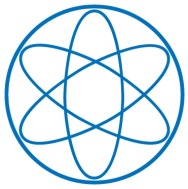
PHYSIK-DEPARTMENT



Model Selection for and Partial-Wave Analysis of a Five-Pion Final State at the COMPASS Experiment at CERN

Dissertation von Karl Alexander Bicker

TECHNISCHE UNIVERSITÄT MÜNCHEN

Technische Universität München
Physik-Department E18

Model Selection for and Partial-Wave Analysis of a Five-Pion Final State at the COMPASS Experiment at CERN

Karl Alexander Bicker

Vollständiger Abdruck der von der Fakultät für Physik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzende(r): Univ.-Prof. Dr. Andreas Weiler

Prüfer der Dissertation:

1. Univ.-Prof. Dr. Stephan Paul
2. Hon.-Prof. Allen C. Caldwell, Ph.D.

Die Dissertation wurde am 19.04.2016 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 08.08.2016 angenommen.

Abstract

The light-meson spectrum is an important ingredient to understand quantum chromodynamics, the theory of strong interaction at low energies, where quarks and gluons are confined into hadrons. However, measuring this spectrum experimentally is very challenging, due to the large number of overlapping resonances it contains. To disentangle this complicated spectrum, partial-wave techniques are used. At the COMPASS experiment at CERN, the light-meson spectrum is studied in diffractive dissociation reactions. One such reaction, $\pi^- + p \to \pi^-\pi^+\pi^-\pi^+\pi^- + p$, puts the analysis methodology to the test, because the large number of final-state particles requires a dedicated model-selection procedure and introduces a complicated background situation. In this thesis, such a model-selection procedure is developed and verified on simulated events. In addition to finding a suitable model, it can be used to assess the reliability of the results. The model-selection procedure is then successfully applied to data from the COMPASS 2008 data-taking campaign. A partial-wave decomposition is performed using the selected model. The results are compared to partial-wave decompositions of other final states and good agreement is found.

Kurzfassung

Das Spektrum der leichten Mesonen ist wichtig um die Quantenchromodynamik, das heißt die Theorie der starken Wechselwirkung, bei niedrigen Energien zu verstehen, wo Quarks und Gluonen in Hadronen gebunden sind. Durch die große Zahl überlappender Resonanzen ist die Messung dieses Spektrums experimentell sehr anspruchsvoll. Partialwellenzerlegungen werden verwendet, um dieses komplizierte Spektrum zu entflechten. Das COMPASS Experiment am CERN untersucht das Spektrum der leichten Mesonen in diffraktiver Dissoziation. Eine dieser Reaktionen, $\pi^- + p \to \pi^-\pi^+\pi^-\pi^+\pi^- + p$, stellt hohe Anforderung an die Analysemethoden, da wegen der großen Zahl an Endzustandsteilchen sowohl der Untergrund kompliziert also auch eine dedizierte Prozedur zur Selektion eines Modells nötig ist. In dieser Arbeit wird eine solche Modellauswahlprozedur entwickelt und auf simulierten Ereignissen getestet. Sie findet nicht nur ein geeignetes Modell, sondern kann auch benutzt werden, um die Zuverlässigkeit der Resultate zu beurteilen. Die Modellauswahl wird erfolgreich für die Daten aus der COMPASS Messkampagne von 2008 eingesetzt. Die Resultate der Partialwellenzerlegung mit dem gefundenen Modell werden mit Partialwellenzerlegungen anderer Endzustände verglichen, wobei eine gute Übereinstimmung festgestellt werden kann.

Contents

CHAPTER 1

Introduction

"[...] you will not go to space today."

Randall Munroe,
http://xkcd.com/1133/

Hadron spectroscopy is the study of the excitation spectrum of bound states of the strong interaction, i.e. the measurement of the masses and decays of hadrons. A branch of hadron spectroscopy, the light-meson spectroscopy, focuses on states made up by quark-antiquark pairs containing u, d and s quarks.

Quantum chromodynamics, abbreviated QCD, is the theory of strong interaction. As such, it should in principle be able to predict the spectrum of hadronic states which can be compared to experimental results. However, the running of the coupling constant α_s of QCD prevents expansions in α_s from converging at the energy scales probed by light-meson spectroscopy. One ab-initio approach to solving QCD, which is not hampered by this problem, is lattice QCD. It uses a discretized grid in space and time to make predictions with numerical Monte Carlo methods. Although in lattice QCD, quark masses are typically chosen heavier than those observed in nature and bound states are assumed to be quasi-stable, it is still possible to make qualitative predictions of the light-meson spectrum from first principles [1]. The main limiting factor are the extraordinary computational resources which are required to perform the calculations. Considering Moore's law [2], it is reasonable to assume that the prediction of a quantitative spectrum of light mesons from first principles is achievable in the near future with lattice QCD.

While the light-meson sector has been under study experimentally for a long time [3], due to the comparatively low energy scales involved, the situation remains challenging to this day. The spectrum is characterized by a large number of broad strongly-overlapping resonances and, in addition, the systems under study are quantum-mechanical in nature, which means interference effects have to be taken into account. Advanced analysis techniques are required to extract the light-meson spectrum from the kinematic distributions measured experimentally. In most cases, some sort of model and prior knowledge are needed for the analysis. In addition, the large data sets required for precise measurements demand considerable computing resources. In this light, it is not surprising that the light-meson spectrum is in parts still heavily debated, with many resonances needing confirmation, in particular in the mass region of $2.0\,\mathrm{GeV}/c^2$ and above [4].

Diffractive dissociation reactions, described in more detail in section 3.1, are one possible way to study the light-meson spectrum. These reactions, among others, have been measured in the 2004 and 2008 data-taking campaigns of the COMPASS experiment, during which large data samples were collected. Due to the large cross section and the very good acceptance for charged particle tracks, the dominant channel is $\pi^- + p \to \pi^-\pi^+\pi^- + p$. From this reaction, resonance parameters can be extracted with high precision and several recent publications based on the 2008 data set focused on it [5–9].

Using the same 2008 data set, the topic of this work is the related reaction $\pi^- + p \to \pi^-\pi^+\pi^-\pi^+\pi^- + p$. It allows to access areas of the spectrum which lie at higher masses, where a number of higher excitations of light mesons are expected. These are particularly challenging to measure, because many reactions are not able to access these masses with large cross sections. The reactions that do probe the high-mass region generally involve complex multi-body final states, which, as will be shown below, pose a number of problems on their own. At the same time, heavier mesons generally have a larger width, aggravating the problem of overlapping resonances, putting the analysis methods currently available to the test and requiring large data samples. Additional insight can be gained by combining the measurements of different decay modes. In this context, the five-pion reaction can make an important contribution to the understanding of the high-mass part of the light-meson iso-vector spectrum.

However, the analysis of final states with many particles poses a challenge in itself. The final-state particles are usually arranged into subsystems, which is described in section 3.3.1. Large numbers of final-state particles lead to large numbers of possible subsystems. For many of the subsystems, imposing constraints is difficult because there is only little knowledge about them. Due to the combinatorics of the many possible subsystems and the lack of constraints on them, data-driven model selection is a prerequisite for extracting resonances and their parameters. Such a model-selection procedure was part of an earlier analysis of the five-pion channel, based on the COMPASS data set from 2004 [10]. However, several open questions remained in that work and for the analysis of the significantly larger 2008 data set, it was deemed important to revisit and thoroughly test the model-selection procedure. As a consequence, this work focuses heavily on the development and verification of analysis methodology.

After a short overview of the experimental setup, chapter 2 will discuss the selection of an exclusive five-pion sample from the 2008 COMPASS data set and present kinematic distributions. In chapter 3, the basic principles of the analysis method are laid out. Chapter 4 deals with the model-selection procedure. It is split into two main parts, the first one showing Monte Carlo studies of the model-selection procedure suggested in [10] and the second one introducing an alternative approach to model selection. For the latter, a number of studies are presented. The chapter concludes with a discussion of the model which has been selected for the description of the measured data set. With a model selected, chapter 5 discusses the results of the analysis in terms of their physical interpretation. Finally, chapter 6 summarizes the results and presents the conclusions and the outlook.

CHAPTER 2

Experimental Setup and Event Selection

> "How about instead of epigraphs you just put a bunch of emoticons?"
>
> Dr. Daniel Greenwald

This work is based on data measured by the COmmon Muon and Proton Apparatus for Structure and Spectroscopy, abbreviated to COMPASS, which is a fixed-target experiment located at CERN. A brief overview of its experimental setup is given in the first part of this chapter. The second part discusses the selection of exclusive events with 5 charged pions in the final state.

2.1 The COMPASS Experiment

The COMPASS experiment has been described in great detail in [11] and [12]. Therefore, only a very brief description of the components relevant to this analysis will be included here. A schematic view of the spectrometer setup is shown in figure 2.1.

2.1.1 Beam and Target

For the dataset analyzed here, COMPASS was supplied with a negatively-charged secondary hadron beam with an energy of 190 GeV from CERN's Super Proton Synchrotron, SPS. The beam consists mainly of pions (96.8 %) with a small admixture of kaons (2.4 %) and anti-protons (0.8 %). The beam impinges on a liquid-hydrogen target with a length of 40 cm, corresponding to 5.5 % of one nuclear interaction length. The target is surrounded by a recoil-proton detector, RPD, which consists of two concentric barrels of scintillator slabs.

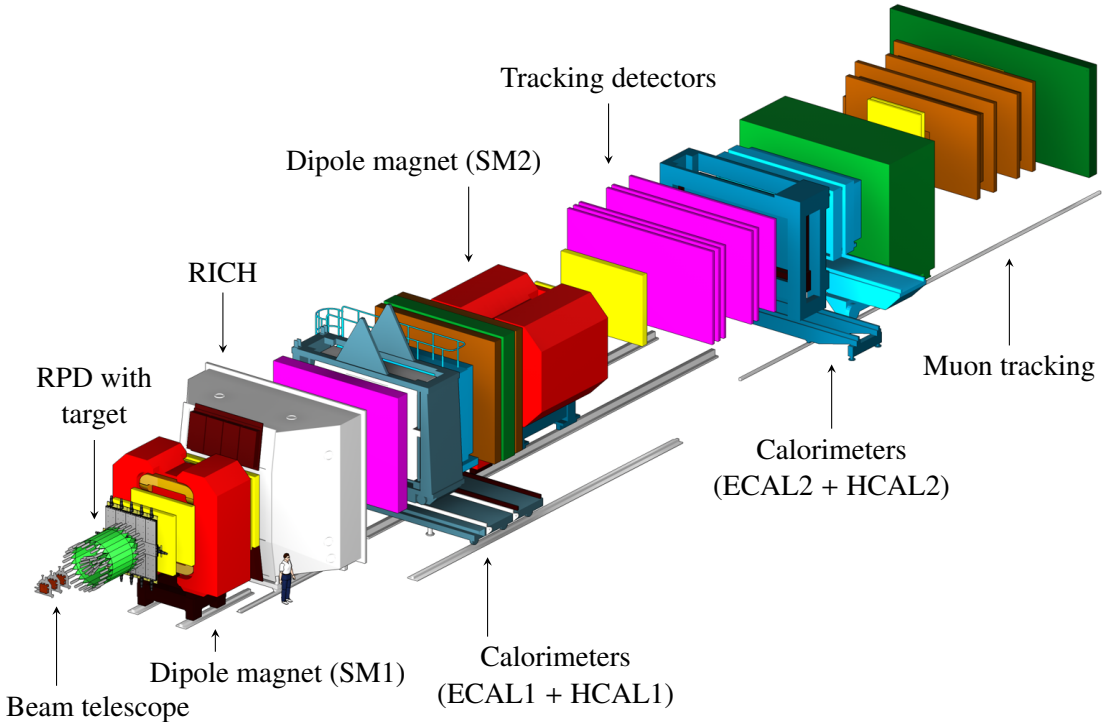


Figure 2.1: Schematic view of the COMPASS spectrometer setup in 2008.

2.1.2 Tracking System and Particle Identification

For beam definition, a combination of silicon-microstrip and scintillating-fiber detectors are used. Behind the target, four silicon-microstrip detector modules constitute the vertex detector. After the target region, a two-stage magnetic spectrometer measures charged final-state particles. Apart from two spectrometer magnets, it is made up from several types of gas detectors. GEM and Micromega detectors measure small angles while multi-wire proportional and drift chambers cover medium and large-angle tracking. In the first stage of the spectrometer, a ring-imaging Cherenkov detector, RICH, provides pion and kaon separation in the momentum range of about 9 to 40 GeV/c.

2.1.3 Trigger

A schematic view of the trigger system used for this analysis is shown in figure 2.2. Two detectors, the scintillating-fiber detector SciFi 1 and the Beam Counter scintillator define the incoming beam. Both scintillator rings of the RPD are required to register a recoiling proton, ensuring an interaction in the target volume. To clean up the event sample, several veto detectors in front of the target prevent the recording of events with beam particles not passing the target volume, while the Beam Killer scintillators placed at the end of the spectrometer filter out events with a non-interacting beam particle. The Sandwich calorimeter ensures that all outgoing particle tracks lie within the geometrical acceptance of the spectrometer.

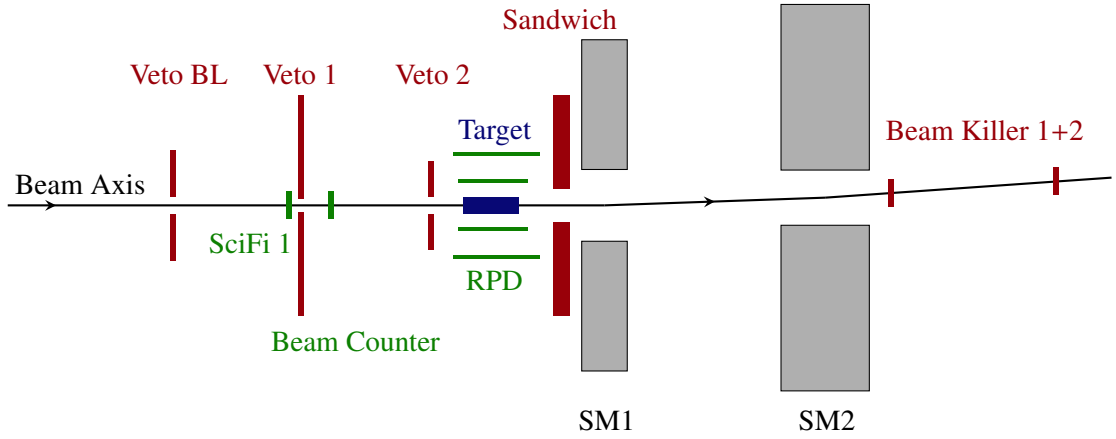


Figure 2.2: Elements of the trigger used in this analysis (not to scale). The target is shown in blue, trigger components in green and the veto system in red. The spectrometers magnets are drawn in gray.

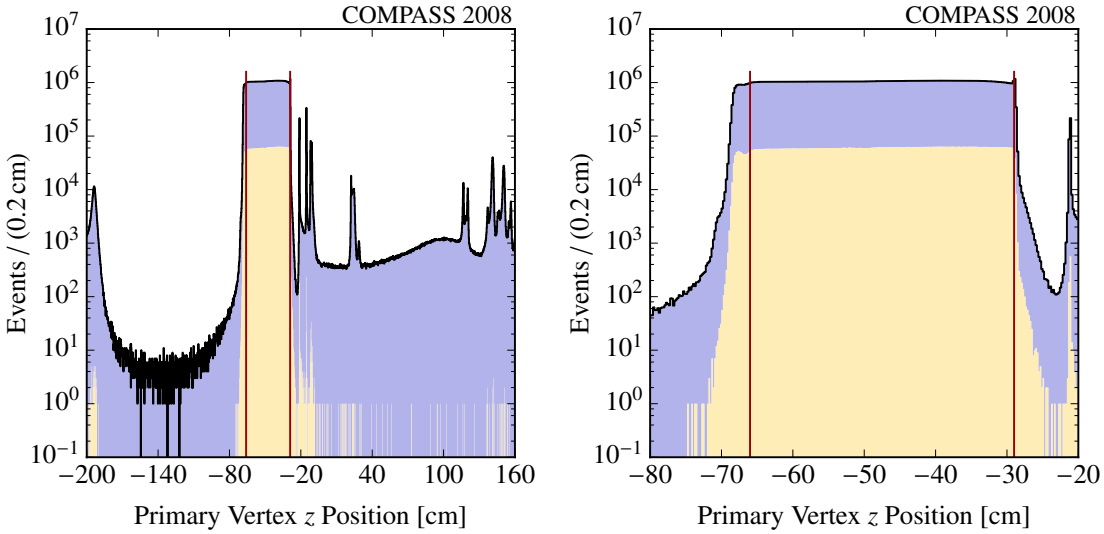
2.1.4 Detector Simulation

The COMGEANT software package, which is based on GEANT-3.21, is used to simulate the passage of particles through the COMPASS experimental setup. To achieve this, it comes with a detailed geometrical description of the material distribution within the spectrometer and uses precise magnetic field maps of the bending magnets. Events from an event generator, i.e. the particle momenta resulting from a physics interaction, are passed to COMGEANT, which will then simulate the trajectories of the particles and their interaction with the spectrometer material and detectors. The resulting pseudo data are then analyzed like the measured data and are used to estimate the acceptance and the resolution of the spectrometer.

2.2 Event Selection

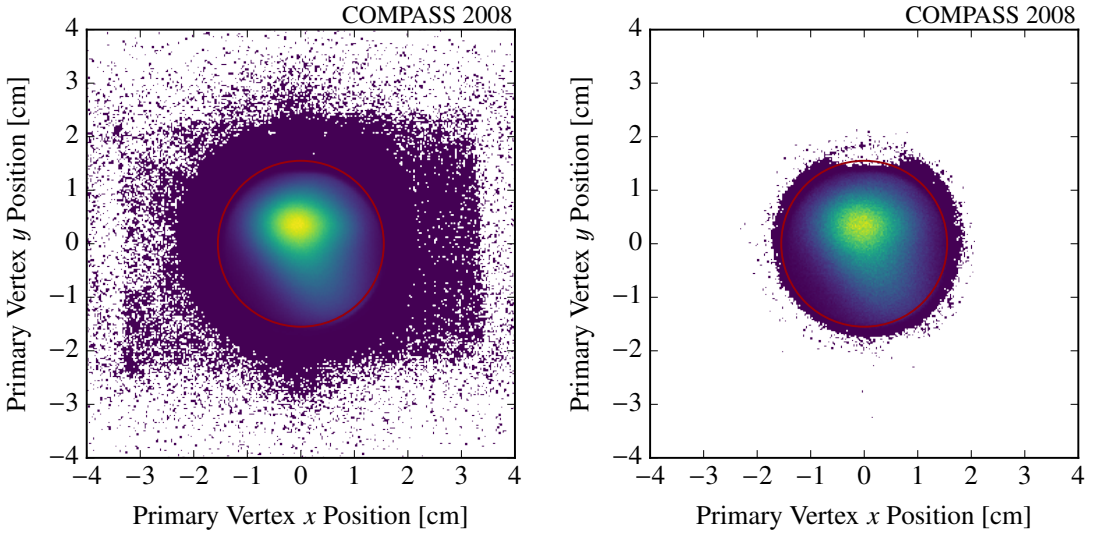
For both technical and practical reasons, the even selection is split into two parts. In the first part, called preselection, the events are filtered by requiring (1) exactly one primary vertex, (2) the position of this primary vertex in z, i.e. along the beam axis, roughly in the target area and (3) five tracks with total charge of -1 measured by the spectrometer. This reduces the size of the event sample by a factor of ~ 35.

For the second part of the event selection, the trigger described in section 2.1.3 is selected. Because of condition (1) from above, only events with exactly one RPD track are allowed. The distribution of the primary vertex position and the corresponding cuts are shown in figures 2.3 to 2.5.



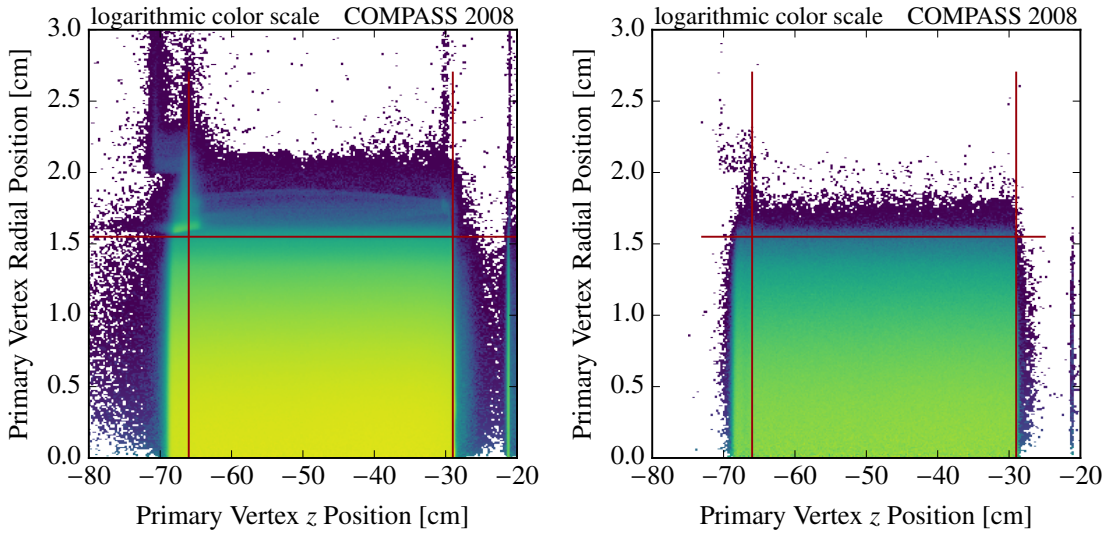
(a) (b)

Figure 2.3: Distribution of the primary vertex position along the beam axis. Shown in blue are all events after the preselection, in yellow after all cuts.



(a) (b)

Figure 2.4: Distribution of the primary vertex position perpendicular to the beam axis. In (a), all events after the preselection are shown while (b) shows the distribution after all cuts.



(a) (b)

Figure 2.5: Distribution of the primary vertex position in the (r, z) plane. The bin content has been corrected for the area element of the cylindrical coordinates. In (a), all events after the preselection are shown while (b) shows the distribution after all cuts.

With the four-momentum p_{beam} of the beam, the total four-momentum p_X of the outgoing five-pion system and the scattering angle θ in the center-of-mass frame, the four-momentum transfer squared

$$\begin{aligned} t &= (p_{\text{beam}} - p_X)^2 \\ &= p_{\text{beam}}^2 + p_X^2 - 2\, p_{\text{beam}}\, p_X \\ &= m_{\text{beam}}^2 + m_X^2 - 2\left(E_{\text{beam}}\, E_X - |\vec{p}_{\text{beam}}|\,|\vec{p}_X|\,\cos\theta\right) \end{aligned} \tag{2.1}$$

is defined. One can see that $|t|$ is minimal for $\theta = 0$ if every other variable is kept constant. This yields

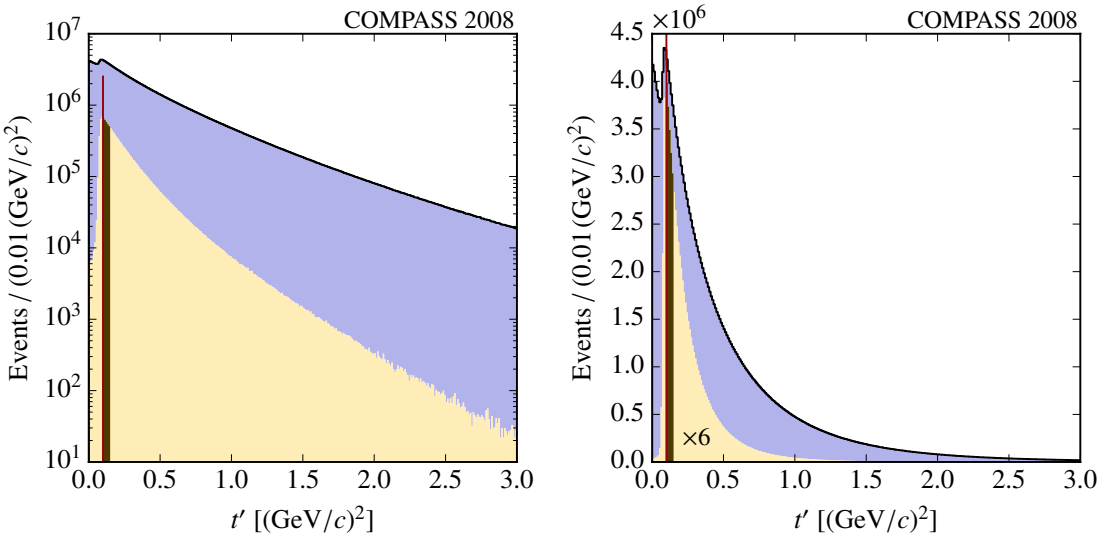
$$|t|_{\text{min}} = -\,m_{\text{beam}}^2 - m_X^2 + 2\left(E_{\text{beam}}\, E_X - |p_{\text{beam}}|\,|p_X|\right) \tag{2.2}$$

which we use to define the reduced four-momentum transfer squared

$$t' = |t| - |t|_{\text{min}} \tag{2.3}$$

This removes the kinematic effect due to the creation of an intermediate object with mass $m_X > m_{\text{beam}}$.

The measured t' distribution is shown in figure 2.6. Because the recoiling proton needs a minimal amount of momentum to be able to leave the target cell and produce a signal in both rings of the RPD, the RPD trigger is only fully efficient for $t' > 0.1\,(\text{GeV}/c)^2$. Events below $0.1\,(\text{GeV}/c)^2$ are not used. Since this analysis aims at a proof of principle, only a narrow bin in t' from 0.1 to $0.15\,(\text{GeV}/c)^2$, indicated in the figure, is used for the partial-wave decomposition. Within the bin, the t' dependences are neglected. The t' region was chosen because it has the lowest non-exclusive background.



(a) (b)

Figure 2.6: Distribution of t' in logarithmic scale (a) and linear scale (b). Shown in blue are all events after the preselection, in yellow after all cuts and in brown the t' bin used for the partial-wave analysis. The yellow and brown distributions have been scaled by a factor of 6 in (b).

The selection of exclusive events is complicated by the fact that the COMPASS apparatus does not measure the magnitude of the beam momentum, but only its direction. However, with the scattering angle θ and the four-momentum p_X of the five-pion system, it is possible to calculate the beam energy by solving the quadratic equation [13]

$$2a_2 E_{\text{beam}}^2 - 2a_1 E_{\text{beam}} - a_0 = 0, \quad \text{where} \quad \begin{aligned} a_0 &= m_{\text{beam}}^2 \, |\vec{p}_X| \cos\theta \\ a_1 &= m_{\text{target}} \, E_X - \frac{1}{2}\left(m_X^2 + m_{\text{beam}}^2\right) \\ a_2 &= m_{\text{target}} - E_X + |\vec{p}_X| \cos\theta \end{aligned} \tag{2.4}$$

The distribution of the calculated beam energy is shown in figure 2.7, together with the cut applied. One can see a clear peak at the nominal beam energy already in the preselection sample. The non-exclusive background is reduced considerably after all cuts are applied.

