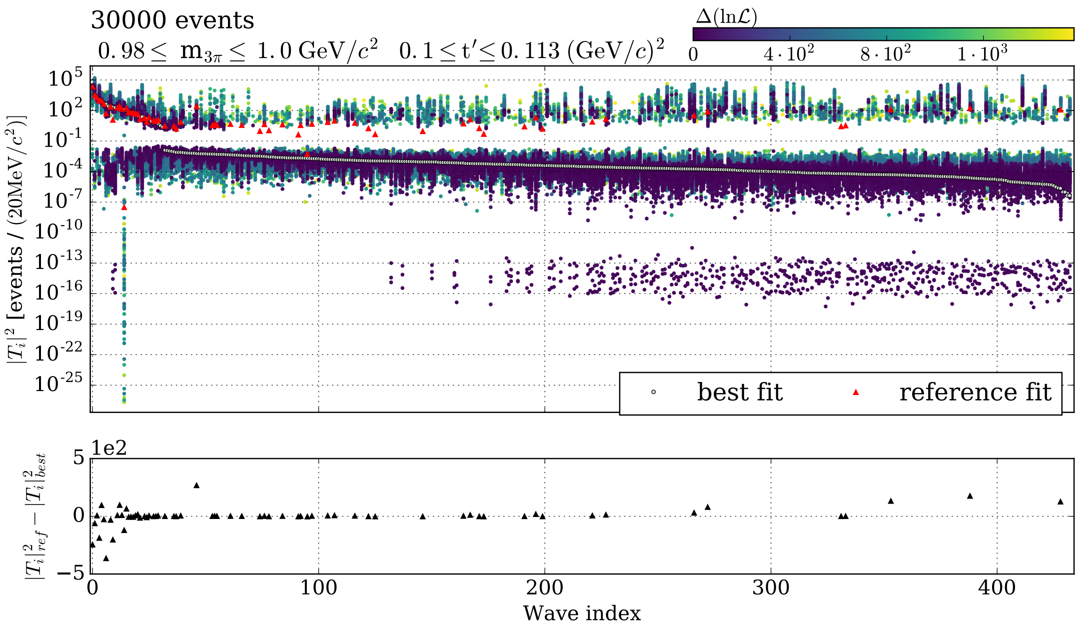


Figure 5.10: Intensities of a fit with the complete wave-pool of 432 waves to simulated data at $0.99\,\mathrm{GeV}/c^2$. The result with the largest likelihood is shown in grey. The intensities have been sorted in descending order. Additional results are ordered according to the best result. The result of the reference fit is superimposed in red. The other attempts are color-coded according to their absolute difference in likelihood with respect to the best result.

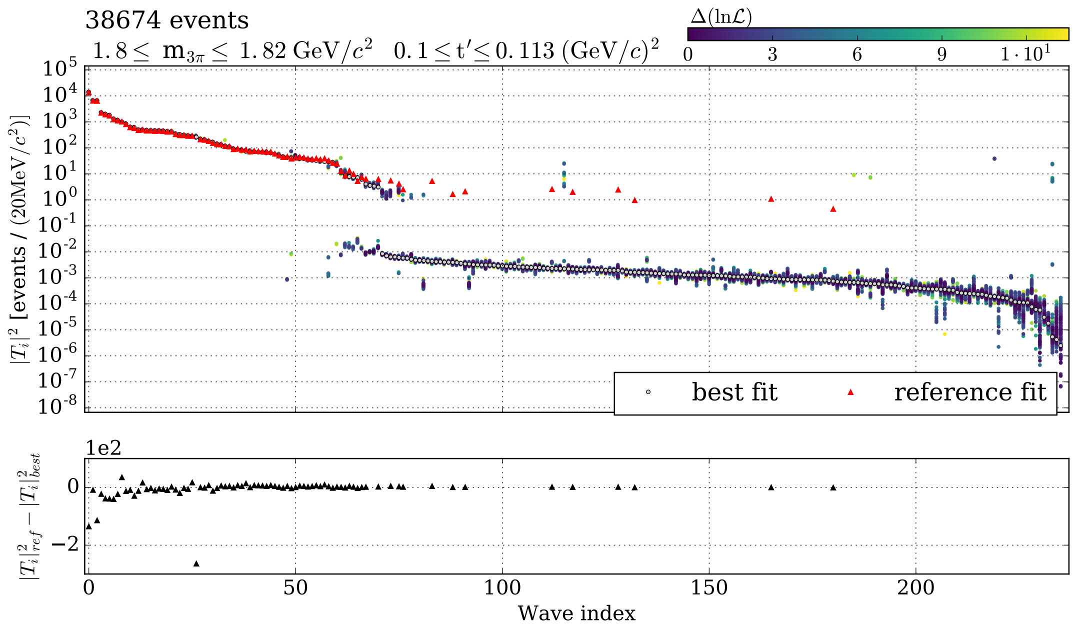


Figure 5.11: Fit of the BCM method with the reduced wave pool to simulated data in the high-mass bin. $\Gamma = 0.2$

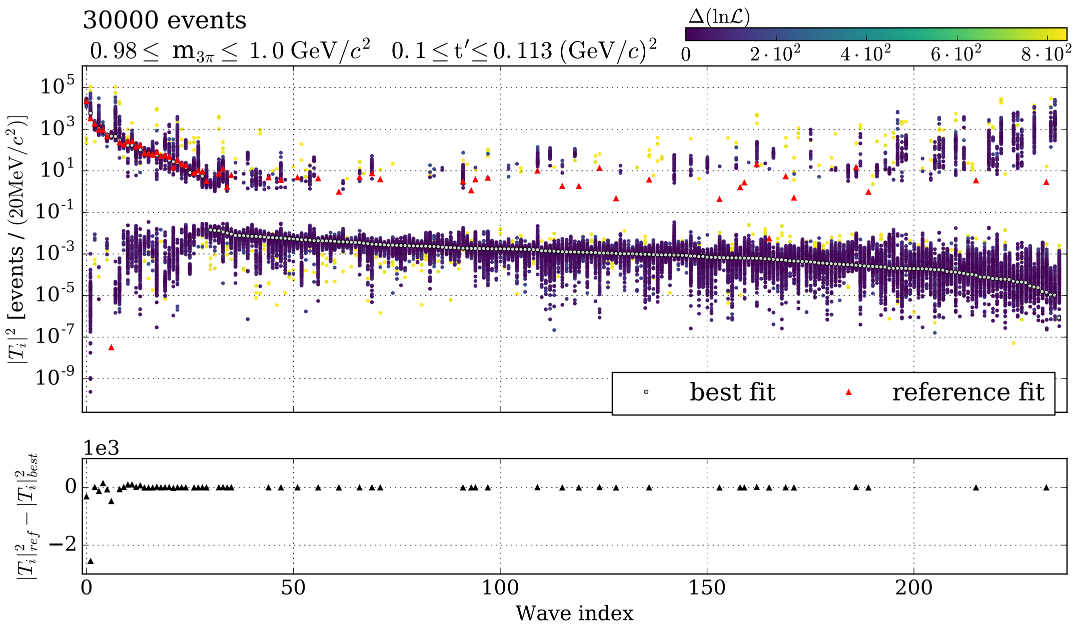


Figure 5.12: Fit of the BCM method with the reduced wave pool to simulated data in the low-mass bin. $\Gamma = 0.2$

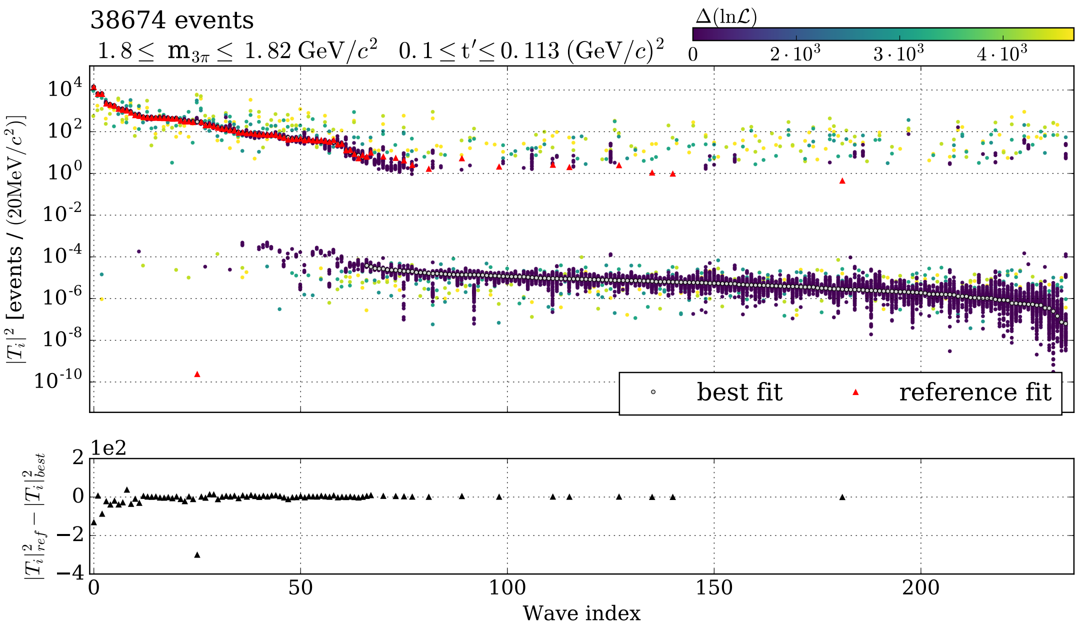


Figure 5.13: Fit of the BCM method with the reduced wave pool to simulated data in the high-mass bin. $\Gamma = 0.05$

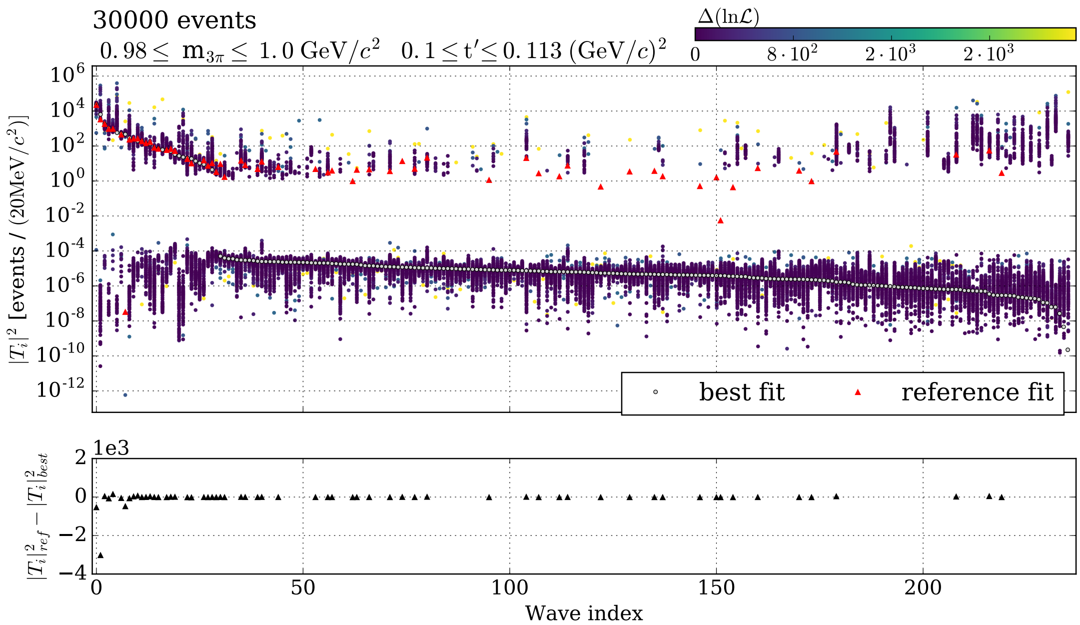


Figure 5.14: Fit of the BCM method with the reduced wave pool to simulated data in the low-mass bin. $\Gamma = 0.05$

5.4 Diagnosis of the Multimodality

In sections 5.2 and 5.3 it has been demonstrated that the fit suffers from many local minima found by the optimization procedure. This problems occurs especially for lower masses. This section provides more insight on the nature of the multimodality and its mass-dependence.
In section 3.1 the decay amplitudes have been introduced. For the same $J^{PC}M^{\epsilon}$, waves of different isobar ξ are only distinguishable by the isobar line-shape Δ_{ξ}. For masses far below the typical threshold of the isobar resonance, these line shapes flatten and become more similar. As an example for this effect, simulated data for single waves has been generated in the high- and the low-mass region. In fig. 5.16 the histograms of the neutral two-pion subsystem are plotted for different waves at different masses. In fig. 5.16a the $0^{-+}0^{+}f_0(1500)\pi S$ wave has been generated at a three-pion mass of $1.81\,\mathrm{GeV}/c^2$. The $f_0(1500)$ isobar is clearly visible in the spectrum of the two-pion system. The same wave has been generated in the low-mass bin of $0.99\,\mathrm{GeV}/c^2$, which is plotted in fig. 5.16b. In this sub-threshold region, only phase space like structure is visible. This can be compared for data of a different wave, namely the $0^{-+}0^{+}f_0(980)\pi S$ wave, in the same mass region. Even though a different isobar is used, the two-pion mass spectrum looks similar. It is clear that these waves are expected to suffer from strong cross-talk.

A hint for the cross-talk between two waves can also be found by looking at the overlap integral of the waves:

$$\mathcal{P}_{ij} = \int \mathrm{d}\varphi_3(\tau')\Psi_i(\tau')\Psi_j(\tau')^* \tag{5.1}$$

The absolute value of the normalized overlap integral $\left|\frac{\mathcal{P}_{ij}}{\sqrt{\mathcal{P}_{ii}}\sqrt{\mathcal{P}_{jj}}}\right|$ is a measure between 0 and 1. Where a value of 1 hints towards strong correlation of the waves. In fig. 5.15 the maximum absolute value of the overlap integral of any wave is plotted for a specific wave. The dashed line corresponds to the $0^{-+}0^{+}[\pi\pi]_S\pi S$ wave and the solid line to the $0^{-+}0^{+}f_0(1500)\pi S$ wave. For low masses both waves show a maximum overlap near 1 with some other wave. For higher masses the value decreases, but remains finite. This is consistent with the problematic cross-talk of these two waves.

The consequence of the strong overlap between waves is the possibility for them to get large intensities that interfere destructively, such that the expected number of events remains about the same.

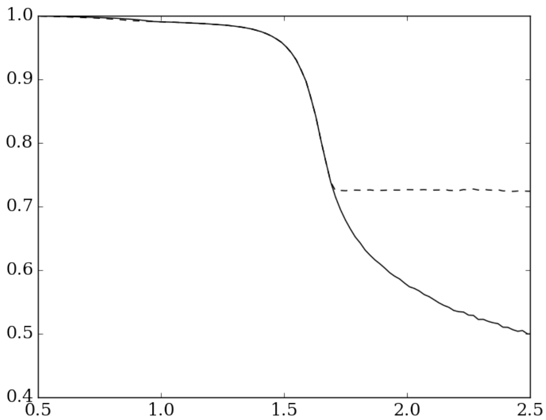


Figure 5.15: Maximum absolute value of the overlap integral with any other wave. Dashed line: $0^{-+}0^{+}[\pi\pi]_S\pi S$ wave; Solid line: $0^{-+}0^{+}f_0(1500)\pi S$ wave

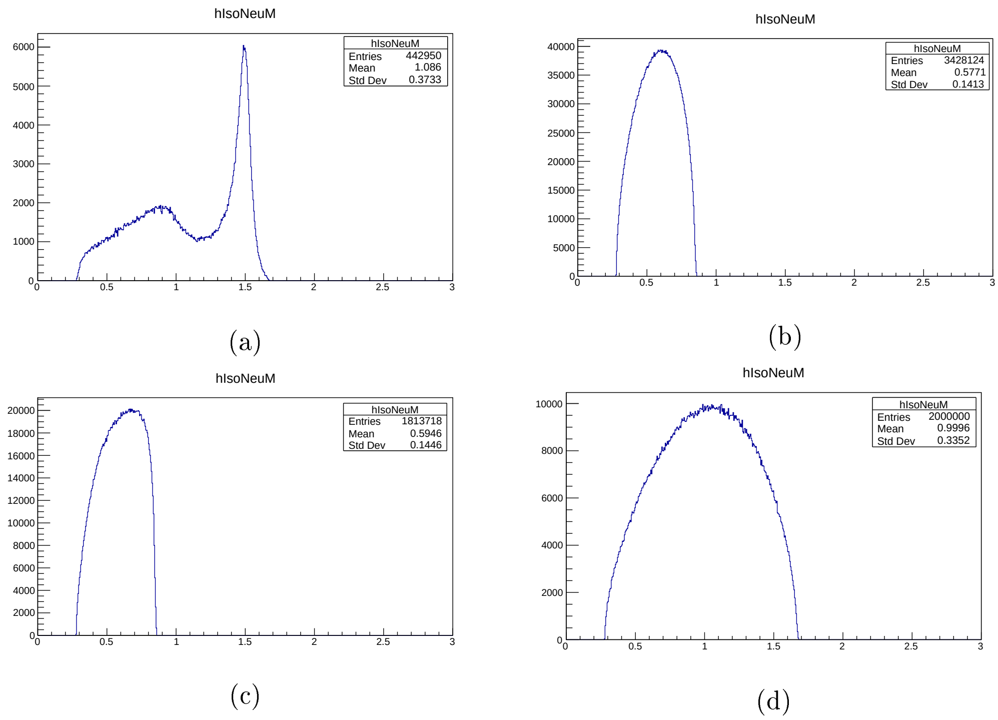


Figure 5.16: Histograms of the neutral two-pion subsystem.

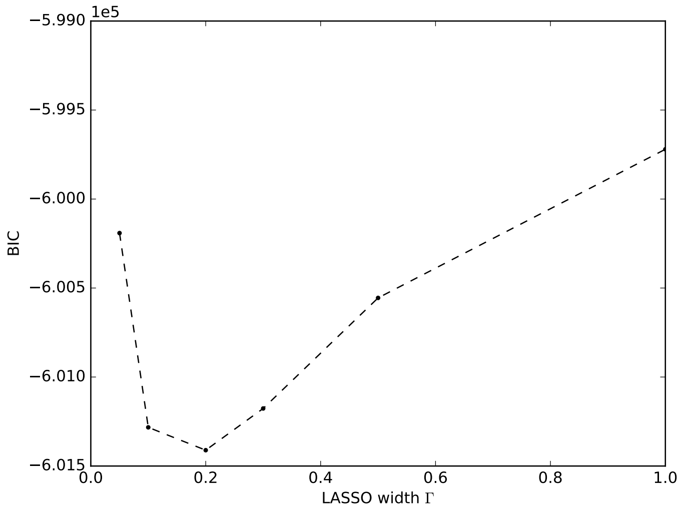
5.5 Application of the LASSO method

In this section the LASSO regularization is applied to simulated data. Here only the reduced wave pool has been used. In fig. 5.18 the effects of the LASSO penalty can be seen. For this fit no smoothing has been applied to the absolute value and the width parameter was set to $\Gamma = 0.2$. For this choice of parameter the wave set seems to be recovered quite well.
The intensity drop is bigger than for the BCM method, but at the same time the separation of the large and small intensity sectors is not so well defined. The drop is happening in a smooth fashion, making the exact selection of waves harder. It can also be seen that the non-differentiability at 0 is a problem for the minimizer, as the solutions show a wide spread in intensity for the deselected waves. The waves that are actually in the reference model are found without these fluctuations.
Looking at the intensity bias shown in the lower subplot it is obvious that the fit with the LASSO introduces some bias towards lower intensities.

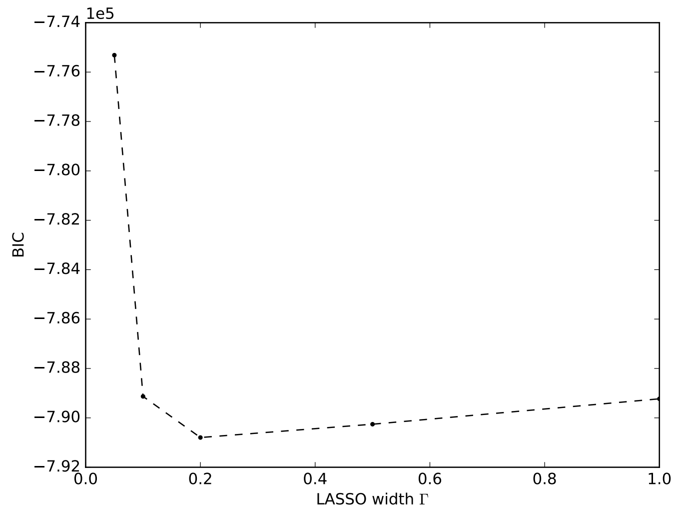
The tuning of Γ has been performed by performing a BIC scan, which is shown in fig. 5.17. The degrees of freedom r have been estimated by taking into account waves with an intensity over 10^{-3} events. For both bins a value of $\Gamma = 0.2$ has been chosen.

The LASSO fit for $\Gamma = 0.2$ in the low-mass bin is shown in fig. 5.19. Similarly to the high-mass bin, a drop in the intensity is visible, which is not as clearly separated as the BCM drop. The interesting feature of this plot is visible for the waves with the highest intensities. For the BCM fits, large spreads over the different fit attempts appeared.

The smoothing of the absolute value has been introduced to have a better behaved function fed into the minimizer. The effect of the smoothing can be seen in fig. 5.20. The strong numerical instabilities can be resolved and the scattering over the fit attempts decreases. Of course the results should be consistent as long as the smoothing is small. This has been checked in fig. 5.21 where instead of the reference fit the best fit without smoothing is overlaid in red. It is clear that the smoothing only affects the lowest intensities, while the large intensities coincide. It is concluded that the smoothing is an appropriate way to deal with the non-differentiability at zero.



(a)



(b)

Figure 5.17: Scan of the BIC criterion for simulated data without smoothing of the penalty. (a) scan for the low-mass bin (b) scan for the high-mass bin. In both cases a value of $\Gamma = 0.2$ is chosen.

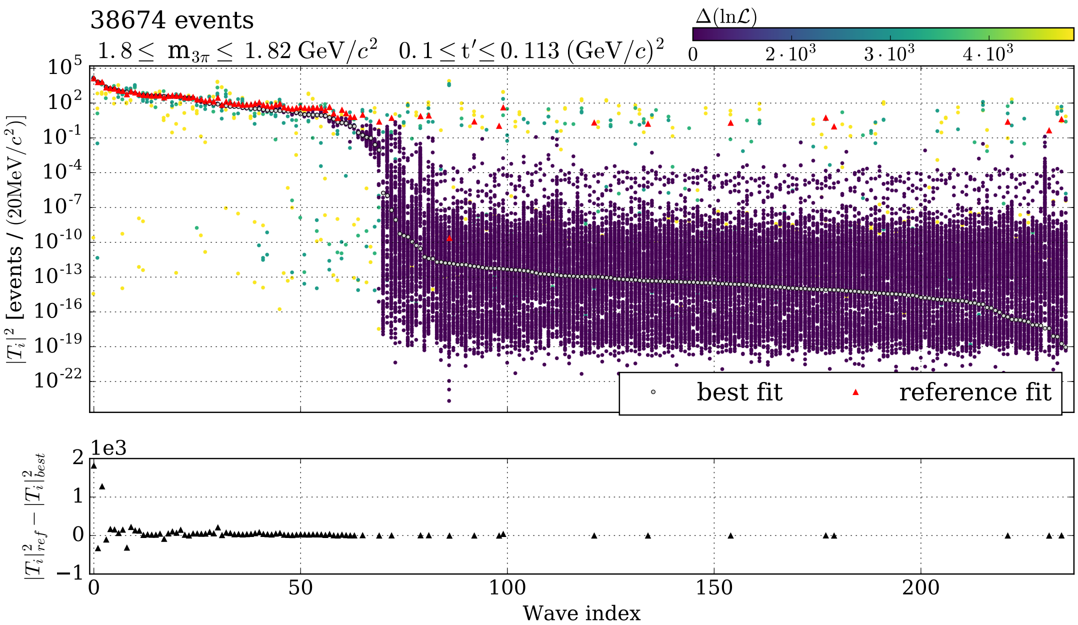


Figure 5.18: Fit result for the LASSO penalty without smoothing on simulated data in the high-mass bin. $\Gamma = 0.2$

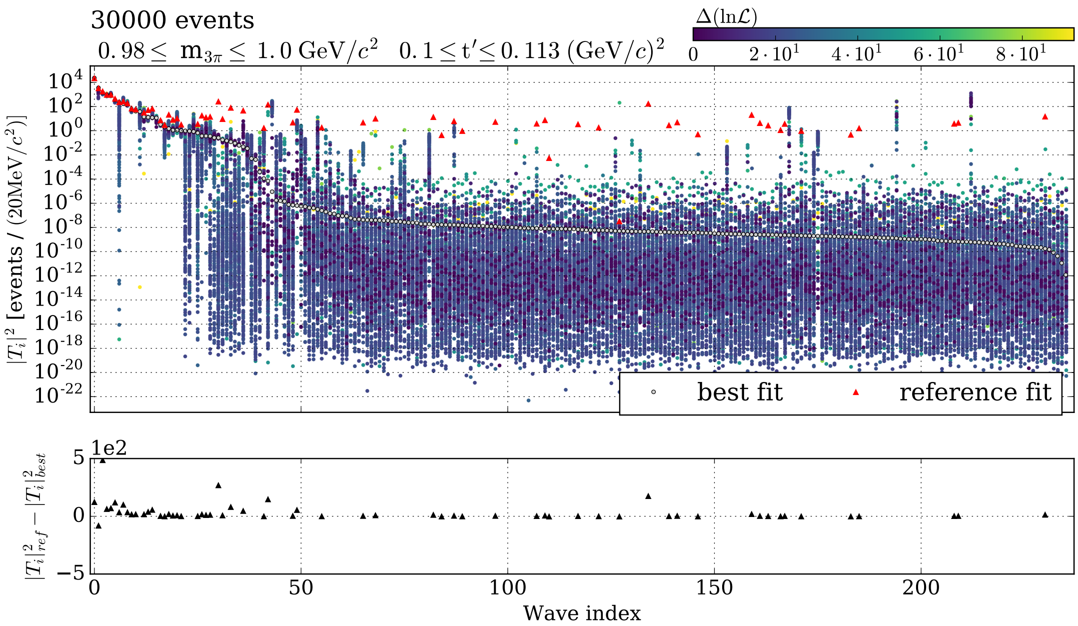


Figure 5.19: Fit result for the LASSO penalty without smoothing on simulated data in the low-mass bin. $\Gamma = 0.2$

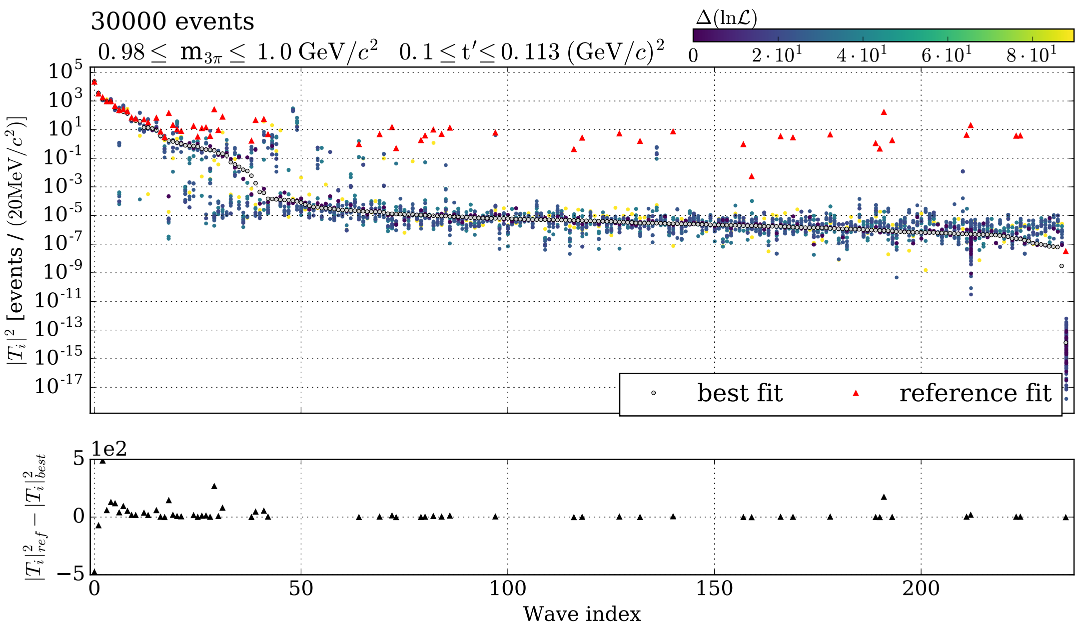


Figure 5.20: Example for the smoothed version of the LASSO penalty.