Because of momentum conservation, the momenta of the beam pion, the recoiling proton and the five-pion system have to be coplanar. It is advantageous to define the angular difference

$$\Delta\phi = 180^\circ - \angle\left({}_\perp\vec{p}_X, {}_\perp\vec{p}_{\text{recoil}}\right) \tag{2.5}$$

where ${}_\perp\vec{p}_X$ and ${}_\perp\vec{p}_{\text{recoil}}$ are the projections of the five-pion momentum and the recoil-proton momentum, respectively, onto the plane perpendicular to the beam momentum. For an exclusive event, $\Delta\phi$ should be zero. The geometrical angular resolution of the RPD depends on the combinations of scintillator segments which are hit and can either be $\pm 7.5^\circ$ or $\pm 3.75^\circ$, as is shown in figure 2.8. To take into account multiple scattering of the recoiling protons in the scintillator material, the cut windows have been opened to $\pm 8.432^\circ$ and $\pm 5.377^\circ$. The $\Delta\phi$ distribution is shown in figure 2.9a.

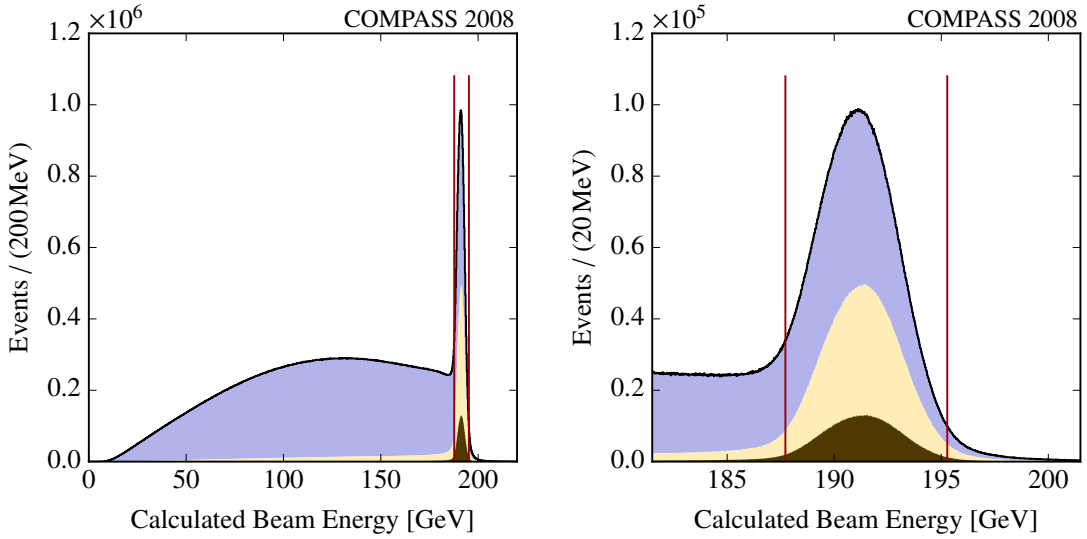


Figure 2.7: Distribution of the calculated beam energy. Shown in blue are all events after the preselection, in yellow after all cuts and in brown the t' bin used for the partial-wave analysis.

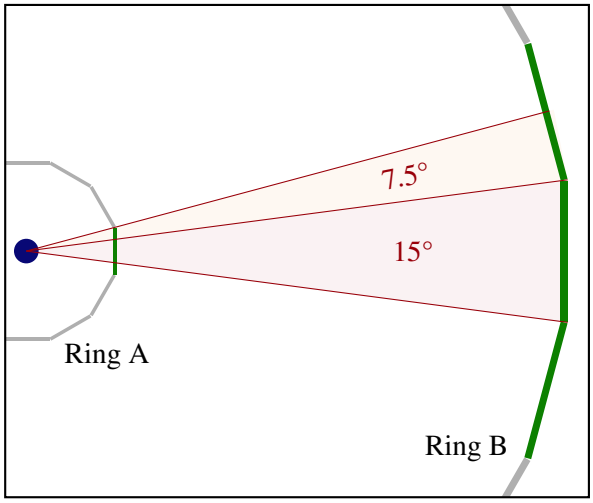
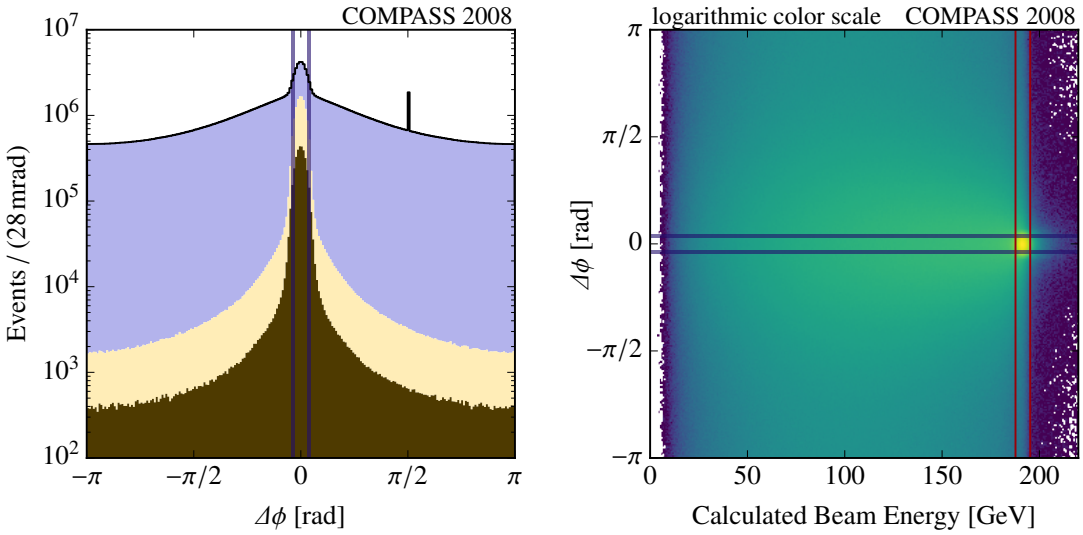


Figure 2.8: Geometry of the RPD showing the two possible combinations of scintillator segments with their opening angles (to scale).



(a) (b)

Figure 2.9: (a) distribution of $\Delta\phi$. In blue are all events after the preselection, in yellow after all cuts and in brown the t' bin used for the partial-wave analysis. The narrow peak visible at $\pi/2$ in the blue distribution is due to errors in the reconstruction and is removed in the event selection. In (b), $\Delta\phi$ is shown against the calculated beam energy with all cuts applied. In both plots, the width of the blue lines correspond to the difference between the two cuts applied to $\Delta\phi$.

The cut on the calculated beam energy and on $\Delta\phi$ heavily influence each other. In fact, the drop in non-exclusive background which can be seen in figure 2.7a when going from the preselection events to the ones with all cuts applied comes almost entirely from the cut on $\Delta\phi$ and the converse is true for the $\Delta\phi$ distribution. This can also be seen when the two quantities are plotted against each other in figure 2.9b.

After applying all cuts, the size of the preselected event sample is reduced further by a factor of ~ 19, giving a total reduction factor of the event selection of ~ 660 and leaving a data set containing 11 243 496 events. The t' bin which is subsequently used in the partial-wave decomposition contains 2 914 914 events.

2.3 Kinematic Distributions

The invariant mass distribution of the five-pion system after all cuts is shown in figure 2.10a. Apart from a small peak-like structure with the maximum at about $1.9\,\mathrm{GeV}/c^2$, the distribution is rather featureless. In figure 2.10b, the mass range from $1.36\,\mathrm{GeV}/c^2$ to $2.56\,\mathrm{GeV}/c^2$ used for the partial-wave analysis is highlighted, containing 1 345 046 events .

Shown in figure 2.11 are the distributions of the mass of the neutral two- and charged three-pion subsystems for the full $m_{5\pi}$ range. The peaks from $\rho(770)$ and $f_2(1270)$ are visible in the $m_{2\pi}$

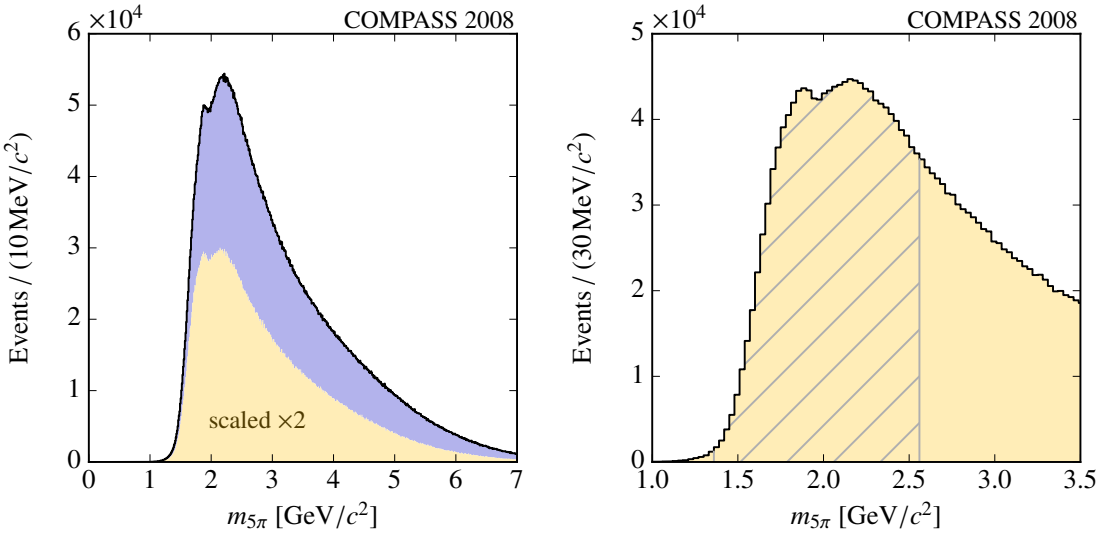


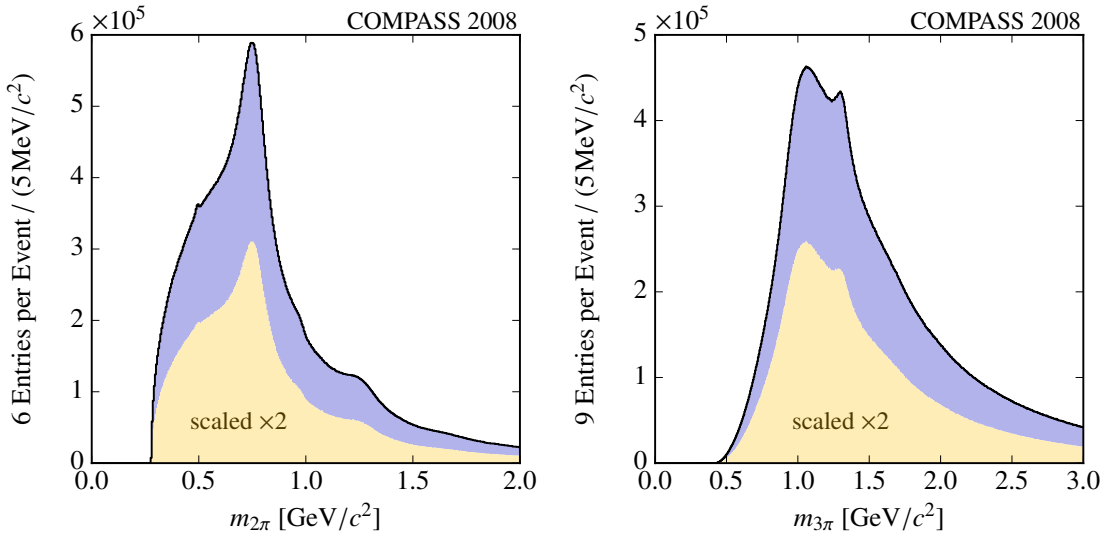
Figure 2.10: The five-pion mass distribution (a), in blue after all cuts and in yellow the t' bin used for the partial-wave analysis scaled up by a factor of 2. The yellow distribution from (a) is shown again in (b), with the hatched region indicating the range studied in the partial-wave analysis, which uses the same $30\,\mathrm{MeV}/c^2$ mass binning.

distribution, as well as a shoulder from $f_0(980)$. The relative strength of the $\rho(770)$ peak decreases in the lower t' region. The small enhancement at $\sim 0.5\,\mathrm{GeV}/c^2$ comes from K^0_s and $\overline{K}^0_s \to \pi^+\pi^-$ originating from $\pi^- p \to \pi^- K^0_s \overline{K}^0_s$. In the $m_{3\pi}$ distribution, all combinations forming a positively or negatively charged three-pion system have been included. A peak at $\sim 1.3\,\mathrm{GeV}/c^2$ is visible which most likely corresponds to the $a_2(1320)$. Its relative strength also decreases in the lower t' region.

In the mass distribution of the neutral four-pion subsystem shown in figure 2.12a, a very sharp peak is visible from $f_1(1285)$, owed to its small width of $\sim 24\,\mathrm{MeV}/c^2$. Two more peak-like structure are visible around $1.5\,\mathrm{GeV}/c^2$, which cannot be clearly attributed to established resonances. Possible candidates are $f_1(1420)$, $f_0(1500)$, $f_2(1565)$ or ρ-like states. When the mass distribution of the two-pion subsystem originating from a four-pion system is plotted against the mass of that four-pion system, as shown in figure 2.12b, one can see rich structure. The sharp vertical line comes from the $f_1(1285)$, which does not seem to favor $\rho(770)$ as a decay mode. The peaks at higher $m_{4\pi}$ have strong $\rho(770)$ contribution and a small $f_2(1270)$ shadow is visible.

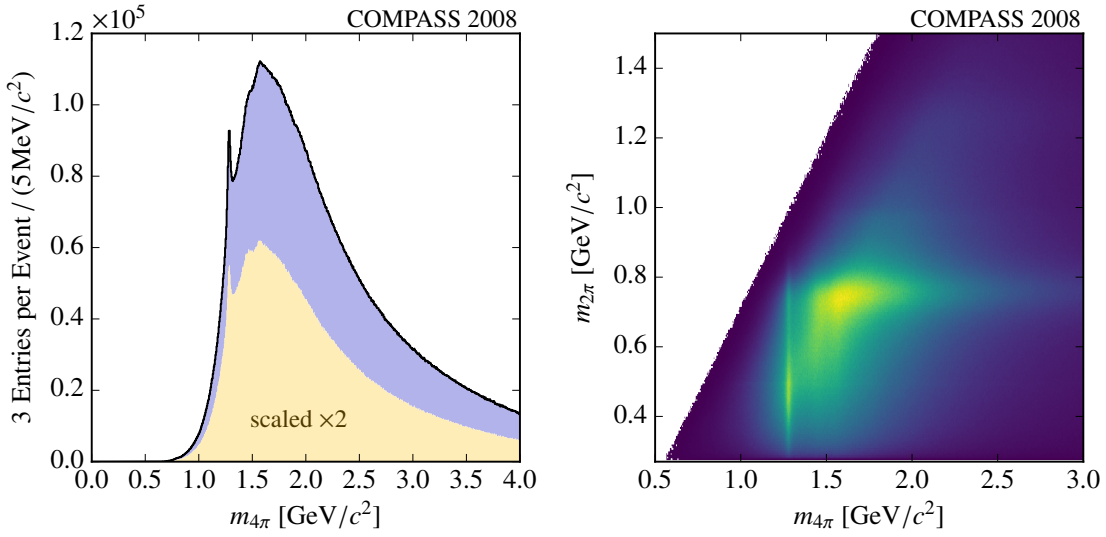
In similar fashion, figure 2.13 shows the masses of the two two-pion subsystems from a four-pion system against each other, for different regions of four-pion mass. As soon as the four-pion mass is large enough, double-$\rho(770)$ contribution is clearly visible, followed by $\rho(770)$-$f_2(1270)$ and even double-$f_2(1270)$ modes.

Because the five-pion final state prefers high masses, a great number of states are contributing to the spectra discussed above. Combined with the large combinatorial background for the subsystems, only very limited conclusions can be drawn from these purely kinematic distributions. Most importantly, there are structures appearing in the mass distributions of all three subsystems, which hints at the appearance of resonances. To infer physical meaning from the measured data, they have to be decomposed into partial waves, which will be discussed in the following chapters.



(a) (b)

Figure 2.11: Distribution of (a) $m_{2\pi}$ and (b) $m_{3\pi}$, in blue after all cuts and in yellow the t' bin used for partial-wave analysis. The latter is scaled by a factor of 2. In (b), both the positively and negatively charged three-pion systems have been included.



(a) (b)

Figure 2.12: The mass distribution of the four-pion system in (a), in blue after all cuts and in yellow the t' bin used for partial-wave analysis. In (b), the mass distribution of two pions from a four-pion system is shown against the mass distribution of the four-pion system, after all cuts, with 12 entries per event.

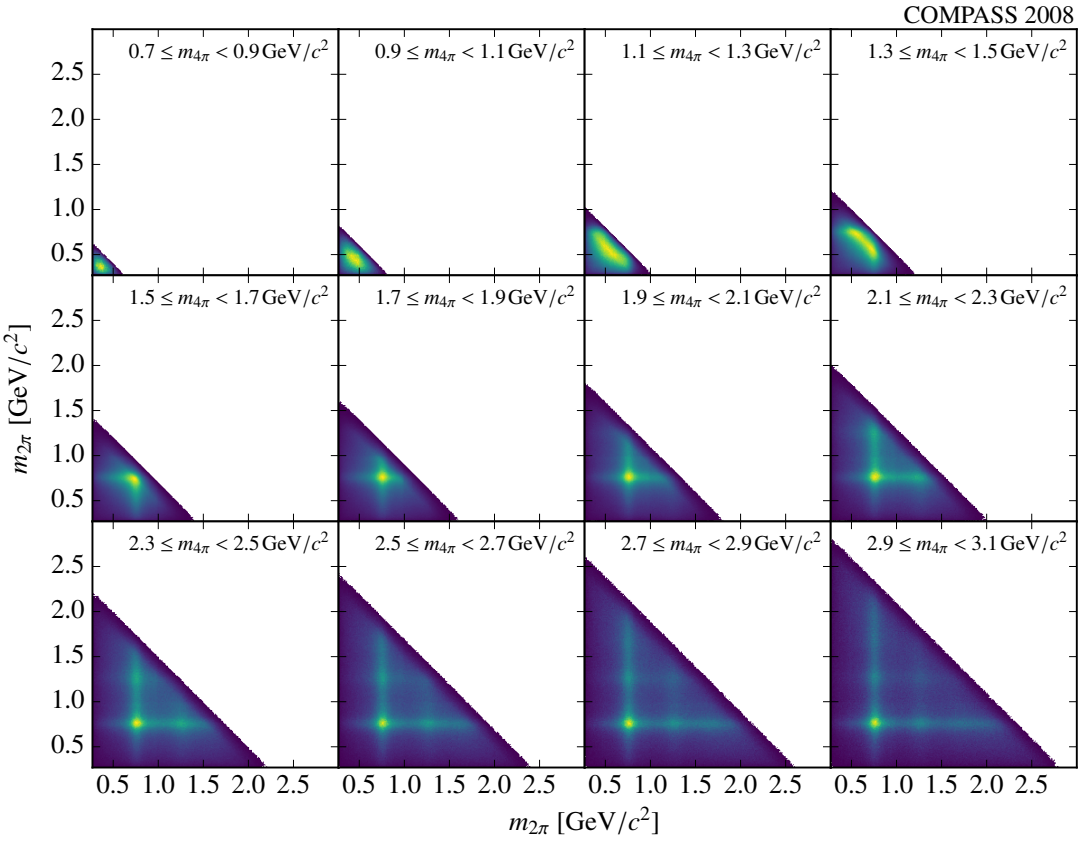


Figure 2.13: Distribution of two-pion masses from a four-pion system for different mass regions of the four-pion system. Depending on how many of the three combinations forming a neutral four-pion system fulfill the mass constraint, there are up to 6 entries per event.

CHAPTER 3

Analysis Technique

> "error: 'long long long' is too long for GCC"
>
> GCC 4.8.2 error message

The ultimate goal of the analysis technique described in this chapter is the identification of resonances in a given set of events and the extraction of their resonance parameters. For conceptual as well as technical reasons, the analysis is conducted in two steps. In the first step, a decomposition of the data set in terms of partial waves is performed. This means, one extracts the complex production amplitudes for given quantum states. In the second step, a resonance model is fitted to the extracted amplitudes to extract resonance parameters.

Several names exist for the individual steps of the analysis as well as the analysis as a whole. The analysis is often called "partial-wave analysis" or, more generally, "amplitude analysis" although this name technically only applies to the first step of the analysis. The first step is sometimes distinguished by using the suffix "... in mass bins" or by calling it "mass-independent fit", while the second step is often called "mass-dependent fit". Here, for clarity, the first step of the analysis will be referred to as "partial-wave decomposition" and the second step will be called "resonance extraction".

As this analysis technique is a crucial part of this work, it will be described here in detail, even though it can be traced back to the 1960s and 1970s [14] and is well-covered in literature [6, 15–18].

3.1 Production of Excited Mesons in Diffractive Reactions

Consider the scattering process

$$1 + 2 \to 3 + 4 \tag{3.1}$$

i.e. particles 1 and 2 react in some way and the result of this reaction are the particles 3 and 4. The four-momentum of particle i is denoted by p_i and the corresponding invariant mass by m_i. The

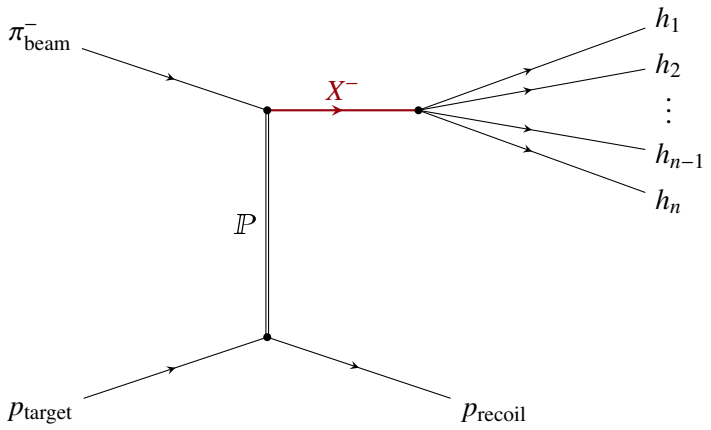


Figure 3.1: Diffractive dissociation of a π^- beam impinging on a proton target into n final-state hadrons h_i.

reaction can now be characterized by the three Mandelstam variables [19]

$$\begin{aligned} s &= (p_1 + p_2)^2 \\ t &= (p_1 - p_3)^2 \\ u &= (p_1 - p_4)^2 \end{aligned} \tag{3.2}$$

Note that because

$$s + t + u = \sum_i m_i^2 \tag{3.3}$$

one of the three can be expressed by the other two. The sum is running over all four particles in equation (3.1).

In the process which is the focus of this work, there are two incoming and six outgoing hadrons. This can be understood in the framework of equation (3.1) if one assumes an intermediate state which decays into 5π. At COMPASS, the center-of-mass energy $\sqrt{s}$ is 18.9 GeV and thereby far above the resonance region, while the typical squared four-momentum transfer t is $O[1\,(\text{GeV}/c)^2]$. In this constellation, which is commonly referred to as diffraction, one expects t-channel exchange to be dominant, i.e. the dominant event topology is expected to look like shown in figure 3.1.

As every measured event has the same initial and the same final state, the intermediate state X is actually a superposition of all possible resonances which are allowed by energy and momentum conservation laws. Unfortunately, in the mass region of interest, the resonances in question are plentiful and overlapping due to their finite width. Therefore, the simple practice of identifying resonances as peaks in the final-state mass distribution

$$m^2_{\text{FS}} = \left(\sum_i p_i \right)^2 \tag{3.4}$$

is usually impossible. The summation index i here runs over the decay products of X.

However, fortunately, the measured event carries more information than just this mass: the actual three-momenta of all decay products are measured. This means one has access to the kinematic distribution of the decay products, which is crucial in decomposing the data in terms of partial waves.

Every intermediate state X is characterized by its isospin I, the G and C parities, its spin J, parity P, the spin projection M and reflectivity ϵ. This set of quantum numbers is usually written as $I^G J^{PC} M^\epsilon$, e.g. $1^-1^{++}0^+$.

As charge-exchange reactions are strongly suppressed at COMPASS energies, the particle X has to carry the charge of the incoming beam particle, i.e. -1, which implies $I > 0$. Furthermore, since there are no known mesons with isospin $I \geq 2$, only $I = 1$ is considered for X^-. In addition, assuming that the exchange particle is a Pomeron $\mathbb{P}$ with $I = 0$ and $G = P = C = +1$, X^- can only have $G = -1$ because of the G-parity of the incoming pion. The reflectivity quantum number ϵ results from a transformation to the reflectivity basis, where states with positive and negative spin projections M are combined, leading to M always being positive and the appearance of $\epsilon = \pm 1$ as a new quantum number [17, 18]. While charged particles are not eigenstates of the C-parity operator, it is customary to give the C eigenvalue of the corresponding neutral partner state.

The superposition of X states is described by a spin-density matrix $\varrho_{\alpha\beta}$ [17]. It is a Hermitian and positive-definite $N_s \times N_s$ matrix, with N_s being the number of contributing states. Therefore, one can write its Cholesky decomposition

$$\begin{aligned} \varrho_{\alpha\beta} = \sum_r^{N_s} {}^r\mathcal{T}_\alpha \, {}^r\mathcal{T}^*_\beta, \quad \text{where} \quad {}^r\mathcal{T}_{\alpha,\beta} = 0 \quad \text{if} \quad \alpha,\beta < r \\ \text{and} \quad {}^r\mathcal{T}_{\alpha,\beta} \in \mathbb{R} \quad \text{if} \quad \alpha,\beta = r \end{aligned} \tag{3.5}$$

meaning that ${}^r\mathcal{T}_\alpha$ is a lower triangular matrix. The unusual index placement and nomenclature is chosen to be consistent with the following sections. The structure of ${}^r\mathcal{T}_\alpha$ is more easily visible in matrix notation:

$$\begin{pmatrix} \varrho_{11} & \cdots & \varrho_{1N_s} \\ \vdots & & \vdots \\ \varrho_{N_s1} & \cdots & \varrho_{N_sN_s} \end{pmatrix} = \begin{pmatrix} {}^1\mathcal{T}_1 & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ {}^1\mathcal{T}_{N_s-1} & \cdots & {}^{N_s-1}\mathcal{T}_{N_s-1} & 0 \\ {}^1\mathcal{T}_{N_s} & {}^2\mathcal{T}_{N_s} & \cdots & {}^{N_s}\mathcal{T}_{N_s} \end{pmatrix} \cdot \begin{pmatrix} {}^1\mathcal{T}^*_1 & {}^1\mathcal{T}^*_2 & \cdots & {}^1\mathcal{T}^*_{N_s} \\ 0 & {}^2\mathcal{T}^*_2 & \cdots & {}^2\mathcal{T}^*_{N_s} \\ \vdots & & & \vdots \\ 0 & \cdots & 0 & {}^{N_s}\mathcal{T}^*_{N_s} \end{pmatrix} \tag{3.6}$$

Decomposing $\varrho_{\alpha\beta}$ has the advantage that the matrix ${}^r\mathcal{T}_\alpha$ is a minimal representation of $\varrho_{\alpha\beta}$, i.e. all entries are linearly independent.

3.2 Partial-Wave Decomposition

According to Fermi's golden rule [20], the differential cross-section for scattering reactions can be written as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\tau}(\tau; m_{\mathrm{FS}}, t', s) = \widetilde{\sigma}_0(m_{\mathrm{FS}}, t', s) \cdot \left|\mathcal{M}(\tau; m_{\mathrm{FS}}, t', s)\right|^2 \tag{3.7}$$

where m_{FS} is the invariant mass of the final state, i.e. the mass of X in the picture of diffractive reactions. As COMPASS typically runs at a fixed beam energy, s is constant and will be dropped from here on. The dependence on t', which is as defined in equation (2.3), will be discussed below. The kinematic variables τ describe the final state, i.e. they are the four-momenta of the n final-state particles. Because (1) three-momentum conversation holds, (2) the final-state mass m_{FS} is explicitly separated from τ and (3) the masses of the final-state particles are assumed to be known, the number of linearly independent degrees of freedom in τ is

$$\begin{aligned} \dim \tau &= 4 \cdot n - 3 - 1 - n \\ &= 3 \cdot n - 4 \end{aligned} \tag{3.8}$$

The actual choice of the variables in τ will be discussed in section 3.3.1. Coming back to equation (3.7), $\mathcal{M}(\tau; m_{\mathrm{FS}})$ is the scattering-matrix element for the process in question. As COMPASS does not measure absolute cross-sections in the relevant channels, all the normalization constants are absorbed in $\widetilde{\sigma}_0(m_{\mathrm{FS}})$.

The reason for the explicit separation of the m_{FS} dependence is one of the pillars of the analysis method: in the partial-wave decomposition, the data are subdivided into narrow bins of the final-state mass, meaning that over the bin width, $\widetilde{\sigma}_0$ and $\left|\mathcal{M}\right|^2$ can be assumed constant with respect to m_{FS}. Due to this, the first step of the analysis is sometimes labeled to be "model-independent", but it must be stressed that this only applies to the resonance content of the intermediate state, while indeed a number of model assumptions still have to be made, as we will see in section 3.3.

There are two approaches to resolve the dependency on t'. While the older approach parameterized the t' dependence [16, 21], larger data sets allowed to subdivide the data not only in m_{FS}, but simultaneously also in t' [6]. The latter approach is assumed in this discussion.

We now proceed to write the cross section as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\tau}(\tau; m_{\mathrm{FS}}, t') = \sigma_0(m_{\mathrm{FS}}) \cdot \mathcal{I}(\tau; m_{\mathrm{FS}}, t') \tag{3.9}$$

with the intensity[a]

$$\begin{aligned}\mathcal{I}(\tau; m_{\mathrm{FS}}, t') &= \sum_{\epsilon=\pm1} \sum_{\alpha,\beta} \sum_{k} {}_{\bullet}\Psi^{\epsilon}_{\alpha,k}(\tau; m_{\mathrm{FS}}, t')\, {}_{\bullet}\varrho^{\epsilon}_{\alpha\beta}(m_{\mathrm{FS}}, t')\, {}_{\bullet}\Psi^{\epsilon *}_{\beta,k}(\tau; m_{\mathrm{FS}}, t') \\ &= \sum_{\epsilon=\pm1} \sum_{\alpha,\beta} \sum_{k} {}_{\bullet}\Psi^{\epsilon}_{\alpha,k}(\tau; m_{\mathrm{FS}}, t') \sum_{r} \left[{}^{r}_{\bullet}\mathcal{T}^{\epsilon}_{\alpha}(m_{\mathrm{FS}}, t')\, {}^{r}_{\bullet}\mathcal{T}^{\epsilon *}_{\beta}(m_{\mathrm{FS}}, t') \right] {}_{\bullet}\Psi^{\epsilon *}_{\beta,k}(\tau; m_{\mathrm{FS}}, t') \\ &= \sum_{\epsilon=\pm1} \sum_{r} \left| \sum_{\alpha} {}^{r}_{\bullet}\mathcal{T}^{\epsilon}_{\alpha}(m_{\mathrm{FS}}, t') \sum_{k} {}_{\bullet}\Psi^{\epsilon}_{\alpha,k}(\tau; m_{\mathrm{FS}}, t') \right|^2 \end{aligned} \tag{3.10}$$

where the spin-density matrix ${}_{\bullet}\varrho_{\alpha\beta}$[b] from equation (3.5) has been used and split into two separate parts ${}_{\bullet}\varrho^{\epsilon}_{\alpha\beta}$ for the reflectivities $\epsilon = \pm 1$. This is possible because different reflectivities do not interfere, leading to a block diagonal form of ${}_{\bullet}\varrho_{\alpha\beta}$ with respect to to ϵ. The determination of the complex-valued production amplitudes ${}^{r}_{\bullet}\mathcal{T}^{\epsilon}_{\alpha}(m_{\mathrm{FS}})$ for the states X, and thereby also of ${}_{\bullet}\varrho^{\epsilon}_{\alpha\beta}$, is the goal of the partial-wave decomposition. The ${}^{r}_{\bullet}\Psi^{\epsilon}_{\beta,\alpha}(\tau) \in \mathbb{C}$ are amplitudes that describe the decay of the states X into the final state. They can be calculated and will be discussed in section 3.3. The indices α and β run over the possible quantum numbers $J^{PC}M$ (but not the reflectivity ϵ), while the index k runs over all possibilities in the parameterization of the decay of X into the final state (which always stays the same). At this point, it is not obvious why there should be more than one possibility for k, but as will become clear in section 3.3, different intermediate states in the decay cause this to happen. All three indices have to be summed coherently, while the sum over the two reflectivity states $+1$ and -1 is incoherent.

The incoherent sum running over r takes into account different non-interfering production processes, e.g. spin-flip or non-flip processes at the proton vertex or the integration over a large t' range. It is important to note that the conditions on ${}^{r}_{\bullet}\mathcal{T}^{\epsilon}_{\alpha}$ from equation (3.5) still apply here, i.e. certain production amplitudes are either real or zero. For reactions where the number of non-interfering production processes N^{ϵ}_{r} is smaller than N^{ϵ}_{s}, one usually imposes

$$ {}^{r}_{\bullet}\mathcal{T}^{\epsilon}_{\alpha} = 0, \quad \text{if} \quad r > N^{\epsilon}_{r} \tag{3.11}$$

as an additional condition. Because this lowers the rank of the spin-density matrix ${}_{\bullet}\varrho^{\epsilon}_{\alpha\beta}$ from N^{ϵ}_{s} to N^{ϵ}_{r}, the quantity N^{ϵ}_{r} is also called the rank of the partial-wave decomposition. Note that due to the block-diagonal nature of ${}_{\bullet}\varrho_{\alpha\beta}$ with respect to ϵ, the two ranks N^{+1}_{r} and N^{-1}_{r} need not be the same. The analysis presented in the following chapters will always use $N^{+1}_{r} = N^{-1}_{r} = 1$; however, the description in this chapter will remain without any assumptions on N^{ϵ}_{r}.

In the intensity given in equation (3.10), the coupling of the X states to the decay modes is included in the decay amplitudes, which is a problem as these couplings are generally unknown. However, because they are assumed to only depend on m_{FS} and not on τ, the couplings can be easily moved

[a] While the conclusion that $\mathcal{I} \equiv |\mathcal{M}|^2$ might be easily drawn from the given formulas, it would be wrong, because $\mathcal{I}$ and $|\mathcal{M}|^2$ are normalized differently, which will become clear in equation (3.16). This difference of normalization is also reflected in the transition $\widetilde{\sigma}_0(m_{\mathrm{FS}}) \to \sigma_0(m_{\mathrm{FS}})$.

[b] There is no difference between $\varrho_{\alpha\beta}$ from equation (3.5) and ${}_{\bullet}\varrho_{\alpha\beta}$ from equation (3.10): the "$\bullet$" has been introduced to facilitate the notation after the upcoming transformation in equation (3.12).

from the decay amplitudes to the production amplitudes giving new variables

$$\begin{aligned} {}_{\bullet}\Psi^{\epsilon}_{\alpha,k}(\tau; m_{\mathrm{FS}}, t') &\rightarrow \Psi^{\epsilon}_{i}(\tau; m_{\mathrm{FS}}, t') \\ {}^{r}_{\bullet}\mathcal{T}^{\epsilon}_{\alpha}(m_{\mathrm{FS}}, t') &\rightarrow {}^{r}\mathcal{T}^{\epsilon}_{i}(m_{\mathrm{FS}}, t') \\ {}_{\bullet}\varrho^{\epsilon}_{\alpha\beta}(m_{\mathrm{FS}}, t') &\rightarrow \varrho^{\epsilon}_{ij}(m_{\mathrm{FS}}, t') \end{aligned} \tag{3.12}$$

The new index i now includes both the quantum numbers α of X and the specific decay mode k into the final state. With this, one obtains a slightly simpler form of the intensity

$$\mathcal{I}(\tau; m_{\mathrm{FS}}, t') = \sum_{\epsilon=\pm1} \sum_{r=1}^{N^{\epsilon}_{r}} \left| \sum_{i} {}^{r}\mathcal{T}^{\epsilon}_{i}(m_{\mathrm{FS}}, t')\, \Psi^{\epsilon}_{i}(\tau; m_{\mathrm{FS}}, t') \right|^2 \tag{3.13}$$

with which we can write the probability to observe an event $\mathscr{E}_{m_{\mathrm{FS}}}$ from the set of measured events $\mathbb{E}$ with coordinates $\tau_{\mathscr{E}}$

$$P(\mathscr{E}_{m_{\mathrm{FS}}}) = \frac{\mathcal{I}(\tau_{\mathscr{E}}; m_{\mathrm{FS}}, t')}{\int \mathrm{d}\phi_n(\tau)\, \mathcal{I}(\tau; m_{\mathrm{FS}}, t')\, \eta(\tau; m_{\mathrm{FS}}, t')} \tag{3.14}$$

The integral in the denominator runs over the whole n-body phase space of the final-state particles, while $\eta(\tau)$ is the product of the acceptance of the measuring device and the reconstruction efficiency for the final state, abbreviated in the following to "acceptance". In the numerator, the acceptance does not appear because for measured events, it has already been applied by the measurement itself. With this probability, we are now able to formulate an extended likelihood function we can maximize in order to find the maximum-likelihood estimate for ${}^{r}\mathcal{T}^{\epsilon}_{i}$:

$$\mathcal{L}(\mathcal{T}; \mathbb{E}) = \frac{e^{-\overline{N}} \overline{N}^{N}}{N!} \prod_{\mathfrak{t}=1}^{N} P(\mathscr{E}_{\mathfrak{t}}) \tag{3.15}$$

N is the number of measured events and $\overline{N}$ is the number of expected measured events. The factor $e^{-\overline{N}}\overline{N}^{N}/N!$ is the Poisson probability to observe N events, which ensures that the maximum of $\mathcal{L}$ is at a point in parameter space where the condition $\overline{N} = N$ is fulfilled within the expected Poissonian fluctuations. To declutter the equations, m_{FS} and t' have been dropped and will be omitted from this point onwards. However, it is important to keep in mind that all the quantities, including N and $\overline{N}$, are given under the assumption that they are constant with respect to the final state mass m_{FS}. Using

$$\overline{N} = \int \mathrm{d}\phi_n(\tau)\, \mathcal{I}(\tau)\, \eta(\tau) \tag{3.16}$$

imposes a normalization constraint which leads to the intensity $\mathcal{I}$ being expressed in terms of number of events. Using the fact that $\overline{N}$ is the denominator in equation (3.14), one can cancel the factor $\overline{N}^{N}$ in equation (3.15) and the likelihood simplifies slightly to

$$\mathcal{L}(\mathcal{T}; \mathbb{E}) = \frac{e^{-\overline{N}}}{N!} \prod_{\mathfrak{t}=1}^{N} \mathcal{I}(\tau_{\mathfrak{t}}) \tag{3.17}$$

Now, because of numerical stability, we transition to the log-likelihood

$$\ln \mathcal{L}(\mathcal{T}; \mathbb{E}) = -\overline{N} - \ln N! + \sum_{\mathfrak{k}=1}^{N} \ln \mathcal{I}(\tau_{\mathfrak{k}}) \tag{3.18}$$

As the term $\ln N!$ does not depend on ${}^{r}\mathcal{T}_{i}^{\epsilon}$, it can be dropped from the log-likelihood without changing its shape:

$$\ln \mathcal{L}(\mathcal{T}; \mathbb{E}) = \sum_{\mathfrak{k}=1}^{N} \ln \mathcal{I}(\tau_{\mathfrak{k}}) - \overline{N} \tag{3.19}$$

For the purpose of describing measured events, which unfortunately are usually not background-free, the intensity given in equation (3.13) is extended by adding an isotropic background amplitude $\mathcal{T}_{\text{flat}} \in \mathbb{R}$, called flat wave, giving

$$\mathcal{I}(\tau) = \sum_{\epsilon=\pm 1} \sum_{r=1}^{N_r^{\epsilon}} \left| \sum_{i} {}^{r}\mathcal{T}_{i}^{\epsilon} \, \Psi_{i}^{\epsilon}(\tau) \right|^2 + \mathcal{T}_{\text{flat}}^2 \tag{3.20}$$

To evaluate $\mathcal{L}$ from equation (3.19), one needs to calculate $\overline{N}$. For clarity, let's first note that

$$\begin{aligned} \left| \sum_{i} {}^{r}\mathcal{T}_{i}^{\epsilon} \, \Psi_{i}^{\epsilon}(\tau) \right|^2 &= \left(\sum_{i} {}^{r}\mathcal{T}_{i}^{\epsilon} \, \Psi_{i}^{\epsilon}(\tau) \right) \left(\sum_{j} {}^{r}\mathcal{T}_{j}^{\epsilon} \, \Psi_{j}^{\epsilon}(\tau) \right)^{*} \\ &= \left(\sum_{i} {}^{r}\mathcal{T}_{i}^{\epsilon} \, \Psi_{i}^{\epsilon}(\tau) \right) \left(\sum_{j} {}^{r}\mathcal{T}_{j}^{\epsilon *} \, \Psi_{j}^{\epsilon *}(\tau) \right) \\ &= \sum_{i,j} {}^{r}\mathcal{T}_{i}^{\epsilon} \, {}^{r}\mathcal{T}_{j}^{\epsilon *} \, \Psi_{i}^{\epsilon}(\tau) \, \Psi_{j}^{\epsilon *}(\tau) \end{aligned} \tag{3.21}$$

Therefore, $\overline{N}$ from equation (3.16) can be written as

$$\begin{aligned} \overline{N} &= \int \mathrm{d}\phi_n(\tau) \left(\sum_{\epsilon=\pm 1} \sum_{r=1}^{N_r^{\epsilon}} \left| \sum_{i} {}^{r}\mathcal{T}_{i}^{\epsilon} \, \Psi_{i}^{\epsilon}(\tau) \right|^2 + \mathcal{T}_{\text{flat}}^2 \right) \eta(\tau) \\ &= \int \mathrm{d}\phi_n(\tau) \left(\sum_{\epsilon=\pm 1} \sum_{r=1}^{N_r^{\epsilon}} \sum_{i,j} {}^{r}\mathcal{T}_{i}^{\epsilon} \, {}^{r}\mathcal{T}_{j}^{\epsilon *} \, \Psi_{i}^{\epsilon}(\tau) \, \Psi_{j}^{\epsilon *}(\tau) + \mathcal{T}_{\text{flat}}^2 \right) \eta(\tau) \\ &= \sum_{\epsilon=\pm 1} \sum_{r=1}^{N_r^{\epsilon}} \sum_{i,j} {}^{r}\mathcal{T}_{i}^{\epsilon} \, {}^{r}\mathcal{T}_{j}^{\epsilon *} \int \mathrm{d}\phi_n(\tau) \; \Psi_{i}^{\epsilon}(\tau) \, \Psi_{j}^{\epsilon *}(\tau) \, \eta(\tau) + \mathcal{T}_{\text{flat}}^2 \int \mathrm{d}\phi_n(\tau) \, \eta(\tau) \end{aligned} \tag{3.22}$$

which is easy to calculate provided one can determine the normalization integrals

$$\mathfrak{N}_{ij}^{\epsilon} = \int \mathrm{d}\phi_n(\tau) \; \Psi_{i}^{\epsilon}(\tau) \, \Psi_{j}^{\epsilon *}(\tau) \, \eta(\tau) \tag{3.23}$$

These high-dimensional integrals are calculated using Monte Carlo methods [22], which approximate the integral over a multivariate, real- or complex-valued function f over the volume Ω by sampling random points $\vec{x}_\mathrm{t}$ from Ω and averaging the function values $f(\vec{x}_\mathrm{t})$.

$$\begin{aligned} f : \mathbb{R}^n \to \mathbb{C}, \quad \Omega \subset \mathbb{R}^n, \quad \vec{x}_\mathrm{t} \in \Omega, \quad V_\Omega = \int_\Omega \mathrm{d}\vec{x} \\ \int_\Omega \mathrm{d}\vec{x}\, f(\vec{x}) = \lim_{N\to\infty} \frac{V_\Omega}{N} \sum_{\mathrm{t}=1}^{N} f(\vec{x}_\mathrm{t}) \end{aligned} \tag{3.24}$$

We apply equation (3.24) to equation (3.23) by generating a set of Monte Carlo events $\Omega = \left\{\tau_1^\mathrm{MC}, \ldots, \tau_{N^{MC}}^\mathrm{MC}\right\}$ which are distributed isotropically in the final state's phase space. These events are then processed through the detector simulation described in section 2.1.4 and subjected to the same selection cuts as the measured data (see section 2.2). The events remaining after this procedure are $\Omega^\mathrm{Acc} = \left\{\tau_1^\mathrm{Acc}, \ldots, \tau_{N^\mathrm{Acc}}^\mathrm{Acc}\right\}$. The acceptance $\eta(\tau)$, which is a continuous function of the phase space variables τ, only assumes two values on Ω:

$$\eta : \Omega \to \{0, 1\}, \quad \eta(\tau_\mathrm{t}) = \begin{cases} 1, & \text{if } \tau_\mathrm{t} \in \Omega^\mathrm{Acc} \\ 0, & \text{otherwise} \end{cases} \tag{3.25}$$

The integrals from equation (3.23) can now be approximated by

$$\mathfrak{N}_{ij}^\epsilon \approx \frac{V_\Omega}{N^\mathrm{MC}} \sum_{\mathrm{t}=1}^{N^\mathrm{Acc}} \Psi_i^\epsilon(\tau_\mathrm{t}^\mathrm{Acc})\, \Psi_j^{\epsilon *}(\tau_\mathrm{t}^\mathrm{Acc}) \tag{3.26}$$

The quantity $V_\Omega = \int \mathrm{d}\phi_n(\tau)$ is not easily calculable, however, one can eliminate it from the formulas by a variable transformation. To this end, we normalize our variables

$$\begin{aligned} \Psi_i^\epsilon(\tau) \quad &\to \quad \psi_i^\epsilon(\tau) = \Psi_i^\epsilon(\tau)\, \frac{\sqrt{V_\Omega}}{\sqrt{\mathcal{P}_{ii}^\epsilon}} \\ {}^r\mathcal{T}_i^\epsilon(m_\mathrm{FS}) \quad &\to \quad {}^rT_i^\epsilon(m_\mathrm{FS}) = {}^r\mathcal{T}_i^\epsilon(m_\mathrm{FS}) \cdot \sqrt{\mathcal{P}_{ii}^\epsilon} \\ \mathfrak{N}_{ij}^\epsilon \quad &\to \quad \mathcal{N}_{ij}^\epsilon = \frac{\mathfrak{N}_{ij}^\epsilon}{\sqrt{\mathcal{P}_{ii}^\epsilon \mathcal{P}_{jj}^\epsilon}} \\ \mathcal{T}_\mathrm{flat} \quad &\to \quad T_\mathrm{flat} = \mathcal{T}_\mathrm{flat} \cdot \sqrt{V_\Omega} \\ \varrho_{ij}^\epsilon \quad &\to \quad \rho_{ij}^\epsilon = \varrho_{ij}^\epsilon \sqrt{\mathcal{P}_{ii}^\epsilon\, \mathcal{P}_{jj}^\epsilon} = \sum_{r=1}^{N_r^\epsilon} {}^rT_i^\epsilon\, {}^rT_j^{\epsilon *} \end{aligned} \tag{3.27}$$

using the phase-space integrals

$$\begin{aligned}\mathcal{P}^{\epsilon}_{ij} &= \int \mathrm{d}\phi_n(\tau)\; \Psi^{\epsilon *}_i(\tau)\, \Psi^{\epsilon}_j(\tau) \\ &\approx \frac{V_\Omega}{N^{\mathrm{MC}}} \sum_{\mathfrak{t}=1}^{N^{\mathrm{MC}}} \Psi^{\epsilon *}_i(\tau^{\mathrm{MC}}_{\mathfrak{t}})\, \Psi^{\epsilon}_j(\tau^{\mathrm{MC}}_{\mathfrak{t}})\end{aligned} \tag{3.28}$$

which do not include the acceptance $\eta(\tau)$. As the decay amplitudes ψ^{ϵ}_i are now normalized and the intensity $\mathcal{I}$ is expressed in terms of number of events, the diagonal element of the spin-density matrix

$$\rho^{\epsilon}_{ii} = \sum_{r=1}^{N^{\epsilon}_r} \left| {}^{r}T^{\epsilon}_i \right|^2 \tag{3.29}$$

is the acceptance-corrected expected number of events in wave i.

Because ${}^{r}\mathcal{T}^{\epsilon}_i \Psi^{\epsilon}_i(\tau) = {}^{r}T^{\epsilon}_i \psi^{\epsilon}_i(\tau) \,/\, \sqrt{V_\Omega}$ and ${}^{r}\mathcal{T}^{\epsilon}_i {}^{r}\mathcal{T}^{\epsilon *}_j \mathfrak{N}^{\epsilon}_{ij} = {}^{r}T^{\epsilon}_i {}^{r}T^{\epsilon *}_j \mathcal{N}^{\epsilon}_{ij}$, the likelihood from equation (3.19) transforms to

$$\begin{aligned}\ln \mathcal{L}(T;\mathbb{E}) &= \sum_{\mathfrak{t}=1}^{N} \ln \left(\sum_{\epsilon=\pm 1} \sum_{r=1}^{N^{\epsilon}_r} \left| \sum_i \frac{{}^{r}T^{\epsilon}_i \psi^{\epsilon}_i(\tau_{\mathfrak{t}})}{\sqrt{V_\Omega}} \right|^2 + \frac{T^2_{\mathrm{flat}}}{V_\Omega} \right) \\ &\qquad - \sum_{\epsilon=\pm 1} \sum_{r=1}^{N^{\epsilon}_r} \sum_{i,j} {}^{r}T^{\epsilon *}_i {}^{r}T^{\epsilon}_j \mathcal{N}^{\epsilon}_{ij} - T^2_{\mathrm{flat}} \frac{\int \mathrm{d}\phi_n(\tau)\, \eta(\tau)}{V_\Omega} \\ &= \sum_{\mathfrak{t}=1}^{N} \ln \left(\sum_{\epsilon=\pm 1} \sum_{r=1}^{N^{\epsilon}_r} \left| \sum_i {}^{r}T^{\epsilon}_i \psi^{\epsilon}_i(\tau_{\mathfrak{t}}) \right|^2 + T^2_{\mathrm{flat}} \right) - N \ln V_\Omega \\ &\qquad - \sum_{\epsilon=\pm 1} \sum_{r=1}^{N^{\epsilon}_r} \sum_{i,j} {}^{r}T^{\epsilon *}_i {}^{r}T^{\epsilon}_j \mathcal{N}^{\epsilon}_{ij} - T^2_{\mathrm{flat}} \frac{\int \mathrm{d}\phi_n(\tau)\, \eta(\tau)}{V_\Omega}\end{aligned} \tag{3.30}$$

As before, additive constants which do not depend on ${}^{r}T^{\epsilon}_i$ can be dropped from the log-likelihood, which means we can disregard $N \ln V_\Omega$. The factor

$$\frac{\int \mathrm{d}\phi_n(\tau)\, \eta(\tau)}{V_\Omega} \approx \frac{N^{\mathrm{Acc}}}{N^{\mathrm{MC}}} \equiv \mathcal{A} \tag{3.31}$$

is the total acceptance. Putting everything together, we arrive at

$$\ln \mathcal{L}(T;\mathbb{E}) = \sum_{\mathfrak{t}=1}^{N} \ln \left(\sum_{\epsilon=\pm 1} \sum_{r=1}^{N^{\epsilon}_r} \left| \sum_i {}^{r}T^{\epsilon}_i \psi^{\epsilon}_i(\tau_{\mathfrak{t}}) \right|^2 + T^2_{\mathrm{flat}} \right) - \sum_{\epsilon=\pm 1} \sum_{r=1}^{N^{\epsilon}_r} \sum_{i,j} {}^{r}T^{\epsilon *}_i {}^{r}T^{\epsilon}_j \mathcal{N}^{\epsilon}_{ij} - T^2_{\mathrm{flat}} \mathcal{A} \tag{3.32}$$

This is the log-likelihood formulation as it is used to determine the maximum-likelihood estimate for the free parameters ${}^{r}T^{\epsilon}_i$. Because the decay amplitudes Ψ^{ϵ}_i and the two integral matrices $\mathcal{P}^{\epsilon}_{ij}$ and $\mathfrak{N}^{\epsilon}_{ij}$ do not depend on ${}^{r}T^{\epsilon}_i$, they can be pre-calculated once and then reused in every iteration of the minimization process.

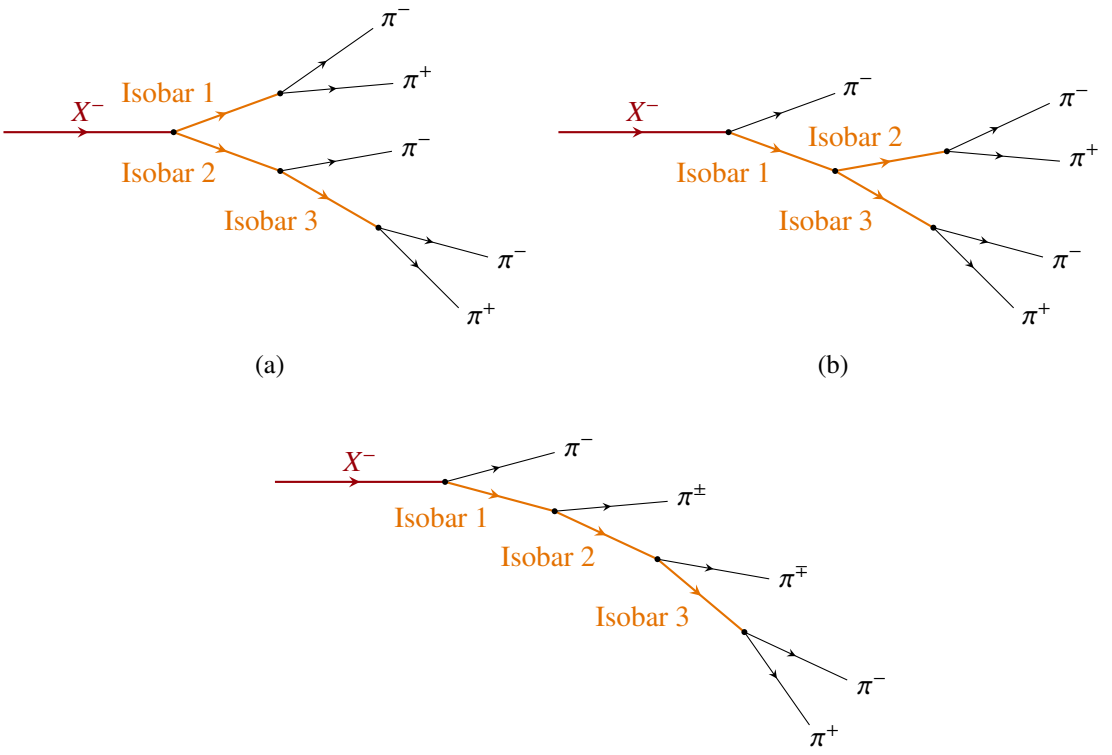


Figure 3.2: Isobar decay topologies for $X^- \to \pi^-\pi^+\pi^-\pi^+\pi^-$.