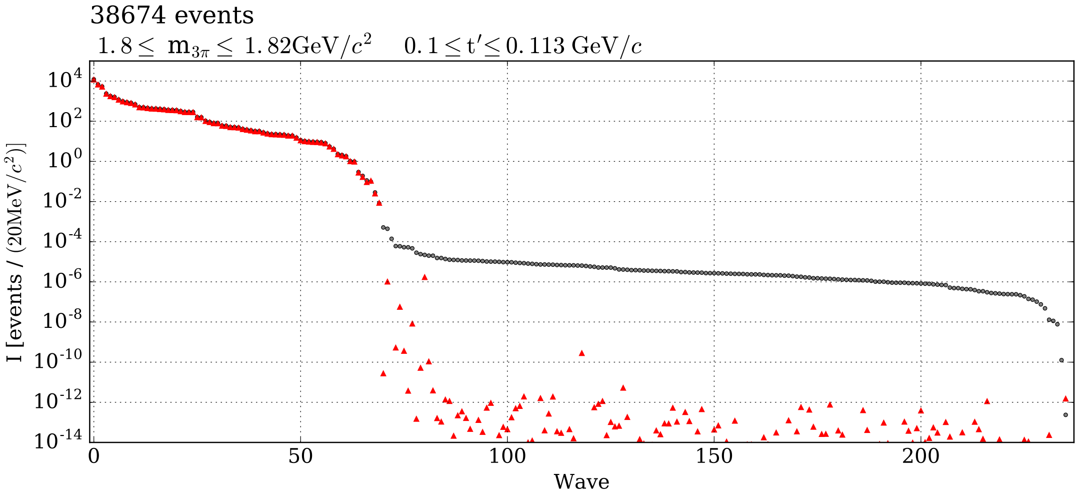


Figure 5.21: Comparison of smoothed LASSO with non-smoothed LASSO.

5.6 Application of the MBCM Regularization

The application of the LASSO method revealed that the bias on large intensities, introduced by the light tails of the penalty term, have the property of prohibiting the large destructive interferences that are visible in the different fit attempts of the fit without regularization or with the BCM method. Nevertheless the BCM method still performs reasonable well in recovering the true model for the best fit result. Because the waves selected by the BCM method are not very dependent on the free scale parameter of the method and in general the bias on the large waves is smaller, it would be desirable to improve this method in such a way that it keeps these properties while being able to resolve the ambiguities of the large destructive interferences.

The MBCM method, introduced in section 4.2.3.3, intends to combine the advantages of both BCM and LASSO methods. It has two free parameters. The scale parameter Γ directly corresponds to the scale parameter of the BCM method. The decay parameter λ controls the tails of the prior or in other words the bias on large intensities. For $\lambda = \infty$ the MBCM is equal to the BCM method, for smaller values the BCM tail-behavior gets modified by an exponential decay, the tail are less heavy and the bias on large intensities increases. The intention is effectively to impose a soft limit on large intensities and therefore requiring more support from the data to overcome the prior and make a wave large.
λ is chosen such that it is large enough to ensure the same selection behavior as the BCM method, while resolving ambiguities of large waves similar to the LASSO.
To be comparable to the BCM method, $\Gamma = 0.2$ was chosen for the scale parameter. For λ values of $\mathcal{O}(1)$ were found to be adequate and $\lambda = 5$ is chosen for the studies presented in this section.

In fig. 5.22 the results for the low-mass bin are shown. The most obvious change as compared to the BCM method, shown in fig. 5.12, is the vanishing of solutions that show a large spread in intensity over the different solutions.
The $0^{-+}0^{+}f_0(1500)\pi S$ wave, that was found with an intensity of about 453 events by the BCM method, is successfully removed from the selected sector. With it, the increase of the intensity of the $0^{-+}0^{+}[\pi\pi]_S\pi S$ wave also vanishes. This shows how the additional bias makes it possible to resolve ambiguities of similar waves at low masses.
The $1^{++}1^{+}\rho\pi D$ wave and the $1^{++}1^{+}[\pi\pi]_S\pi P$ wave are swapped, this has also been observed for the LASSO fits. This is most likely an artifact due to the fit result used to generate the data.
All other waves that could not be recovered correctly are of the order $\mathcal{O}(10)$ events or below.

To illustrate that a value of $\lambda = 5$ is introducing only a weak bias, the MBCM fits are plotted with the results of the BCM method superimposed in red as the reference model. For both the high-mass, shown in fig. 5.23, and the low-mass bin, shown in fig. 5.24, the results mostly coincide. For the high-mass bin the only difference can be seen for two small waves in the region around the intensity drop. For the low-mass the effects are more insightful. The waves discussed in the previous paragraph differ from the BCM result, while the rest of the found waves are on top of the BCM results.

It is evident that the MBCM method is capable of extending the BCM method such that regularization properties similar to the LASSO are introduced. The selection works like for the BCM method, a drop in the intensity spectrum is induced. Because the scale parameter is related the one of the BCM method, the position of the drop is only weakly dependent on the exact parameter value. The decay-parameter is only used as an upper limit for the waves. This should make the method less dependent on expensive parameter tuning like it is required for the LASSO.
The promising effects seen on simulated data justify the application to bins of measured data in section 6.3.

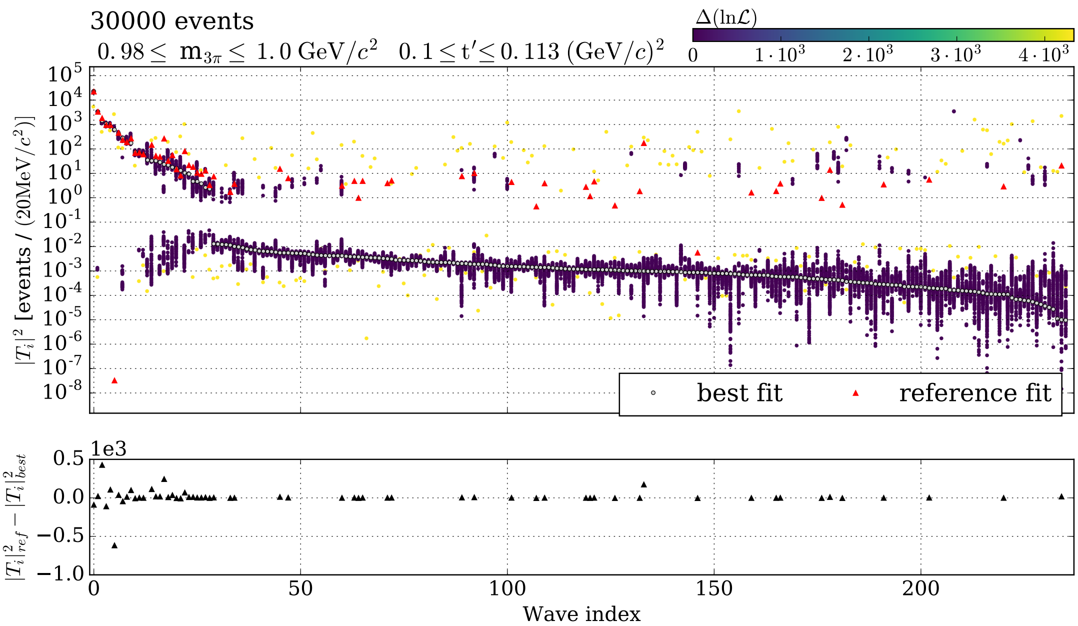


Figure 5.22: Intensity plot for the application of the MBCM method on the low-mass bin for simulated data. The larger bias as compared to the BCM method is able to resolve issues of destructive interference.

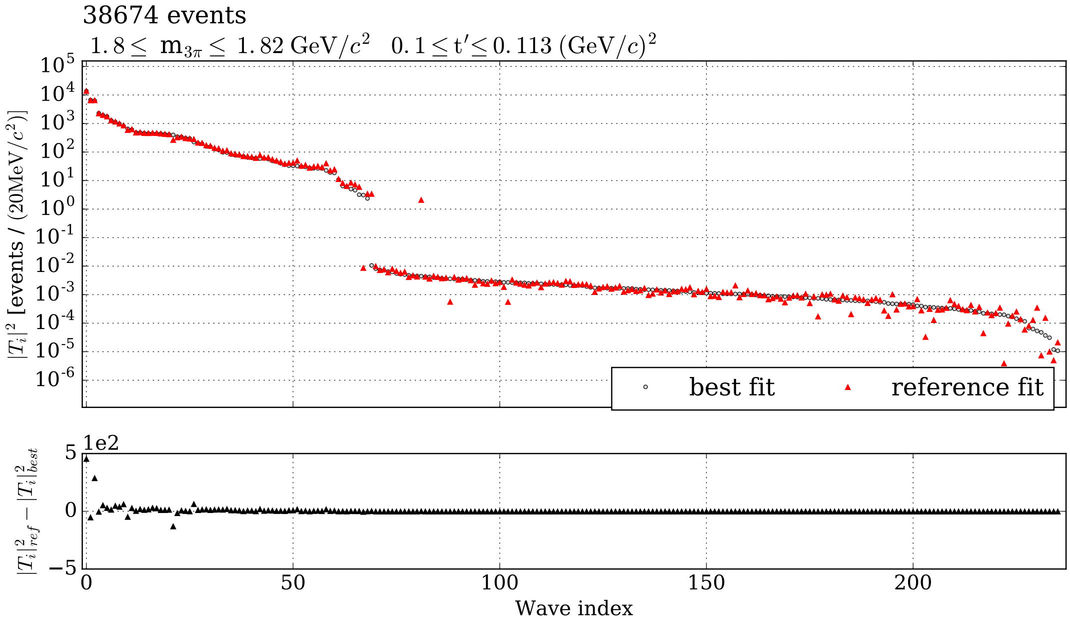


Figure 5.23: Comparison of the MBCM best fit in the high-mass bin with the BCM fit results superimposed as the reference fit in red.

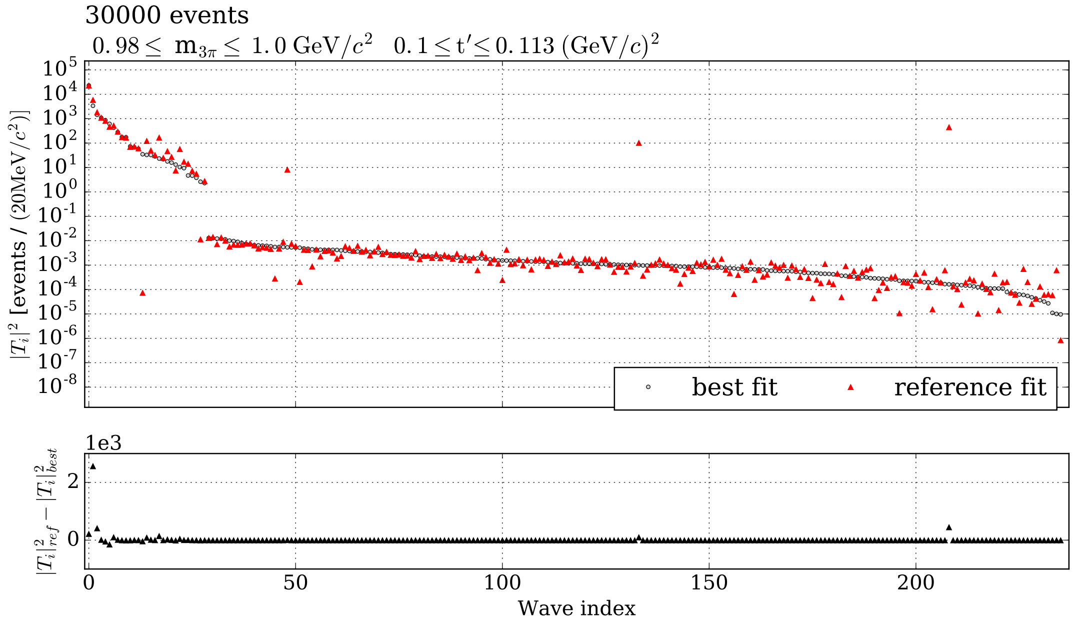


Figure 5.24: Comparison of the MBCM best fit in the low-mass bin with the BCM fit results superimposed as the reference fit in red.

Chapter 6

Model Component Selection on Measured Data

In this chapter studies, analogue to the ones conducted on simulated data, will be repeated on measured data in the same two mass-bins. Only the most important results will be shown here as more insight can be gained by extending these fits to the whole mass-range, as done in chapter 7. Nevertheless the study on the two single bins serves as a justification for the application on the rest of the data.

Again the fit with the hand-selected 88-waves model is used for comparison of the fit results. It is important to stress that this 'reference' model is of course not necessarily the correct one anymore. Still, the biggest waves should be reasonable well described by it and differences are expected for lower intensity waves.
The focus will lie mostly on the three $0^{-+}0^{+}$ waves with isobars $f_0(1500)$, $f_0(980)$ and $\pi\pi_S$ as these waves showed unusually large intensities for lower masses.
Of these waves, only the one with the $\pi\pi_S$ isobar is considered reasonable in the low-mass bin. In the hand-selected model the other two were thresholded below $1.7\,\mathrm{GeV}/c^2$, for the $f_0(1500)$ wave, and below $1.2\,\mathrm{GeV}/c^2$, for the $f_0(980)$ wave.

The study of Ref. [Dro15] showed that the BCM method was not able to exclude these waves in the low-mass region. Instead the intensities of these waves were large and the thresholds, used in the hand-selected model, had to be reimplemented.

6.1 Fit without regularization

It has already been seen that the unregularized fit was not able to correctly infer the waves in the low-mass bin. It is no surprise that the situation is even worse on measured data. As an example the fit with the reduced wave pool at a mass of $0.99\,\mathrm{GeV}/c^2$ is shown in fig. 6.1. The result looks similar to the simulated data in this region. All waves pick up large intensities.
The two largest waves are the $0^{-+}0^{+}f_0(1500)$ wave with an intensity of 408735 events and the $0^{-+}0^{+}f_0(980)$ wave with an intensity of 258082 events. Both of these

waves have been excluded in the hand-selected model because their appearance is not considered reasonable at 0.99 GeV/c^2 final state mass. The effect of cross-talk within this sector is expected to be large due to destructive interference of the similar waves. One can conclude that the naive MLE fit of the wave pool does not lead to useful results just as expected from the observations on simulated data.

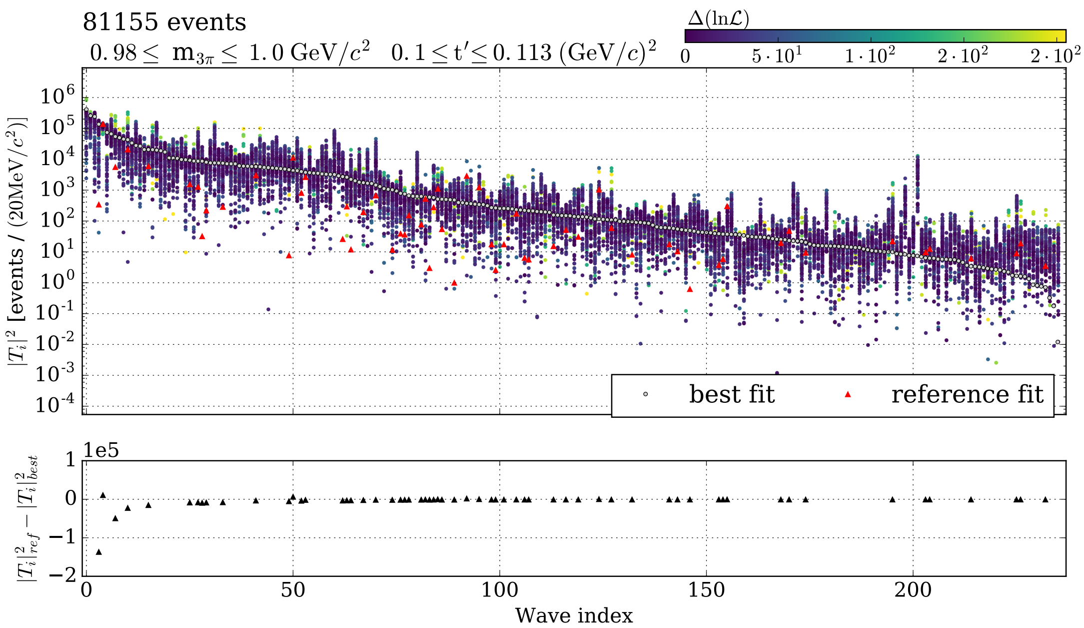


Figure 6.1: MLE fit of the reduced wave pool to measured data in the 0.99 GeV/c^2 bin.

6.2 BCM regularization

In this section the issues with the BCM method will be presented. The focus lies on explaining the problems discovered in Ref. [Dro15].

In fig. 6.2 the BCM result for the high-mass bin and a width-parameter of $\Gamma = 0.5$ is shown. For this choice of parameter value the intensity drop does not appear on simulated data [1]. For the measured data however the intensity drop is visible. The reasons for the appearance of the drop is considered to be related to the acceptance correction that is applied for measured data. This can be understood in the following way. The acceptance correction, applied to fits of measured data,

[1] compare to section 5.3.1

effectively scales up the intensity of the waves. For this finale state the acceptance is roughly $\eta \approx 0.47$. The scaling of the intensity can also be interpreted as a scaling of the width-parameter by $\Gamma_{\text{acceptance}} \approx \frac{\Gamma_{\text{ideal}}}{\sqrt{0.47}}$. It is therefore possible to calculate a width-parameter that does not lead to an intensity drop on measured data from the value used for simulated data. Indeed for a value of $\Gamma_{\text{acceptance}} \approx \frac{0.5}{\sqrt{0.47}} \approx 0.73$ no such drop occurs, as can be seen in fig. 6.3.

The results of the simulated data show that in general smaller values of the parameter do not have a large effect on the intensities and the selected wave set. For measured data $\Gamma = 0.5$ is close to the value at which the intensity drop appears. It has been found reasonable to chose $\Gamma = 0.2$, which removes more of the smallest waves, but keeps the multimodality at an acceptable level.

In the high-mass bin, shown in fig. 6.2, the large-intensity waves coincide mostly with the waves of the 88-wave set. The size of the selected wave set would be larger, but no arguments can be made about whether the additional small waves are physical. For the $0.99\,\text{GeV}/c^2$ bin the results are shown in fig. 6.4. The intensities of the 0^{-+} waves are again large. This is not surprising, considering that first signs of such unphysical behavior are visible even on simulated data.

Different values of the scale parameter Γ have been tried, but none of them was able to resolve the issues with unphysical high-intensity waves at low masses. One can conclude that the BCM method is not able to resolve the issues, discovered in Ref. [Dro15], by decreasing the scale-parameter or using a reduced wave pool of waves with only positive reflectivity.

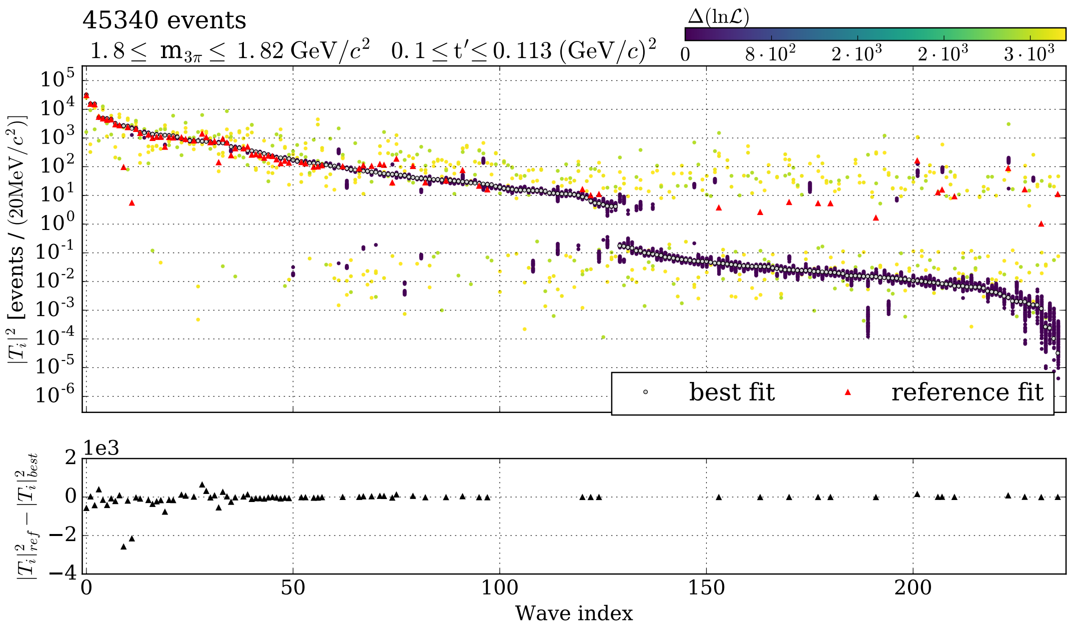


Figure 6.2: Fit of the reduced wave pool with the BCM method and $\Gamma = 0.5$ at $1.81\,\mathrm{GeV}/c^2$.

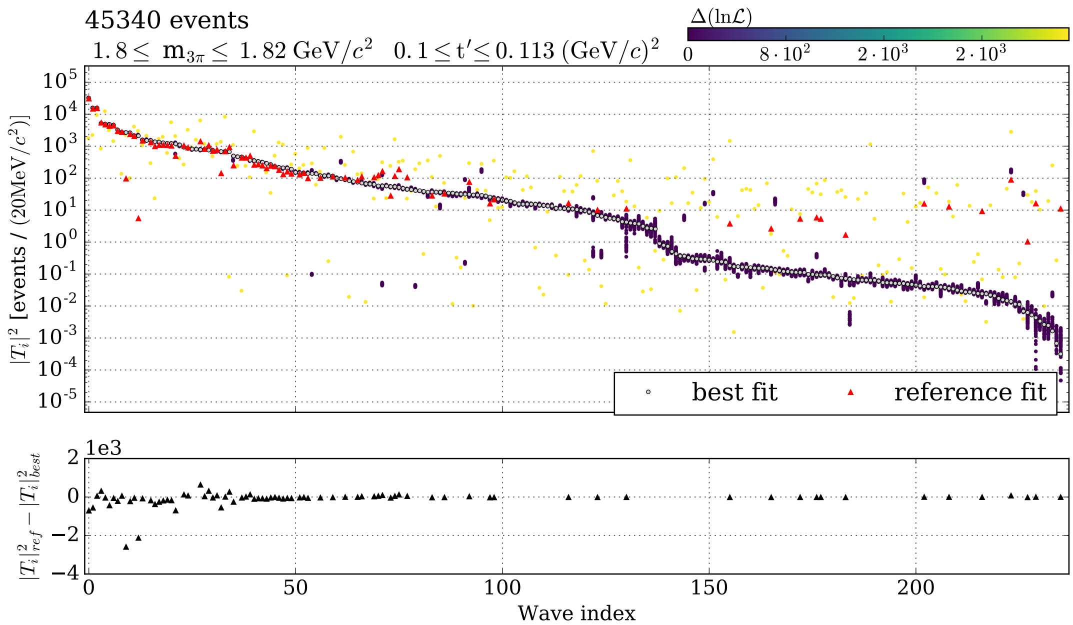


Figure 6.3: Fit of the reduced wave pool with the BCM method and $\Gamma = 0.73$ at $1.81\,\mathrm{GeV}/c^2$.

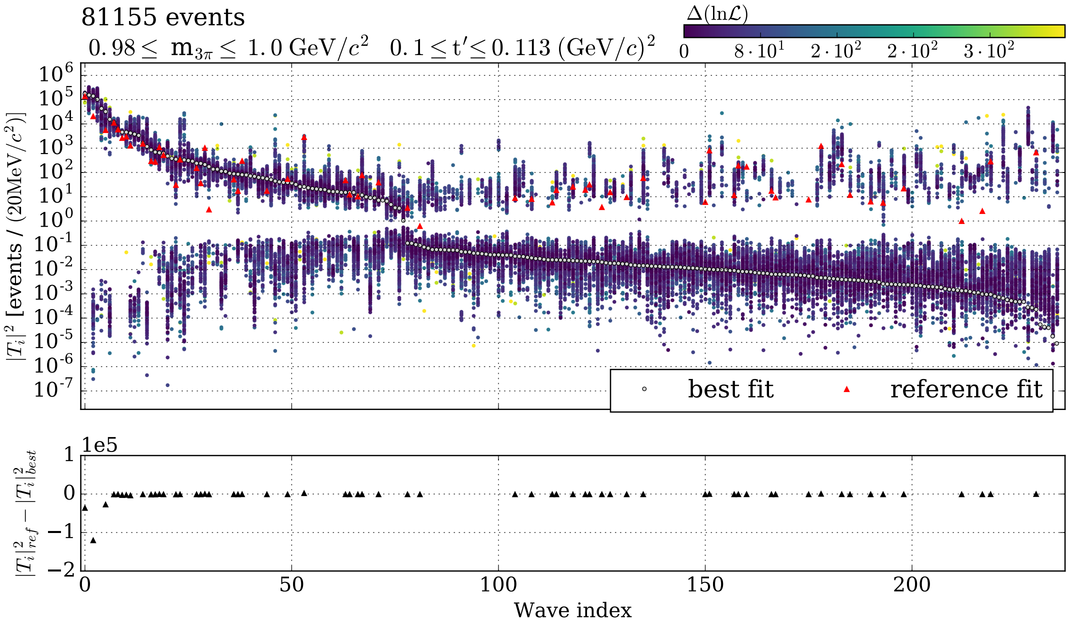


Figure 6.4: Fit of the reduced wave pool with the BCM method with $\Gamma = 0.5$ at $0.99\,\mathrm{GeV}/c^2$.

6.3 MBCM regularization

For simulated data the MBCM method has been introduced in order to improve the results obtained by the BCM method. A larger bias on the intensity proved to be advantageous in suppressing unwanted solutions with large destructive interference. The effect on the best fit was not very strong on simulated data.
The width-parameters of BCM and MBCM correspond to each other and $\Gamma = 0.2$ has been chosen for the MBCM method as well. To control the strength of the bias $\lambda = 5.1$ [2] has been chosen. This value is considered large enough to not introduce a strong bias on waves that are considered well established, but strong enough such that waves that do not show a strong support by the data are suppressed.

Again inference of the intensities is not problematic for the high-mass bin and leads to similar results as the BCM method. The focus lies on the low-mass bin to explain the properties of the method.

For the low-mass bin the effects are more prominent. Looking at the 0^{-+} sector it is evident that the issues can be resolved by introducing a more demanding prior.
Both the $0^{-+}0^{+}f_0(980)\pi S$ and $0^{-+}0^{+}f_0(1500)\pi S$ waves are removed from the set of selected waves in this case. The intensities of the largest waves resemble the hand-selected wave set much more. Comparing the result of the MBCM method in fig. 6.5 with the result of the BCM method in fig. 6.4, one immediately notices that for the MBCM method the largest waves coincide with the waves of the hand-selected 88-wave model. This proves the advantage of additional bias on large intensity. Waves with large intensities that are not strongly supported by the data can easily be removed, while the effect on the other waves is small.
Furthermore, the BCM method showed strong scattering over the fit attempts, even for the largest waves. The MBCM fit suppresses such behavior almost completely for large-intensity waves. Uniquely found waves are therefore easier to interpret.

[2] $\lambda = 5.1$ has been chosen for a study conduced in section 7.5. This does not make a significant difference to $\lambda = 5$, which has been chosen for the studies on simulated data.

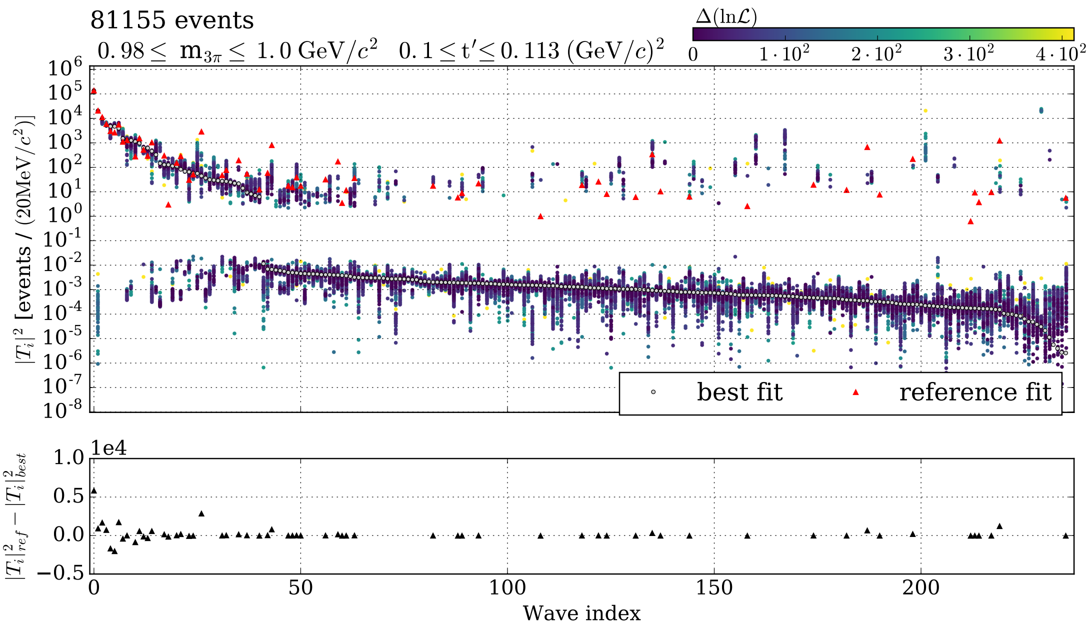


Figure 6.5: Fit of the reduced wave pool with the MBCM method with $\Gamma = 0.2$ and $\lambda = 5.1$ at $0.99\,\mathrm{GeV}/c^2$.

6.4 LASSO regularization

The MBCM method has been motivated by a minimal modification of the BCM method in order to improve it such that solutions with large intensities are suppressed on simulated data. The effect on measured data is the removal of the largest waves that were considered unphysical. The motivation of introducing more bias is based on the success of the LASSO method in removing large spreads in intensity for the low-mass bin. It is thus reasonable to also apply the LASSO method and look at the results.

The results shown here are for different values of the tuning parameter of the LASSO on the reduced wave set. Both, an application with penalized and without penalized flat wave have been tried. Because on measured data contributions from background are expected, especially in the case when waves with negative reflectivity are removed from the wave pool, here the results without penalization of the flat wave are shown. This method is then also more comparable with the BCM and MBCM methods.

To select an appropriate width for the LASSO the information criteria, presen-

ted in section 4.2.3.2, can be invoked. It has already been discussed that the different criteria also select different wave sets. It is not clear whether the application of these criteria can be well justified for this analysis. Also, one has to decide on one of those criteria. Because of this, different values have to be considered for systematic studies in the end. On simulated data the parameter choice of the BIC seemed to be able to best recover the input model. On measured data this does not have to be the case. Nevertheless the sparsest result is obtained by the parameter choice of the BIC. The BIC-scan of the parameters are shown in fig. 6.8 for the low- and high-mass bins.
For the low-mass bin a value of $\Gamma = 0.3$ is found. For the high-mass bin the value was $\Gamma = 0.5$. Here the regularization is weaker for the high-mass bin than for the low-mass bin, which makes sense given the experience that inference seems to work reasonable well in this regime. This can be examined in more detail by extension of the method to the full mass range in section 7.3.

The intensity plot for the fit results of the $1.81\,\mathrm{GeV}/c^2$ bin is shown in fig. 6.6. A drop in intensity can be observed at about the 130th wave. Similar to the findings on simulated data, this drop is not as pronounced as for the BCM or MBCM methods. It is important to point out that the fit attempts show less scattering of the solutions.

In the low-mass region the BIC criterion choses a LASSO parameter of $\Gamma = 0.3$. The result is plotted in fig. 6.7. Similar to the MBCM method the large-intensity waves are removed. It is remarkable that the scattering of the fit attempts is minimal also for the low-mass bin.

For the LASSO method one can conclude that it is capable of solving issues with huge destructive cancellations of waves in the low mass region. The method requires computationally expensive tuning and the selected wave set is strongly dependent on the choice of the tuning parameter. The fitting procedure shows only little scattering of the found solutions.
Both the MBCM method and the LASSO method are superior to the BCM method in removing unphysical waves at low masses. While the LASSO method also provides a more stable fit, the tuning of the parameter and the selection of the wave set are not as straight forward as for the MBCM method. Additionally, the choice of $\lambda = 5.1$ for the MBCM prior introduces less bias on large intensities than the choices of Γ_{LASSO} by the BIC procedure for the LASSO method. The similarity of the MBCM method to the BCM method makes it possible to select a wave set even for less bias than the LASSO.

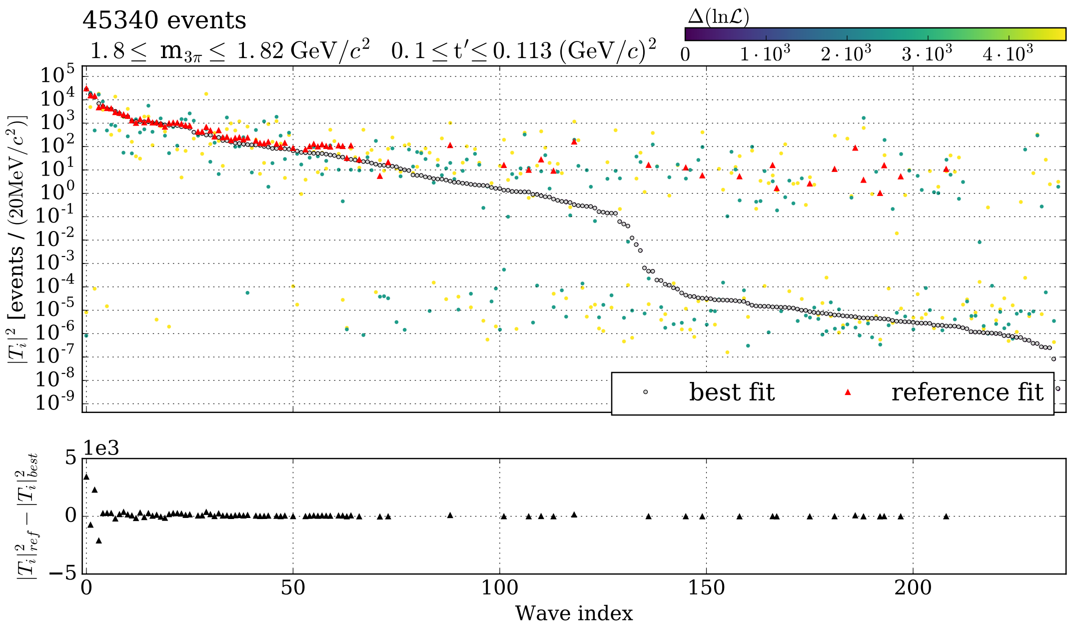


Figure 6.6: Fit of the reduced wave pool with the LASSO method with $\Gamma = 0.5$ at $1.81\,\mathrm{GeV}/c^2$.

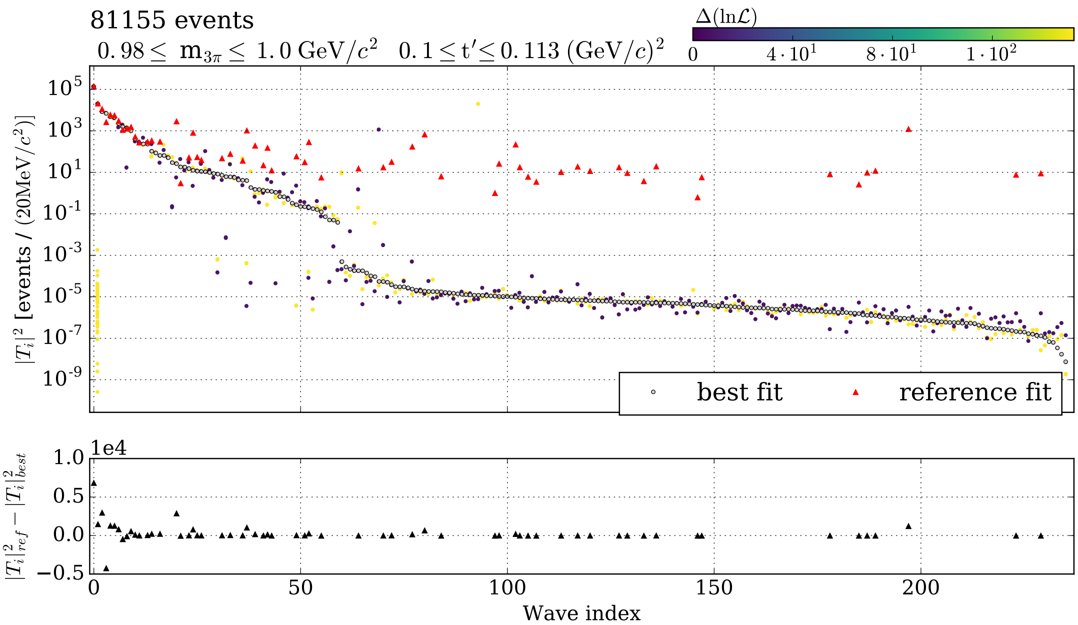


Figure 6.7: Fit of the reduced wave pool with the LASSO method with $\Gamma = 0.3$ at $0.99\,\mathrm{GeV}/c^2$.

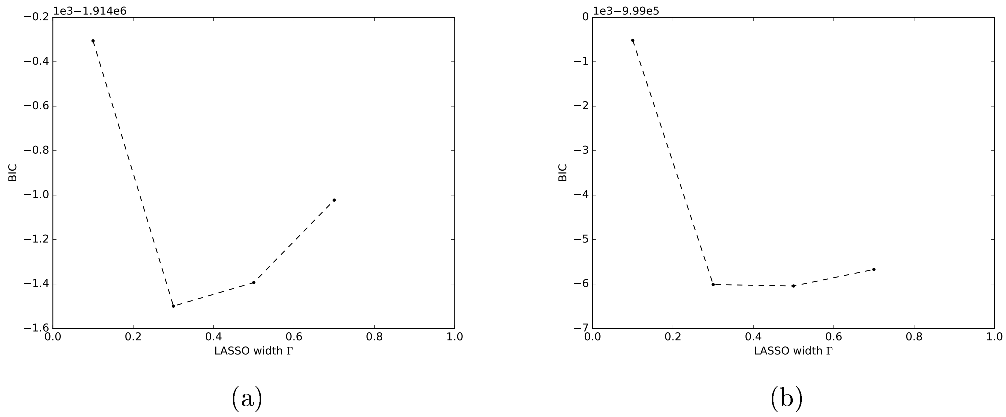


Figure 6.8: Scan of the BIC criterion for the LASSO parameters $\Gamma_{\mathrm{LASSO}} = 0.1, 0.3, 0.5, 0.7$. (a) Scan for the $0.99\,\mathrm{GeV}/c^2$. A value of $\Gamma = 0.3$ is chosen. (b) Scan for the $1.81\,\mathrm{GeV}/c^2$. A value of $\Gamma = 0.5$ is chosen.

Chapter 7

Model Component Selection over complete Mass Range

In chapters 5 and 6 the different model selection methods have been applied to single bins in mass. The effects of the methods in the higher and lower mass region have been studied in these bins. This chapter extends the methods to the full mass range between $0.5\,\mathrm{GeV}/c^2$ and $2.5\,\mathrm{GeV}/c^2$ while still being restricted to the lowest t'-bin. This allows a direct comparison to the case-study of the BCM method performed in [Dro15].

With the extension to the full mass range, much more information is available and the results will be more insightful than for single bins only, because resonance structures are recovered in the mass-spectrum and can be identified. An example for this has been shown in section 3.3. The intensity of a wave should show a peak like structure and a simultaneous phase motion should be visible. The phase motion however can only be measured relative to a reference wave, which itself may show resonant behavior. The relative phase motion can therefore serve as an additional hint for resonant behavior in a wave. A full resonance fit is required to verify this. Like in the previous analyses [Dro15; Haa13], the $1^{++}0^{+}\rho(770)\pi S$ wave has been chosen to measure the relative phase, because it shows intensity over the complete mass range and little cross-talk with other waves.

The different model selection methods, introduced in the previous chapters, will be applied and their properties will be discussed. In section 7.5, a modification of the MBCM method will be introduced that makes use of the phase space of the decay amplitudes to further improve the results.

7.1 Unregularized MLE Fit

Before turning to the model selection procedures, the results of the unregularized MLE fit with the full and the reduced wave pool are presented. For this case, the whole wave pool is equivalent to the final wave set. It is known from the findings in the single low-mass bin, that this will probably produce undesired results.

From a Bayesian viewpoint the Frequentist MLE fit is equivalent to a MAP estimate with a flat prior. This can be considered a non-informative way of obtaining estimates of the amplitudes. If this approach were to succeed, meaning that the results show no unphysical large intensity enhancements or discontinuities, then no additional regularization of the fit would be needed or even justified and the data alone is provide enough information to identify all the waves.

Like for the single mass bins, first the complete wave pool will be used. The intensity of several waves with positive reflectivity are shown in fig. 7.1. The waves show a strong increase in intensity for the lower was region. This behavior is expected from the findings of the unregularized fit on simulated data in the low-mass bin. The increase in the intensity of individual waves can be understood as cancellations due to destructive interference within one J^{PC}-sector. As discussed above section 5.4 in the lower mass region the decay amplitudes become more and more similar as their only distinguishably criterion is the mass shape of the isobar, which becomes less characteristic in the far low-mass tail.

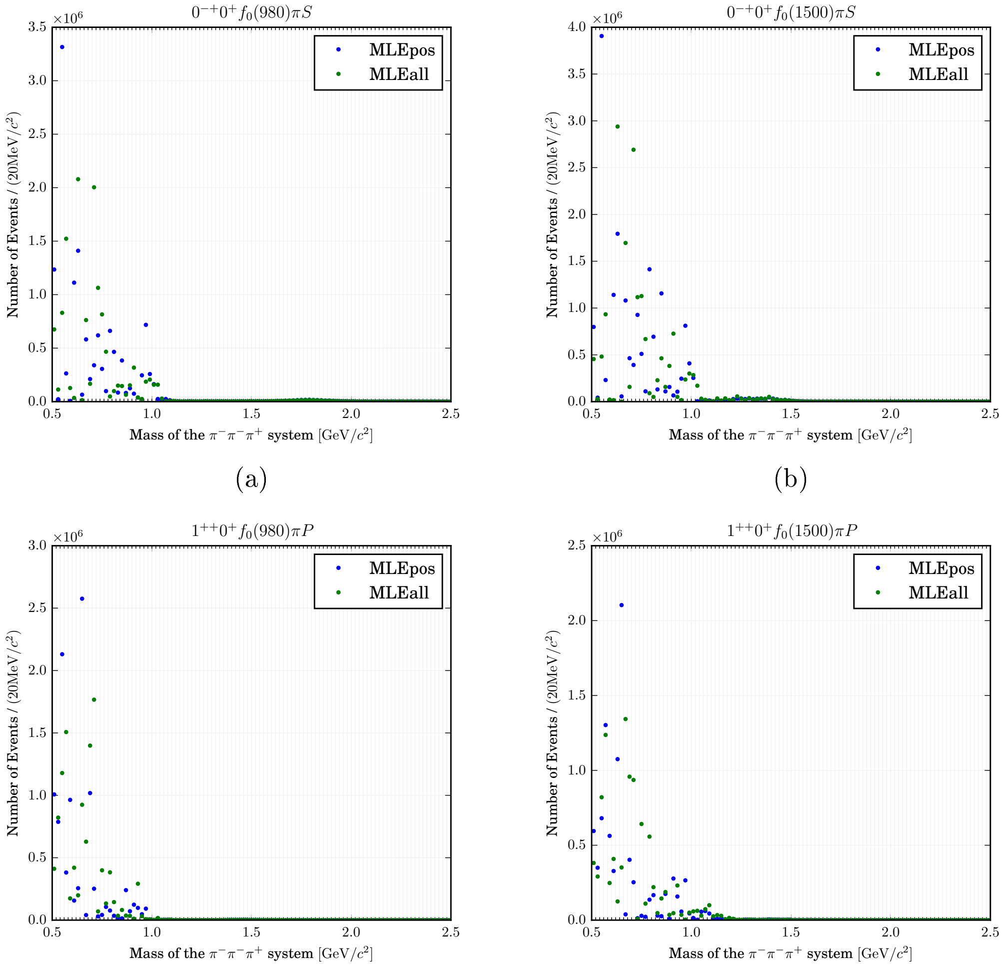
In the higher-mass regions starting from $1.5\,\mathrm{GeV/c^2}$ and for other isobars than the three 0^{++} ones, the effects of strongly correlated decay amplitudes diminishes and the inferred intensities look far more reasonable as can be seen in fig. 7.2. Additionally signals are visible in the relative phases. Again this is consistent with the observation of better behaved fits on simulated and measured data in the single high-mass bin.

For waves with negative reflectivity no such structures could be discovered in neither the intensity nor the phase. In fig. 7.3 two waves with negative reflectivity are shown, which were found to be large in the analysis of Ref. [Dro15] or Ref. [Haa13]. Here an enhancement of the intensities towards lower masses can be observed, much like for waves with positive reflectivity. Nevertheless neither for the higher masses, nor for different isobars any smooth structure could be observed. Regarding the discussion of the flat wave, given in the next paragraph, the waves with negative reflectivity are considered to describe artifacts and background.

An interesting effect can be seen in the isotropic background wave fig. 7.4. For the fit with the complete wave pool, which is plotted in green, the intensity vanishes completely over the full mass range. The fit with the reduced wave pool, plotted in blue, shows a broad intensity spectrum, starting a bit below a mass of $1.0\,\mathrm{GeV}/c^2$. Even though it may seem desirable to have a perfect model for the data that makes any type of background description by the isotropic flat wave superfluous, it is certainly not expected that the flat wave disappears completely. A similar observation was made for the BCM model selection performed in Ref. [Dro15]. The BCM fits were also performed for the complete wave pool. While the flat wave did not vanish completely, it did so for most of the mass range. For tests of rank 2 fits, performed in Ref. [Haa13], also a similar result was obtained. Fits with higher rank are partly comparable to fits with negative and positive reflectivity waves, as has been explained in section 3.1.

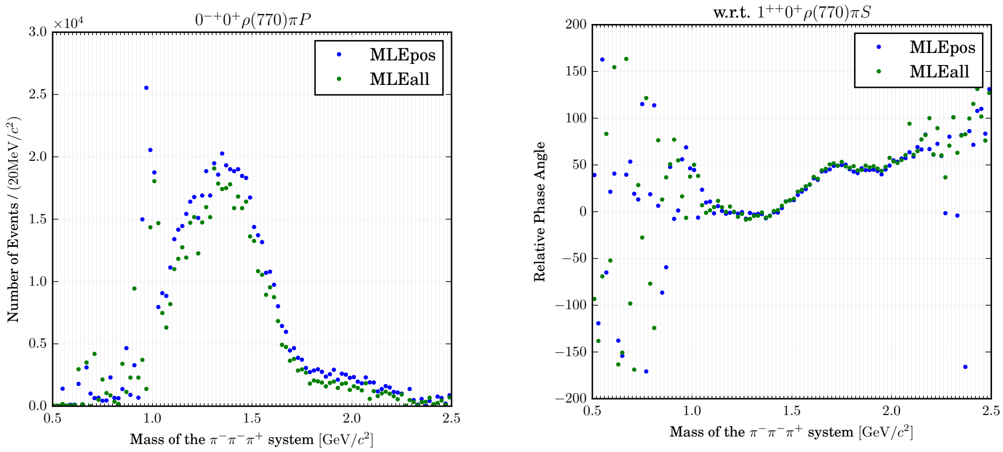
Regarding this effect in the context of the vast intensity increase in the negative reflectivity sector is interesting. It seems that the three different incoherent sectors available to the fit are strongly concurring to describe the data. While there is certainly structure in the dominant sector of positive reflectivity the negative reflectivity sector shows practically no smooth structure at all. This suggests that the model has problems describing the data accurately and/or that strong structure in the background is present. The inclusion of the many degrees of freedom of the negative sector allows the fit to better adapt to the structure in the data.