3.3 Decay Amplitudes

In the derivation of the formalism to determine the production amplitudes ${}^{r}T_i^{\epsilon}$ in section 3.2, it was assumed that the decay amplitudes Ψ_i^{ϵ}, which describe the decay of the intermediate state X to the n-body final state, are calculable. Here, we present the model which has been used as a basis for this calculation in this analysis. For simplicity, we sacrifice the generality of section 3.2 by assuming that the final state consists of five charged pions.

3.3.1 The Isobar Model

The isobar model presupposes that the decay into the final state takes place via a chain of subsequent two-particle decays. For the decay into five charged pions, there are three possible ways this can be realized, as shown in figure 3.2. For a decay into n particles, this requires the introduction of $n-2$ additional intermediate states which are called isobars. The total decay amplitude Ψ_i^{ϵ} then essentially becomes a product of the amplitudes for the two-body decays.

The reference frames which are used to calculate the two-body decay amplitudes are shown in figure 3.3. The coordinate system used to describe the decay of the X particle is called the Gottfried-

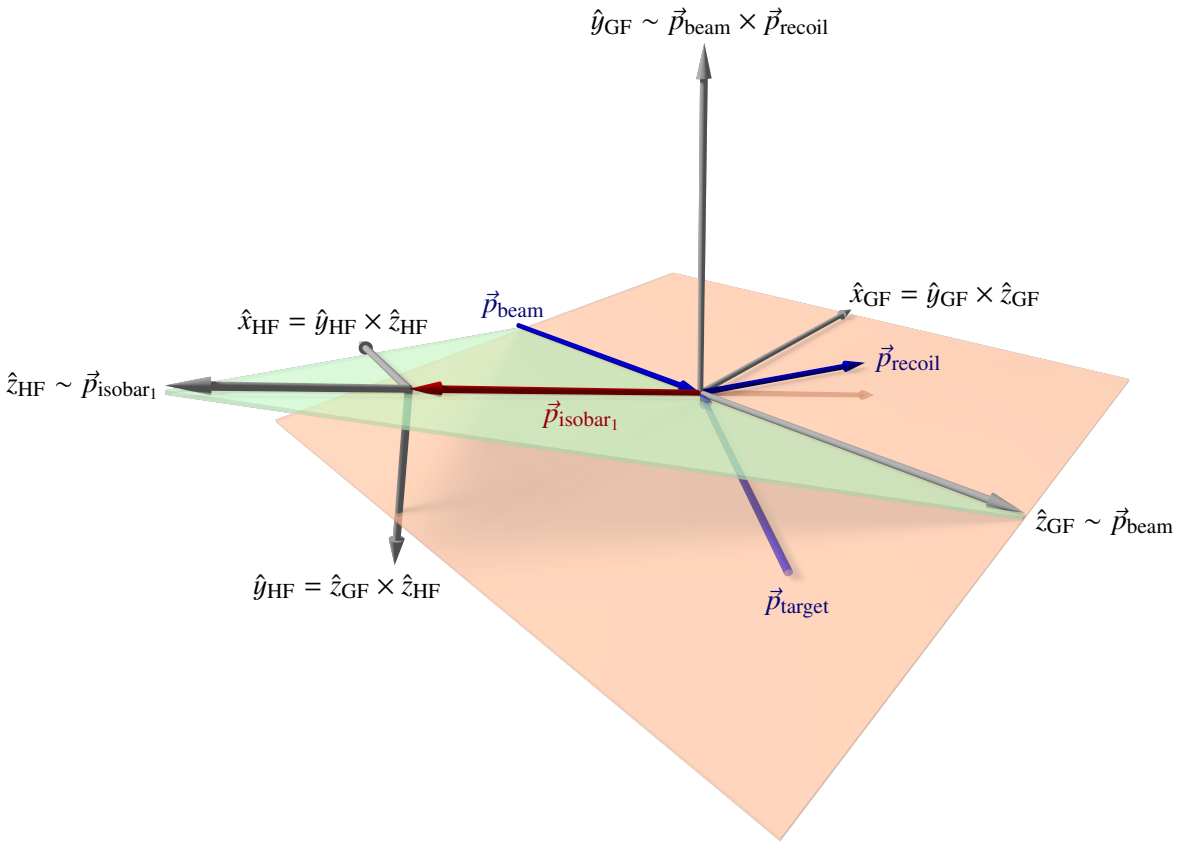


Figure 3.3: The Gottfried-Jackson and helicity reference frames.

Jackson frame (GF), which is a rest frame of X. The z-axis $\hat{z}_{GF}$ is defined along the beam momentum $\vec{p}_{beam}$, while the y-axis $\hat{y}_{GF}$ is along the normal of the X production plane, given by $\vec{p}_{beam} \times \vec{p}_{recoil}$, where the latter is the momentum of the recoil particle. The x-axis $\hat{x}_{GF}$ is chosen such that the resulting coordinate system is right-handed.

For the description of the decay of the first isobars in the decay chain, the helicity frame (HF) is used, which is a rest frame of the decaying isobar. The z-axis $\hat{z}_{HF}$ points along the direction of motion of the isobar; therefore, the spin projection with respect to this quantization axis is the isobar helicity λ. Even though the direction of motion is ill defined, because the isobar is obviously at rest in its own rest frame, the axis' direction can be easily preserved through the Lorentz boost. The y-axis $\hat{y}_{HF}$ is then defined by the cross product of the two z-axes $\hat{z}_{GF} \times \hat{z}_{HF}$ and as before, the x-axis $\hat{x}_{HF}$ is given by $\hat{y}_{HF} \times \hat{z}_{HF}$ to fulfill right-handedness.

This definition of the helicity frame can be applied analogously for the decay of isobars further down the decay chain, i.e. isobars where the parent particle is an isobar itself. In this case, the y-axis is defined as the cross product of the z-axes of the helicity frames of the parent isobar and the daughter isobar.

It is important to note that we use active Lorentz transformation to transform the relevant four-momenta to the corresponding coordinate systems. Because Lorentz transformations do not commute

in general, their order has to be specified. We always start by transforming into the Gottfried-Jackson frame and then proceed sequentially along the decay chain.

Let us now consider the decay of a parent particle into two daughter particles. To facilitate notation, we collect the quantum numbers J and M/λ as well as the nominal mass m_0 and width Γ_0 of the respective particles into one index. For the particle X, we also include the reflectivity ϵ. This index shall be $\mathfrak{P}$ for the parent particle and $\mathfrak{D}_{1,2}$ for the two daughters. If any of the variables is given explicitly, the subscript P is used for the parent particle, while the subscripts 1 or 2 denote the daughter particles. Additionally, L is the orbital angular momentum between the two daughters, and S is the total intrinsic spin to which the spins of the daughters J_1 and J_2 couple. Using these definitions and the reference frames specified earlier, one can write down the two-body decay amplitude

$$A^{\mathfrak{P}}_{\mathfrak{D}_1\mathfrak{D}_2}(\varphi, \vartheta, m_\mathrm{P}, m_1, m_2) = \sum_{\lambda_1,\lambda_2} D^{J_\mathrm{P}*}_{M_\mathrm{P}\lambda}(\vartheta, \varphi, 0)\, f^{\mathfrak{P}}_{\mathfrak{D}_1\mathfrak{D}_2}(m_\mathrm{P}, m_1, m_2) \tag{3.33}$$

with $\lambda = \lambda_2 - \lambda_1$. The angles φ and ϑ are the spherical coordinates of the momentum of one of the decay particles, while m_P, m_1 and m_2 are the measured masses for the respective event. The amplitude factorizes in an angular part, described by the Wigner D-function $D^{J*}_{M\lambda}$ [23] and a dynamic part $f^{\mathfrak{P}}_{\mathfrak{D}_1\mathfrak{D}_2}(m_\mathrm{P}, m_1, m_2)$. As the momenta of the daughters are oriented back to back, only two angles are needed and the third angle is set to zero. Furthermore, if $\mathfrak{P}$ is an isobar, M_P in $D^{J_\mathrm{P}}_{M_\mathrm{P}\lambda}$ is replaced by λ_P, the helicity of the parent particle.

The dynamic part describes the strong interaction in the decay and is written as

$$f^{\mathfrak{P}}_{\mathfrak{D}_1\mathfrak{D}_2}(m_\mathrm{P}, m_1, m_2) = \overbrace{\sqrt{2L+1}}^{\text{normalization}}\ \overbrace{(J_1\,\lambda_1\ J_2\,{-\lambda_2}\,|\,S\ \lambda)\,(L\,0\ S\ \lambda\,|\,J_\mathrm{P}\,\lambda)}^{\text{Clebsch-Gordan coefficients}}$$
$$\underbrace{F_L(m_\mathrm{P}, m_1, m_2)\, \Delta^{\mathfrak{P}}_{\mathfrak{D}_1\mathfrak{D}_2}(m_\mathrm{P}, m_1, m_2)\, \alpha^{\mathfrak{P}}_{\mathfrak{D}_1\mathfrak{D}_2}}_{\text{dynamics}}\ \underbrace{A^{\mathfrak{D}_1}\, A^{\mathfrak{D}_2}}_{\text{daughter decays}} \tag{3.34}$$

where two Clebsch-Gordan coefficients appear, the first for the coupling of the two daughter spins J_1 and J_2 to the total intrinsic spin S, the second for the coupling of S and the orbital angular momentum L to the spin J_P of the parent. The real-valued functions F_L are the Blatt-Weisskopf angular-momentum barrier factors [24] in the parameterization by von Hippel and Quigg [25]. The propagator of the parent particle $\Delta^{\mathfrak{P}}_{\mathfrak{D}_1\mathfrak{D}_2}$ describes the m_P dependence of the amplitude. For most of the cases, it is parameterized by a Breit-Wigner function. This will be discussed in detail in section 3.3.3. The complex number $\alpha^{\mathfrak{P}}_{\mathfrak{D}_1\mathfrak{D}_2}$ is the coupling of the parent particle to the two decay products. The decay amplitudes $A^{\mathfrak{D}_1}$ and $A^{\mathfrak{D}_2}$ are set to unity if the corresponding decay product is quasi-stable, i.e. a pion. If $\mathfrak{D}$ is an isobar, then $A^{\mathfrak{D}}$ is the full two-body decay amplitude including all the dependencies as given in equation (3.33), which were left out in equation (3.34) for clarity. This means that to get the total decay amplitude Ψ^{ϵ}_i for X going into five pions, equation (3.33) has to be evaluated recursively starting with $A^{X}_{\mathfrak{D}_1\mathfrak{D}_2}$ and then propagating down the decay chain.

The couplings $\alpha^{\mathfrak{P}}_{\mathfrak{D}_1\mathfrak{D}_2}$ are generally unknown; however, because they do not depend on λ_1 and λ_2, they can be pulled in front of the sum in equation (3.33). If, in addition, they do not depend on m_P, m_1 and m_2, which is an assumption that might not be true for non-resonant parts of the decay, they can be moved in front of the total decay amplitude and combined to a total coupling of the particle X to the

particular decay mode. As described in equation (3.12), this coupling is then moved to the production amplitudes.

The amplitude for the decay of X has to include the reflectivity ϵ as well. This is reflected in a modified Wigner D-function

$$ {}^{\epsilon}D^{J_X}_{M_X\lambda}(\vartheta,\varphi,0) = c(M)\left[D^{J_X}_{(+M_X)\lambda}(\vartheta,\varphi,0) - \epsilon P_X(-1)^{J_X-M_X} D^{J_X}_{(-M_X)\lambda}(\vartheta,\varphi,0)\right] \tag{3.35} $$

with

$$ c(M) = \begin{cases} 1/2, & \text{if } M = 0 \\ 1/\sqrt{2}, & \text{otherwise} \end{cases} \tag{3.36} $$

which replaces the one used in equation (3.33) if $\mathfrak{P} = X$.

Now we can specify the quantities constituting τ: for every isobar in the decay chain, there are two angles and the mass of the isobar. The decay of X contributes only two angles as the data have been binned in the mass of X. Because there are $n-2$ intermediate states, τ consists of $(n-2)\cdot 3+2 = 3\cdot n-4$ quantities, which we have found already in equation (3.8).

It is useful to define a notation to identify isobar decay chains. To write the decay of one particle, we use

$$ \mathfrak{P} \to \mathfrak{D}_1 \begin{bmatrix} L \\ S \end{bmatrix} \mathfrak{D}_2 \tag{3.37} $$

Should one or both of the daughter particles happen to be unstable, i.e. an isobar, we apply equation (3.37) recursively and enclose the subdecay in round brackets, e.g.

$$ \mathfrak{P} \to \left(\text{Isobar} \to \mathfrak{D}_1 \begin{bmatrix} L \\ S \end{bmatrix} \mathfrak{D}_2\right) \begin{bmatrix} L \\ S \end{bmatrix} \mathfrak{D}_3 \tag{3.38} $$

Using this notation, any decay chain can be written. If a full wave is referenced, the quantum numbers of X are added on the left. The wave

$$ 1^-1^{++}0^+ \to \left(f_1(1285) \to \left(a_1(1260) \to \rho(770) \begin{bmatrix} 0 \\ 1 \end{bmatrix} \pi^-\right) \begin{bmatrix} 1 \\ 1 \end{bmatrix} \pi^+\right) \begin{bmatrix} 1 \\ 1 \end{bmatrix} \pi^- \tag{3.39} $$

can serve as an example for the decay of an X with $I^G J^{PC} M^{\epsilon} = 1^-1^{++}0^+$ into five charged pions via the topology shown in figure 3.2c, where Isobar 1, 2 and 3 are the resonances $f_1(1285)$, $a_1(1260)$ and $\rho(770)$, respectively. Because the decay of isobars into two charged pions is unambiguous, it is not explicitly included in the notation for brevity. Since I^G is always equal to 1^- in diffractive reactions with a pion beam, it is sometimes omitted.

3.3.2 Symmetrization

Because pions of the same charge are indistinguishable particles, the decay amplitude Ψ_i^{ϵ} has to be invariant under the exchange of the four-vectors of a same-charge pion pair. This property is called

Bose symmetry and is ensured by writing the decay amplitude as

$$\Psi_i^\epsilon(\tau) = \frac{1}{\sqrt{N_\mathfrak{B}}} \sum_\mathfrak{B} \Psi_i^\epsilon(\tau_\mathfrak{B}) \tag{3.40}$$

The 3 π^- and 2 π^+ contribute 3! = 6 and 2 possible permutations, respectively, leading to a total of 12 permutations which are indexed by $\mathfrak{B}$. The arguments $\tau_\mathfrak{B}$ are the kinematic variables for permutation $\mathfrak{B}$ applied.

Similarly, if there is a constellation where permuting pions of opposite charge only changes the sign of the charge of an isobar, an additional symmetrization has to be introduced, because the respective decay chains are indistinguishable. However, in this case, the isospin coupling in the form of Clebsch-Gordan coefficients may lead to relative signs $\varsigma_\mathfrak{I}$ in the sum of symmetrization terms. Analogously to equation (3.40), the isospin symmetry is introduced by writing the decay amplitude as

$$\Psi_i^\epsilon(\tau) = \frac{1}{\sqrt{N_\mathfrak{I}}} \sum_\mathfrak{I} \varsigma_\mathfrak{I} \Psi_i^\epsilon(\tau_\mathfrak{I}) \tag{3.41}$$

To determine $\varsigma_\mathfrak{I}$, the sign of the total isospin Clebsch-Gordan coefficient is taken

$$\varsigma_\mathfrak{I} = \operatorname{sgn}\left[\prod_\mathfrak{k} (I^{\mathfrak{D}_1^\mathfrak{k}}{}_\mathfrak{I} I_3^{\mathfrak{D}_1^\mathfrak{k}}\, I^{\mathfrak{D}_2^\mathfrak{k}}{}_\mathfrak{I} I_3^{\mathfrak{D}_2^\mathfrak{k}} \,|\, I^{\mathfrak{P}^\mathfrak{k}}{}_\mathfrak{I} I_3^{\mathfrak{P}^\mathfrak{k}})\right] \tag{3.42}$$

The index $\mathfrak{k}$ runs over the two-body decay vertices, including the decay vertex of X. The product includes all the Clebsch-Gordan coefficients for the coupling of the total isospins I and their third component I_3 of the decaying particles $\mathfrak{P}^\mathfrak{k}$ and the daughters $\mathfrak{D}_1^\mathfrak{k}$ and $\mathfrak{D}_2^\mathfrak{k}$. The permutations $\mathfrak{I}$ swap pions of opposite charge, which means the third components of the isospin depend on the given permutation.

As is evident from figure 3.2c, there is only one constellation in the decay into five pions where this situation may arise and there are only two symmetrization terms which have to be considered. In these cases, both Bose and isospin symmetrization have to be respected, leading to a total of 24 terms in the symmetrization.

3.3.3 Parameterizing the Mass Dependence of Isobars

In this analysis, the mass dependence of the amplitude of the parent particle $\Delta^\mathfrak{P}_{\mathfrak{D}_1\mathfrak{D}_2}(m_\mathrm{P}, m_1, m_2)$ in equation (3.34) is in most cases parameterized by a relativistic Breit-Wigner form [26]

$$\Delta(m_\mathrm{P}, m_1, m_2; m_{0,\mathrm{P}}, \Gamma_{0,\mathrm{P}}) = \frac{m_{0,\mathrm{P}}\,\Gamma_{0,\mathrm{P}}}{m_{0,\mathrm{P}}^2 - m_\mathrm{P}^2 - i\,m_{0,\mathrm{P}}\,\Gamma_\mathrm{P}(m_\mathrm{P}, m_1, m_2; m_{0,\mathrm{P}}, \Gamma_{0,\mathrm{P}})} \tag{3.43}$$

with the mass-dependent width

$$\Gamma_{\mathrm{P}}(m_{\mathrm{P}}, m_1, m_2; m_{0,\mathrm{P}}, \Gamma_{0,\mathrm{P}}) = \Gamma_{0,\mathrm{P}} \frac{m_{0,\mathrm{P}}}{m_{\mathrm{P}}} \sum_i \alpha_i \frac{\overbrace{q(m_{\mathrm{P}}, m_{1,i}, m_{2,i})}^{q_i}}{\underbrace{q(m_{0,\mathrm{P}}, m_{1,i}, m_{2,i})}_{q_{0,i}}} \frac{F_L^2(q_i)}{F_L^2(q_{0,i})} \tag{3.44}$$

where $m_{0,\mathrm{P}}$ and $\Gamma_{0,\mathrm{P}}$ are the nominal mass and width of the parent particle, respectively, and $q(m_{\mathrm{P}}, m_1, m_2)$ is the two-body breakup momentum. The Breit-Wigner form is a good approximation only for narrow resonances far away from thresholds. The sum in equation (3.44) runs over all possible decay channels i, but, because the branching ratios α_i are not known for many of the isobars used in this analysis, all the summands except the one for the decay into $\mathfrak{D}_1 + \mathfrak{D}_2$ are dropped. This approximation is only valid if the decay channel in question is the dominant one for the parent particle, which may not always be the case.

The breakup momentum is calculated by taking the square root of

$$q^2(m_{\mathrm{P}}, m_1, m_2) = \frac{\left[m_{\mathrm{P}}^2 - (m_1 + m_2)^2\right] \cdot \left[m_{\mathrm{P}}^2 - (m_1 - m_2)^2\right]}{4 \cdot m_{\mathrm{P}}^2} \tag{3.45}$$

Equation (3.44) is only valid if the particles $\mathfrak{D}_1$ and $\mathfrak{D}_2$ are stable, i.e. $m_i \equiv m_{0,i}$ for $i = 1, 2$. If this is not the case, it is easy to see that q^2 can become smaller than zero if the measured masses are such that either $m_1 + m_2 > m_{\mathrm{P}}$ or $m_1 - m_2 > m_{\mathrm{P}}$. While the latter case is unlikely from a kinematic point of view, the former may very easily occur. In the past [10, 27], the workaround

$$q(m_{\mathrm{P}}, m_1, m_2) = \sqrt{|q^2(m_{\mathrm{P}}, m_1, m_2)|} \tag{3.46}$$

was applied. Subsequently, q_0 was calculated using the nominal parent mass $m_{0,\mathrm{P}}$, but the measured daughter masses m_1 and m_2, as it is written in equation (3.44). Unfortunately, this makes the Breit-Wigner function multivariate, even though it should only depend on one variable, m_{P}. Furthermore, if the measured masses fulfill the condition $m_1 + m_2 = m_{\mathrm{P}}$ or $|m_1 - m_2| = m_{\mathrm{P}}$, q_0 will be zero leading to $A^{\mathfrak{P}}_{\mathfrak{D}_1\mathfrak{D}_2}$ being zero.

Assuming one of the daughter particles to be stable, it is possible to visualize the effect of equation (3.46). As a showcase, the decay $a_2(1320) \to \pi + \rho(770)$ is shown in figure 3.4. The two bands where the intensity goes to zero are easily visible.

To have a parameterization of $\Gamma_{\mathrm{P}}(m_{\mathrm{P}}, m_1, m_2; m_{0,\mathrm{P}}, \Gamma_{0,\mathrm{P}})$ which correctly takes into account the case where one or both daughter particles have finite width, equation (3.44) is rewritten as

$$\Gamma_{\mathrm{P}}(m_{\mathrm{P}}; \Gamma_{0,\mathrm{P}}) = \Gamma_{0,\mathrm{P}} \frac{\int \mathrm{d}\phi_{n'}(\tau)\, |\Psi(\tau; m_{\mathrm{P}})|^2}{\int \mathrm{d}\phi_{n'}(\tau)\, |\Psi(\tau; m_{0,\mathrm{P}})|^2} \tag{3.47}$$

The n'-particle decay amplitudes $\Psi(\tau; m_{\mathrm{P}})$ for the respective decay subtree are integrated over the phase space $\mathrm{d}\phi_{n'}(\tau)$ of the final-state particles originating from the parent particle where the mass of the parent particle m_{P} is fixed. The integrals can be calculated in the same way as the ones from

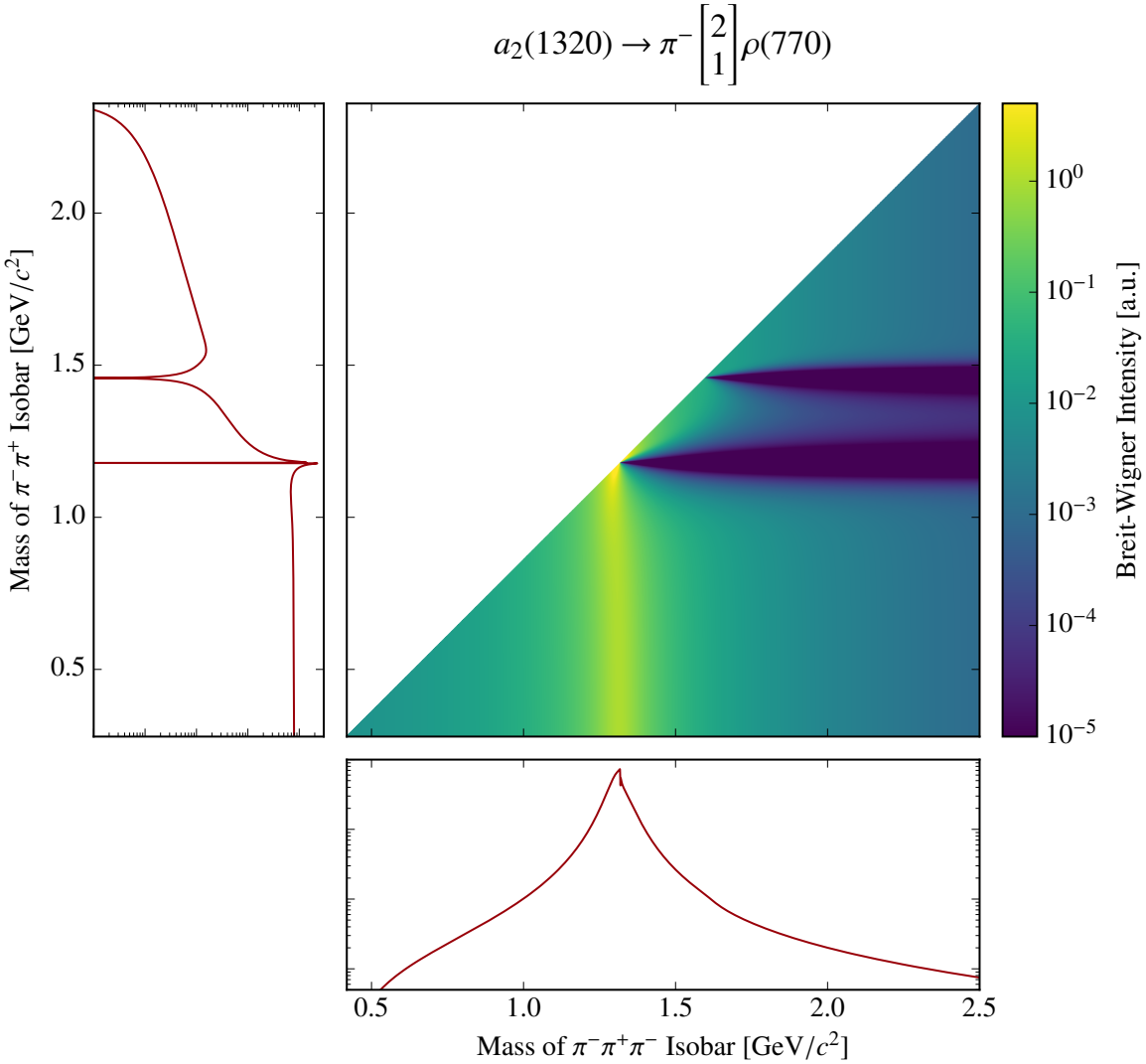


Figure 3.4: Effects of equation (3.46) on the Breit-Wigner intensity.

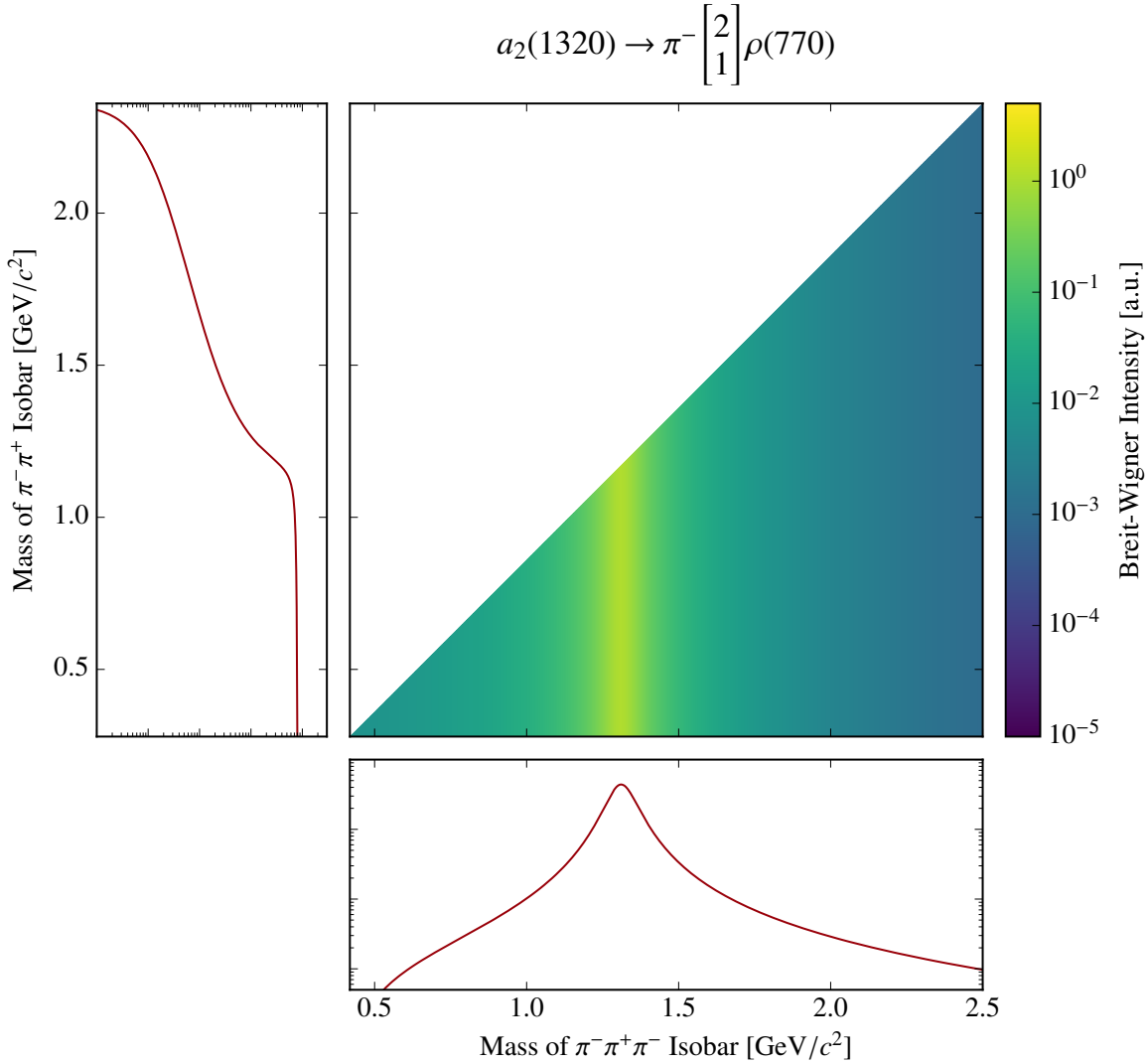


Figure 3.5: Breit-Wigner intensity when using equation (3.47).

equation (3.23), i.e. by using the Monte Carlo integration method as written in equation (3.24). To calculate the decay amplitudes, equation (3.33) is applied to the decay subtree. Note that the same sum over the decay channels as in equation (3.44) should appear here, but has been dropped with the same argumentation as before.

The Breit-Wigner function $\Delta^{\mathfrak{P}}_{\mathfrak{D}_1\mathfrak{D}_2}(m_\mathrm{P}, m_1, m_2)$ and therefore $\Gamma_\mathrm{P}(m_\mathrm{P}; \Gamma_{0,\mathrm{P}})$ from equation (3.47) have to be evaluated for every isobar decay vertex of every measured event and all the Monte Carlo events used for calculation of the various integrals. This means that, to keep the computing requirements on a realistic level, it is necessary to tabulate the integrals from equation (3.47). It is worth to point out that in the case of five final-state particles, to calculate $\Psi(\tau; m_\mathrm{P})$ when integrating, unstable decay products can themselves decay into unstable particles again, meaning that the integration has to be nested recursively.

Figure 3.5 shows how the Breit-Wigner intensity looks like for an unstable daughter particle if equation (3.47) is used instead of equation (3.44). As one can see, the bands with zero intensity have vanished, the function now correctly depends on only one variable and the projections do not feature discontinuities anymore.

While the prerequisites for treating isobars as Breit-Wigner shaped resonances are not always strictly met, there usually are no better alternatives available. In two cases relevant to this analysis, however, superior options exist. One is the two-pion isobar where the pions are in a $L = 0$, i.e. S-wave, configuration, which is referred to as the σ isobar in this work. There, a parameterization originating from a fit to $\pi\pi$-scattering data is available [28]. Because the relative strengths and the phases of the broad σ resonance and the $f_0(980)$ are source dependent, a modified parameterization is used with the $f_0(980)$ removed, as described in [6].

The parameterization of the $f_0(980)$ itself has to be modified as well, as the $K\overline{K}$-threshold sits close to its mass peak. Therefore, a Flatté parameterization is used in the same form as in [6].

3.4 Resonance Extraction

The partial-wave decomposition described in section 3.2 does not make any assumption about the resonance content in the intermediate state by means of binning in the final-state mass $m_\mathrm{FS} \equiv m$, where we drop the subscript from here on for simplicity. We now describe the m dependence of $\rho^\epsilon_{\alpha\beta}(m)$ with a model [18], using the same decomposition as in equation (3.6):

$$\begin{aligned} \hat{\rho}^\epsilon_{\alpha\beta}(m) = \sum_r {}^r\hat{T}^\epsilon_\alpha(m)\, {}^r\hat{T}^{\epsilon *}_\beta(m), \quad \text{where} \quad {}^r\hat{T}^\epsilon_{\alpha,\beta}(m) \equiv 0 \quad \text{if} \quad \alpha,\beta < r \\ \text{and} \quad {}^r\hat{T}^\epsilon_{\alpha,\beta}(m) \in \mathbb{R} \quad \text{if} \quad \alpha,\beta = r \end{aligned} \tag{3.48}$$

The production amplitudes are parameterized as

$${}^r\hat{T}^\epsilon_\alpha(m) = \left[\sum_{k_\alpha} {}^r_{k_\alpha}C^\epsilon_{\alpha\ k_\alpha}\mathcal{S}_\alpha(m)\right] \sqrt{\mathcal{P}^\epsilon_{\alpha\alpha}(m)}\, \exp\left[i\,\varphi^\epsilon_r(m)\right] \tag{3.49}$$

which takes into account the normalization in equation (3.27)[c]. The sum runs over all the model components ${}_{k_\alpha}\mathcal{S}_\alpha(m)$, i.e. all the resonances and non-resonant components, contributing to describe the production amplitude ${}^{r}T^{\epsilon}_{\alpha}(m)$ of wave with index α. A resonance can contribute to multiple waves, given that they have the same quantum numbers $I^G J^{PC}$. The shape functions depend neither on rank index r nor on reflectivity ϵ because we assume the shape to be independent of both the production process and the reflectivity. The complex-valued couplings ${}^{r}_{k_\alpha}C^{\epsilon}_{\alpha}$ contain the relative strengths and phases between the components and are set to zero for wave indices $\alpha < r$, thereby ensuring that the first condition in equation (3.48) is fulfilled.

By applying the phase factor $\exp\left[i\,\varphi^{\epsilon}_{r}(m)\right]$ where

$$\varphi^{\epsilon}_{r}(m) = -\arg\left[\sum_{k_r} {}^{r}_{k_r}C^{\epsilon}_{r}\;{}_{k_r}\mathcal{S}_r(m)\right] \tag{3.50}$$

we make sure that ${}^{r}\hat{T}^{\epsilon}_{\alpha} \in \mathbb{R}$ if $\alpha = r$.

The structure is more obvious if the model predictions ${}^{r}\hat{T}^{\epsilon}_{\alpha}$ are written in matrix form:

$$\begin{aligned}
&\begin{pmatrix} {}^{1}\hat{T}^{\epsilon}_{1}(m) & \cdots & 0 \\ \vdots & & \vdots \\ {}^{1}\hat{T}^{\epsilon}_{N^{\epsilon}_{s}}(m) & \cdots & {}^{N^{\epsilon}_{s}}\hat{T}^{\epsilon}_{N^{\epsilon}_{s}}(m) \end{pmatrix} \\
&= \begin{pmatrix} \sqrt{\mathcal{P}^{\epsilon}_{11}(m)} \\ \vdots \\ \sqrt{\mathcal{P}^{\epsilon}_{N^{\epsilon}_{s}N^{\epsilon}_{s}}(m)} \end{pmatrix} \cdot \begin{pmatrix} \sum\limits_{k_1} {}^{1}_{k_1}C^{\epsilon}_{1}\;{}_{k_1}\mathcal{S}_1(m) & \cdots & 0 \\ \vdots & & \vdots \\ \sum\limits_{k_{N^{\epsilon}_{s}}} {}^{1}_{k_{N^{\epsilon}_{s}}}C^{\epsilon}_{N^{\epsilon}_{s}}\;{}_{k_{N^{\epsilon}_{s}}}\mathcal{S}_{N^{\epsilon}_{s}}(m) & \cdots & \sum\limits_{k_{N^{\epsilon}_{s}}} {}^{N^{\epsilon}_{s}}_{k_{N^{\epsilon}_{s}}}C^{\epsilon}_{N^{\epsilon}_{s}}\;{}_{k_{N^{\epsilon}_{s}}}\mathcal{S}_{N^{\epsilon}_{s}}(m) \end{pmatrix} \\
&\qquad \cdot \begin{pmatrix} \exp\left[i\,\varphi^{\epsilon}_{1}(m)\right] & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \exp\left[i\,\varphi^{\epsilon}_{N^{\epsilon}_{s}}(m)\right] \end{pmatrix}
\end{aligned} \tag{3.51}$$

Instead of actually applying the phase transformation, we move back to the spin-density matrix

$$\begin{aligned}
\hat{\rho}^{\epsilon}_{\alpha\beta}(m) = \sum_{r} &\left(\left[\sum_{k_\alpha} {}^{r}_{k_\alpha}C^{\epsilon}_{\alpha}\;{}_{k_\alpha}\mathcal{S}_\alpha(m)\right] \;\sqrt{\mathcal{P}^{\epsilon}_{\alpha\alpha}(m)}\,\exp\left[i\,\varphi^{\epsilon}_{r}(m)\right] \right) \\
&\cdot \left(\left[\sum_{k_\beta} {}^{r}_{k_\beta}C^{\epsilon}_{\beta}\;{}_{k_\beta}\mathcal{S}_\beta(m)\right]^{*} \sqrt{\mathcal{P}^{\epsilon}_{\beta\beta}(m)}\,\exp\left[-i\,\varphi^{\epsilon}_{r}(m)\right] \right)
\end{aligned} \tag{3.52}$$

[c] Note that, by the definition in equation (3.28), $\mathcal{P}^{\epsilon}_{\alpha\alpha}$ contains the phase-space volume V_Ω, which is not easy to calculate. This will be discussed in detail in section 5.3.

where the factors $\exp\left[\pm i\,\varphi_r^\epsilon(m)\right]$ cancel each other. If the rank N_r^ϵ was chosen to be smaller than the number of waves N_s^ϵ, the condition in equation (3.11) applies analogously to ${}^{r}_{k_\alpha}C^\epsilon_\alpha$:

$$ {}^{r}_{k_\alpha}C^\epsilon_\alpha = 0, \quad \text{if} \quad r > N_r^\epsilon \tag{3.53} $$

Applying this in equation (3.52), the sum over r runs to N_r^ϵ and we obtain

$$ \hat{\rho}^\epsilon_{\alpha\beta}(m) = \sum_{r=1}^{N_r^\epsilon} \left[\sum_{k_\alpha} {}^{r}_{k_\alpha}C^\epsilon_\alpha \; {}_{k_\alpha}\mathcal{S}_\alpha(m) \sqrt{\mathcal{P}^\epsilon_{\alpha\alpha}(m)} \right] \left[\sum_{k_\beta} {}^{r}_{k_\beta}C^\epsilon_\beta \; {}_{k_\beta}\mathcal{S}_\beta(m) \sqrt{\mathcal{P}^\epsilon_{\beta\beta}(m)} \right]^* \tag{3.54} $$

There is an additional ambiguity appearing in this formulation of $\hat{\rho}^\epsilon_{\alpha\beta}(m)$: if we multiply all couplings of a given rank with $\exp[i\,{}_\bullet\varphi_r^\epsilon]$, then the spin-density matrix does not change. Note that ${}_\bullet\varphi_r^\epsilon$ is different from $\varphi_r^\epsilon(m)$ appearing before, in that it does not depend on m. To remove this ambiguity, we require

$$ {}^{r}_{1}C^\epsilon_\alpha \in \mathbb{R}, \quad \text{if} \quad \alpha = r \tag{3.55} $$

Usually, it is not practical to try to describe the production amplitudes of all partial waves which were used during the partial-wave decomposition. Fortunately, it is possible to consider just a submatrix of $\rho_{\alpha\beta}$, corresponding to a selected subset of partial waves.

Resonances are parameterized using Breit-Wigner amplitudes of the form

$$ \mathcal{S}^{\text{res}}(m; m_0, \Gamma_0) = \frac{m_0\,\Gamma_0}{m_0^2 - m^2 - i\,m_0\,\Gamma_0} \tag{3.56} $$

where the parameters m_0 and Γ_0 of ${}_{k_\alpha}\mathcal{S}_\alpha(m)$ are determined by the indices α and k_α. If two $\mathcal{S}^{\text{res}}$ describe the same resonance, they share one common set of parameters. As mentioned in section 3.3.3, the Breit-Wigner amplitude is a good approximation only for narrow resonances far away from thresholds. In addition, summing Breit-Wigner amplitudes generally violates unitarity. For a hypothetical resonance with $m_0 = 1.5\,\text{GeV}/c^2$ and $\Gamma_0 = 0.1\,\text{GeV}/c^2$, the Breit-Wigner amplitude is shown exemplarily in figure 3.6.

A non-resonant component with the parameterization

$$ \mathcal{S}^{\text{bkg}}(m; g, m_1, m_2) = \exp\left\{ -g \left[q(m, m_1, m_2) \cdot F_L^2\left(q(m, m_1, m_2)\right) \right]^2 \right\} \tag{3.57} $$

is added for every wave. In the exponent, the barrier factors F_L and the breakup momenta q, as defined in sections 3.3.1 and 3.3.3, appear.

By writing the differences between the resonance model $\hat{\rho}^\epsilon_{\alpha\beta}(m)$ and the extracted spin-density matrix $\rho^\epsilon_{\alpha\beta}(m)$ as vectors

$$ \delta^\epsilon_{\alpha\beta}(m) = \begin{pmatrix} \text{Re}\left[\hat{\rho}^\epsilon_{\alpha\beta}(m) - \rho^\epsilon_{\alpha\beta}(m)\right] \\ \text{Im}\left[\hat{\rho}^\epsilon_{\alpha\beta}(m) - \rho^\epsilon_{\alpha\beta}(m)\right] \end{pmatrix} \tag{3.58} $$

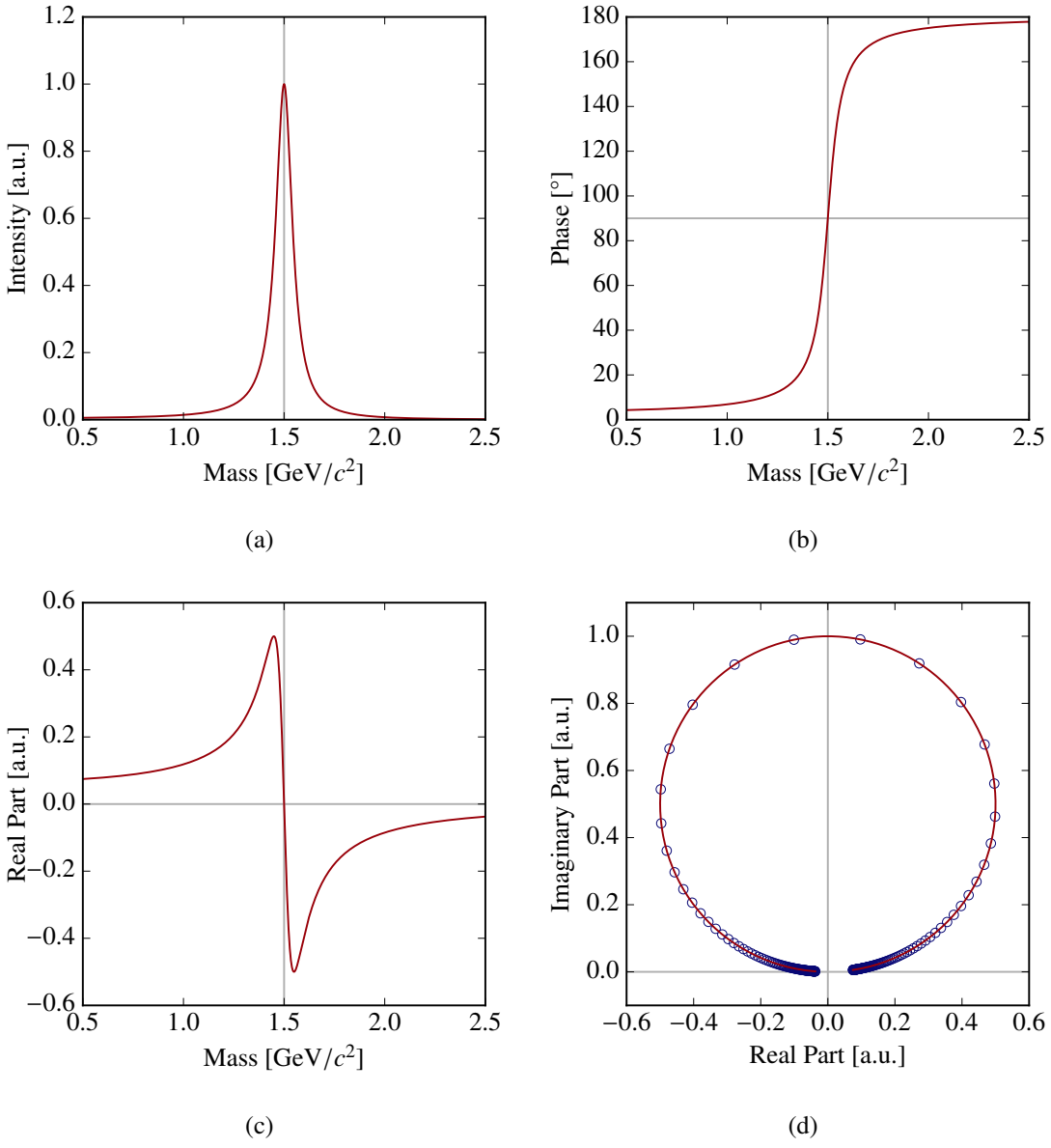


Figure 3.6: The Breit-Wigner amplitude from equation (3.56) for a hypothetical resonance with $m_0 = 1.5\,\mathrm{GeV}/c^2$ and $\Gamma_0 = 0.1\,\mathrm{GeV}/c^2$. The points in (d) are spaced in mass by $10\,\mathrm{MeV}/c^2$.

one can formulate the χ^2 function

$$\chi^2 = \sum_{m_i} \sum_{\epsilon=\pm1} \sum_{\alpha,\beta\leq\alpha} \left[\delta^{\epsilon}_{\alpha\beta}(m_i)\right]^T \mathfrak{E}^{\epsilon\ -1}_{\alpha\beta}(m_i)\, \delta^{\epsilon}_{\alpha\beta}(m_i) \tag{3.59}$$

with the 2×2 covariance matrix $\mathfrak{E}^{\epsilon}_{\alpha\beta}$ for the real and imaginary part of every element of the spin-density matrix, which is calculated from the covariance matrix of the productions amplitudes ${}^{r}T^{\epsilon}_{\alpha}$ obtained from the partial-wave decomposition in mass bins. Because the spin-density matrix is a Hermitian matrix, only the lower triangle of it is included in the sum. The squared Pearson residuals of all mass bins with masses m_i are then summed up.

By minimizing the χ^2 function from equation (3.59), the resonance parameters $m_{0,k}$, $\Gamma_{0,k}$, the couplings ${}^{r}C^{\epsilon}_{\alpha k}$ and the non-resonant parameter g_k are then fitted to the production amplitudes.

Several caveats of this method should be mentioned. Because the parameters for many decay modes are not known, the Breit-Wigner amplitudes from equation (3.56) which are fitted to the spin-density (sub)matrix do not take into account the growth of the phase space over the width of the resonance, as shown in equations (3.43) and (3.47). This might slightly change the values of the extracted parameters. Another simplification is that, in addition to the uncertainties of the real and imaginary parts themselves, the χ^2 function includes only the correlations between the real and imaginary parts of the same amplitudes. No correlations between the amplitudes of different partial waves are taken into account, even though this information would be available from the partial-wave decomposition. The reason for this, apart from increasing complexity, is that for $N^{\epsilon}_{r} < N^{\epsilon}_{s}$, there are more spin-density matrix elements than production amplitudes, leading to covariance matrices for the spin-density matrix with zero eigenvalues so that the uncertainties on the fit parameters lose meaning. Additionally, in the formulation in equation (3.59), Gaussian uncertainties are assumed, which, as is shown in figure 4.2 on page 47, is not a good approximation. There are efforts underway to improve these areas [29, 30], but as a precision extraction of resonance parameters is not the goal of this work, including them in the analysis presented was not deemed warranted.

CHAPTER 4

Model Selection

"Pouf, cinq!"

Hugh Laurie

In the previous chapter, at the end of section 3.1, the dimension N_s of the spin-density matrix has been introduced. It is the number of combinations of X quantum numbers and corresponding decay modes which are considered in the partial-wave decomposition, i.e. the number of waves the cross section is decomposed into. So far, however, neither the actual value of N_s nor the choice of waves included in $\rho_{\alpha\beta}$ have been discussed.

In principle, there is an infinite number of possible intermediate states X. However, even if it was possible to use an infinite number of waves in the partial-wave decomposition, it would not be a sound approach because the number of measured events is limited. It is possible to describe any measured distribution perfectly if the number of parameters of the describing model is equal or higher than the number of measured data points. However, such a description would have no predictive power and would therefore be without use.

On the other hand, too small a selection of partial waves will also have negative impact. In the best case, waves which are contributing significantly to the measured data but are omitted in the decomposition will lead to part of the data not being described well. In the worst case, events will be wrongly attributed to other waves, a phenomenon called model leakage.

In the end, the task of model selection is to find the best possible description of the data with the minimal number of parameters, which is what this chapter will explore. For the purpose of this discussion, let us call a set of waves used for partial-wave decomposition wave set or model.

4.1 The Wave Pool

A model is constructed by selecting waves from a pool of all possible waves to form a wave set. The wave pool should contain all the waves for which a significant contribution to the data is conceivable.

To avoid bias, the wave pool should be constructed systematically from a set of rules which are motivated by physics considerations. In practice, the rules are also influenced by practical reasons, as the wave pool size is an important factor in the complexity of the model-selection process.

The rules for constructing the wave pool are (1) limits for the spin J of the states X, its spin projection M and its reflectivity ϵ, (2) limits for the spin S of the isobar and the orbital angular momentum L of the two-body decays and (3) the set of isobars which are included and their allowed decay modes. Following the argumentation in section 3.1, only states with $I^G = 1^-$ are allowed.

Limiting J, S and L is motivated by the fact that higher angular momenta require more energy to produce them, so that they are suppressed at lower masses. For the current analysis, we use $L \leq 2$ while not limiting J and S. Because the spin of the isobars is limited and because of angular momentum conservation, the highest J than can appear is 6 and the highest S is 4. In diffractive reactions, there is a kinematic suppression of higher M values, which is why only partial waves with $M \leq 1$ are used. Also, as reflectivity coincides with naturality at COMPASS energies, negative reflectivity is expected to be suppressed, so we require $\epsilon = +1$.

Selecting the isobars and their decay modes is a difficult task, because there are a number of mesons for which their existence, nature and resonance parameters are not well known. This is particularly problematic for states decaying into four pions. For this analysis, it was decided to use only isobars for which at least the existence was generally well established. While the decay modes for three-pion isobars, i.e. isobars decaying into three pions, are generally well known, the four-body isobars are again problematic. For most of them, the decay modes are not known at all and several judgment calls had to be made in specifying them. The complete list of isobars used to construct the wave pool, together with their Breit-Wigner parameters and decay modes is given in table 4.1. Note that the decay $\rho_3(1690) \to \pi^+ + \pi^-$ is allowed according to the table, but is excluded by the $L \leq 2$ limit above.

Using the rules above, a wave pool consisting of 1546 waves is produced.

Table 4.1: Isobars used in the analysis with their Breit-Wigner parameters (if applicable) and decay modes.

Isobar	Mass [MeV/c^2]	Width [MeV/c^2]	Decay Products	
$\sigma^{[*]}$			π	π
$\rho(770)$	769	150.9	π	π
$f_0(980)^{[*]}$			π	π
$b_1(1235)$	1229.5	142	$\rho(770)$	σ
			$a_1(1260)$	π
$a_1(1260)$	1230	425	$\rho(770)$	$\pi^{[\dagger]}$
			σ	π

Continued on next page

[*] Not parameterized by a Breit-Wigner function, see section 3.3.3

[†] Only the decay mode with $L = 0$ is allowed

Table 4.1: Isobars used in the analysis with their Breit-Wigner parameters (if applicable) and decay modes (continued).

Isobar	Mass [MeV/c^2]	Width [MeV/c^2]	Decay Products	
$f_2(1270)$	1275.1	185.1	π	π
			$\rho(770)$	$\rho(770)^{[\dagger]}$
			σ	σ
			$a_1(1260)$	π
$f_1(1285)$	1282.1	24.2	$\rho(770)$	$\rho(770)^{[\dagger]}$
			$a_1(1260)$	π
$\pi(1300)$	1300	400	$\rho(770)$	π
			σ	π
$a_2(1320)$	1318.3	107	$\rho(770)$	π
$f_0(1370)$	1350	350	$\rho(770)$	$\rho(770)^{[\dagger]}$
			σ	σ
			$a_1(1260)$	π
			$\pi(1300)$	π
$f_1(1420)$	1426.4	54.9	$a_1(1260)$	π
			$a_2(1320)$	π
$\rho(1450)$	1465	400	$\rho(770)$	σ
			$a_1(1260)$	π
			$\pi(1300)$	π
$f_0(1500)$	1505	109	$\rho(770)$	$\rho(770)^{[\dagger]}$
			σ	σ
			$a_1(1260)$	π
			$\pi(1300)$	π
$f_2(1565)$	1562	134	$\rho(770)$	$\rho(770)$
			σ	σ
			$f_2(1270)$	σ
			$a_1(1260)$	π
			$\pi(1300)$	π
			$a_2(1320)$	π
$\eta_2(1645)$	1617	181	$\rho(770)$	$\rho(770)$
			$f_2(1270)$	σ
			$a_1(1260)$	π
			$a_2(1320)$	π

[*] Not parameterized by a Breit-Wigner function, see section 3.3.3

[†] Only the decay mode with $L = 0$ is allowed

Continued on next page

Table 4.1: Isobars used in the analysis with their Breit-Wigner parameters (if applicable) and decay modes (continued).