It has already pointed out in section 2.2 that resonances with positive reflectivity should be dominant in this analysis. Test with and without the inclusion of waves with negative reflectivity, performed in Refs. [Ado+17; Haa13; Dro15], proved that the effect on the waves with positive reflectivity is negligible. For these reasons, in most of the following studies the reduced wave pool, containing only positive reflectivity waves and the flat wave, will be used for most studies as long as not otherwise stated.



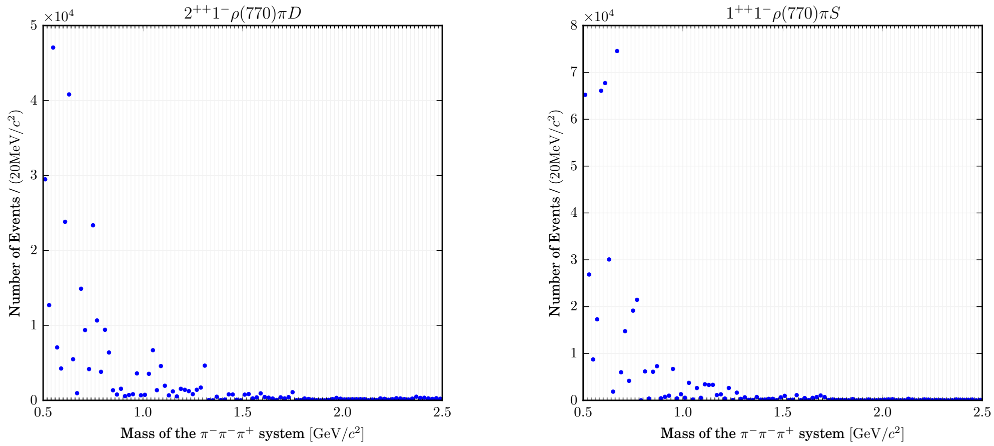
(a) (b) (c) (d)

Figure 7.1: Mass spectrum for selected 0^{-+} waves for the unregularized MLE fit with the complete wave pool. An extreme increase in intensity is visible towards lower masses.



(a) (b)

Figure 7.2: Example of a large signal that is visible even without regularization. (a) intensity spectrum (b) phase motion.



(a) (b)

Figure 7.3: Example of two waves with negative reflectivity for the unregularized MLE fit. (a) selected by the BCM method in Ref. [Dro15] (b) included in the hand-selected wave set of Ref. [Haa13].

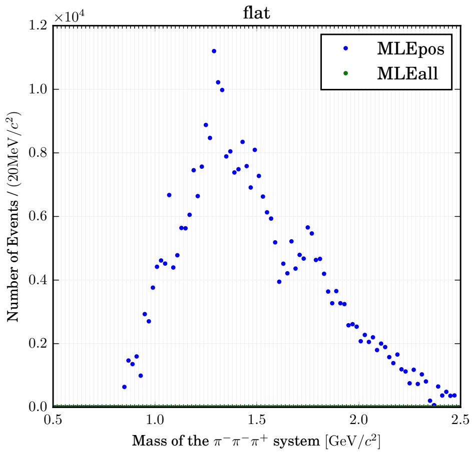


Figure 7.4: Intensity of the isotropic wave for the unregularized MLE fit with the reduced and complete wave pools. While the flat wave shows a broad intensity spectrum for the fit with the reduced wave pool, shown in blue, the flat wave vanishes completely for fit with the complete wave pool.

7.2 BCM Regularization

In [Dro15] the BCM method has been applied with the full wave pool of 432 waves. The scale parameter was fixed to $\Gamma = 0.5$. Similar to the unregularized MLE fit in the above section, large increases in intensity of several waves were observed for lower masses. The effect on the flat wave was also similar. Over most of the mass range its intensity was practically zero. Some residual intensity was observed up to a mass of about $1.5\,\mathrm{GeV}/c^2$. Also there, no insight could be gained on the nature of waves with negative reflectivity.

Because the BCM method has only been applied with the full wave pool of 432 waves, in this section the method will be applied to the reduced wave pool containing only the positive reflectivity waves. This is again to check if the inclusion of the negative reflectivity sector is causing the intensity increase of the positive reflectivity waves. For the fixed parameter $\Gamma = 0.5$ the results for the same waves as in fig. 7.1 are shown in fig. 7.5. The intensity enhancements are smaller than for the unregularized fit by about a factor of 10. Nevertheless, it is obvious that even with the reduced wave pool the unphysical behavior persists.

In order to check if a decrease in the scale to $\Gamma = 0.2$ and $\Gamma = 0.05$ is capable of resolving the issues, these fits have been performed as well. The decrease of the scale-parameter requires a longer time for the fits to converge. Unfortunately in a few bins, with large numbers of events, the fits did not converge in an appropriate time. Because the decrease of the scale-parameter did not solve the problems in the rest of the bins, it is not expected to gain any insight from this mass region and no attempt has been made in repeating these fits.

The studies performed with the BCM method, presented in this section and the previous work, reveal that this method is not practically applicable for selection of a wave set and further manual intervention is needed. This was expected from the conclusions of the previous work and the results obtained in the single mass-bins.

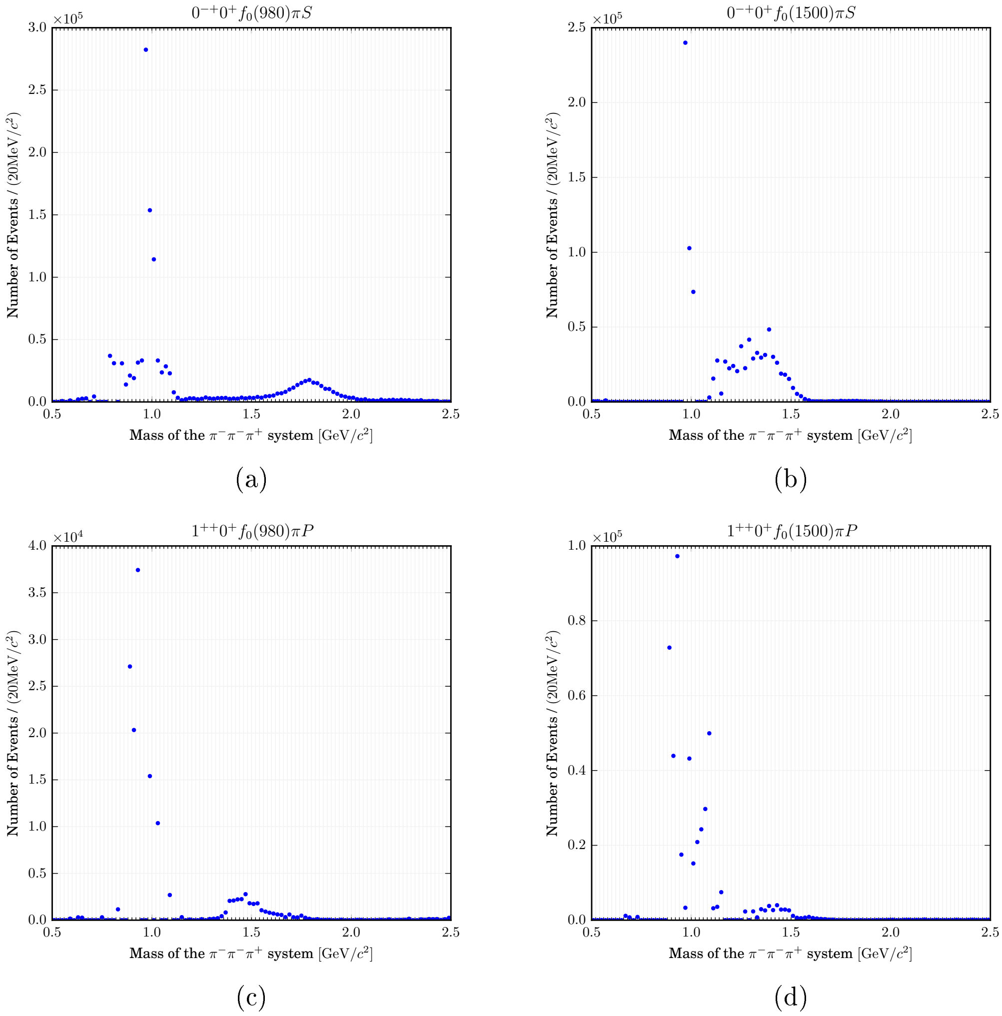
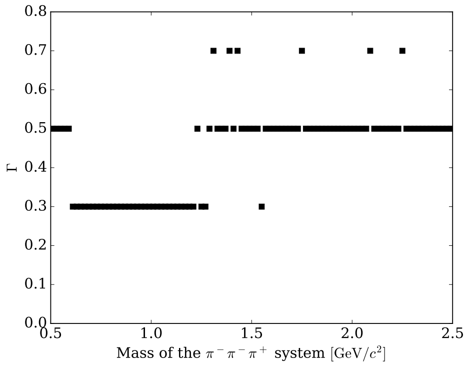


Figure 7.5: Mass spectrum for selected 0^{-+} waves for the unregularized MLE fit with the complete wave pool. An extreme increase in intensity is visible towards lower masses.

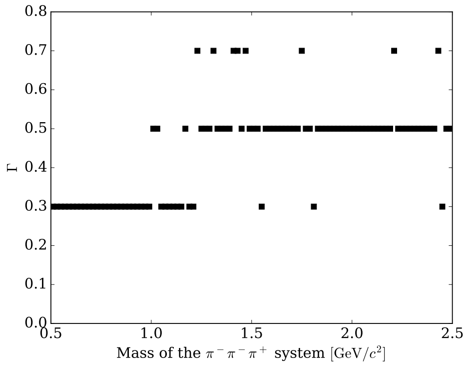
7.3 LASSO Regularization

In contrast to the BCM method the LASSO regularization showed good properties in inference for the single mass bins for both simulated and measured data. As mentioned above the method requires the tuning of its parameter as there is no limiting behaviour. For this data the BIC, given in eq. (4.15), produced promising results in the single mass bins whereas selection via the AIC tended to select very large wave sets and is considered less suitable for model selection in this context. These large wave sets would not be suitable for a refit of the selected model.
The authors of [Gue+15] suggested the penalization of the isotropic flat wave in their method. Because the negative reflectivity sector is left out and other sources of background are expected, a variant with and without penalization of the flat wave has been tried and preferred.
Due to the computational effort required in parameter scanning the smoothed version of the LASSO prior has been used. It has been demonstrated in fig. 5.20 that the smoothing has little influence on the actual result while greatly improving fit stability and convergence time.
The choice of prior value over the masses is shown in fig. 7.6b for the fits with penalized flat wave and in fig. 7.6a for the ones without. The scan has been conducted for the following four values of the parameter $\Gamma = 0.1, 0.3, 0.5, 0.7$.
The BIC tends to select smaller widths and therefore a stronger bias with lower masses. This behavior is sensible when considering that the fits without penalty or the ones with the BCM method applied become more reasonable at higher masses. The values are similar in both cases, with penalized flat wave and without.

The results for the LASSO fit are shown in fig. 7.7 for all four Γ values. Returning to the waves chosen to benchmark the method it can be seen that the issues of large destructive interference can be resolved in the low-mass region for all of the problematic waves. An interesting property is the thresholding of the $0^-0^+f_0(980)$ wave in the region around $1.1\,\mathrm{GeV}/c^2$. This is similar to the threshold that has been chosen for this wave by hand.
Also the $0^-0^+\sigma_0$ has been found to follow the results of the hand-selected model in the intensity spectrum.
In fig. 7.7b and fig. 7.7d waves with the $f_0(1500)$ isobar are shown. They still show a strong intensity increase below $1.5\,\mathrm{GeV}/c^2$.



(a)



(b)

Figure 7.6: BIC choices of LASSO parameter for (a) no penalization of the flat wave (b) with penalization of the flat wave

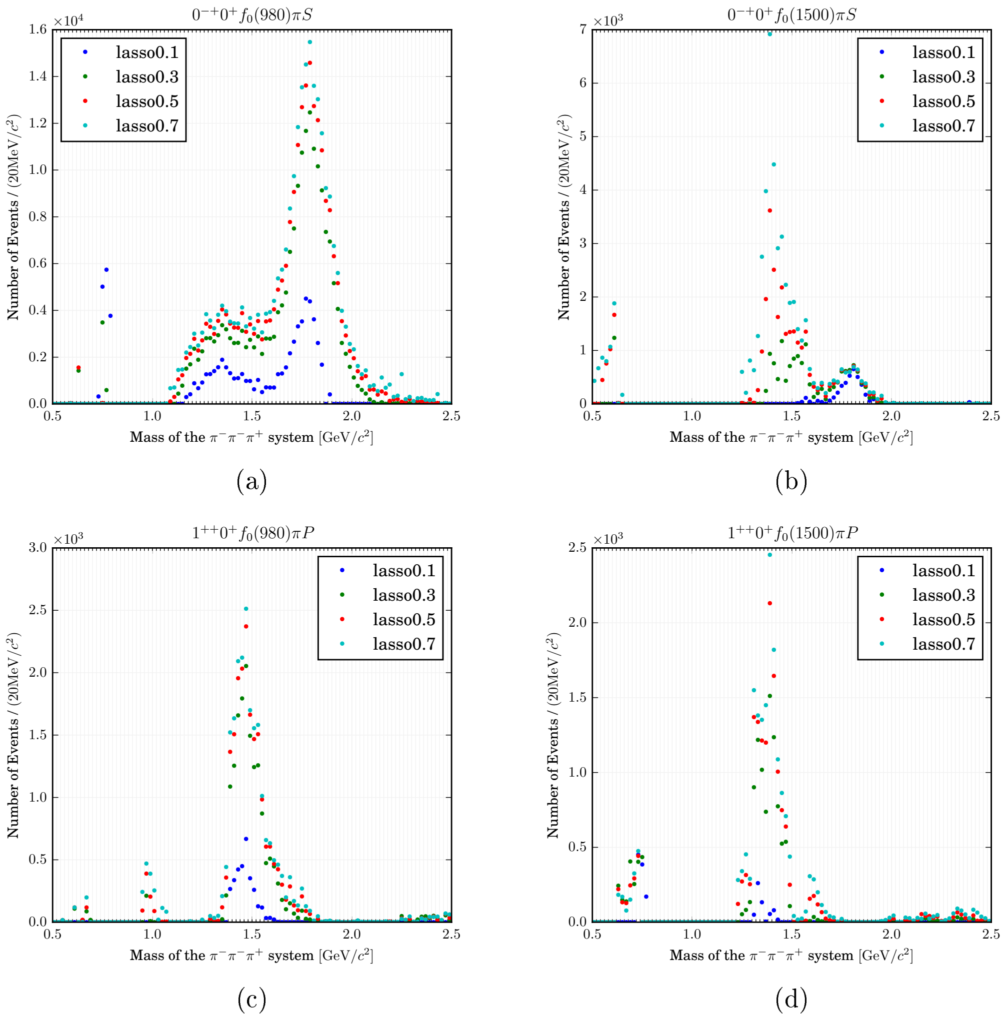


Figure 7.7: Fits with LASSO regularization using the reduced wave pool. The flat wave has not been penalized for these fits. Four different values for $\Gamma = 0.1, 0.3, 0.5, 0.7$ have been used to be able to scan the parameter. For all values the LASSO is able to provide reasonable thresholds for the $f_0(980)$ wave.

7.4 MBCM Regularization

In the single mass bins, the MBCM method showed a significant improvement over the BCM method, in The idea of the MBCM method introduced for the single mass bins was to combine the advantages of both BCM and LASSO. For the single mass bins it was able to resolve issues with too large intensities of the $0^-0^+f_0(980)$ and $0^-0^+f_0(1500)$ waves.
The idea is to keep the tail of the prior as heavy and thereby as weakly informative as possible, but as informative as necessary. The benchmark was the disappearance of the $0^-0^+f_0(980)$ and $0^-0^+f_0(1500)$ waves in the low-mass bin.
Here the MBCM method has been applied with a scale parameter of $\Gamma = 0.2$ and a decay parameter of $\lambda = 5.1$. These values have been chosen for the following reason. The scale parameter is related to the scale parameter of the BCM method. The choice of $\Gamma = 0.2$ was found to be large enough for fast convergence of the fits while simultaneously being small enough to come closer to the limit of the BCM prior. As argued previously smaller scale-parameters seem to be desirable, at least on simulated data.
For the decay-parameter $\lambda = 5.1$ was chosen because no strong influence in the high-mass bin was observed. It was concluded the the influence of the lighter tails of the MBCM penalty should be weak enough to be acceptable. This can also be understood when the decay-parameter is related to the scale-parameter of the LASSO via the relation introduced in section 4.2.3.3. It is therefore expected to have a weaker bias than the LASSO for any of the LASSO scale-parameters chosen by the information criteria in section 7.3.

In fig. 7.8a the intensity spectrum of the $0^{-+}0^+f_0(980)$ wave is shown for the MBCM fit. Like for the LASSO methods the threshold of the wave is nicely recovered.
Another similarity with the results obtained with the LASSO method is the strong peak of the $0^{-+}f_0(1500)$ wave, shown in fig. 7.8b. The increase in intensity is bigger, which can be related to the argument about the relation of the decay-parameter and the LASSO scale-parameter made above.
There seems to be two kinds of intensity increases that appear for lower masses that can nicely by demonstrated by the two 0^{-+} waves discussed so far. Both the LASSO and the MBCM method were able to resolve issues with the $0^-0^+f_0(980)$. Because the $f_0(980)$ is very narrow and explicitly separated from the $[\pi\pi]_S$. Only little additional bias is required to resolve the cross-talk between these two waves. The separation is not so clear for the $f_0(1500)$ and the $[\pi\pi]_S$ isobars. The peak structure found by MBCM and LASSO methods is scaling strongly with the strength of the suppression. In contrast to the other wave however it is persistent for all but the most extreme biases induced by the penalties. Which is $\Gamma = 0.1$ for the LASSO or

$\lambda = 0.1$ for the MBCM penalty. These small values are not considered reasonable because the effect on well-established waves is by far to strong.

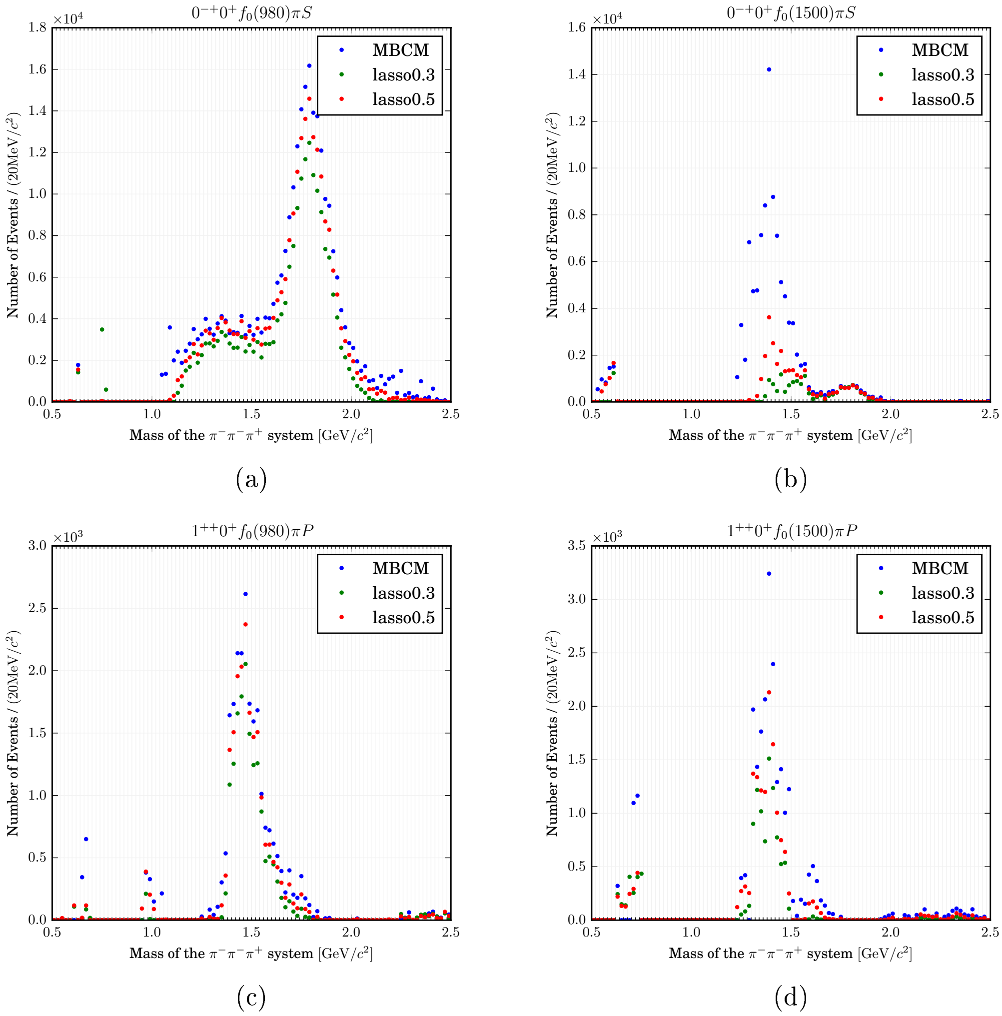


Figure 7.8: Mass spectrum for selected 0^{-+} and 1^{++} waves for the MBCM and LASSO fits. (a) both MBCM and LASSO fits recover the threshold of the $0^{-+}0^{+}f_0(980)$ wave. (b) and (d) show an enhancement of waves with $f_0(1500)$ isobar below the expected threshold.

7.5 Weighting according to decay Phase Space

The approach of a binned fit in mass bears the advantage of recovering the mass dependence of the transition amplitudes without any parametrisation. It therefore also provides a cross check for the appearance of artefacts in the data. With a parametrised fit unwanted fluctuations in one bin or a certain region of bins may translate to structures that look physical over a larger region of the fit but are instead artefacts.
Ignoring the connection between the bins however means that not all information at hand is also available to the fit. For example whether a certain decay represented by a partial wave is sensible in a certain mass region.
Nevertheless this information is later used to check ones believes in the results. In other words, there is prior knowledge available that is not being used in the fits.
Not including any of this information several situations arise in which the model becomes unstable, meaning that waves have no continuous intensity over a certain mass region or become extremely large, exceeding the number of expected events by far.
With the usual hand-selection approach this information was included in the form of thresholds. These thresholds were implemented to forbid problematic waves below certain masses. The physics reason for this is in the available phase space to the decay. Certainly a decay which has practically no available phase space is not expected. This is for example the case for decays that contain a high-mass isobar like the $f_0(1500)$ or $\rho_3(1690)$ at low masses. In the extreme low mass regions of the fit it is therefore reasonable to exclude certain waves altogether. Because the phase space of a wave is continuous over the mass range it is hard to tell where to put the threshold that resolves ambiguous behavior.
Here a way of including this knowledge in a continuous and systematic fashion via the decay parameter λ of the MBCM method is introduced. The underlying assumption is that for a very small phase space for the decay large productions rates are necessary for the resonance to be visible. So the prior knowledge would disfavor such decays and hence in an informative setting the prior should suppress those waves much more strongly than the ones with a reasonable large phase space of the decay channel. For large masses the suppression should be equal for all possible waves considered in the wave pool, because their phase spaces should be sufficiently open or otherwise it would not make sense including them in the wave pool in the first place.
This is also consistent with the property of the fit being more stable for higher masses.
The strength of the suppression can be controlled by the decay parameter of the prior. Where a smaller value corresponds to stronger suppression and a vice versa. For every wave λ becomes a function of the mass. To get the desired behavior described above the diagonal elements of the phase-space integrals $\mathcal{P}_{ii}$, given in

eq. (5.1), can be used. A simple linear form eq. (7.1) was chosen in such a way that the value of the decay-parameter becomes equal at a large normalization mass m_X^{norm}. A small offset-parameter o stops the suppression from becoming to strong so it does not become impossible for a wave to be included, effectively giving an upper limit for the bias. The global scale s defines the minimum bias acting on the waves in the limit of large masses.

$$\lambda\left(i, m_X\right) = o + s \cdot \mathcal{P}_{ii}\left(m_X\right) / (\mathcal{P}_{ii}(m_X^{\text{norm}})) \tag{7.1}$$

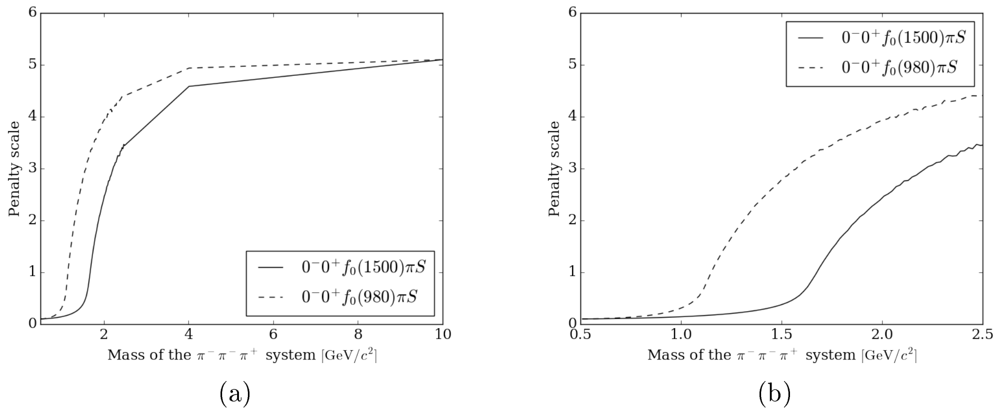
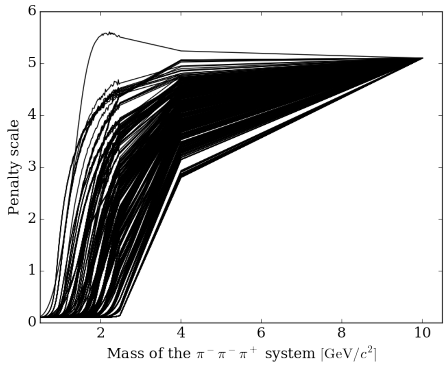


Figure 7.9: Example for the mass behavior of the suppression parameter $\lambda\left(i, m_X\right)$ for two 0^-0^+ waves.

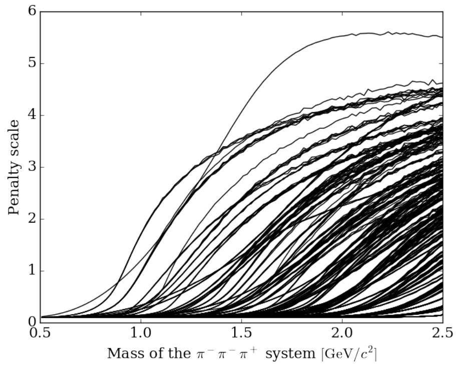
The decay-parameter of the prior is visualized in fig. 7.9 for the two waves $0^-0^+f_0(1500)\pi S$ and $0^-0^+f_0(980)\pi S$ which have both shown undesired behaviour in the region around $1.3\,\text{GeV}/c^2$ as discussed above. The parameters of eq. (7.1) have been chosen to be $o = 0.1$ for the offset, which should be small enough for strong suppression whilst being large enough to keep possible strong signals, and $s = 5$ which seems to be a reasonable choice in the high mass region when a fixed parameter for the bessel-prior is used.

The normalization at $10\,\text{GeV}/c^2$ is shown in fig. 7.9a. Additionally a point at $4\,\text{GeV}/c^2$ has been included. For larger masses the decay-parameters converge to the common normalization value, which is 5.1 in this case, while for lower masses the thresholding behavior of the weighting becomes apparent as seen in more detail in fig. 7.9b which is constrained to the mass regions of the fit.

The curves rise steeply at about the same region as their respective thresholds in the hand-selected wave-set. The idea is that the fit can now find thresholds that are compatible with ones believes in a systematic way for all waves while having enough



(a)



(b)

Figure 7.10: Example for the mass behaviour of the suppression parameter for all waves.

freedom to determine the exact position of the thresholds by itself.
Effectively the choice of thresholds has been reduced to the choice of two parameters and one mass for the normalization.

The effect of this phase-space weighted MBCM method can best be seen for the $0^{-+}f_0(1500)$ wave. It showed an unphysical increase in intensity below $1.5\,\mathrm{GeV}/c^2$ for both the MBCM and LASSO method. The result of the new weighted MBCM method is capable of using the information about the phase space, such that a physical threshold is found. The result for the $0^{-+}f_0(1500)$ wave is presented in fig. 7.12. Only a single mass bin is showing an increased intensity. This can later be fixed manually without problem and only appears in the first t' bin, as will be shown in chapter 8.

7.5.1 Effects of the weighted MBCM Method in single Bins

The properties of the weighted MBCM have also been studied in single mass bins for both simulated and measured data. The results for the high- and low-mass bin for measured data are shown in fig. 7.13 and fig. 7.14. The weighted MBCM method provides much more unique solutions, as almost not scattering of the fit attempts appears, even in the low-mass bin. A similar result has been found for simulated data. It has to be remarked that the penalty is quite restrictive at low masses. The wave sets are therefore small in the low-mass region.

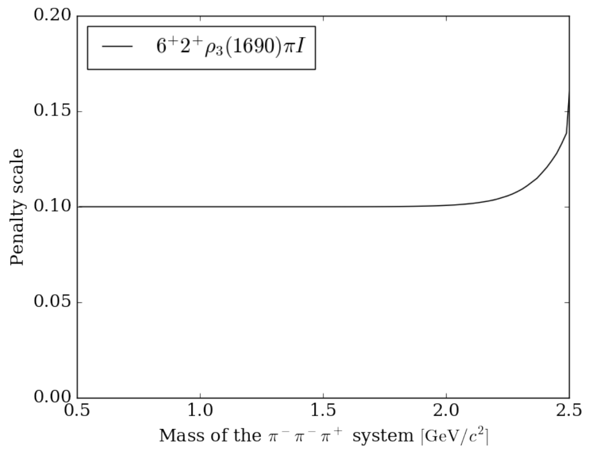


Figure 7.11: Example of a wave for which the suppression is very strong over the complete mass range. The high orbital momentum, spin and spin projection in combination with the heavy $\rho_3(1690)$ isobar.

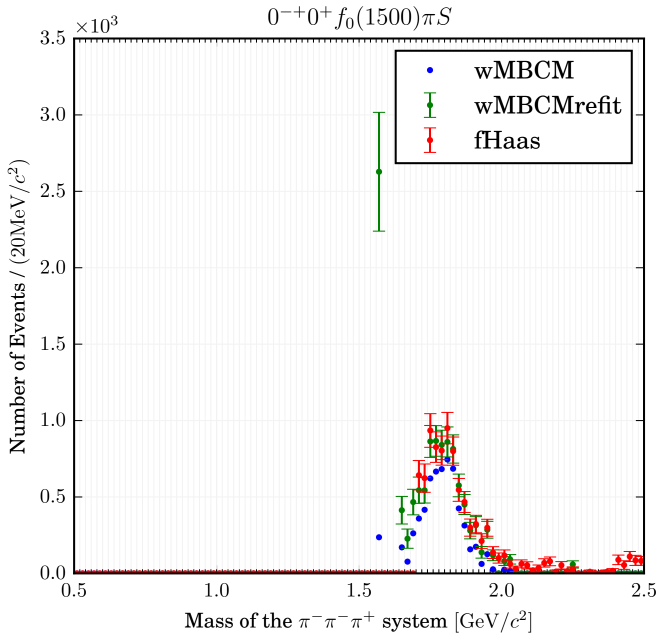


Figure 7.12: Comparison of the weighted MBCM method with the hand-selected 88-wave set. The result of the weighted MBCM fit is shown in blue and its refit in green. The hand-selected model is overlaid in red. Apart from a single bin, the threshold is recovered correctly.

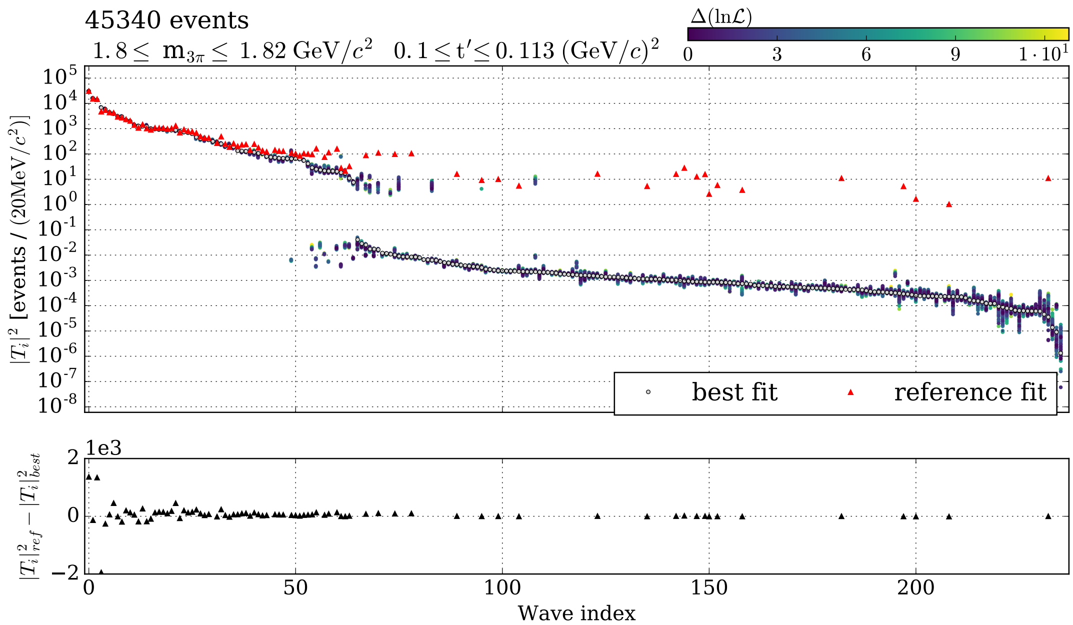


Figure 7.13: Fit result for the weighted MBCM method in the high-mass bin.

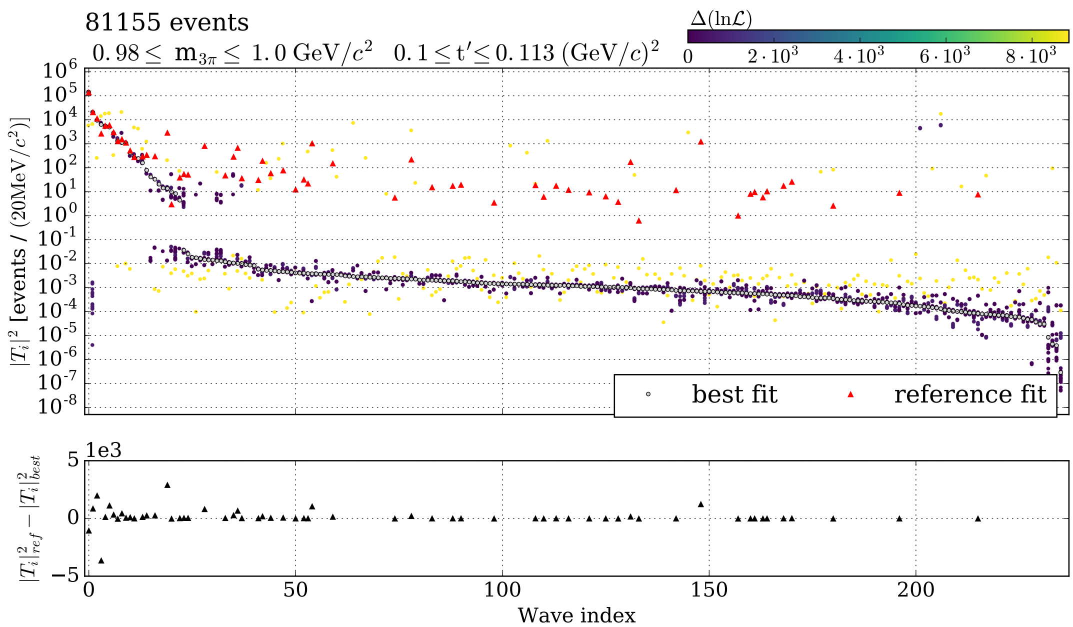


Figure 7.14: Fit result for the weighted MBCM method in the low-mass bin.

Chapter 8

Model-Component Selection in m_X and t'

The MBCM method in combination with the phase-space weighting of λ is able to produce results that are physically sensible and resemble the hand-selected model in the choice of thresholds. Additional waves are found with this approach, some of which were previously discovered by the BCM method [Dro15] and some, which are entirely new.
Additionally the hierarchy in the penalization of the waves reduces the effects of the multimodality. This has two advantages. The first one being the interpretability of the found modes. Different modes that are close in likelihood are found to also produce similar results. The selected model is therefore much more interpretable because it is clearly separated from the solutions with smaller likelihood and recovered more than once.
The second is a significant decrease in the time required for the fit to converge. This speed-up is extremely helpful in performing studies of the wave set.

Together with the above mentioned advantages the promising results obtained in the lowest t'-bin justified the application of the method to the rest of the available data. This study takes the step from a case-study performed on the single t'-bin towards a realistic analysis. The additional information in the complete dataset should also help to generate reliable wave sets.

In this chapter the weighted MBCM method will be applied to the 11 t'-bins available. The choice of parameters is the same as in section 7.5, also the normalization of λ was chosen to be at $10\,\mathrm{GeV}/c^2$. The reduced wave pool, containing only positive reflectivity waves, has been used. These settings were found empirically to be a good choice, but different ones are possible for further systematic studies in the future.

In section 8.1, the model selection results will be discussed, using the problematic 0^{-+} partial waves as examples. The individual wave sets, obtained for each

two-dimensional bin in mass and t', will then be refitted. The effects of the refit and the procedure of refitting itself will be critically discussed.
The previous analyses relied on a single wave set with lower thresholds in mass. Here a simple method for building a combined wave set will be presented. This wave set is then fitted to the complete dataset once more. Selecting a single combined wave set will also be discussed in the context of this analysis. The results will be compared with the refit of the individual wave sets in each bin and the results without refit.

8.1 Fit results in m_X and t'

In this section the results of the weighted MBCM fit are shown for a selection of waves. A typical representation of the results obtained over all t'-bin is the so-called t'-summed plot. For these type of plots the intensities for a wave at a given mass are summed over all 11 t'-bins. These plots usually show a smoother behavior over the mass range and help to identify resonances by looking for peaks in the spectrum. The disadvantage is the loss of the relative phase information, which cannot be summed up.
In fig. 8.1 and fig. 8.2 the t'-summed intensity plots for the $0^{-+}0^{+}f_0(980)$ and $0^{-+}0^{+}f_0(1500)$ waves are shown. Both waves show a smooth intensity spectrum and no discontinuities. The $0^{-+}0^{+}f_0(1500)$ wave shows a peak structure starting from $1.5\,\mathrm{GeV}/c^2$. This wave was thresholded at $1.7\,\mathrm{GeV}/c^2$ in the hand-selected 88-wave set.
The t'-dependence of the intensity spectrum can be made visible in form of a heat map. For these plots it is very important to stress that size of the bins is not equal over t'. The non-equidistant binning in t' is listed in table 3.1. The heat-maps can use a linear scale for the color, because the number of events in each t' bin is similar. In fig. 8.3 and fig. 8.4 the plots for the 0^{-+} waves with isobars $f_0(980)$ and $f_0(1500)$ are resolved in t'. The results are smooth in both mass and t'.

The third 0^{-+} wave, that usually caused problems with fit stability and large intensities, is the $0^{-}0^{+}[\pi\pi]_S\pi S$ wave. This wave also shows smooth behavior over the dataset. In the t'-summed intensity spectrum, shown in fig. 8.5, a pronounced shoulder is visible at about $1.3\,\mathrm{GeV}/c^2$.

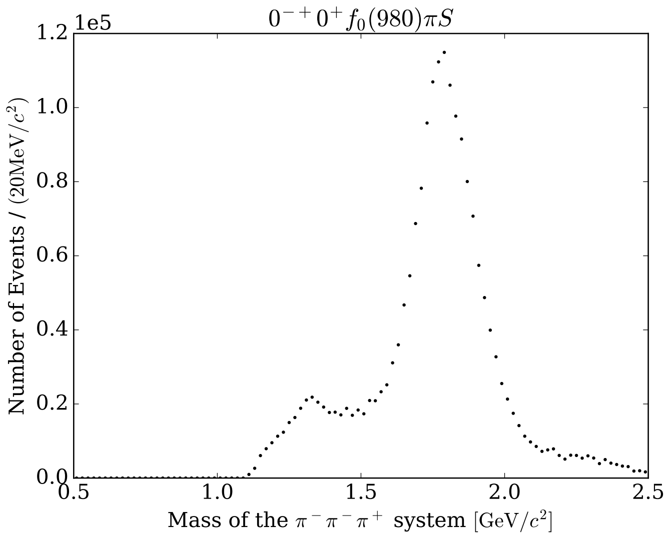


Figure 8.1: t'-summed intensity spectrum of the $0^{-+}0^{+}f_0(980)\pi S$ wave for the model selection fit. The intensity smoothly approaches 0 for masses around 1.1 GeV/c^2.

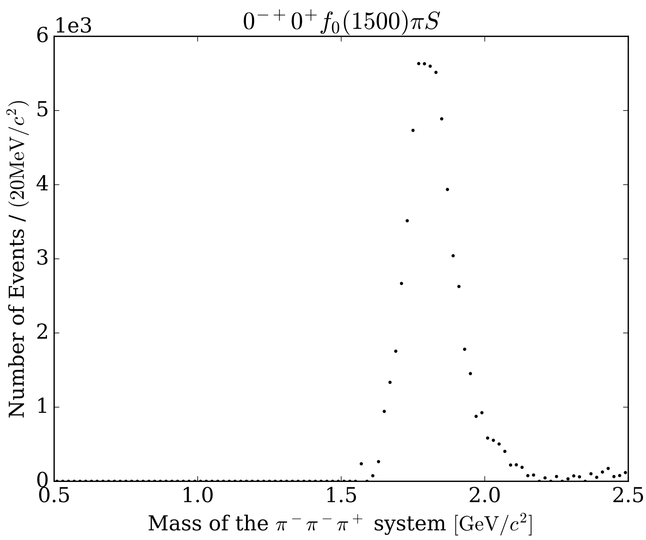


Figure 8.2: t'-summed intensity spectrum of the $0^{-+}0^{+}f_0(1500)\pi S$ wave for the model selection fit. The intensity spectrum shows a clear peak starting at about 1.5 GeV/c^2. For the hand-selected wave set the threshold was at 1.7 GeV/c^2. The threshold suggested by the weighted MBCM method is lower, such that the full peak structure could be recovered.

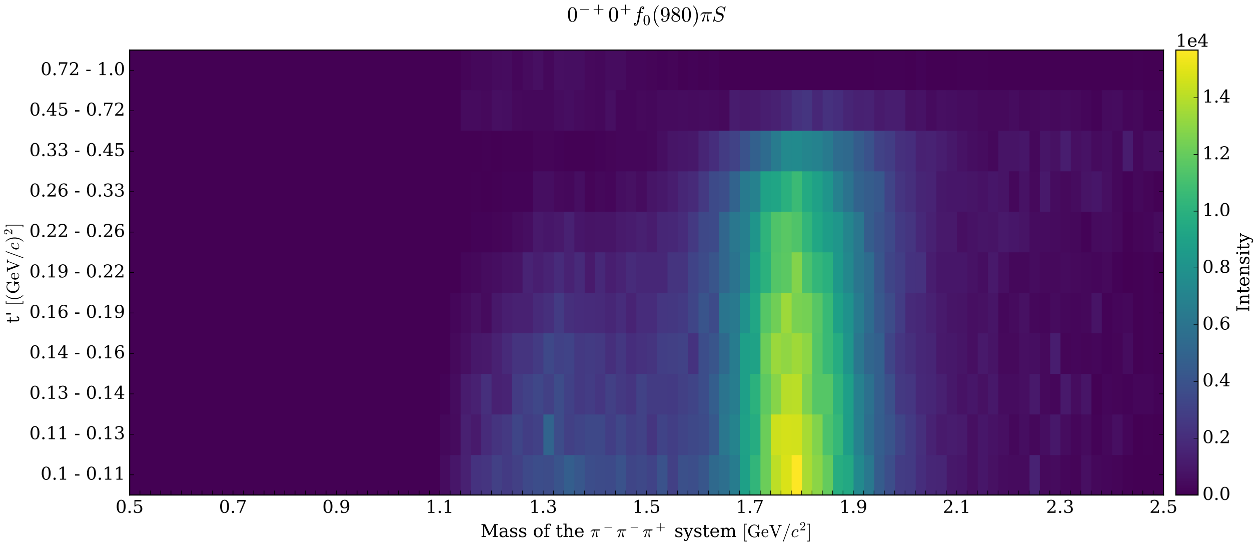


Figure 8.3: mass and t' plot of the $0^{-+}0^{+}f_0(980)\pi S$ wave for the model selection fit. The different t'-dependence of the dominant peak at about $1.8\,\mathrm{GeV}/c^2$ and the shoulder at $1.3\,\mathrm{GeV}/c^2$ is visible.

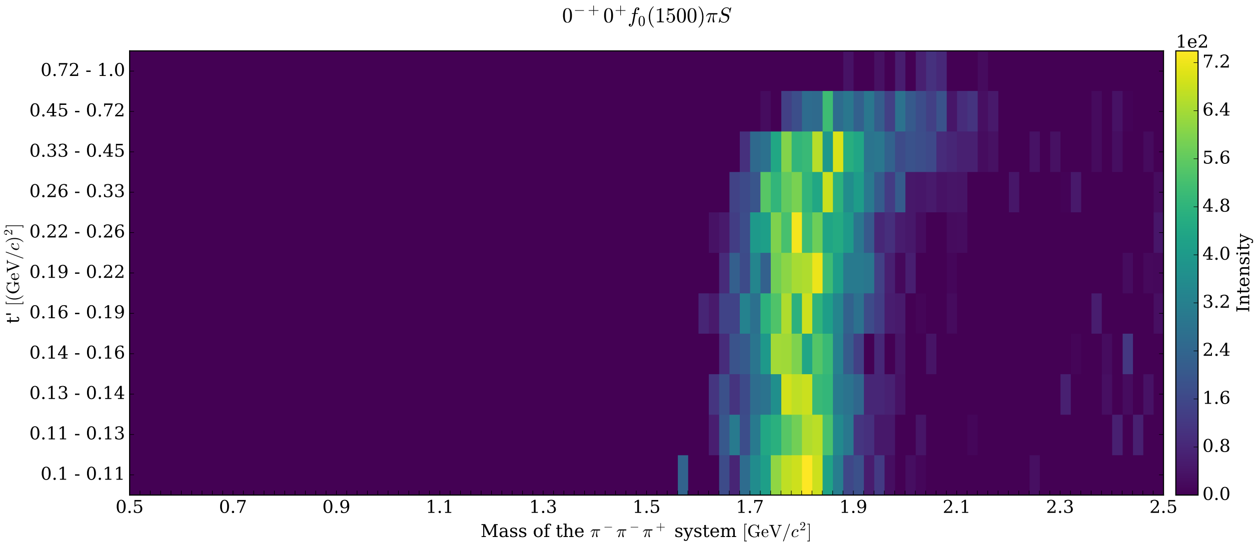


Figure 8.4: mass and t' plot of the $0^{-+}0^{+}f_0(1500)\pi S$ wave for the model selection fit. The weighted MBCM method nicely recovers the peak structure at about $1.8\,\mathrm{GeV}/c^2$. Only for the lowest t' bin a single mass bin below the thresholding region is found. The thresholds show a t'-dependence.

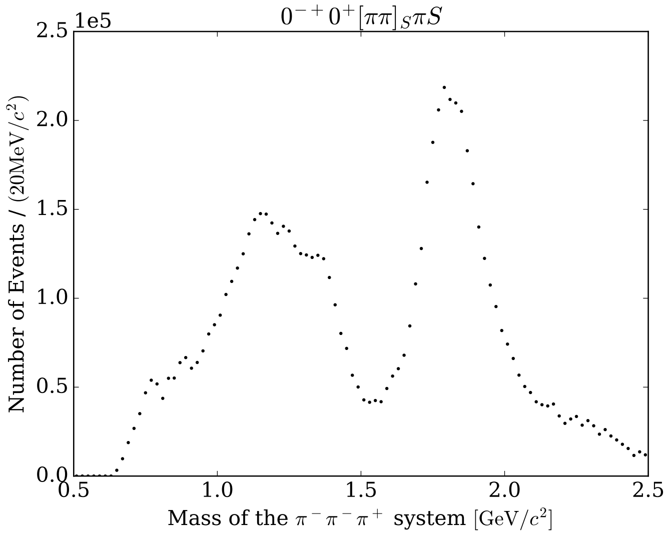


Figure 8.5: t'-summed intensity spectrum of the $0^{-+}0^{+}[\pi\pi]_S\pi S$ wave for the model selection fit. At $1.3\,\mathrm{GeV}/c^2$ a shoulder is visible that was not observed in the previous analysis of Ref. [Haa13].

8.2 Refit of the models

The intention of the model selection was the selection of relevant waves over the two-dimensional bins, which are then joint into a single wave set. The refitting of the model is then performed in order to get error estimates that can be used for the resonance parameter fit. Whether a refit should be performed or not can be regarded controversial. A discussion of this can be found at the end of this chapter in section 8.7.

The weighted MBCM method is capable of finding the required thresholds similar to the hand-selected model, as already mentioned above in section 7.5.
Because the model selection procedure was applied to each bin individually, individual models are available for the different bins. This means that waves may appear and disappear from bin to bin. For a 'classical' wave set only lower thresholds in mass were given, such that in each t'-bin the wave set was the same.

In this section the individual wave sets have been refitted in their respective two-dimensional bins. No additional post-processing of the wave sets has been applied.

The t'-summed intensity plots are again shown for the 0^{-+} waves with the isobars $f_0(980)$, $f_0(1500)$ and $[\pi\pi]_S$ in fig. 8.6, fig. 8.7 and fig. 8.8.
Without the regularization the intensities of the waves can increase, so in general the refits show a larger intensity than their corresponding selection fits.
For the waves with the $f_0(980)$ and $f_0(1500)$ isobars the refit keeps the shapes of the model selection fits for almost all bins, meaning that no extreme increase in intensity or distortion is visible. Only in the first bin after the threshold this behaviour appears for both waves. This can easily be removed manually if desired.
The cross talk is stronger for the $[\pi\pi]_S$ wave. At $1.1\,\mathrm{GeV}/c^2$ a discontinuity in the spectrum is visible. The increase in intensity can be understood as cross-talk between the $[\pi\pi]_S$ wave and the $f_0(980)$ wave. The discontinuity appears exactly at the threshold of $f_0(980)$ wave. Whether the shoulder observed at about $1.3\,\mathrm{GeV}/c^2$ can also be related to the cross-talk is not clear. One has to point out that this mass region suffers from the cross-talk of these three waves and without further improvement of the model this has to be included in the systematics of the resonance parameter fit.