Isobar	Mass [MeV/c^2]	Width [MeV/c^2]	Decay Products	
$\pi_2(1670)$	1672.2	260	$\rho(770)$	π
			σ	π
			$f_2(1270)$	π
$\rho_3(1690)$	1688.8	161	π	π
			$\rho(770)$	σ
			$f_2(1270)$	$\rho(770)^{[\dagger]}$
			$a_1(1260)$	π
			$a_2(1320)$	π
$\rho(1700)$	1720	250	$\rho(770)$	σ
			$a_1(1260)$	π
			$\pi(1300)$	π
$\pi(1800)$	1812	208	$f_0(980)$	π
			σ	π

[*] Not parameterized by a Breit-Wigner function, see section 3.3.3
[†] Only the decay mode with $L = 0$ is allowed

4.2 Genetic Algorithm

Genetic algorithms, a subset of evolutionary algorithms, are a class of optimization algorithms. In [10], such an algorithm was applied to COMPASS 5π data measured in 2004. While the results looked promising, it was decided to thoroughly test the method before applying it to the much larger data set from 2008. Unfortunately, the result was that such algorithms are not suitable for model selection in this context.

Having already stated the conclusion, this section will be kept brief. Derivations already explained in [10] will not be repeated. Instead, the focus will lay on highlighting the adaptations and improvements developed in the course of this work. Rather than detailing all the exhaustive investigations, only the final results contributing to the conclusion will be given. However, the description of the algorithms used to arrive at this conclusion will be thorough.

4.2.1 Evidence

In the Bayesian formalism, the quality of two competing models can be compared using the evidence

$$\mathcal{E}_k = \int \mathrm{d}\mathbf{T}_k \, \mathcal{L}(\mathbf{T}_k) \, \Pi(\mathbf{T}_k) \tag{4.1}$$

is used. The likelihood $\mathcal{L}(\mathbf{T})$ as defined in equation (3.32) appears, albeit not in logarithmic form. The vector

$$\mathbf{T}_k = \begin{pmatrix} \mathrm{Re}\left[{}^1T_1^1\right] \\ \mathrm{Re}\left[{}^1T_2^1\right] \\ \mathrm{Im}\left[{}^1T_2^1\right] \\ \vdots \\ \mathrm{Re}\left[{}^1T_{N_s^k}^1\right] \\ \mathrm{Im}\left[{}^1T_{N_s^k}^1\right] \\ T_{\text{flat}} \end{pmatrix}_k \tag{4.2}$$

contains all the free parameters of the model, i.e. the production amplitudes ${}^rT_i^\epsilon$ introduced in the previous chapter. As discussed after equation (3.11) and in section 4.1, respectively, we limited ourselves to rank $r = 1$ and reflectivity $\epsilon = 1$. From here on, we will therefore drop the rank and reflectivity indices and only write T_i. Note that in equation (4.2), we jump directly from $\mathrm{Re}(T_1)$ to $\mathrm{Re}(T_2)$ because, according to equation (3.5), $\mathrm{Im}(T_1) = 0$ and is therefore not a free parameter. This means that $\dim(\mathbf{T}_k) = 2N_s^k$. The index k denotes the model we are investigating, i.e. the selection of partial waves in our wave set, and N_s^k is the number of waves in the set. The prior $\Pi(\mathbf{T}_k)$ is the probability density function for $\mathbf{T}_k$ before taking into account the measured data. It provides a means to include preexisting knowledge into the parameter estimation and as such, there is some freedom in choosing it. The actual choice of $\Pi(\mathbf{T}_k)$ will be discussed later in this section.

Note that in the derivation of the likelihood in section 3.2, several terms not depending on T_i have been dropped. This implies that the actual value of $\mathcal{E}_k$ does not carry meaning. Instead, only the Bayes factor

$$B_{kl} = \frac{\mathcal{E}_k}{\mathcal{E}_l} \tag{4.3}$$

is of significance for the comparison of two models k and l. While there are recipes for interpreting the value of B_{kl} [31], the genetic algorithm only uses B_{kl} to determine if model k or l is preferred. Therefore, this topic will not be discussed here.

The integral in equation (4.1) is expensive to calculate: it is high-dimensional, with $\dim(\mathbf{T}_k) = O(100)$, and, because $\mathcal{L}$ appears instead of $\ln \mathcal{L}$, numerically challenging. With nested sampling[a] [35], an algorithm exists that should be able to calculate such evidence integrals. Unfortunately, it turned out that it was unable to cope with the large number of parameters in typical models and did not converge within a practical timescale.

Therefore, a different approach was chosen to calculate the evidence. In the limit of a large data set, we can assume that the likelihood is strongly peaked at the maximum-likelihood estimate $\widetilde{\mathbf{T}}_k$ for the production amplitudes. If we assume that this peak has the form of a multivariate Gaussian and that

[a] Here, the MultiNest implementation [32–34] is used.

the prior is constant, the integral can be solved analytically[b] [36] and we obtain

$$\mathcal{E}_k \approx \mathcal{L}\left(\widetilde{\mathbf{T}}_k\right) \Pi\left(\widetilde{\mathbf{T}}_k\right) \sqrt{(2\pi)^{2N_s^k} \det(\mathfrak{E}_k)} \tag{4.4}$$

where we used the covariance matrix $\mathfrak{E}_k$ at the maximum-likelihood estimate. This approximation is called Laplace's method or saddle-point approximation. By taking the logarithm of the evidence, we can transition from $\mathcal{L}$ to $\ln \mathcal{L}$, which solves the problem of numerical stability.

For $\Pi(\mathbf{T}_k)$, we choose a minimum-bias prior

$$\Pi(\mathbf{T}_k) = \begin{cases} \dfrac{1}{V_{\mathbf{T}_k}}, & \text{if } \sum\limits_{i,j} T_i\, T_j^*\, \mathcal{N}_{ij} + T_{\text{flat}}^2 \mathcal{A} = N \\ 0, & \text{otherwise} \end{cases} \tag{4.5}$$

which is constant if the production amplitudes are valid, i.e. if the number N of measured events is equal to the number $\overline{N}$ of expected events as expressed in equation (3.22) and using equations (3.23), (3.27) and (3.31). Note that in the formulation of the condition, we write $T_{i,j}$ without explicitly adding the index k for the model. As single production amplitudes of different models never appear together, we drop the index k for better readability. The volume $V_{\mathbf{T}_k}$ is chosen such that

$$\int \mathrm{d}\mathbf{T}_k\, \Pi(\mathbf{T}_k) = \frac{1}{V_{\mathbf{T}_k}} \int\limits_{\Omega_k} \mathrm{d}\mathbf{T}_k \equiv 1 \tag{4.6}$$

$$\Leftrightarrow$$

$$V_{\mathbf{T}_k} = \int\limits_{\Omega_k} \mathrm{d}\mathbf{T}_k \tag{4.7}$$

For now, we define Ω_k as the hypersurface in the space of production amplitudes where the condition in equation (4.5) is true. This is a rather sloppy definition, because $\mathrm{d}\mathbf{T}_k \in \mathbb{R}^{2N_S^k} \supset \Omega_k$ in equation (4.6), but $T_i \in \mathbb{C}$ in equation (4.5). During the calculation of $V_{\mathbf{T}_k}$, a better definition will present itself and we will remedy the situation.

Solving equation (4.7) is equivalent to determining the volume of Ω_k. In [10], this was achieved by assuming

$$N \approx \mathcal{A} \left[\sum_i \left|T_i\right|^2 + T_{\text{flat}}^2 \right] \tag{4.8}$$

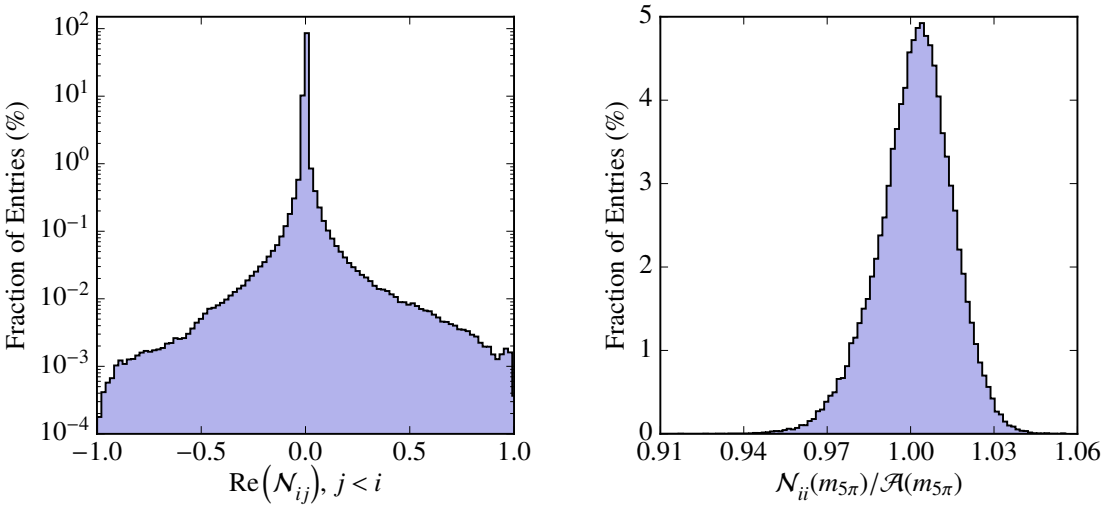
[b] Approximating the likelihood function with a Gaussian is equivalent to writing

$$\mathcal{L}(\mathbf{T}) \approx \mathcal{L}(\widetilde{\mathbf{T}})\, \exp\left[(\mathbf{T} - \widetilde{\mathbf{T}})^T \mathfrak{E}\, (\mathbf{T} - \widetilde{\mathbf{T}})\right]$$

where we used the covariance matrix $\mathfrak{E}$. Using this in the integral from equation (4.1) and assuming that the prior is constant, i.e. $\Pi(\mathbf{T}) \equiv \Pi(\widetilde{\mathbf{T}})$, we obtain

$$\begin{aligned} \mathcal{E} &\approx \mathcal{L}(\widetilde{\mathbf{T}})\, \Pi(\widetilde{\mathbf{T}}) \int \mathrm{d}\mathbf{T}\, \exp\left[(\mathbf{T} - \widetilde{\mathbf{T}})^T \mathfrak{E}\, (\mathbf{T} - \widetilde{\mathbf{T}})\right] \\ &= \mathcal{L}(\widetilde{\mathbf{T}})\, \Pi(\widetilde{\mathbf{T}})\, \sqrt{(2\pi)^{2N_s} \det(\mathfrak{E})} \end{aligned}$$

where the integral is the inverse of the well-known normalization constant of a multivariate Gaussian distribution. As pointed out in [36], this is equivalent to approximating $\ln \mathcal{L}$ by a Taylor series up to the quadratic term.



(a) Off-diagonal real parts of $\mathcal{N}_{ij}$, $j < i$.

(b) Relative variation of $\mathcal{N}_{ii}$ with respect to $\mathcal{A}$.

Figure 4.1: Histograms of the elements of the integral matrix $\mathcal{N}_{ij}$ for the whole 5π mass region and the whole wave pool.

which is the equation of the surface of a $2N_s^k$-dimensional sphere, the volume of which is easily calculated. Two assumptions enter in equation (4.8): the off-diagonal elements of the integral matrix $\mathcal{N}_{ij}$ are assumed to all be zero and the diagonal elements $\mathcal{N}_{ii}$ are assumed to all be equal to the total acceptance $\mathcal{A}$. Figure 4.1a shows the normalized distribution of $\mathrm{Re}(\mathcal{N}_{ij})$ for all $j < i$ for 40 mass bins ranging from 1.36 to 2.56 GeV/c^2 for the whole wave pool. Most of the off-diagonal elements are indeed small, however, there is also a small fraction with values significantly different from zero. The second assumption allows to pull $\mathcal{A}$ in front of the sum and then dropped entirely, because it does not depend on the model and therefore cancels in the Bayes factor. Figure 4.1b shows the normalized distribution of the diagonal elements $\mathcal{N}_{ii}$. To remove their $m_{5\pi}$ dependence, the diagonal elements of every mass bin have been normalized by the acceptance in this bin. One can see that there is a spread of several percent. In combination, one can expect an effect in the order of several percent of $V_{\mathbf{T}_k}$ because of the assumption (4.8). Therefore, improving the calculation of $V_{\mathbf{T}_k}$ seemed justified.

Let us start by noting that $\mathcal{N}_{ij}$ is as Hermitian matrix by construction. Because all partial-wave decay amplitudes Ψ_i are linearly independent and the integration in equation (3.23) is an inner product, $\mathcal{N}_{ij}$ is a Gram matrix and therefore positive definite. While the condition in equation (4.5) is formulated in terms of $\mathcal{N}_{ij,k} \in \mathbb{C}^{N_s^k \times N_s^k}$ and $T_{i,k} \in \mathbb{C}^{N_s^k}$, we are interested in a formulation using $\mathbf{T}_k \in \mathbb{R}^{2N_s^k}$. To that end, we first consider a vector $\mathbf{V} \in \mathbb{C}^N$ and a Hermitian, positive-definite matrix $\mathbf{M} \in \mathbb{C}^{N \times N}$ and

write them as

$$\mathbf{V} = \mathbf{a} + i\,\mathbf{b} = \begin{pmatrix} a_1 + i\,b_1 \\ \vdots \\ a_N + i\,b_N \end{pmatrix}$$
$$\mathbf{M} = \mathbf{m} + i\,\mathbf{n} = \begin{pmatrix} m_{11} + i\,n_{11} & \dots & m_{1N} + i\,n_{1N} \\ \vdots & & \vdots \\ m_{N1} + i\,n_{N1} & \dots & m_{NN} + i\,n_{NN} \end{pmatrix} \tag{4.9}$$

Using Einstein notation for the following part, we can see that

$$\begin{aligned} \mathbf{V}^*\mathbf{M}\,\mathbf{V} &= (a_k - i\,b_k)\,(m_{kl} + i\,n_{kl})\,(a_l + i\,b_l) \\ &= a_k\,m_{kl}\,a_l + b_k\,m_{kl}\,b_l - a_k\,n_{kl}\,b_l + b_k\,n_{kl}\,a_l \\ &\qquad + i\left[\underbrace{a_k\,n_{kl}\,a_l}_{=0} + \underbrace{b_k\,n_{kl}\,b_l}_{=0} + \underbrace{a_k\,m_{kl}\,b_l - b_k\,m_{kl}\,a_l}_{=0}\right] \\ &= \begin{pmatrix} \mathbf{a} & \mathbf{b} \end{pmatrix} \begin{pmatrix} \mathbf{m} & -\mathbf{n} \\ \mathbf{n} & \mathbf{m} \end{pmatrix} \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix} \\ &\equiv \hat{\mathbf{V}}^T \hat{\mathbf{M}}\,\hat{\mathbf{V}} \end{aligned} \tag{4.10}$$

The terms $a_k\,n_{kl}\,a_l$ and $b_k\,n_{kl}\,b_l$ are zero because $n_{kl} = -n_{lk}$ and $n_{kk} = 0$ for Hermitian matrices. For the same reason, $m_{kl} = m_{lk}$ and the two other terms in the imaginary parts cancel each other. Since $\mathbf{V}^*\mathbf{M}\mathbf{V} > 0$ for all $\mathbf{V} \neq 0$[c], the same is true for $\hat{\mathbf{V}}^T\hat{\mathbf{M}}\,\hat{\mathbf{V}}$, which is equivalent to $\hat{\mathbf{M}}$ being positive definite as well. Because $\mathbf{m}$ is symmetric and $\mathbf{n}$ is antisymmetric, $\hat{\mathbf{M}}$ is symmetric, so that one can write its Cholesky decomposition

$$\hat{\mathbf{V}}^T\hat{\mathbf{M}}\,\hat{\mathbf{V}} = \hat{\mathbf{V}}^T\hat{\mathbf{D}}^T\hat{\mathbf{D}}\,\hat{\mathbf{V}} = \left(\hat{\mathbf{D}}\,\hat{\mathbf{V}}\right)^T \hat{\mathbf{D}}\,\hat{\mathbf{V}} \tag{4.11}$$

We now reorder the vector $\mathbf{T}_k$ of production amplitudes from equation (4.2), and define $\boldsymbol{\mathcal{N}}_k$ such that

$$\mathbf{T}_k = \begin{pmatrix} \mathrm{Re}\left(T_1\right) \\ \mathrm{Re}\left(T_2\right) \\ \vdots \\ \mathrm{Re}\left(T_{N_s^k}\right) \\ \mathrm{Im}\left(T_2\right) \\ \vdots \\ \mathrm{Im}\left(T_{N_s^k}\right) \\ T_{\mathrm{flat}} \end{pmatrix}_k \qquad \text{and} \qquad \boldsymbol{\mathcal{N}}_k = \begin{pmatrix} \mathrm{Re}\left(\mathcal{N}_{ij}\right) & -\mathrm{Im}\left(\mathcal{N}_{ij}\right)^{\ddagger} & \vdots \\ \mathrm{Im}\left(\mathcal{N}_{ij}\right)^{\ddagger} & \mathrm{Re}\left(\mathcal{N}_{ij}\right)^{\ddagger} & 0 \\ \cdots \quad 0 \quad \cdots & & \mathcal{A} \end{pmatrix}_k \equiv \mathbf{L}_k^T\mathbf{L}_k \tag{4.12}$$

[c] This is equivalent to $\mathbf{M}$ being positive definite.

Because $\mathrm{Im}(T_1)$ does not appear in $\mathbf{T}_k$, care has to be taken to remove the corresponding rows and columns in the construction of $\boldsymbol{\mathcal{N}}_k$. To that end, we write $-\mathrm{Im}(\mathcal{N}_{ij})^{\ddagger}$, where the column $j = 1$ has been removed. In $\mathrm{Im}(\mathcal{N}_{ij})^{\ddagger}$, the row $i = 1$ and in $\mathrm{Re}(\mathcal{N}_{ij})^{\ddagger}$, both the row $i = 1$ and the column $j = 1$ are dropped. Using equations (4.9) to (4.11), we reformulate the condition from equation (4.5)

$$N = \sum_{i,j} T_i T_j^* \mathcal{N}_{ij} + T_{\text{flat}}^2 \mathcal{A} = \mathbf{T}^T \boldsymbol{\mathcal{N}}\, \mathbf{T} \tag{4.13}$$

which also allows us to properly define

$$\Omega_k = \left\{ \mathbf{T}_k \in \mathbb{R}^{2N_s^k} : \mathbf{T}_k^T \boldsymbol{\mathcal{N}}_k\, \mathbf{T}_k = N \right\} \tag{4.14}$$

and then finally calculate

$$\begin{aligned} V_{\mathbf{T}_k} &= \int_{\Omega_k} \mathrm{d}\mathbf{T}_k \\ &= \int \mathrm{d}\mathbf{T}_k\, \delta\left(\mathbf{T}_k^T \boldsymbol{\mathcal{N}}_k\, \mathbf{T}_k - N\right) \\ &= \int \mathrm{d}\mathbf{T}_k\, \delta\left(\left[\mathbf{L}_k \mathbf{T}_k\right]^T \mathbf{L}_k \mathbf{T}_k - N\right) \end{aligned} \tag{4.15}$$

Performing the variable transformation

$$\begin{aligned} \mathbf{T}_k \quad &\rightarrow \quad \hat{\mathbf{T}}_k \;= \mathbf{L}_k^{-1}\, \mathbf{T}_k \\ \mathrm{d}\mathbf{T}_k \quad &\rightarrow \quad \mathrm{d}\hat{\mathbf{T}}_k = \mathrm{d}\mathbf{T}_k \det(\mathbf{L}_k) \end{aligned} \tag{4.16}$$

yields

$$\begin{aligned} V_{\mathbf{T}_k} &= \int \mathrm{d}\hat{\mathbf{T}}_k \left[\det(\mathbf{L}_k)\right]^{-1} \delta\left(\hat{\mathbf{T}}_k^T \hat{\mathbf{T}}_k - N\right) \\ &= \left[\det(\mathbf{L}_k)\right]^{-1} \int \mathrm{d}\hat{\mathbf{T}}_k\, \delta\left(\hat{\mathbf{T}}_k^T \hat{\mathbf{T}}_k - N\right) \\ &= \left[\det(\mathbf{L}_k)\right]^{-1} \cdot S_{\text{sphere}}\left(R = \sqrt{N}; d = 2N_s^k\right) \end{aligned} \tag{4.17}$$

Because

$$\begin{aligned} \det\left(\boldsymbol{\mathcal{N}}_k\right) &= \det\left(\mathbf{L}_k^T \mathbf{L}_k\right) \\ &= \det\left(\mathbf{L}_k^T\right) \det\left(\mathbf{L}_k\right) \\ &= \left[\det\left(\mathbf{L}_k\right)\right]^2 \end{aligned} \tag{4.18}$$

we arrive at the final expression

$$V_{\mathbf{T}_k} = \left[\det(\boldsymbol{\mathcal{N}}_k)\right]^{-1/2} \cdot S_{\text{sphere}}\left(R = \sqrt{N}; d = 2N_s^k\right) \tag{4.19}$$

where

$$S_{\text{sphere}}(R;d) = \frac{2\pi^{d/2}}{\Gamma\left(\frac{d}{2}\right)} R^{d-1} \tag{4.20}$$

is the surface of the d-dimensional sphere.

In [10], another factor

$$s_i = \int_{5\sqrt{\mathfrak{E}_{ii}}}^{\infty} \mathrm{d}x \, \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x - \left|T_i\right|^2}{2\,\mathfrak{E}_{ii}}\right] \tag{4.21}$$

was added to the evidence for every wave, in order to introduce a penalty for waves which have intensities compatible with zero. This included the flat wave, leading to the situation that a larger background component in the description of the data led to a better evidence, which is not the desired behavior. While the s_i factors were kept in this analysis, it was not included for the flat wave, i.e. $s_{\text{flat}} \equiv 1$.

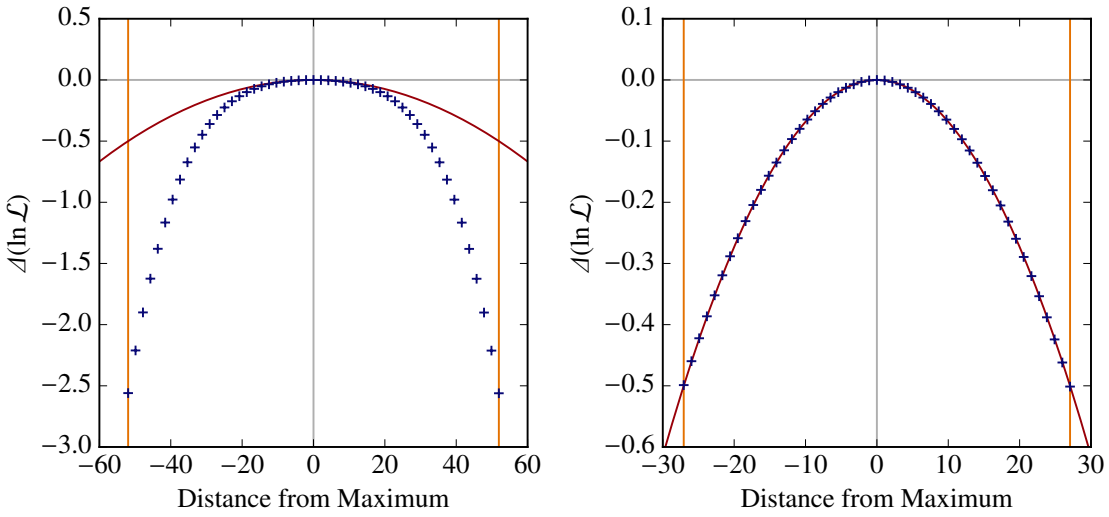
Putting equations (4.4), (4.5), (4.19) and (4.21) together, the log-evidence is

$$\ln \mathcal{E}_k = \ln \mathcal{L}\left(\widetilde{\mathbf{T}}_k\right) + \frac{1}{2} \ln\left[\det\left(\boldsymbol{\mathcal{N}}_k\right)\right] - \ln S^k_{\text{sphere}} + \ln \sqrt{(2\pi)^{2N_s^k} \det(\mathfrak{E}_k)} + \sum_i \ln s_i \tag{4.22}$$

This formula was derived under the assumption that the shape of the likelihood is approximately Gaussian. This assumption can be tested: first, one diagonalizes the covariance matrix, which results in a basis where the eigenvalues of $\mathfrak{E}_k$ are locally uncorrelated variances in the direction of the eigenvectors. Then, one can calculate the values of the likelihood along the direction of the eigenvectors and compare them to a Gaussian with the corresponding variance. Figure 4.2 shows such a comparison, for the biggest and second-biggest eigenvalue. For the second-biggest eigenvalue, the Gaussian approximation works very well and the same is true for all smaller eigenvalues. However, for the biggest eigenvalue, the likelihood shows a bathtub-like shape, which is not well described by a Gaussian. In the example shown, this leads to an overestimation of the evidence. Even though only one eigenvalue is affected, because it is the biggest one, the effect will not be negligible. It is not easy to estimate to what extend this invalidates the conclusions drawn from the comparison of the evidences of two different models, but it seems likely that, if the models have similar evidence, their ranking could be swapped. In this light, the evidence as written in equation (4.22) should only be used with caution.

Independent of this, there is an additional problem of different nature: in the implementation of the improved formula for the parameter volume $V_{\mathbf{T}}$ from equation (4.19), $\det(\mathcal{N}_{ij})$ was used instead of $\det(\boldsymbol{\mathcal{N}})$, i.e. the determinant of the matrix $\mathcal{N}_{ij}$ with complex entries instead of the matrix $\boldsymbol{\mathcal{N}}$ with real entries without the column and row corresponding to $\mathrm{Im}(T_1)$. Assuming that the off-diagonal elements of $\mathcal{N}_{ij}$ are small, the inclusion of this row and column leads to an additional factor in the determinant which has the same spread as shown in figure 4.1b, i.e. an error of a few percent is introduced. In addition[d], $\det(\boldsymbol{\mathcal{N}}) = \det(\mathcal{N}_{ij})^2$ meaning a factor $1/2$ was omitted in equation (4.22).

[d] This relation is only valid if the row and column corresponding to $\mathrm{Im}(T_1)$ are included in $\boldsymbol{\mathcal{N}}$. The derivation of this relation is rather lengthy and has been moved to appendix A.



(a) (b)

Figure 4.2: Slice through the likelihood surface along the direction of the eigenvector of the covariance matrix $\mathfrak{E}_k$ corresponding to the biggest (a) and second-biggest (b) eigenvalue at the likelihood maximum. The likelihood values are shown as blue crosses, the Gaussian approximation is shown in red and its standard deviation is indicated by the orange lines.

As the effect of the parameter volume on the evidence is a penalty on models with bigger number of parameters, i.e. larger wave sets, this tends to penalize bigger wave sets more than appropriate.

This mistake was only noticed after all tests with the genetic algorithm were already completed. While it might influence which model is found by the algorithm, the problems described in section 4.2.3, which lead to the transition to the better model selection method presented in section 4.3, are independent of the exact formulation of the evidence. Because the advantages of the new method would warrant its use even if the genetic algorithm had not exhibited problems, investing resources to rerun the latter was not deemed justified. Therefore, all results presented here have been obtained using the slightly incorrect formulation of the evidence, with $\det(\mathcal{N}_{ij})$ instead of $\det(\mathcal{N})$ in equation (4.22).