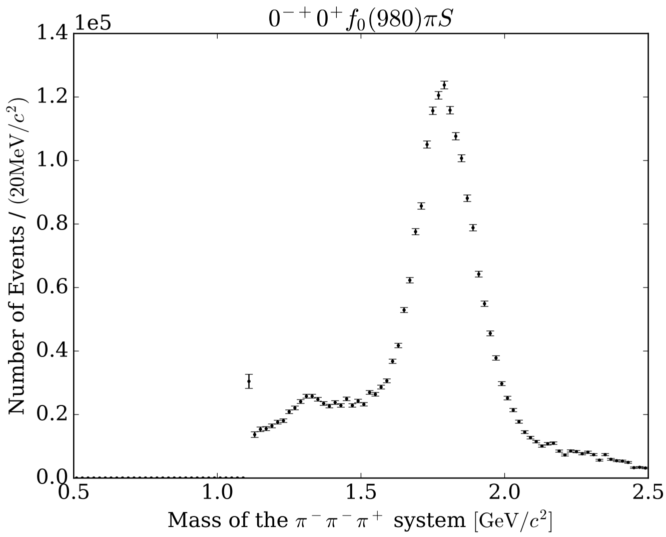


Figure 8.6: t'-summed intensity spectrum of the $0^{-+}0^{+}f_0(980)\pi S$ wave for the refit. The intensity spectrum is stable, but shows a discontinuity at the thresholding region of about 1.1 GeV/c^2 similar to the hand-selected model.

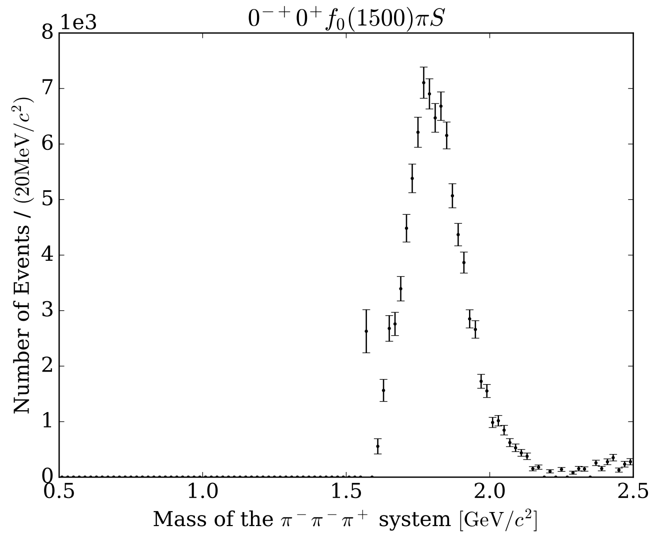


Figure 8.7: t'-summed intensity spectrum of the $0^{-+}0^{+}f_0(1500)\pi S$ wave for the refit. The intensity peak does not show any unphysical increases in intensity. The first bin, in which the wave was found, however does behave discontinuous.

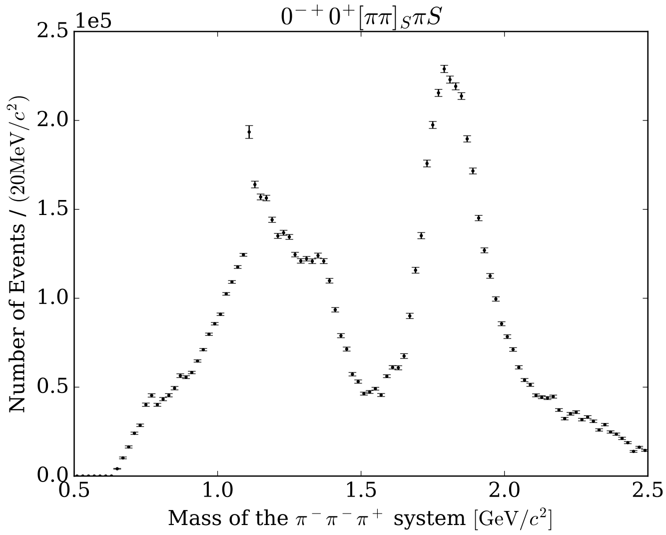


Figure 8.8: t'-summed intensity spectrum of the $0^{-+}0^+[\pi\pi]_S\pi S$ wave for the refit. After the refit, the shoulder $1.3\,\text{GeV}/c^2$ persists. At $1.1\,\text{GeV}/c^2$ cross-talk, in the form of an increased intensity, with the $0^{-+}0^+f_0(980)\pi S$ wave is visible (compare fig. 8.6).

8.3 Important Waves

In this section four waves, shown in fig. 8.9, are discussed. These are either dominant in the intensity spectrum or important for the further analysis.

In fig. 8.9a the t'-summed intensity plot for the $a_1(1260)$ resonance in the $1^{++}0^+\rho(770)\pi S$ wave is shown. This is the wave with the largest intensity in this analysis. Both the hand-selected wave set and the individual wave sets selected for each bin in general recover the same shape. A slight modification of the top part of the peak appears for the selected models. To further investigate this a resonance fit is required.

Another large wave is shown in fig. 8.9b. The $2^{++}1^+\rho(770)\pi D$ is dominated by the $a_2(1320)$ resonance and represents one of the cleanest signals in this analysis. Again both the hand-selected and the model-selection wave sets recover the same shape, while minor deviations need to be considered in systematic studies of the resonance fit.

The studies of three-pion diffractive dissociation data uncovered a new state, known as the $a_1(1420)$ [Ado+17]. This resonance is exclusively in the $1^{++}0^+\rho(770)\pi P$ wave, which is shown in fig. 8.9c.

One important study conducted in Ref. [Haa13] was concerned with the spin-exotic $1^{-+}1^+\rho(770)\pi P$ wave. The results of the intensity fits are shown in fig. 8.9d. In contrast to the previous waves deviations in the double-peak structure are visible. The first peak was identified as thresholding effect of the hand-selected model in Ref. [Haa13]. It is interesting to see that the same peak for the model-selection wave set is smoother and broader, with a smaller intensity and a shift to lower masses. The second peak shows a shift to higher masses and higher intensities. These effects are worth investigating in the future and should be included in the systematics of the analysis.

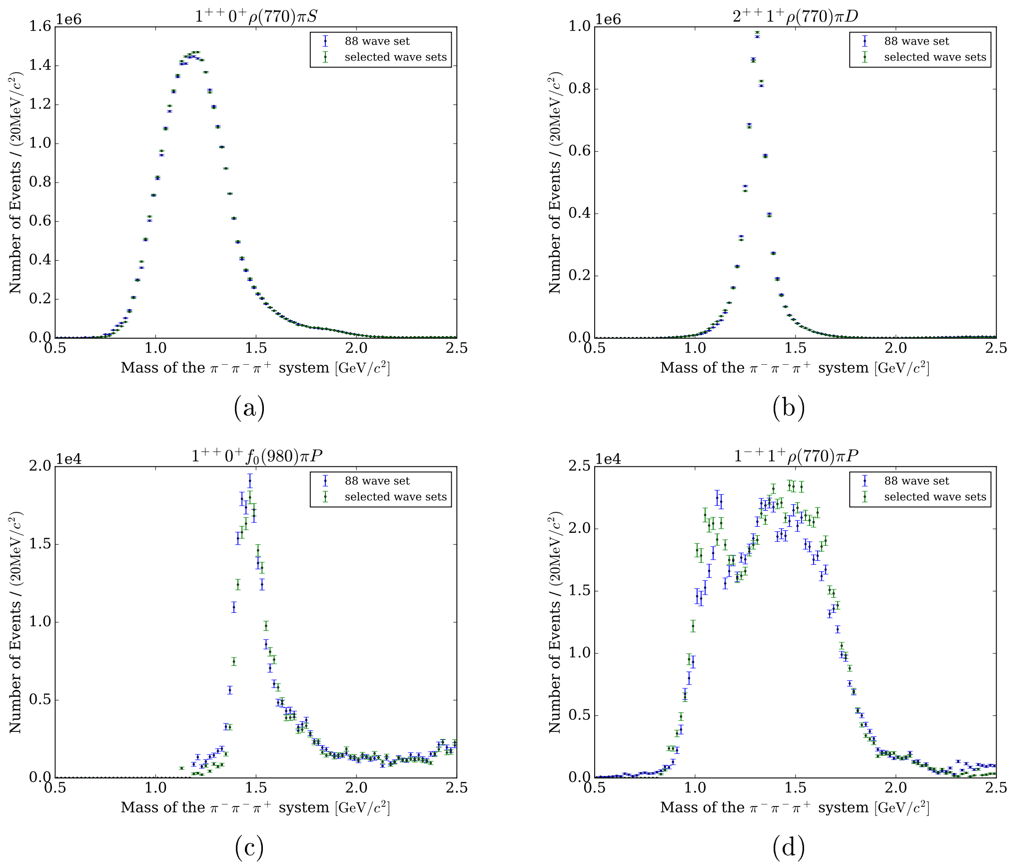


Figure 8.9: t'-summed intensity plots for four important waves. The solutions of the refit of individual wave sets in each two-dimensional bin are plotted in green. The solutions of the fit with the hand-selected wave set is plotted in blue.

8.4 Waves discovered in the BCM Case-Study

In the first case-study with the BCM penalty several new waves have been discovered as potentially relevant [Dro15]. In this section two of the interesting newly discovered waves will be discussed. The application of the weighted MBCM model selection to the complete dataset is able to resolve the t'-dependency of these waves and can therefore give insight on their possible relevance.

A promising peak was observed in the $1^{++}1^{+}f_2(1270)\pi F$ wave at about $2.2\,\mathrm{GeV}/c^2$. This structure was connected with a strong phase motion. The wave was also found with the weighted MBCM method. Its t'-summed plot is shown in fig. 8.10.

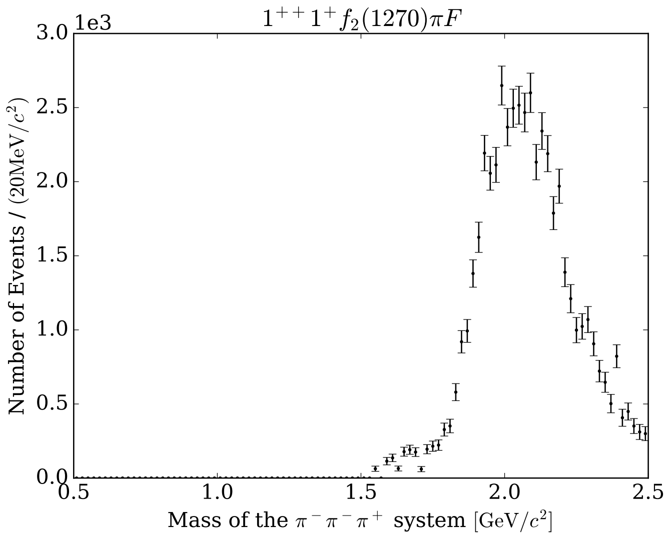


Figure 8.10: t'-summed intensity spectrum of the new $1^{++}1^{+}f_2(1270)\pi F$ wave that was first discovered in the lowest t' bin by Ref. [Dro15].

The heat-map of the t'-resolved fits is shown in fig. 8.11. In the two-dimensional representation it can be seen that the wave shows intensity over almost the complete t'-region. The extent in mass increases towards higher regions in t'. In the highest t'-bin it is not found, probably because there are less events in this bin.

The intensity and relative phase-motion in the lowest t'-bin are shown in fig. 8.12a and fig. 8.12b. Like in the previous analysis a clear phase-motion is visible. The same is still true for higher t'-bins. In fig. 8.12c and fig. 8.12d the intensity and phase plots are shown for the t'-range between $0.45\,(\mathrm{GeV}/c)^2$ and $0.72\,(\mathrm{GeV}/c)^2$. In this t'-region the peak-structure and phase-motion are still visible.
The stable peak-structure together with the corresponding phase motion suggest that this wave may indeed contain a resonance and should be included in further studies. One can speculate at this point, whether this wave contains a signal related to the resonance candidates $a_1(1930)$ or $a_1(2095)$, which are listed as further states in the PDG [Oli+14].

Another interesting structure was visible in the $5^{++}1^{+}\rho(770)\pi G$ wave. Its mass and t' spectrum is shown in fig. 8.13, its t'-summed spectrum in fig. 8.14. Looking at bins of low and high t', more insight can be gained. For the lowest t'-bin a broad peak-like structure is visible in fig. 8.15a. Also a slow rising phase-motion appears in fig. 8.15b. This phase motion flattens for higher bins in t', as shown in fig. 8.15d.

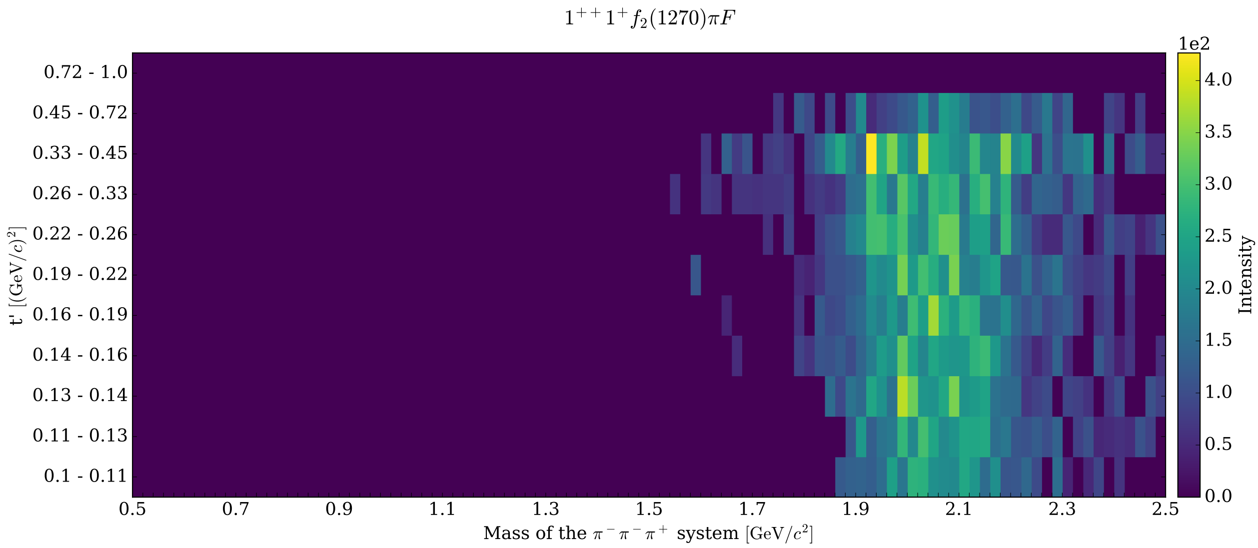
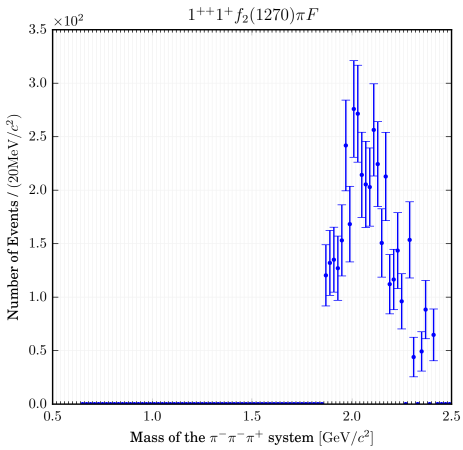
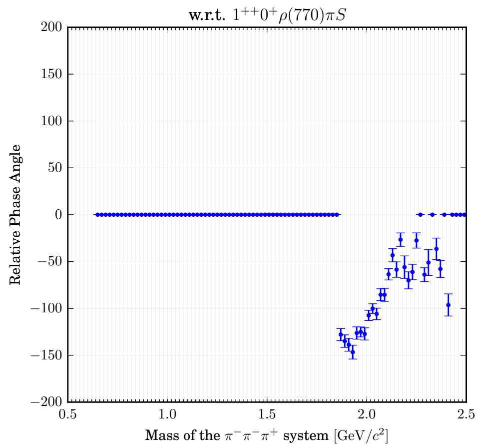


Figure 8.11: mass and t' plot of the $1^{++}1^{+}f_2(1270)\pi F$ wave for the refit with individual wave sets for each bin. The wave is found over all but the highest t'-bin.

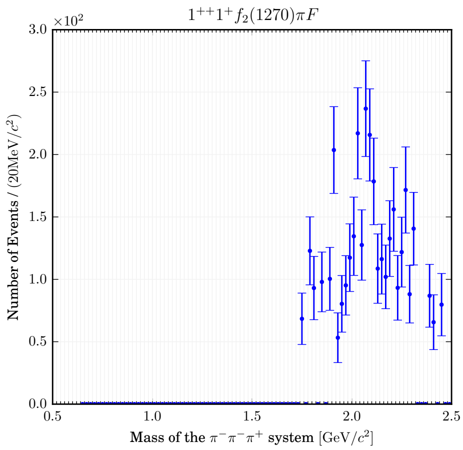
Additionally, the peak structure vanishes, leaving only a rise in intensity towards higher masses, shown in fig. 8.15c. Contributions to high-spin waves are expected as a result of the non-resonant Deck-Effect that tends to project into them. The flattening of the phase-motion and the broad peak support that this wave is most likely such a background contribution. This has already been speculated in Ref. [Dro15].



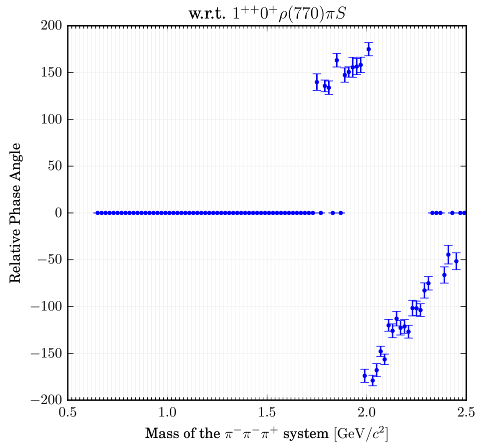
(a)



(b)



(c)



(d)

Figure 8.12: The $1^{++}1^{+}f_2(1270)\pi F$ wave shows a small, but stable peak over the t' range. (a) and (b) show the intensity and phase of the wave at the lowest t' bin. (c) and (d) show the same for the t' bin between $0.45\,(\mathrm{GeV}/c)^2$ and $0.72\,(\mathrm{GeV}/c)^2$. Both regions show a peak and a phase motion.

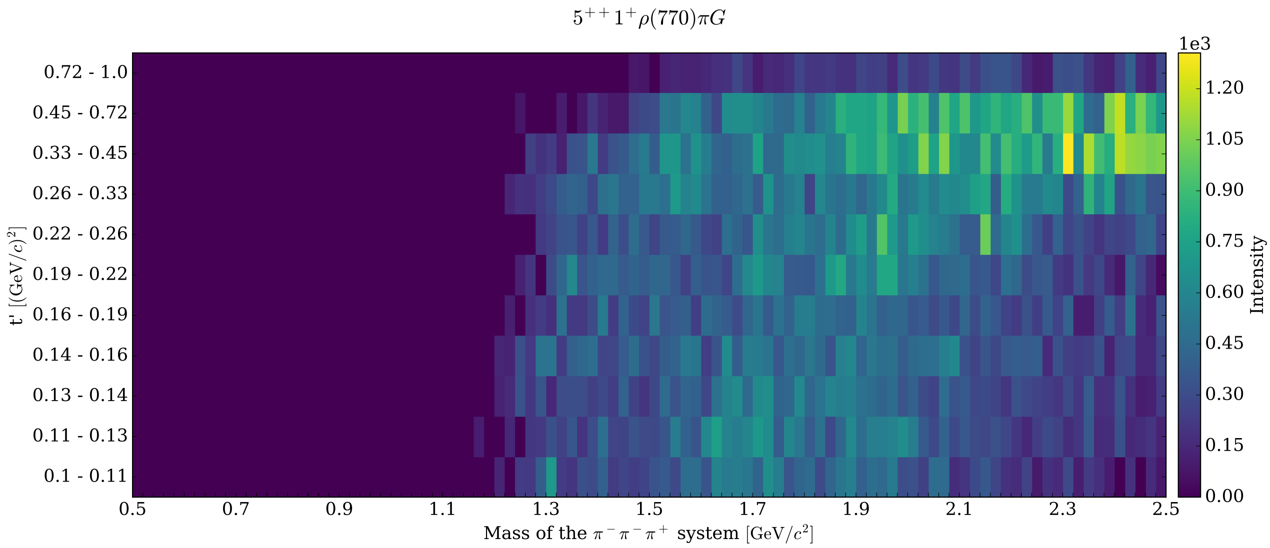


Figure 8.13: mass and t' plot of the $5^{++}1^{+}\rho(770)\pi G$ wave. The wave shows intensity over all t' bins.

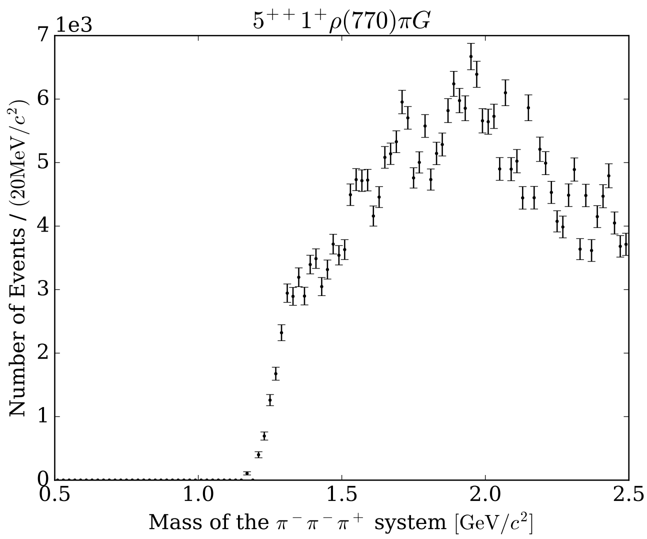


Figure 8.14: t'-summed intensity spectrum of the $5^{++}1^{+}\rho(770)\pi G$ wave. A broad structure, starting from $1.2\,\mathrm{GeV}/c^2$, is visible.

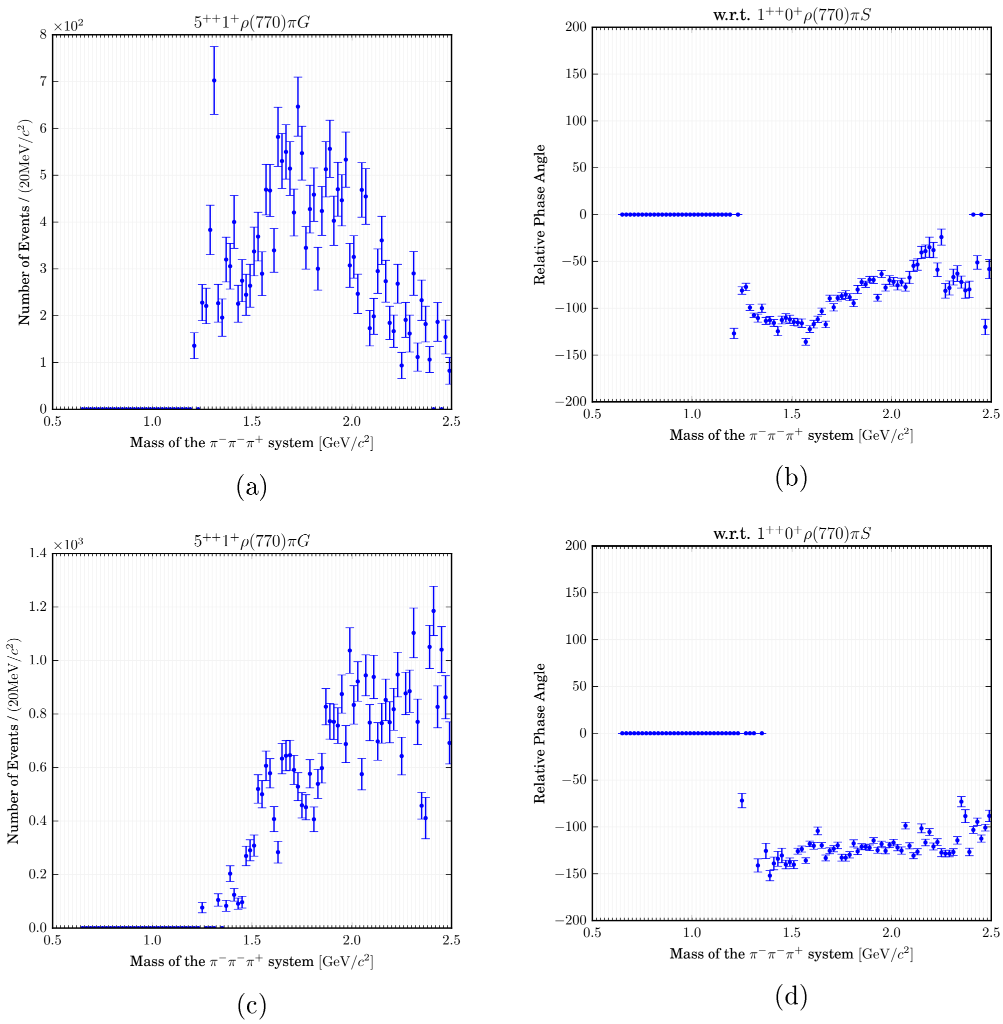


Figure 8.15: The $5^{++}1^{+}\rho(770)\pi G$ wave shows a peak in the low t' region and an increase in intensity for higher t'. (a) and (b) show the intensity and phase of the wave at the lowest t' bin. (c) and (d) show the same for the t' bin between $0.45\,(\mathrm{GeV}/c)^2$ and $0.72\,(\mathrm{GeV}/c)^2$.

8.5 Newly discovered Waves

The application of the weighted MBCM method also revealed some new and possibly interesting waves. Most of these waves are either of high spin or high spin projection and are also small as the high-intensity waves have mostly been found by hand-selection and a few during the application of the BCM method.

Two interesting waves are presented here. First the $2^{-+}2^{+}f_0(980)\pi D$, which was mostly found in the higher t'-regions, is shown in fig. 8.16 in the t'-resolved plot. The t'-summed plot is shown in fig. 8.18a. One can speculate whether this signal is a decay of the $\pi_2(1880)$. Further studies of this small signal are required.

An interesting structure can also be seen in the spin-exotic $3^{-+}2^{+}\rho(770)\pi F$ wave and partly in the $3^{-+}2^{+}f_2(1270)\pi D$ wave. The mass and t' spectrum of this wave is shown in fig. 8.17 and the t'-summed plot in fig. 8.18b.

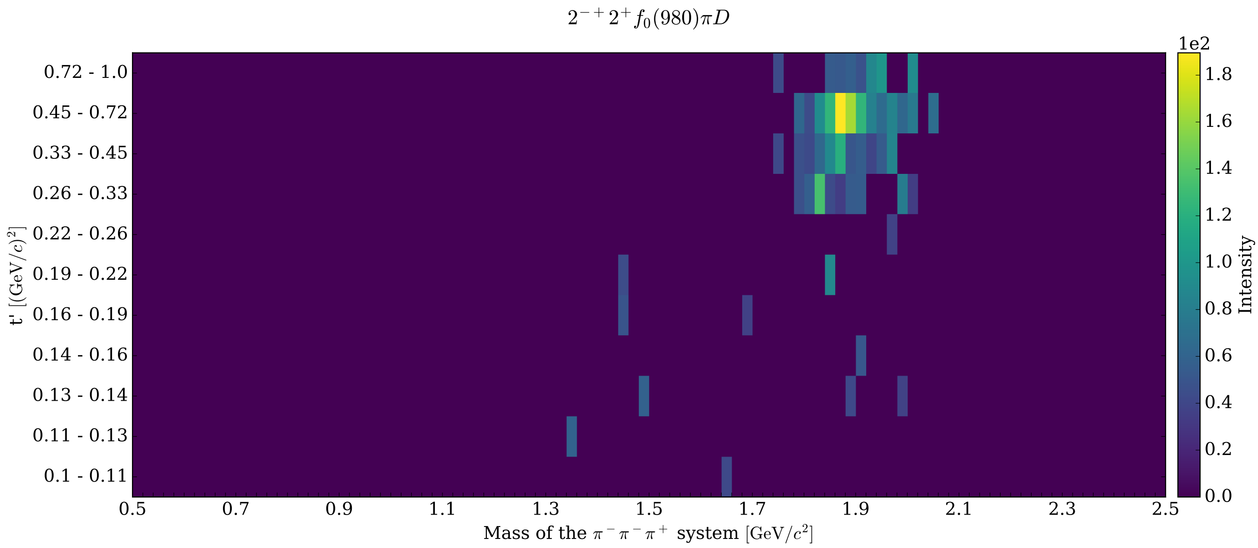


Figure 8.16: Mass and t' plot of the $2^{-+}2^{+}f_0(980)\pi D$ wave. For the higher t' bins the wave is visible as small intensity enhancement.

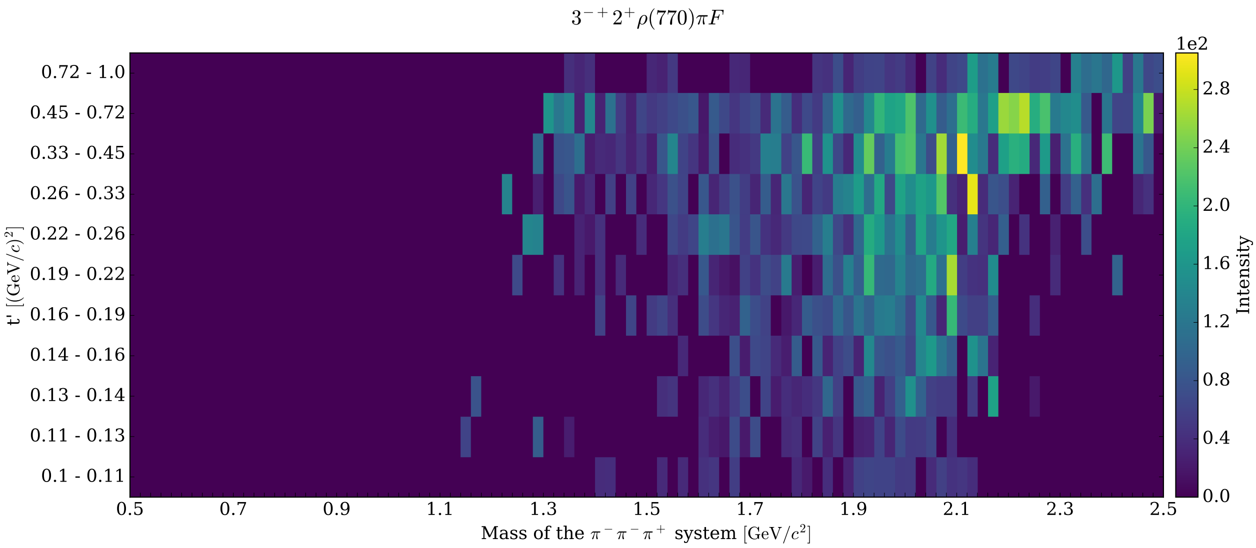


Figure 8.17: Mass and t' plot of the $3^{-+}2^{+}\rho(770)\pi F$ wave. This spin-exotic signal is small, but appears over a large mass and t' range.

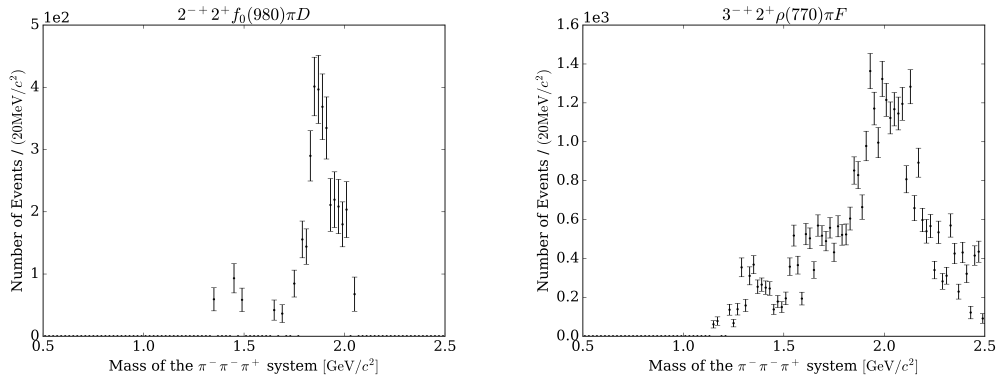


Figure 8.18: t'-summed intensity spectrum of two newly discovered waves. (a) the $2^{-+}2^{+}f_0(980)\pi D$ shows a small peak. (b) Spectrum of the exotic $3^{-+}2^{+}\rho(770)\pi F$ wave.

8.6 Creation of a Combined Wave Set

Any of the analyses so far have relied on a single universal wave set for all of the data. For reliable inference of the transition amplitudes the introduction of thresholds was necessary.

In Ref. [Dro15] the results obtained with the BCM method were very much unstable over the mass spectrum. In order to merge the wave sets that were found for the single mass bins and additional filtering of the sets had to be introduced. The requirement for the inclusion of a wave into the combined wave set was the stable appearance over at least ten bins in mass. The threshold was then set at the first stable appearance [Dro15].

The problem of this method is of course the possible removal of waves which are just suffering from strong cross-talk with other waves. In the worst case two waves can easily be swapped and may therefore both be flagged as unstable even though they are needed to describe the data.

The MBCM method largely suppresses such effects due to the hierarchy of the penalty.

Still some residual fluctuations persist. As they are typically small, a similar criterion on the stability over the bins can be invoked. Here instead of using just one bin in t' the complete dataset has been used. The thresholds were chosen at the first appearance over more than 5 bins in mass in the t'-summed intensity spectrum. This criterion is less restrictive than the 10 bins required in the previous analysis. Being less restrictive is possible as the weighted MBCM method fixes a lot of instabilities already.

8.6.1 Comparison of Wave Sets

Including every wave that was selected in any of the bins, in mass and t', results in a selection of 169 of the 236 waves of the reduced wave pool. The previous study performed in Ref. [Dro15] used the complete wave pool and was only applied on the lowest t'-bin. The BCM method recovered every wave in the wave pool [Dro15]. The simple thresholding algorithm and the inclusion of the thresholds used in the hand-selected wave set reduced the wave set to 118 waves.

The MBCM method does not require so much additional intervention. Thresholds are mostly found automatically and the refits of the single wave sets in each bin, presented in section 8.2, show stable inference.

The thresholding algorithm, requiring stable inference over more than 5 bins in mass, was only used to get a cleaner combined wave set. This resulted in a combined wave set of 134 waves with positive reflectivity, selected over all t'-bins.

Out of these 134 waves, 29 are neither in the hand-selected 88-wave set, nor in

$1^{++}1^{+}\rho_3(1690)\pi D$	$2^{-+}1^{+}f_0(1500)\pi D$	$2^{-+}2^{+}f_0(980)\pi D$	$2^{-+}2^{+}\rho(770)\pi F$
$2^{-+}2^{+}[\pi\pi]_S\pi D$	$3^{++}0^{+}f_0(980)\pi F$	$3^{++}0^{+}f_2(1270)\pi F$	$3^{++}0^{+}f_2(1270)\pi H$
$3^{++}2^{+}f_2(1270)\pi P$	$3^{++}2^{+}\rho_3(1690)\pi S$	$3^{++}2^{+}\rho(770)\pi D$	$3^{++}2^{+}\rho(770)\pi G$
$3^{++}2^{+}[\pi\pi]_S\pi F$	$3^{-+}2^{+}f_2(1270)\pi D$	$3^{-+}2^{+}\rho(770)\pi F$	$4^{++}1^{+}f_2(1270)\pi H$
$4^{-+}0^{+}\rho_3(1690)\pi F$	$4^{-+}1^{+}f_2(1270)\pi G$	$4^{-+}2^{+}f_2(1270)\pi D$	$4^{-+}2^{+}\rho(770)\pi F$
$4^{-+}2^{+}[\pi\pi]_S\pi G$	$5^{++}0^{+}f_0(980)\pi H$	$5^{++}2^{+}\rho(770)\pi G$	$5^{++}2^{+}[\pi\pi]_S\pi H$
$5^{-+}1^{+}\rho(770)\pi H$	$6^{-+}0^{+}f_0(980)\pi I$	$6^{-+}1^{+}f_0(980)\pi I$	$6^{-+}1^{+}\rho_3(1690)\pi F$
$6^{-+}2^{+}\rho(770)\pi H$			

Table 8.1: Newly found waves that were neither part of the hand-selected wave set nor of the one selected in Ref. [Dro15].

$1^{++}1^{+}f_0(1500)\pi P$	$1^{++}1^{-}[\pi\pi]_S\pi P$	$2^{++}1^{-}\rho(770)\pi D$	$2^{-+}1^{+}f_2(1270)\pi G$
$2^{-+}1^{-}[\pi\pi]_S\pi D$	$3^{++}1^{+}f_0(1500)\pi F$	$3^{++}1^{-}f_2(1270)\pi P$	$3^{++}1^{-}\rho_3(1690)\pi S$
$4^{++}1^{-}\rho(770)\pi G$	$4^{-+}0^{+}f_0(1500)\pi G$	$4^{-+}0^{+}f_2(1270)\pi I$	$4^{-+}1^{+}f_0(1500)\pi G$
$4^{-+}1^{-}[\pi\pi]_S\pi G$	$5^{++}0^{+}f_0(1500)\pi H$	$5^{++}0^{+}\rho_3(1690)\pi G$	$5^{++}1^{+}f_0(1500)\pi H$
$5^{++}1^{-}\rho(770)\pi I$	$6^{++}1^{-}\rho(770)\pi I$	$6^{-+}0^{+}f_0(1500)\pi I$	

Table 8.2: Waves that were removed from the wave set selected in Ref. [Dro15]. Waves with negative reflectivity were not included in the reduced wave pool in the first place.

the BCM-selected 118 wave set. Most of these newly found waves are small and have $M \geq 1$. It is expected that the extension of the model selection to higher t'-bins will find waves with higher spin-projections, as these are typically suppressed for lower t' in this reaction. The most interesting newly found waves were already discussed in section 8.5.

It is interesting to remark, that 19 waves of the 118-wave set are excluded by the weighted MBCM method. These include of course all 9 waves with negative reflectivity, as they were not included in the reduced wave pool. The residual 10 waves were either small or contained the $f_0(1500)$ isobar.

Compared to the hand-selected model, all but 3 waves with positive reflectivity have been found. These were typically all small. The negative waves were again of course excluded.

For the combined wave set the lowest threshold for any wave was at a mass of $0.64\,\mathrm{GeV}/c^2$. Below this threshold value, only the flat wave was considered.

8.6.2 Problems of a joined Wave Set

The creation of a single universal wave set over all t'-bins can be regarded controversial. There are two problems that may appear in a combined wave set. The first one is related to the t'-dependence of the waves, the second one is related to the simple way the combined wave set is created here.

In fig. 8.7 it was evident that the thresholds may show a t'-dependence for certain waves. A universal threshold is therefore either too high or too low in mass for all t'-bins. The result of this can be seen in the fit with the combined wave set. In fig. 8.19 the peak structure does not show a slow rise from 0 at the threshold anymore, while the fit with individual wave sets in fig. 8.2 was able to reproduce the peak without such a discontinuity. The same feature is visible in the two-dimensional representation in fig. 8.20.

The method of creating the wave set was chosen to be very simple, so the full two-dimensional information was not taken into account. Using only the t'-summed information may lead to the inclusion of waves that are not stable over all bins but show a continuous intensity in the summed plots. Such a case is shown in fig. 8.21. This problem can be resolved using a more advanced thresholding technique or by removing these waves manually. As these waves were small, no further effort has been made to remove them, opting for a larger wave set.

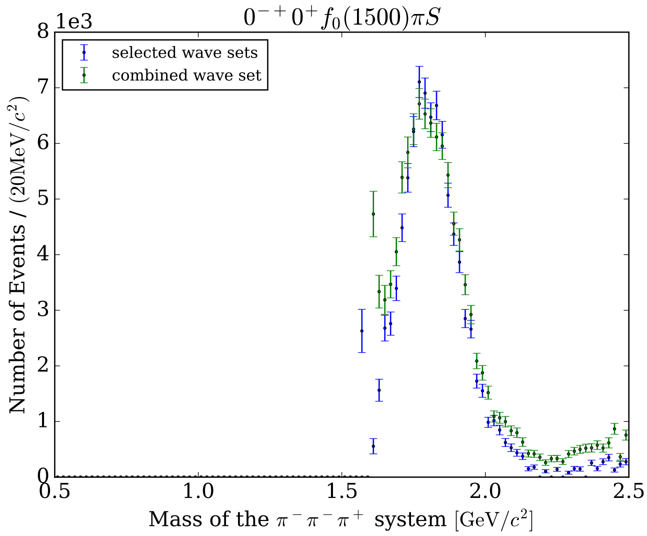


Figure 8.19: t'-summed intensity spectrum of the $0^{-+}0^{+}f_0(1500)\pi S$ wave. In blue the result of the refit with individual wave sets in each two-dimensional bin is shown, in green the refits of the combined wave set.

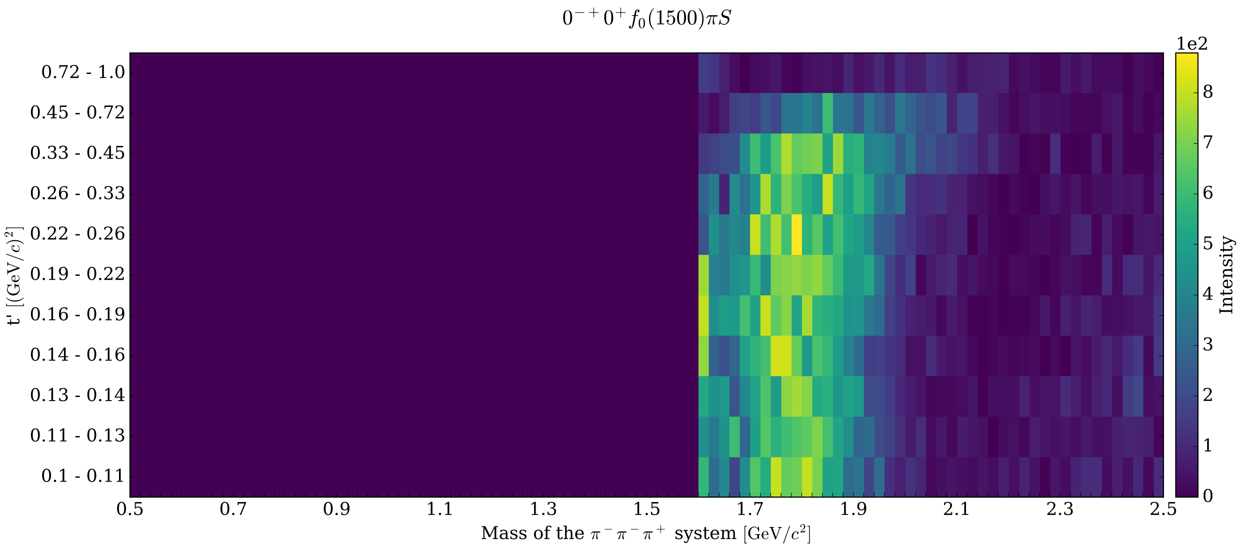


Figure 8.20: Mass and t' plot of the $0^{-+}0^{+}f_0(1500)\pi S$ wave for the refit with the combined wave set. In the combined wave set for all t' the same threshold has been used. The t' dependence, visible in fig. 8.4, is ignored.

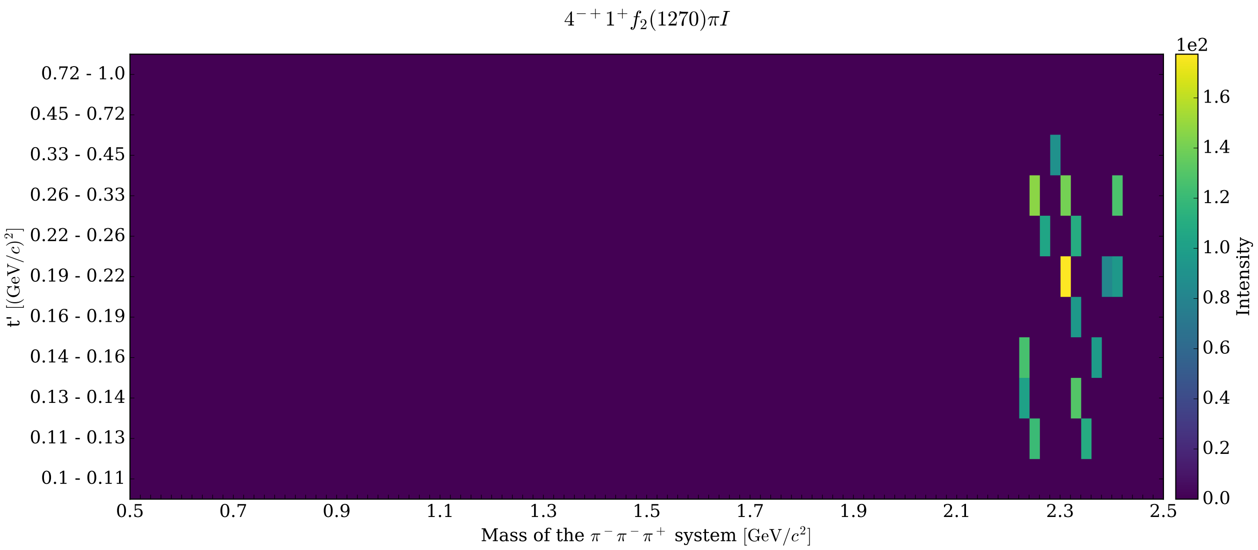


Figure 8.21: Example of a wave that could be excluded from the wave set by using the t'-resolved instead of the t'-summed data.

8.7 Discussion about Refits and Wave Sets

The above results show different behaviour for the intensity spectra of the waves. In previous analyses a simple wave set was used for the partial-wave decomposition. It was required to introduce thresholds in mass to guarantee reasonable results. The weighted MBCM method is able to find such thresholds automatically.
By applying the model selection procedure for each individual bin in mass and t' even more information is available. The thresholds can now also be interpreted as a function of t'. This is for example visible in the $0^{-+}0^{+}f_0(1500)\pi S$ wave. Also upper thresholds or even thresholding regions are possible.
This leads to the question whether it is desirable to have only a single wave set for the complete dataset.
The advantage is certainly the easy handling of such a simple approach. It is only required to find a reasonable lower threshold. For individual wave sets in each bin or at least in each t'-bin this may become much harder as probably every wave set should be reviewed by hand in order to fix possible 'holes' in the mass and t' spectrum.

One can also take this even further and ask if the refit should be applied at all. The reason why one would like to stick with the penalized fit as a final result can be justified from a Bayesian point of view. The penalized MLE is interpreted as a maximum a posteriori estimate. Usually point estimated are not desirable, but the curse of dimensionality makes it impossible to sample in such a high-dimensional setting.
As the only meaningful distribution for a Bayesian is the posterior, the mode estimate would also be the final result.
It is worth mentioning that the result of the penalized fit shows less discontinuities than the refit. This can, for example, be observed in fig. 8.1 where the intensity rises smoothly from 0 in the thresholding region. For the refit in fig. 8.6 the same wave shows a sudden jump to several ten thousands of events.
Continuity is expected in nature and therefore also for the intensity of the waves. This is a further argument for using the model selection fit directly.

Nevertheless there are also arguments in favour of fitting the selected model again without penalty. Currently the refit is a requirement for the analysis chain. The results of the partial-wave decomposition are the input for the resonance-parameter fit.
For this second step of the analysis Gaussian errors are assumed. These error estimates are obtained by approximating the Likelihood as a Gaussian and using the inverse of the Hessian. It has been shown in the works of Refs. [Bic16; Dro15] that the Gaussian approximation may not hold in all cases. It is not clear to what

extent this influences the error estimates.
It is expected that then inclusion of a penalty term would further break gaussianity. This has also been remarked by Ref. [Gue+15] for the LASSO fits. They suggest to either perform a refit as done in this thesis or, like they do themselves, use a bootstrap approach to estimate the errors.
A second reason is the bias introduced to the number of total expected events $\bar{N}$. By biasing the individual waves towards zero also $\bar{N}$ gets biased towards zero. As long as this results only in scaling of the intensities no effect on the resonance parameters should be introduced in the resonance parameter fit. However if one is interested in branching fraction this may have some negative effects.

The author of this thesis thinks that the weighted MBCM method is capable of producing results that are well suited for the current analysis chain. He suggests using the results in the following way. The weighted MBCM method provides a fast way of obtaining wave sets that are suitable for partial-wave decomposition. As each bin is fitted independently some waves may jump in and out of the wave sets in the different bins. The analyst should then gauge the relevance of each wave by looking at the full available information over the mass- and t'-bins and fix potential issues with the waves.
Large waves can then be included in certain two-dimensional regions. This approach does not ignore the t'-dependence of the waves and also allows upper thresholds. These individual, cleaned wave sets should then be used for the analysis.
The results of this method are then comparable with what has been done in previous analyses.

In the long run a more sophisticated method is proposed. The author believes that only the posterior should be used for inference. The weighted MBCM method already introduced different priors or penalties for different waves. The result was not only the correct derivation of thresholds, but also a speed-up in convergence of the fits. This leads to the hope that the individual priors for the waves may produce a simpler posterior that may make sampling of the posterior possible with the help of Hamilton Monte Carlo methods like NUTS [HG11].
The partial-wave decomposition could then make a first step towards a fully Bayesian analysis. Of course this is not an easy task and requires the restructuring of the complete analysis chain. It might well be that this is not even possible due to computational limitations. However if possible, the results obtained from sampling the posterior could provide a much better understanding of the errors and a better justification of the obtained values due to using the posterior.

Chapter 9

Conclusions and Outlook

The goal of this thesis was the development and verification of a procedure that is capable of systematically selecting a set of relevant model components in the context of partial-wave analysis. It aims at replacing less reproducible methods like selection of partial waves by hand, which was used to built the currently used 88-wave set. The method has been applied to both simulated and measured diffractive dissociation data for the $\pi^-\pi^-\pi^+$ final state.

The previously developed BCM method [Bic16; Dro15] showed promising results on measured data but was not ready for application due to instabilities in the selection process that required manual intervention. Continuing this previous work, at first the effect of the BCM method has been tested in the idealized case of simulated data for the $\pi^-\pi^-\pi^+$ final state. The behavior of the BCM method, as observed on measured data, could be reproduced on simulated data by decreasing its free scale-parameter from $\Gamma = 0.5$ to $\Gamma = 0.2$. The reason for this smaller parameter value could be related to the acceptance correction applied on measured data that acts as scaling of the intensities. Additionally a separation of the incoherent sectors could be reproduced on simulated data. Waves with negative reflectivity may be completely removed from the selected waves. The same preference has been observed on measured data. Furthermore, instabilities of the selection procedure were found at low 3π masses on simulated data. These instabilities are similar to the ones that appeared on measured data. The reason for the worse inference at low masses could be related to increasing similarity of some of the decay amplitudes that serve as basis functions for the statistical model. These similar waves can then interfere destructively, so that their intensities increase. The destructive interference then leads to overfitting of the model and multiple unstable solutions.

The LASSO method suggested in Ref. [Gue+15] has been transferred to our 3π analysis. The authors tested their method on simulated data. Applying it to simulated and measured data for our analysis showed similar model-selection properties as the BCM method, but additionally suppressed most of the aforementioned instabilities. The application of the LASSO method, however, requires the tuning of its free parameter as the selection is strongly sensitive to its choice. For this, the authors of Ref. [Gue+15] suggested the usage of information criteria like AIC [Aka74] and BIC [Sch78]. The application of these criteria showed that the parameter estimate provided by BIC comes closest to the desired wave set. However, at the same time a stronger bias on large-intensity waves was observed as compared to the BCM method.