4.2.2 Implementation

The genetic algorithm has evolved considerably during testing with respect to the one described in [10]. It has been reimplemented from scratch with only the basic principles remaining the same. Therefore, we will give an in depth description of the algorithm. All results presented here have been obtained with this version.

The most important change is the departure from an algorithm which tries to produce a common model for all mass bins. The algorithm described here is applied to every mass bin independently.

The central element of the genetic algorithm is the breeding pool. It is a list of the models with the highest evidence encountered by the algorithm. New models are produced from models in the

Quantity	Variable	Value
Number of models	N_{models}	
Number of cluster jobs	N_{jobs}	350
Breeding pool size	N_{BP}	5000
Initial wave sets size mean	μ_{WS}	40
Initial wave sets size sigma	σ_{WS}	10
Ratio crossover/mutation	ρ	0.5
Maximum mutation fraction	l	0.5
Slope parameter for breeding selection	ϵ	10

Table 4.2: Parameters of the genetic algorithm.

breeding pool by either combining ("mating") two models or by altering ("mutating") a model. For this, models are selected from the breeding pool according to a probability distribution which depends on their evidence. The parameters of the genetic algorithm are listed in table 4.2.

At the start, the algorithm fills the breeding pool by randomly selecting waves from the wave pool to form wave sets whose sizes follow a normal distribution with mean μ_{WS} and standard deviation σ_{WS}. Then, the evidence as written in equation (4.22) is calculated. For this, the maximum-likelihood estimate has to be determined for every model, which is the most (computing-)time intensive part of the algorithm. To keep the runtime within practical limits, the maximum-likelihood estimation for the various models is run in parallel on a computing cluster. To use the cluster in an efficient way, the number of jobs being run is kept constant at N_{jobs}. As soon as the evidence has been determined for all models in the initial breeding pool, the algorithm starts breeding new models and calculates the evidence for them. The number of new models is determined by the requirement that N_{jobs} jobs are always being run.

To breed new models, the models in the breeding pool are sorted by their evidence and two models are randomly drawn according to the probability distribution

$$P(i;\epsilon,N_{\text{BP}}) = \left(\frac{1}{i+\epsilon} - \frac{1}{N_{\text{BP}}+\epsilon}\right)\cdot\left[\sum_{k=1}^{N_{\text{BP}}}\left(\frac{1}{k+\epsilon} - \frac{1}{N_{\text{BP}}+\epsilon}\right)\right]^{-1} \tag{4.23}$$

which is a function of the position i of the models in the sorted breeding pool, with $i = 1$ designating the best model. The probability distribution is shown in figure 4.3. The form of the distribution was chosen such that models with higher evidences are preferred considerably while still allowing for the selection of models from the remainder of the breeding pool. Note that the actual value of the evidence does not enter into the selection process.

Once two models are selected for reproduction, they are either mated with each other or mutated separately. Which of these two actions is performed is used is chosen at random, such that on average, the ratio of mating to mutation is equal to ρ.

Mating two models is done by crossover, more precisely with the cut and splice method. The waves in both models are shuffled randomly and one crossover point is selected at random for each model. The distribution of the crossover points is uniform. The models are then cut at the crossover point, the resulting parts are swapped and merged to form two new models. The process is illustrated in figure 4.4. This method for mating models can change their size drastically: if the crossover points happen to lie near the end of both models, one of the new models will contain almost all waves of both parents, while the other will contain only a small number.

When mutating a model, two random numbers N_{add} and N_{drop} are drawn uniformly from $[0, l \cdot N_s]$, where N_s is the number of waves in the model. Then, N_{drop} randomly chosen waves are removed from the model and N_{add} waves are randomly chosen from the wave pool and added to the model. When adding waves, it is ensured that no wave can appear twice in a model. Mutation is performed independently for the two chosen models. Mutation can also change the size of the resulting models, but typically less drastically.

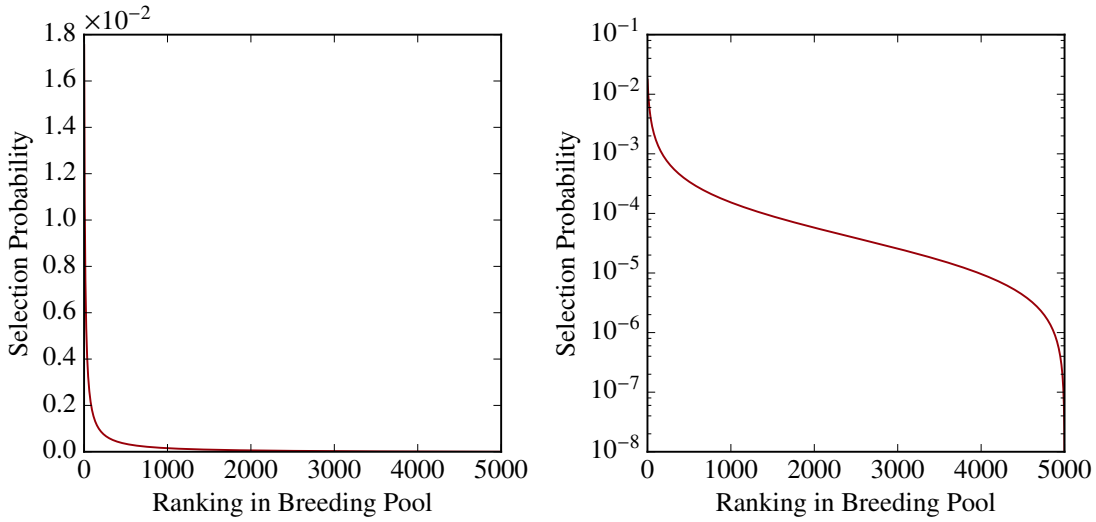
The evidence for the newly created models is then calculated on the computing cluster. If the evidence for a new model turns out to be better than the worst evidence present in the breeding pool, it enters the pool while the worst model in the pool is dropped. It is not forbidden that a newly bred model is identical to one which is already in the breeding pool. If such a situation arises, the evidence for the model is calculated again. If a maximum-likelihood estimate is found which yields a better evidence, which can happen if the likelihood function is multimodal, the pool is updated. It is not possible that the same model appears more than once in the breeding pool.

One can see from equation (3.5) that, for rank 1, there is one wave for which the production amplitude is purely real. This wave is called anchor wave. While in principle, it should not matter which of the waves of the model exhibits this property, it is a practical observation that the result of the maximum-likelihood estimation can be influenced by this choice[e]. To reduce this influence as much as possible, the waves of the newly bred model are sorted by the intensity these waves had in the maximum-likelihood estimation of the model they originated from. If they were added from the wave pool during mutation, the intensity is set to zero in the context of this sorting process. The wave with maximum intensity is then chosen as anchor wave. This ensures a well-defined phase reference.

The genetic algorithm stops after N_{models} have been generated and their evidence calculated. This number differs for the different runs of the genetic algorithm which were performed, but is typically of the order of 10^5 to 10^6.

There are three reasons for departing from the form of the algorithm as used in [10]. First, because resonances are by their nature appearing and disappearing as mass changes, having varying models from mass bin to mass bin was introduced. Second, having the algorithm keep a certain number of jobs running on the computing cluster greatly increases throughput, which allows for longer runs of the algorithm. Third, the change in the reproduction processes was implemented to reduce local convergence, i.e. the situation where the genetic algorithm would get stuck with a breeding pool with low genetic diversity. In this situation, all the models in the pool differ only by insignificant waves

[e] This is connected to the fact that, if the maximum-likelihood estimate for the production amplitude of the anchor wave is very small, the complex phase is not well-determined. This leads to an ambiguity in the likelihood, where a rotation in the complex plane around zero does not change the value of the likelihood, which can influence the result of the estimation.



(a) (b)

Figure 4.3: Breeding selection probability distribution for $\epsilon = 10$ in regular (a) and logarithmic scale (b).

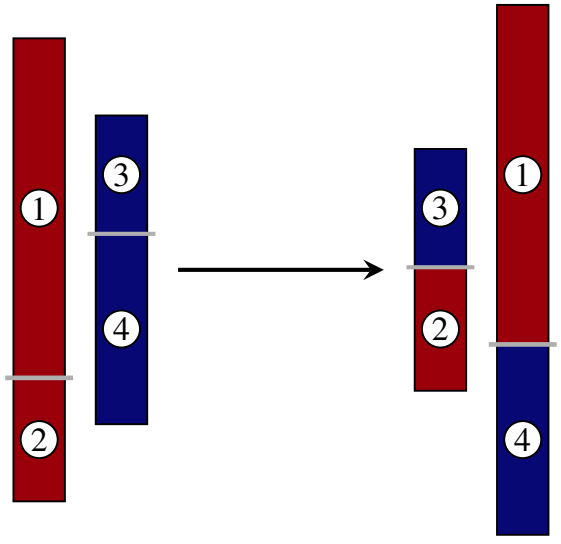


Figure 4.4: Cut and splice crossover mechanism for mating two models.

and have very similar evidence. The more radical crossover mating and higher mutation rate were introduced to make an escape from such a situation more likely. Unfortunately, as we will see in the next section, this did not prove successful.

4.2.3 Verification of Method

To assess the performance of the genetic algorithm, an event sample following a known model was required. Such a test sample was produced by using the intensity from equation (3.13) to weight generated phase-space events using a sample-and-reject procedure. The amplitude normalization given in equation (3.27) was applied. As the intensity itself is not normalized, the weights are normalized to the biggest weight found in the generated event sample. The 65 waves making up the wave set and the corresponding production amplitudes used for generating the event sample are given in appendix B. This is called the input model, which originated from a preliminary run of the algorithm on measured data. The Monte Carlo data sample contains 12 688 events, which is the typical sample size of a single $m_{5\pi}$ bin in the COMPASS 2004 five-pion data set. To facilitate visualization of the model-selection process, every wave has been given an index, with the lowest index designating the wave with the highest intensity and vice versa.

Figure 4.5 shows a run of the genetic algorithm generating $2 \cdot 10^6$ models on the test sample[f]. To plot the progression of the algorithm in time, all the models which were created during the run were ordered by their time of creation. For consistency, always fifty newly created models are combined into one time unit, which is then an analogy to one generation from [10].

As expected, the evidence distribution of newly created models moves towards larger values of the evidence with the passage of in-algorithm time. This is shown in figure 4.5a. However, a considerable amount of new models have evidence values significantly worse than those of the best models. This behavior is a consequence of the procreation mechanisms described in the previous sections and shows that the algorithm does a good job of exploring the model space. The red line in figure 4.5a represents the evidence calculated for the input model on the generated data sample. One can see that the genetic algorithm does not succeed in finding a model that reaches the input model's evidence.

The evolution of the wave set sizes of new models is shown in figure 4.5b. After producing many large models in the early stages of the run, the algorithm settles for wave sets with the appropriate size, i.e. close to 65 waves. The spread of model sizes is considerable, which is expected due to the procreation mechanisms. The transients which are visible are correlated with new waves permeating the breeding pool.

Correlating evidences and wave set sizes of all the models in the run, one obtains at figure 4.5c. It is evident that the maximum evidence is reached if the wave set contains around $60 - 70$ waves, which is the desired and expected behavior.

The degree to which the input model has been found by the algorithm is illustrated by figure 4.5d. The wave index on the y-axis corresponds to the one in table B.1. The linear color scale shows how often a

[f] This corresponds to several 10^5 CPU hours.

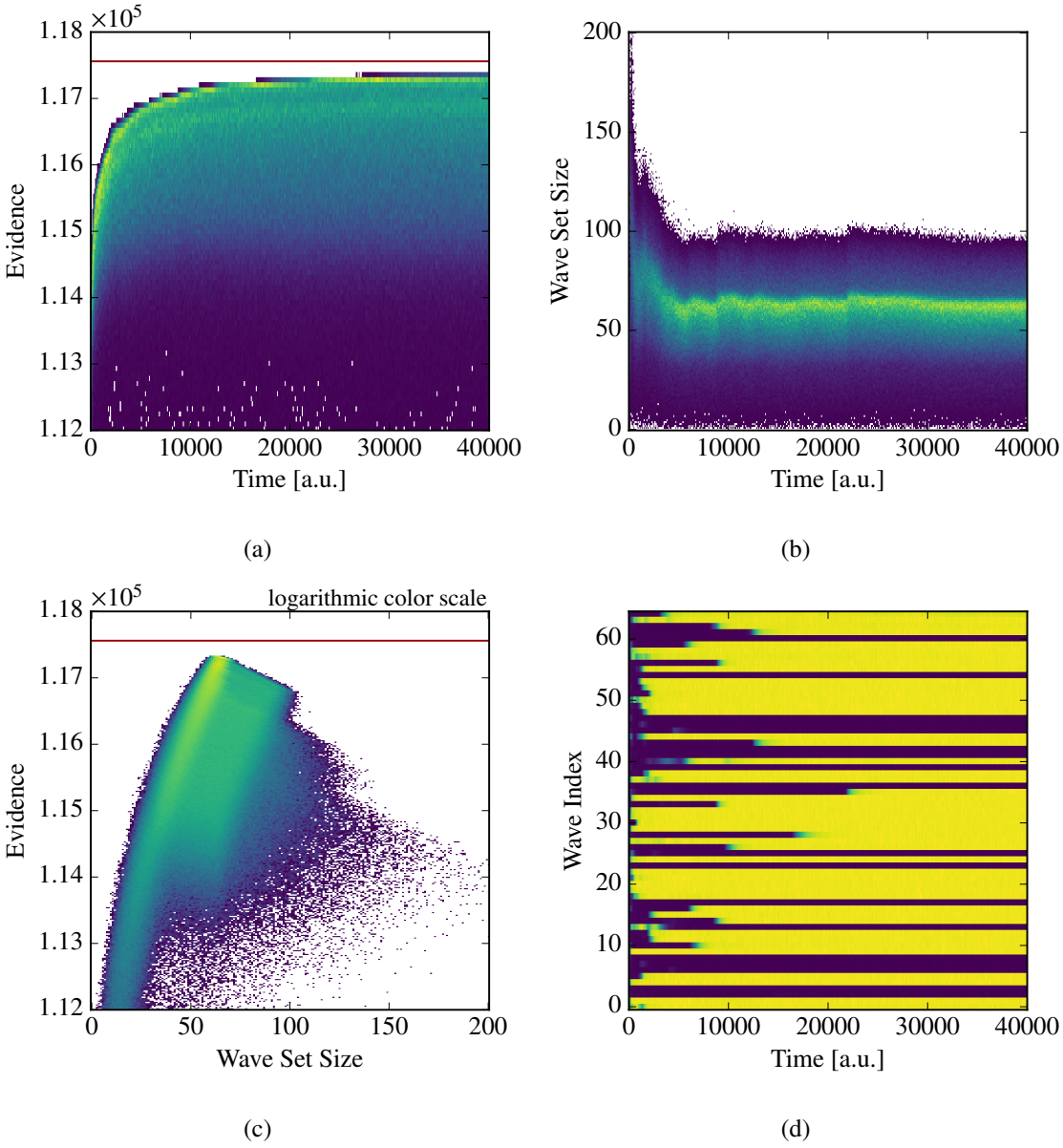


Figure 4.5: Visualization of the genetic algorithm for a $N_{\text{models}} = 2 \cdot 10^6$MLE-long run on Monte Carlo events. See text for details.

wave with a given index appears in one of the models at a given time. A yellow horizontal band means that the corresponding wave is pervasive in the models in the breeding pool, appearing consistently in newly created models and only occasionally being dropped in the procreation processes. In contrast, a violet band indicates that the models in the breeding pool do not contain the wave consistently, so it appears in new models only at random by mutation. Note that this plot only shows the presence of waves which were used in the input model: every model plotted can contain any number of additional waves. Unfortunately, the algorithm fails to pick up several of the larger waves in the input model.

The behavior seen in figure 4.5 points to local convergence: the genetic algorithm has reached a state where the models in the breeding pool are very similar and significant improvements of the evidence are only seen on very long time scales.

To investigate the convergence behavior of the algorithm in more detail, it was started on the same event sample four times, but with different seeds for the random number generator. This means that the initial breeding pool is different in every run. The algorithms were then run until each accumulated $0.4 \cdot 10^6$ models.

Figures 4.6 and 4.7 show how the algorithm picks up waves of the input model and how evidence and wave set size correlate for the four runs[g]. It is evident that the genetic algorithm has a significant dependence on the starting point. This lends further support to the hypothesis of local convergence.

In summary, there is strong evidence that the genetic algorithm exhibits local convergence behavior to such an extent that the input model is not found even when increasing the run time to the practical limit. This leads to the conclusion that the genetic algorithm as presented here is not a suitable tool for model selection in the context of this analysis.

4.3 Biggest-Conceivable-Model Method

The partial-wave decomposition formalism exhibits two interesting properties in the context of model selection.

First, models are built from components, the partial waves. Instead of a number of fixed models, there are a number of fixed model components, i.e. the waves in the wave pool, from which models can be built at will. The genetic algorithm exploited this property already by treating waves similar to genes.

Second, the absolute value of the production amplitude of a wave being zero is equivalent to the wave not being present in the model. This also holds asymptotically, i.e. the smaller the value, the less important the wave. There is an easy way to define a scale for what "small" means in this context: owing to the normalization in equation (3.16), the intensity of a wave, i.e. the absolute square of its production amplitude, is expressed in terms of number of events corrected for acceptance. Therefore, considering waves with intensities of the order of 1 to 10 events and below to be small suggests itself.

[g] The corresponding plots for the evidence and wave set size evolution can be found in appendix C.

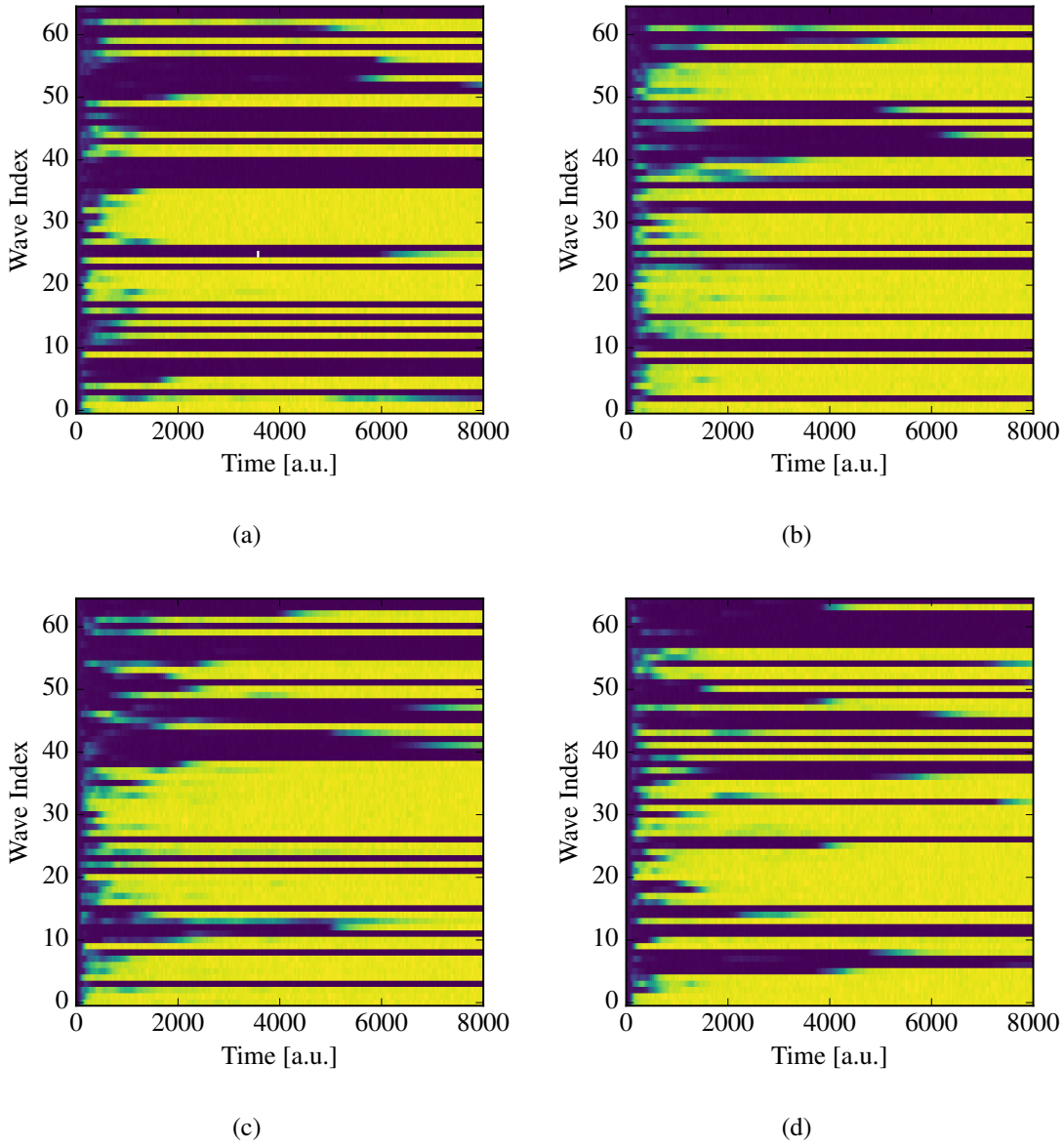


Figure 4.6: Comparison of the wave pickup for four different runs of the genetic algorithm on the same generated data set. See text for details.

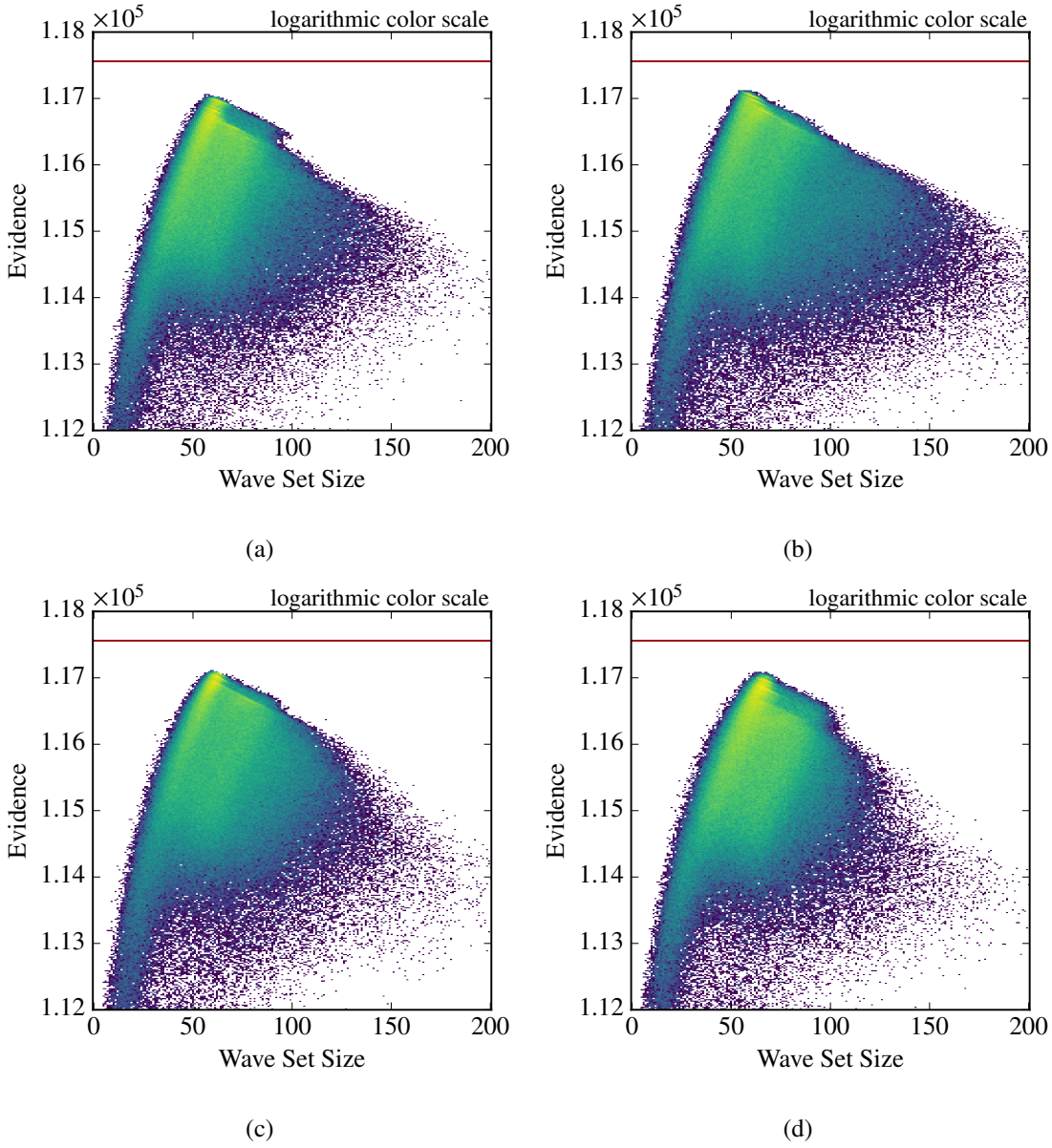


Figure 4.7: Comparison of the evidence vs. wave set size distributions for four different runs of the genetic algorithm on the same generated data set. See text for details.

Together, these two properties can be exploited for model selection. If we treat the whole wave pool as a model, this would be the biggest conceivable model and would, by design, contain all possible models we are able to consider in the selection procedure. Using this model to perform a maximum-likelihood estimation for the production amplitudes, we could then identify the waves in the pool with small intensities, remove them and keep the remaining waves as our selected model.

Apart from the technical considerations of performing maximum-likelihood estimation with over 3000 parameters[h] and assuming that the likelihood function is such that the maximum can be reliably determined, which turns out to be a good assumption, a more fundamental problem remains: due to the large number of free parameters, the point in the parameter space where the likelihood is maximal will invariably be such that not only the underlying distributions present in the event sample are described but also the statistical fluctuations. The problem is exacerbated by the fact that two or more waves with large amplitudes can destructively interfere with each other to describe small features in the data. This is particularly critical as one might not be able to distinguish between waves describing distributions and waves describing fluctuations, because both might exhibit similar intensities. Therefore, a mechanism has to be introduced to counteract this tendency.