The suppression of the instabilities, were related to the stronger bias of the LASSO on large intensities. A novel method, named MBCM, was developed that combines the advantageous properties of both the BCM and LASSO method. This new method introduces a stronger bias on large-intensity waves similar to the LASSO method. This regularization property can be adjusted with a free parameter that is loosely related to the parameter of the LASSO.
The idea is to perform the model selection similar to the BCM method and resolving ambiguities by the introduction of as much bias as needed, while keeping it as small as possible. On simulated data, all methods performed well in recovering the true underlying model.

The methods have then been applied to measured data. The MBCM and LASSO methods were able to resolve most of the instabilities that persisted when the BCM method was used. This could be especially seen in the 0^{-+} waves with the $f_0(1500)$, $f_0(980)$ and $[\pi\pi]_S$ isobars.

However some waves with the $f_0(1500)$ isobars were still found in the mass-region below 1.5 GeV, where the phase space for such decays is small. These sub-threshold decays are expected to be suppressed.
Instead of removing these waves manually and thereby treating them differently compared to the rest of the waves, the MBCM method was extended to include the phase-space information into the suppression parameter of the individual waves. Applying this final method to the data, wave sets are obtained that are not in contradiction with the physical intuition. The method is able to produce thresholds required for stable and sensible inference of the transition amplitudes. These thresholds are similar to the ones introduced by hand.

The promising results of the new weighted MBCM method motivated the application to the full available data. Making the first leap from a case study to the application in a realistic setting.
The extensions to the full dataset available recovers most of the waves from the old 88-wave set obtained form selection by hand. The newly discovered waves of the BCM method were mostly found as well. Two of the interesting signals, first discovered in the case-study of Ref. [Dro15], could be confirmed and automatic thresholds were found. The extension to the full t' range made it possible to have

a more detailed look at possible resonance content. The $1^{++}1^{+}f_2(1270)\pi F$ wave indeed shows peak-like enhancements and a rising phase motion over almost the complete t' range. This signal shows the power of the model-selection procedures over selection by hand. Additional new waves were found, which were neither in the 88-wave set nor in the BCM-selected one. Especially in the region of larger t' those new waves appeared. Some of them contain interesting signals like the $2^{-+}2^{+}f_0(980)\pi D$ or the spin-exotic $3^{-+}2^{+}\rho(770)\pi F$ wave. Both signals are small, but could be interesting for future studies.

From the results obtained on the full dataset a wave set was built in a simple way. This wave set automatically contains thresholds that are comparable with the ones introduced by hand. This wave set has then again been applied to the full available dataset.

With the new weighted MBCM method it is possible to select wave sets with thresholds in a systematic and fast way. The next step would be a resonance extraction of the newly produced results. Apart from that, the method enables a variety of future studies. The most obvious one being the extension of the wave pool to include waves with higher spins or spin projections. But also the effects of improved calculations of the decay amplitudes on the wave set, such as relativistic corrections or different isobar shapes, can be tested without time consuming and hardly reproducible studies.

Appendix A

88 Wave Set

Wave	Threshold [GeV/c^2]	Wave	Threshold [GeV/c^2]
$0^{-+}0^{+}f_0(1500)\pi S$	1.7	$0^{-+}0^{+}f_0(980)\pi S$	1.2
$0^{-+}0^{+}f_2(1270)\pi D$		$0^{-+}0^{+}\rho(770)\pi P$	
$0^{-+}0^{+}[\pi\pi]_S\pi S$			
$1^{++}0^{+}f_0(980)\pi P$	1.18	$1^{++}0^{+}f_2(1270)\pi P$	1.22
$1^{++}0^{+}f_2(1270)\pi F$		$1^{++}0^{+}\rho_3(1690)\pi D$	
$1^{++}0^{+}\rho_3(1690)\pi G$		$1^{++}0^{+}\rho(770)\pi S$	
$1^{++}0^{+}\rho(770)\pi D$		$1^{++}0^{+}[\pi\pi]_S\pi P$	
$1^{++}1^{+}f_0(980)\pi P$	1.14	$1^{++}1^{+}f_2(1270)\pi P$	
$1^{++}1^{+}\rho(770)\pi S$		$1^{++}1^{+}\rho(770)\pi D$	
$1^{++}1^{+}[\pi\pi]_S\pi P$	1.1		
$1^{++}1^{-}\rho(770)\pi S$			
$1^{-+}0^{-}\rho(770)\pi P$			
$1^{-+}1^{+}\rho(770)\pi P$			
$1^{-+}1^{-}\rho(770)\pi P$			
$2^{++}0^{-}f_2(1270)\pi P$	1.18	$2^{++}0^{-}\rho(770)\pi D$	
$2^{++}1^{+}f_2(1270)\pi P$	1.0	$2^{++}1^{+}\rho_3(1690)\pi D$	0.8
$2^{++}1^{+}\rho(770)\pi D$			
$2^{++}1^{-}f_2(1270)\pi P$	1.3		
$2^{++}2^{+}f_2(1270)\pi P$	1.4	$2^{++}2^{+}\rho(770)\pi D$	
$2^{-+}0^{+}f_0(980)\pi D$	1.16	$2^{-+}0^{+}f_2(1270)\pi S$	
$2^{-+}0^{+}f_2(1270)\pi D$		$2^{-+}0^{+}f_2(1270)\pi G$	
$2^{-+}0^{+}\rho_3(1690)\pi P$	1.0	$2^{-+}0^{+}\rho(770)\pi P$	
$2^{-+}0^{+}\rho(770)\pi F$		$2^{-+}0^{+}[\pi\pi]_S\pi D$	
$2^{-+}1^{+}f_2(1270)\pi S$	1.1	$2^{-+}1^{+}f_2(1270)\pi D$	
$2^{-+}1^{+}\rho_3(1690)\pi P$	1.3	$2^{-+}1^{+}\rho(770)\pi P$	
$2^{-+}1^{+}\rho(770)\pi F$		$2^{-+}1^{+}[\pi\pi]_S\pi D$	
$2^{-+}1^{-}f_2(1270)\pi S$			
$2^{-+}2^{+}f_2(1270)\pi S$		$2^{-+}2^{+}f_2(1270)\pi D$	
$2^{-+}2^{+}\rho(770)\pi P$			

Wave	**Threshold** [GeV/c^2]	**Wave**	**Threshold** [GeV/c^2]
$3^{++}0^{+}f_2(1270)\pi P$	0.96	$3^{++}0^{+}\rho_3(1690)\pi S$	1.38
$3^{++}0^{+}\rho_3(1690)\pi I$		$3^{++}0^{+}\rho(770)\pi D$	
$3^{++}0^{+}\rho(770)\pi G$		$3^{++}0^{+}[\pi\pi]_S\pi F$	
$3^{++}1^{+}f_2(1270)\pi P$	1.14	$3^{++}1^{+}\rho_3(1690)\pi S$	1.38
$3^{++}1^{+}\rho(770)\pi D$		$3^{++}1^{+}\rho(770)\pi G$	
$3^{++}1^{+}[\pi\pi]_S\pi F$			
$3^{-+}1^{+}f_2(1270)\pi D$	1.34	$3^{-+}1^{+}\rho(770)\pi F$	
$4^{++}1^{+}f_2(1270)\pi F$		$4^{++}1^{+}\rho_3(1690)\pi D$	1.7
$4^{++}1^{+}\rho(770)\pi G$			
$4^{++}2^{+}f_2(1270)\pi F$		$4^{++}2^{+}\rho(770)\pi G$	
$4^{-+}0^{+}f_2(1270)\pi D$		$4^{-+}0^{+}f_2(1270)\pi G$	1.6
$4^{-+}0^{+}\rho(770)\pi F$		$4^{-+}0^{+}[\pi\pi]_S\pi G$	1.4
$4^{-+}1^{+}f_2(1270)\pi D$		$4^{-+}1^{+}\rho(770)\pi F$	
$5^{++}0^{+}f_2(1270)\pi F$	0.98	$5^{++}0^{+}f_2(1270)\pi H$	
$5^{++}0^{+}\rho_3(1690)\pi D$	1.36	$5^{++}0^{+}\rho(770)\pi G$	
$5^{++}0^{+}[\pi\pi]_S\pi H$			
$5^{++}1^{+}f_2(1270)\pi F$		$5^{++}1^{+}[\pi\pi]_S\pi H$	
$6^{++}1^{+}f_2(1270)\pi H$		$6^{++}1^{+}\rho(770)\pi I$	
$6^{-+}0^{+}f_2(1270)\pi G$		$6^{-+}0^{+}\rho_3(1690)\pi F$	
$6^{-+}0^{+}\rho(770)\pi H$		$6^{-+}0^{+}[\pi\pi]_S\pi I$	
$6^{-+}1^{+}\rho(770)\pi H$		$6^{-+}1^{+}[\pi\pi]_S\pi I$	

Appendix B

Combined Wave Set

Wave	Threshold [GeV/c^2]	Wave	Threshold [GeV/c^2]
$0^{-+}0^{+}f_0(1500)\pi S$	1.6	$0^{-+}0^{+}f_0(980)\pi S$	1.1
$0^{-+}0^{+}f_2(1270)\pi D$	1.38	$0^{-+}0^{+}\rho_3(1690)\pi F$	2.38
$0^{-+}0^{+}\rho(770)\pi P$	0.64	$0^{-+}0^{+}[\pi\pi]_S\pi S$	0.64
$1^{++}0^{+}f_0(1500)\pi P$	1.6	$1^{++}0^{+}f_0(980)\pi P$	1.18
$1^{++}0^{+}f_2(1270)\pi P$	1.12	$1^{++}0^{+}f_2(1270)\pi F$	1.0
$1^{++}0^{+}\rho_3(1690)\pi D$	1.72	$1^{++}0^{+}\rho_3(1690)\pi G$	1.08
$1^{++}0^{+}\rho(770)\pi S$	0.64	$1^{++}0^{+}\rho(770)\pi D$	0.82
$1^{++}0^{+}[\pi\pi]_S\pi P$	0.64		
$1^{++}1^{+}f_0(980)\pi P$	1.26	$1^{++}1^{+}f_2(1270)\pi P$	1.06
$1^{++}1^{+}f_2(1270)\pi F$	1.58	$1^{++}1^{+}\rho_3(1690)\pi D$	2.0
$1^{++}1^{+}\rho(770)\pi S$	0.66	$1^{++}1^{+}\rho(770)\pi D$	0.86
$1^{++}1^{+}[\pi\pi]_S\pi P$	0.64		
$1^{-+}1^{+}f_2(1270)\pi D$	1.68	$1^{-+}1^{+}\rho(770)\pi P$	0.82
$2^{++}1^{+}f_2(1270)\pi P$	1.1	$2^{++}1^{+}f_2(1270)\pi F$	1.26
$2^{++}1^{+}\rho(770)\pi D$	0.7		
$2^{++}2^{+}f_2(1270)\pi P$	1.42	$2^{++}2^{+}\rho(770)\pi D$	0.94
$2^{-+}0^{+}f_0(1500)\pi D$	2.16	$2^{-+}0^{+}f_0(980)\pi D$	1.32
$2^{-+}0^{+}f_2(1270)\pi S$	1.16	$2^{-+}0^{+}f_2(1270)\pi D$	1.52
$2^{-+}0^{+}f_2(1270)\pi G$	1.78	$2^{-+}0^{+}\rho_3(1690)\pi P$	1.62
$2^{-+}0^{+}\rho_3(1690)\pi F$	2.32	$2^{-+}0^{+}\rho(770)\pi P$	0.76
$2^{-+}0^{+}\rho(770)\pi F$	1.14	$2^{-+}0^{+}[\pi\pi]_S\pi D$	0.64
$2^{-+}1^{+}f_0(1500)\pi D$	2.02	$2^{-+}1^{+}f_0(980)\pi D$	1.32
$2^{-+}1^{+}f_2(1270)\pi S$	1.26	$2^{-+}1^{+}f_2(1270)\pi D$	1.42
$2^{-+}1^{+}\rho_3(1690)\pi P$	1.84	$2^{-+}1^{+}\rho(770)\pi P$	0.64
$2^{-+}1^{+}\rho(770)\pi F$	1.02	$2^{-+}1^{+}[\pi\pi]_S\pi D$	0.66
$2^{-+}2^{+}f_0(980)\pi D$	1.78	$2^{-+}2^{+}f_2(1270)\pi S$	1.32
$2^{-+}2^{+}f_2(1270)\pi D$	1.54	$2^{-+}2^{+}\rho(770)\pi P$	0.8
$2^{-+}2^{+}\rho(770)\pi F$	1.08	$2^{-+}2^{+}[\pi\pi]_S\pi D$	1.56
$3^{++}0^{+}f_0(980)\pi F$	1.98	$3^{++}0^{+}f_2(1270)\pi P$	1.22

Wave	**Threshold** [GeV/c^2]	**Wave**	**Threshold** [GeV/c^2]
$3^{++}0^+f_2(1270)\pi F$	2.1	$3^{++}0^+f_2(1270)\pi H$	2.26
$3^{++}0^+\rho_3(1690)\pi S$	1.7	$3^{++}0^+\rho_3(1690)\pi D$	1.82
$3^{++}0^+\rho(770)\pi D$	0.98	$3^{++}0^+\rho(770)\pi G$	1.42
$3^{++}0^+[\pi\pi]_S\pi F$	1.08		
$3^{++}1^+f_0(980)\pi F$	2.3	$3^{++}1^+f_2(1270)\pi P$	1.12
$3^{++}1^+f_2(1270)\pi F$	2.06	$3^{++}1^+\rho_3(1690)\pi S$	1.52
$3^{++}1^+\rho(770)\pi D$	0.88	$3^{++}1^+\rho(770)\pi G$	1.34
$3^{++}1^+[\pi\pi]_S\pi F$	1.0		
$3^{++}2^+f_2(1270)\pi P$	1.58	$3^{++}2^+\rho_3(1690)\pi S$	0.98
$3^{++}2^+\rho(770)\pi D$	0.94	$3^{++}2^+\rho(770)\pi G$	1.46
$3^{++}2^+[\pi\pi]_S\pi F$	1.14		
$3^{-+}1^+f_2(1270)\pi D$	1.66	$3^{-+}1^+f_2(1270)\pi G$	2.08
$3^{-+}1^+\rho(770)\pi F$	1.18		
$3^{-+}2^+f_2(1270)\pi D$	1.56	$3^{-+}2^+\rho(770)\pi F$	1.22
$4^{++}1^+f_2(1270)\pi F$	1.6	$4^{++}1^+f_2(1270)\pi H$	2.24
$4^{++}1^+\rho_3(1690)\pi D$	1.92	$4^{++}1^+\rho(770)\pi G$	1.12
$4^{++}2^+f_2(1270)\pi F$	1.98	$4^{++}2^+\rho(770)\pi G$	1.54
$4^{-+}0^+f_0(980)\pi G$	1.72	$4^{-+}0^+f_2(1270)\pi D$	1.3
$4^{-+}0^+f_2(1270)\pi G$	2.12	$4^{-+}0^+\rho_3(1690)\pi P$	1.8
$4^{-+}0^+\rho_3(1690)\pi F$	2.38	$4^{-+}0^+\rho(770)\pi F$	0.9
$4^{-+}0^+\rho(770)\pi H$	1.54	$4^{-+}0^+[\pi\pi]_S\pi G$	0.96
$4^{-+}1^+f_2(1270)\pi D$	0.8	$4^{-+}1^+f_2(1270)\pi G$	2.02
$4^{-+}1^+f_2(1270)\pi I$	2.22	$4^{-+}1^+\rho_3(1690)\pi P$	0.96
$4^{-+}1^+\rho(770)\pi F$	0.9	$4^{-+}1^+\rho(770)\pi H$	1.58
$4^{-+}1^+[\pi\pi]_S\pi G$	1.62		
$4^{-+}2^+f_2(1270)\pi D$	1.06	$4^{-+}2^+\rho(770)\pi F$	1.1
$4^{-+}2^+[\pi\pi]_S\pi G$	2.14		
$5^{++}0^+f_0(980)\pi H$	1.46	$5^{++}0^+f_2(1270)\pi F$	1.62
$5^{++}0^+\rho_3(1690)\pi D$	2.14	$5^{++}0^+\rho(770)\pi G$	1.06
$5^{++}0^+\rho(770)\pi I$	1.62	$5^{++}0^+[\pi\pi]_S\pi H$	1.46
$5^{++}1^+f_0(980)\pi H$	2.32	$5^{++}1^+f_2(1270)\pi F$	1.54
$5^{++}1^+f_2(1270)\pi H$	2.3	$5^{++}1^+\rho_3(1690)\pi D$	2.06
$5^{++}1^+\rho(770)\pi G$	1.2	$5^{++}1^+\rho(770)\pi I$	1.74
$5^{++}1^+[\pi\pi]_S\pi H$	1.62		
$5^{++}2^+\rho(770)\pi G$	1.54	$5^{++}2^+[\pi\pi]_S\pi H$	2.2
$5^{-+}1^+\rho(770)\pi H$	2.38		
$6^{++}1^+f_2(1270)\pi H$	2.16	$6^{++}1^+\rho(770)\pi I$	1.84
$6^{-+}0^+f_0(980)\pi I$	1.22	$6^{-+}0^+f_2(1270)\pi G$	1.74
$6^{-+}0^+\rho_3(1690)\pi F$	1.18	$6^{-+}0^+\rho(770)\pi H$	1.16

Wave	**Threshold** [GeV/c^2]	**Wave**	**Threshold** [GeV/c^2]
$6^{-+}0^{+}[\pi\pi]_S\pi I$	1.22		
$6^{-+}1^{+}f_0(1500)\pi I$	1.76	$6^{-+}1^{+}f_0(980)\pi I$	1.54
$6^{-+}1^{+}f_2(1270)\pi G$	1.74	$6^{-+}1^{+}\rho_3(1690)\pi F$	2.26
$6^{-+}1^{+}\rho(770)\pi H$	1.38	$6^{-+}1^{+}[\pi\pi]_S\pi I$	1.78
$6^{-+}2^{+}\rho(770)\pi H$	1.92		

Bibliography

[Abb+15] P. Abbon et al. »The COMPASS setup for physics with hadron beams«. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 779 (Apr. 2015), pp. 69–115. DOI: `10.1016/j.nima.2015.01.035` (cited on pages 5, 6, 11).

[Ado+17] C. Adolph et al. »Resonance production and $\pi\pi$ S-wave in $\pi^- + p \to \pi^-\pi^-\pi^+ + p_{\mathrm{recoil}}$ at 190 GeV/c«. In: *Phys. Rev. D* 95 (3 Feb. 2017), p. 032004. DOI: `10.1103/PhysRevD.95.032004`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.95.032004` (cited on pages 5, 7, 8, 10–13, 15–17, 20–22, 27, 29, 89, 115).

[Aka74] H. Akaike. »A new look at the statistical model identification«. In: *IEEE Transactions on Automatic Control* 19.6 (Dec. 1974), pp. 716–723. ISSN: 0018-9286. DOI: `10.1109/TAC.1974.1100705` (cited on pages 35, 132).

[Bic16] Karl Bicker. »Model Selection for and Partial-Wave Analysis of a Five-Pion Final State at the COMPASS Experiment at CERN«. PhD Thesis. Technische Universität München, 19th Apr. 2016 (cited on pages 2, 5, 13, 27, 28, 32, 54, 128, 131).

[Box76] George E. P. Box. »Science and Statistics«. In: *Journal of the American Statistical Association* 71.356 (Dec. 1976), pp. 791–799. DOI: `10.1080/01621459.1976.10480949` (cited on page 43).

[CD08] François Caron and Arnaud Doucet. »Sparse Bayesian nonparametric regression«. In: *Proceedings of the 25th international conference on Machine learning - ICML '08*. ACM Press, 2008. DOI: `10.1145/1390156.1390168` (cited on page 39).

[Dec64] Robert T. Deck. »Kinematical Interpretation of the First $\pi - p$ Resonance«. In: *Phys. Rev. Lett.* 13 (5 Aug. 1964), pp. 169–173. DOI: `10.1103/PhysRevLett.13.169`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.13.169` (cited on page 9).

[Dro+15] Michael Droettboom et al. *matplotlib: matplotlib v1.5.0*. Oct. 2015. DOI: `10.5281/zenodo.32914`. URL: `https://doi.org/10.5281/zenodo.32914`.

[Dro15] Oliver Drotleff. »Model Selection of $\pi^- + p \to \pi^-\pi^+\pi^- + p$ at the COMPASS Experiment at CERN«. Diploma Thesis. Technische Universität München, 2nd Nov. 2015 (cited on pages 2, 5, 20, 28, 30, 32, 44, 53, 55, 75–77, 87–89, 91, 93, 107, 116–118, 124, 125, 128, 131, 132).

[Gel08] Andrew Gelman. *Some thoughts on the saying, "All models are wrong, but some are useful"*. 12th June 2008. URL: `http://andrewgelman.com/2008/06/12/all_models_are/` (visited on 10/04/2017) (cited on page 43).

[Gue+15] Baptiste Guegan et al. »Model selection for amplitude analysis«. In: *JINST* 10.09 (2015), P09002. DOI: `10.1088/1748-0221/10/09/P09002`. arXiv: `1505.05133 [physics.data-an]` (cited on pages 2, 28, 33–36, 95, 129, 131, 132).

[Haa13] Florian Haas. »Two-Dimensional Partial-Wave Analysis of Exclusive 190 GeV π^-p Scattering into the $\pi^-\pi^-\pi^+$ Final State at COMPASS (CERN)«. PhD Thesis. Technische Universität München, 17th Dec. 2013 (cited on pages 2, 5, 7–11, 13, 18, 27, 29, 30, 44, 87–89, 91, 111, 115).

[HG11] Matthew D. Hoffman and Andrew Gelman. »The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo«. In: (18th Nov. 2011). arXiv: `1111.4246v1 [stat.CO]` (cited on page 129).

[LN89] Dong C. Liu and Jorge Nocedal. »On the limited memory BFGS method for large scale optimization«. In: *Mathematical Programming* 45.1-3 (Aug. 1989), pp. 503–528. DOI: `10.1007/bf01589116`.

[Neu12] Sebastian Neubert. »First Amplitude Analysis of Resonant Structures in the 5-Pion Continuum at COMPASS«. PhD Thesis. Technische Universität München, 5th June 2012 (cited on page 27).

[Noc80] Jorge Nocedal. »Updating quasi-Newton matrices with limited storage«. In: *Mathematics of Computation* 35.151 (Sept. 1980), pp. 773–773. DOI: `10.1090/s0025-5718-1980-0572855-7`.

[Oli+14] K. A. Olive et al. »Review of Particle Physics«. In: *Chin. Phys.* C38 (2014), p. 090001. DOI: `10.1088/1674-1137/38/9/090001` (cited on page 117).

[OS09] R. B. O'Hara and M. J. Sillanpää. »A review of Bayesian variable selection methods: what, how and which«. In: *Bayesian Analysis* 4.1 (Mar. 2009), pp. 85–117. DOI: `10.1214/09-ba403` (cited on page 32).

[PC08] Trevor Park and George Casella. »The Bayesian Lasso«. In: *Journal of the American Statistical Association* 103.482 (2008), pp. 681–686. DOI: 10.1198/016214508000000337. eprint: http://dx.doi.org/10.1198/016214508000000337. URL: http://dx.doi.org/10.1198/016214508000000337 (cited on pages 32, 36).

[Sch14] Stephan Schmeing. »Resonance Extraction in Diffractive 3π Production using 190 GeV/c π^- at the COMPASS Experiment (CERN)«. Master Thesis. Technische Universität München, 2014 (cited on page 27).

[Sch78] Gideon Schwarz. »Estimating the Dimension of a Model«. In: *Ann. Statist.* 6.2 (Mar. 1978), pp. 461–464. DOI: 10.1214/aos/1176344136. URL: http://dx.doi.org/10.1214/aos/1176344136 (cited on pages 36, 132).

[Tib96] Robert Tibshirani. »Regression Shrinkage and Selection via the Lasso«. In: *Journal of the Royal Statistical Society. Series B (Methodological)* 58.1 (1996), pp. 267–288. ISSN: 00359246. URL: http://www.jstor.org/stable/2346178 (cited on pages 2, 28, 32, 34).

[Uhl16] Sebastian Uhl. »Photon Reconstruction and Partial-Wave Analysis of Three-Body Final States with Neutral Particles at COMPASS)«. PhD Thesis. Technische Universität München, 2016 (cited on pages 27, 55).

[Wal15] Stefan Wallner. »Extraction of Resonance Parameters of Light Meson Resonances in the Charged Three-Pion Final State at the COMPASS Experiment (CERN)«. Master Thesis. Technische Universität München, 7th Oct. 2015 (cited on page 27).

List of Figures

List of Tables

Acknowledgments

First of all I would like to thank Prof. Dr. Stephan Paul for giving me the opportunity to work on this thesis.

I would also like to thank Boris Grube who was always very inspiring with his passion for physics. I learned a great deal during our discussions.

Furthermore my thanks go to Karl Bicker and Oliver Drotleff who made the first steps toward this type model selection analysis and provided the starting point for my work.

Big thanks also to my colleagues at E18. The atmosphere at our chair is always kind and I enjoyed my time a lot.

I am very grateful for my good friends, my family and of course my wonderful girlfriend who always supported me during stressful times.

I acknowledge the support by the DFG Cluster of Excellence "Origin and Structure of the Universe". Much of the analysis has been carried out on the computing facilities of the Computational Center for Particle and Astrophysics (C2PAP).