One way to accomplish this is to increase the preference for amplitudes which are small, or formulated differently, introduce a penalty for every wave which has significant intensity [44]. This would cause the maximum likelihood estimate to be a compromise between optimal description of the data, including statistical fluctuations, and reducing the number of waves with large intensities. Using a prior function of the form

$$\Pi(\mathbf{T}; \Gamma = 0.5) = \sum_{i}^{N_s} \left(1 + \frac{|T_i|^2}{\Gamma^2} \right)^{-1} \tag{4.24}$$

which is called half-Cauchy prior and is shown in figure 4.8, to extend the likelihood from equation (3.32)

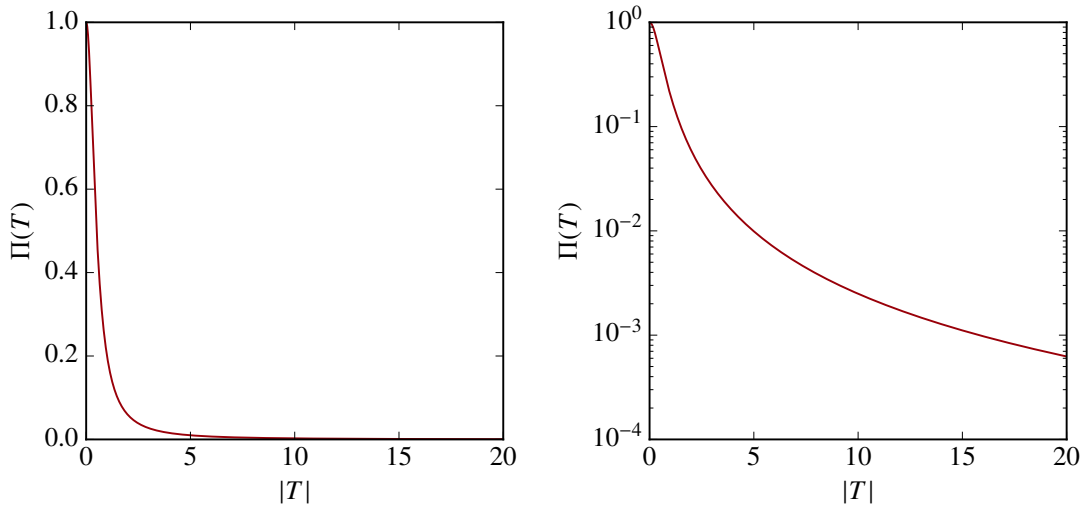
$$\ln \hat{\mathcal{L}}(\mathbf{T}; \mathbb{E}) = \ln \mathcal{L}(\mathbf{T}; \mathbb{E}) + \ln \Pi(\mathbf{T}) \tag{4.25}$$

does exactly that. Note that in this formulation of the prior, the width parameter Γ is measured in units of absolute value of the production amplitude and not in units of intensity. Like before, we limited ourselves to rank $r = 1$ and reflectivity $\epsilon = 1$ and therefore do not write the indices for rank and reflectivity.

The half-Cauchy prior has several advantageous properties: it has a peak at zero where its derivative is zero, which is important because it is used within a gradient-following algorithm. It falls rather sharply around Γ but has a very long tail beyond that. This means that it will shift the position of the likelihood maximum towards zero production amplitude for waves for which the maximum would have been around or below Γ, while it has little to no effect on waves for which the maximum is at $|T| \gg \Gamma$.

Setting $\Gamma = 0.5$ is motivated by defining an intensity of 1 to 10 as small, as was discussed above. A limited study has been performed with different values of Γ and 0.5 has been found to be an adequate value.

[h] At the suggestion of [37], we have moved from ROOT's MINUIT2/MIGRAD minimizer implementation [38, 39] to NLopt's [40] low-storage BFGS [41–43], which is able to cope with the large number of parameters with reasonable runtime: one MLE for a typical event sample size of about $4 \cdot 10^4$ takes $O(1\,\text{h})$.



(a) (b)

Figure 4.8: Half-Cauchy prior with $\Gamma = 0.5$ for one production amplitude.

4.3.1 Verification of Method

The event sample for testing the genetic algorithm, described in section 4.2.3, is also used to test the biggest-conceivable-model (BCM) selection procedure. The maximum-likelihood estimation (MLE) was performed 1000 times with different random start values for the gradient-following algorithm. The results are shown in figures 4.9 to 4.11.

Figure 4.9 shows the distribution of the likelihood values of all 1000 MLEs. As the absolute likelihood value does not carry meaning in itself, the difference with respect to the best found likelihood is shown. Ideally, this distribution should have only entries at zero, indicating that every MLE found the same likelihood which would mean that the result is independent of the randomly chosen start values of the gradient following. The structure which is obtained instead is markedly different: The best likelihood is only found a small number of times and most of the MLE results are rather evenly distributed over a large range of likelihood. This points to a multimodal likelihood function.

In figure 4.10, the intensities for all waves and all MLEs are synthesized. The waves are numbered and shown on the x-axis. They are ordered roughly by intensity, which will be discussed below. To visualize the intensities found in different MLEs, all results have been ordered by their likelihood and divided into ten sets with equal size, shown in figure 4.9 as areas of different color. The range from the minimum to the maximum intensity found in each set is shown in figure 4.10 as a bar of the corresponding color, e.g. the dark violet bar shows the minimal intensity range required to comprise all intensities found in the 10 % best MLEs.

There are two steps visible in the intensity distribution, the first one around wave 90, which is comparable to the size of the input model, and the second one around wave 160. These steps distinguish between the waves which should be included in the final model and those which can be

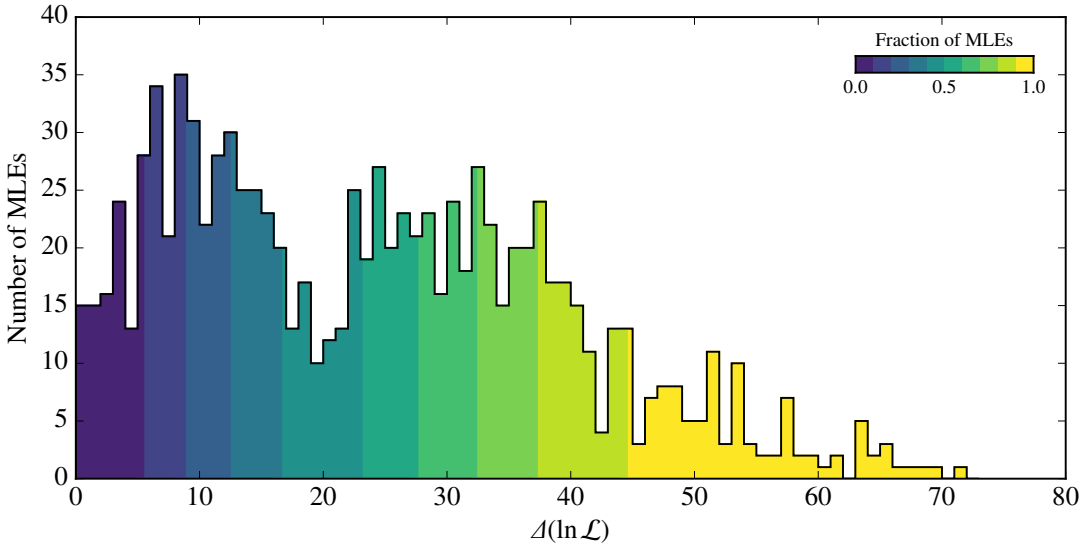


Figure 4.9: Likelihood distribution for all MLEs on a Monte Carlo data sample with 12 688 events.

discarded. The decision were exactly to apply the cut is difficult to make with the given intensity distribution. However, as we will see, the situation is unambiguous when the method is applied to measured data, so we do not discuss this problem further here.

To investigate the waves with large intensities, figure 4.11 shows a zoomed-in view of figure 4.10. In addition to the intensity bars, gray points have been added showing the intensities of the MLE with the highest likelihood. In red, the results of an MLE with the input model performed on the Monte Carlo data sample is shown. The red points are in the same order as in table B.1. There are waves in the best result which have large intensities but were not present in the input model, i.e. they do not have a corresponding red point. These waves have been moved to the left so that a rough order by intensity is established. To identify a wave from the reference fit result, the red points have to be counted starting with wave zero at the very left. To facilitate counting, every fifth reference intensity has been marked with a cross.

One can see that there are several waves which have a remarkably small intensity spread over all 1000 MLEs. Further, a large number of waves from the input model are found in the best MLE, but exhibit large variations throughout all MLEs. There are eight waves which either have large intensities, even though they were not present in the input model, or have a small intensity, even though the intensity in the input model was large. These have been marked with arrows and are listed in table 4.3, ordered as they appear on the plot from left to right.

The normalization given in equation (3.16) and the likelihood formula (3.17) ensure that the likelihood maximum always lies at a place where the total intensity and the acceptance-corrected number of measured events will be compatible considering the Poisson distribution. This means that one cannot remove one wave from the model without one or more other waves compensating. Indeed, except for waves 6 and 8 in table 4.3, the misidentified waves can be arranged in pairs with one wave of the pair being in the input model but missed in the selected model and the other wave being in the selected model but not in the input model. The waves in these pairs are strikingly similar, with waves 1 and 2

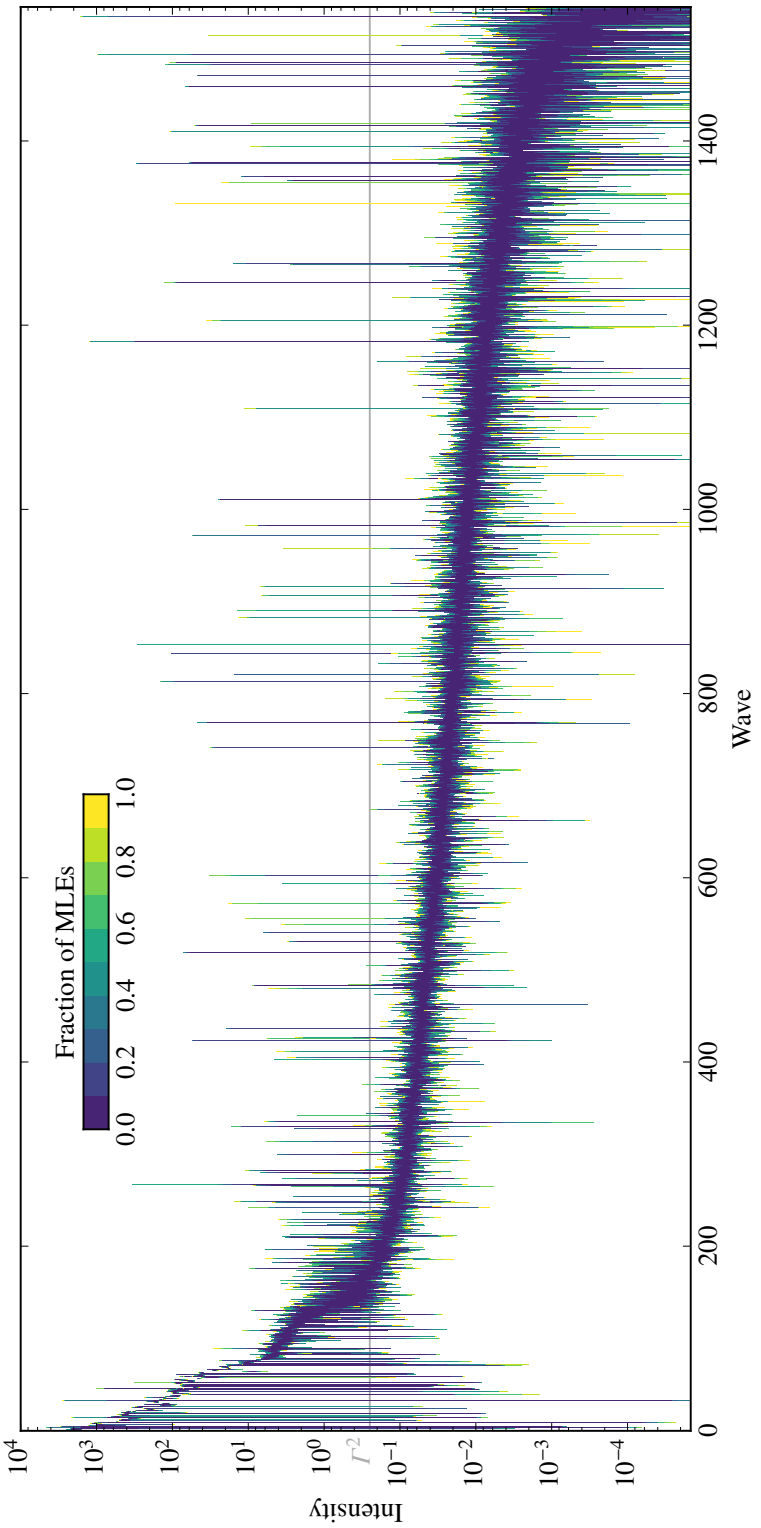


Figure 4.10: Partial-wave intensities for all MLEs on a Monte Carlo data sample with 12 688 events.

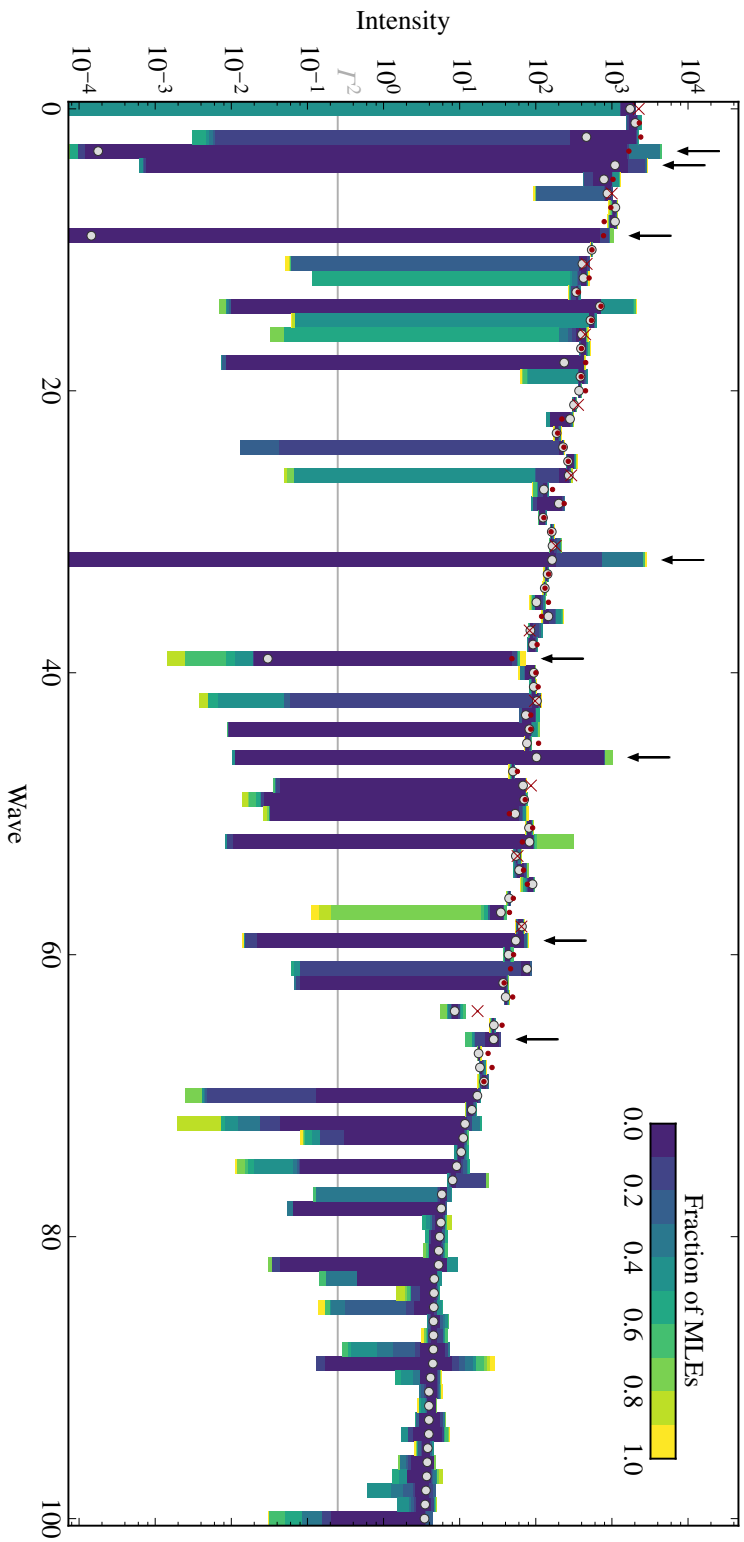


Figure 4.11: Partial-wave intensities for all MLEs on a Monte Carlo data sample with 12 688 events (zoomed version of figure 4.10). Gray points mark the intensities of the MLE with the highest likelihood, red points and crosses an MLE with the input model and the arrows point to misidentified waves.

#	Wave
1	$1^-0^{-+}0^+ \to f_0(980)\begin{bmatrix}0\\0\end{bmatrix}\left(\pi(1800) \to \sigma\begin{bmatrix}0\\0\end{bmatrix}\pi^-\right)$
2	$1^-0^{-+}0^+ \to f_0(980)\begin{bmatrix}0\\0\end{bmatrix}\left(\pi(1300) \to \sigma\begin{bmatrix}0\\0\end{bmatrix}\pi^-\right)$
3	$1^-0^{-+}0^+ \to \sigma\begin{bmatrix}0\\0\end{bmatrix}\left(\pi(1300) \to \sigma\begin{bmatrix}0\\0\end{bmatrix}\pi^-\right)$
4	$1^-0^{-+}0^+ \to \sigma\begin{bmatrix}0\\0\end{bmatrix}\left(\pi(1800) \to f_0(980)\begin{bmatrix}0\\0\end{bmatrix}\pi^-\right)$
5	$1^-1^{++}0^+ \to \sigma\begin{bmatrix}1\\0\end{bmatrix}\left(\pi(1800) \to \sigma\begin{bmatrix}0\\0\end{bmatrix}\pi^-\right)$
6	$1^-1^{++}0^+ \to \left(\rho(1450) \to \rho(770)\begin{bmatrix}0\\1\end{bmatrix}\sigma\right)\begin{bmatrix}0\\1\end{bmatrix}\pi^-$
7	$1^-1^{++}0^+ \to \sigma\begin{bmatrix}1\\0\end{bmatrix}\left(\pi(1300) \to \sigma\begin{bmatrix}0\\0\end{bmatrix}\pi^-\right)$
8	$1^-1^{++}0^+ \to f_2(1270)\begin{bmatrix}0\\1\end{bmatrix}\left(a_1(1260) \to \sigma\begin{bmatrix}1\\0\end{bmatrix}\pi^-\right)$

Table 4.3: Misidentified waves on a Monte Carlo data sample with 12 688 events. See text for details.

as well as 5 and 7 only differing in the mass dependence of one of the isobars. The pair with waves 3 and 4 also differs only in mass shape, but of two isobars. Misidentifications of this kind are not surprising, as the model components feature the same angular distribution and only differ in one or two of the mass distributions, giving little distinctive power. The picture that there are pairs of waves where one wave is replaced by another is too simple however, because the intensities of all waves are related and waves cannot necessarily just be swapped pairwise and still yield a good description of the data. Waves 6 and 8 are superfluous, meaning they are picked up without missing a wave from the input model at the same time.

There is a very important observation to make here: all waves with large intensity and small intensity spread are correctly identified. Waves with large spreads are identified correctly more often than not and waves which are not correctly found always have a large spread. This means that the variation of the partial-wave intensities found in the MLEs is a useful tool to identify reliably selected waves.

The method was also tested on a larger Monte Carlo data sample containing 42 334 events, generated with the same input model. This sample size is typical for a single $m_{5\pi}$ bin in the COMPASS 2008 five-pion data set. The results are shown in figures 4.12 to 4.14.

The likelihood, shown in figure 4.12, is also broadly distributed and extends to higher values (cf. figure 4.9). In the overall intensity distribution in figure 4.13, it is evident that the spread of the intensities is smaller on average (cf. figure 4.10), which is to be expected with a larger event sample. It is interesting to see that the first intensity step stays at the same position while the second one moves to the right. The zoomed view in figure 4.14 shows the same tendency, with smaller spreads in general and more waves with very small spread. The number of misidentified waves has decreased to

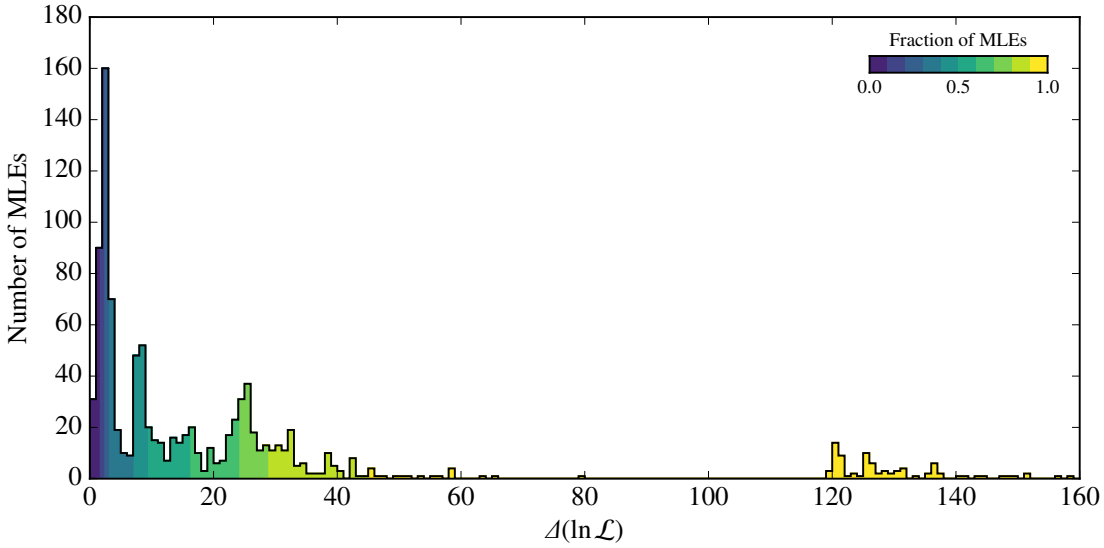


Figure 4.12: Likelihood distribution for all MLEs on a Monte Carlo data sample with 42 334 events.

#	Wave
1	$1^-0^{-+}0^+ \to f_0(980)\begin{bmatrix}0\\0\end{bmatrix}\left(\pi(1800) \to \sigma\begin{bmatrix}0\\0\end{bmatrix}\pi^-\right)$
2	$1^-0^{-+}0^+ \to f_0(980)\begin{bmatrix}0\\0\end{bmatrix}\left(\pi(1300) \to \sigma\begin{bmatrix}0\\0\end{bmatrix}\pi^-\right)$
3	$1^-0^{-+}0^+ \to \sigma\begin{bmatrix}0\\0\end{bmatrix}\left(\pi(1300) \to \sigma\begin{bmatrix}0\\0\end{bmatrix}\pi^-\right)$
4	$1^-0^{-+}0^+ \to \sigma\begin{bmatrix}0\\0\end{bmatrix}\left(\pi(1800) \to f_0(980)\begin{bmatrix}0\\0\end{bmatrix}\pi^-\right)$

Table 4.4: Misidentified waves on a Monte Carlo data sample with 42 334 events. See text for details.

four, listed in table 4.4. The four larger misidentified waves have stayed the same, while the smaller waves are now being picked up correctly.

To investigate the variations in partial-wave intensities in different MLEs in more detail, waves with an intensity of more than 1000 counts in at least one of the MLEs and a large fluctuation of their intensity (ratio between the maximum and minimum found intensity larger than 5) were examined. In order to study the correlations of the intensities of these 15 waves, they are plotted in an Andrews plot [45] in figure 4.15a, where every MLE result is represented by one curve. If all 15 intensities of two different MLEs are similar, then the corresponding curves will end up lying close together. Several classes of intensity patterns are visible around which the results from the MLEs are clustering. A clustering algorithm[i] is able to identify these classes and assign them to the different results. Figure 4.15b shows the same as figure 4.15a, but colored differently based on the cluster assignment found. In total, 17 clusters are found. In light red in the background are the results which were

[i] Here, the DBSCAN algorithm [46] implemented in the scikit-learn package [47, 48] is used.

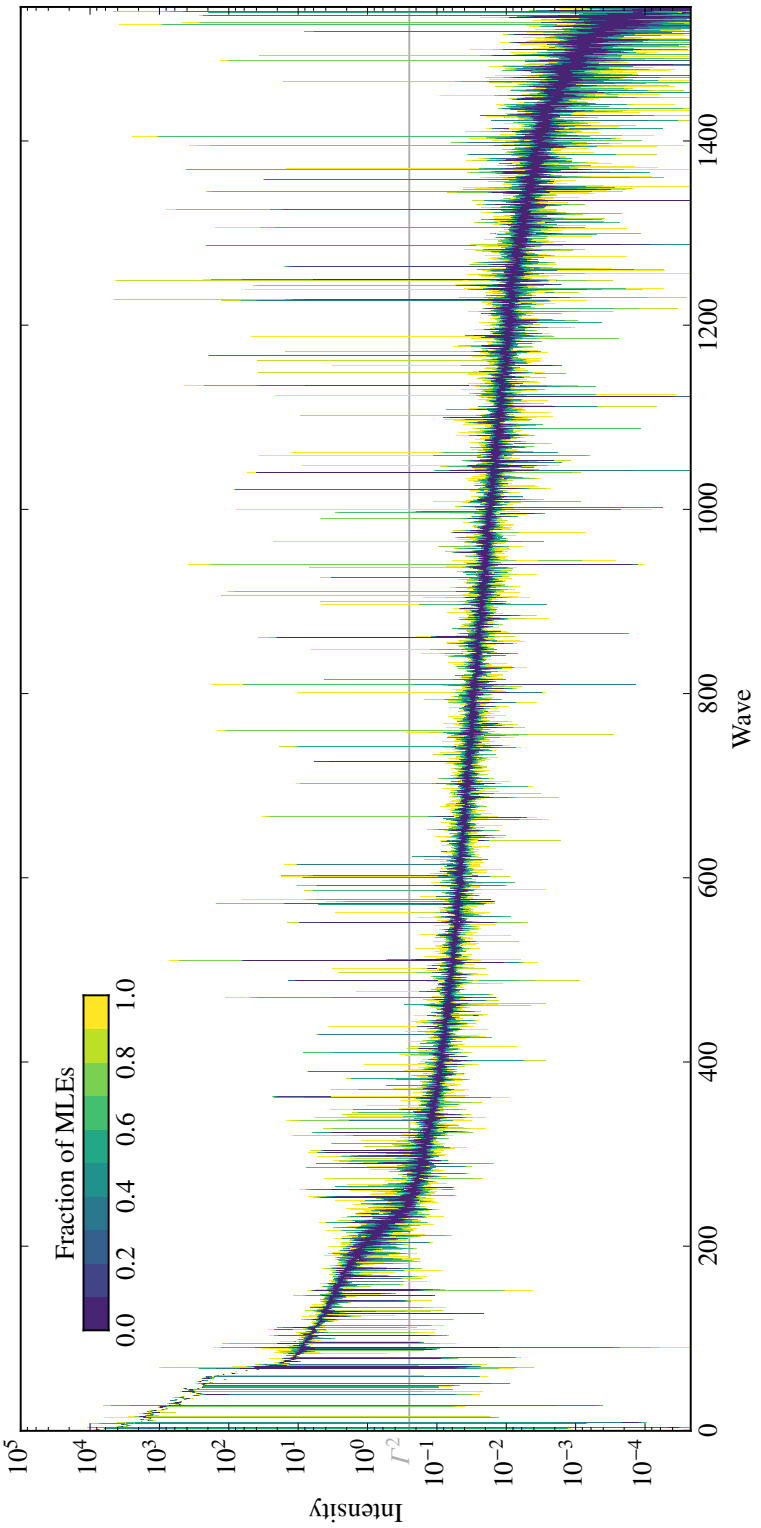


Figure 4.13: Partial-wave intensities for all MLEs on a Monte Carlo data sample with 42 334 events.

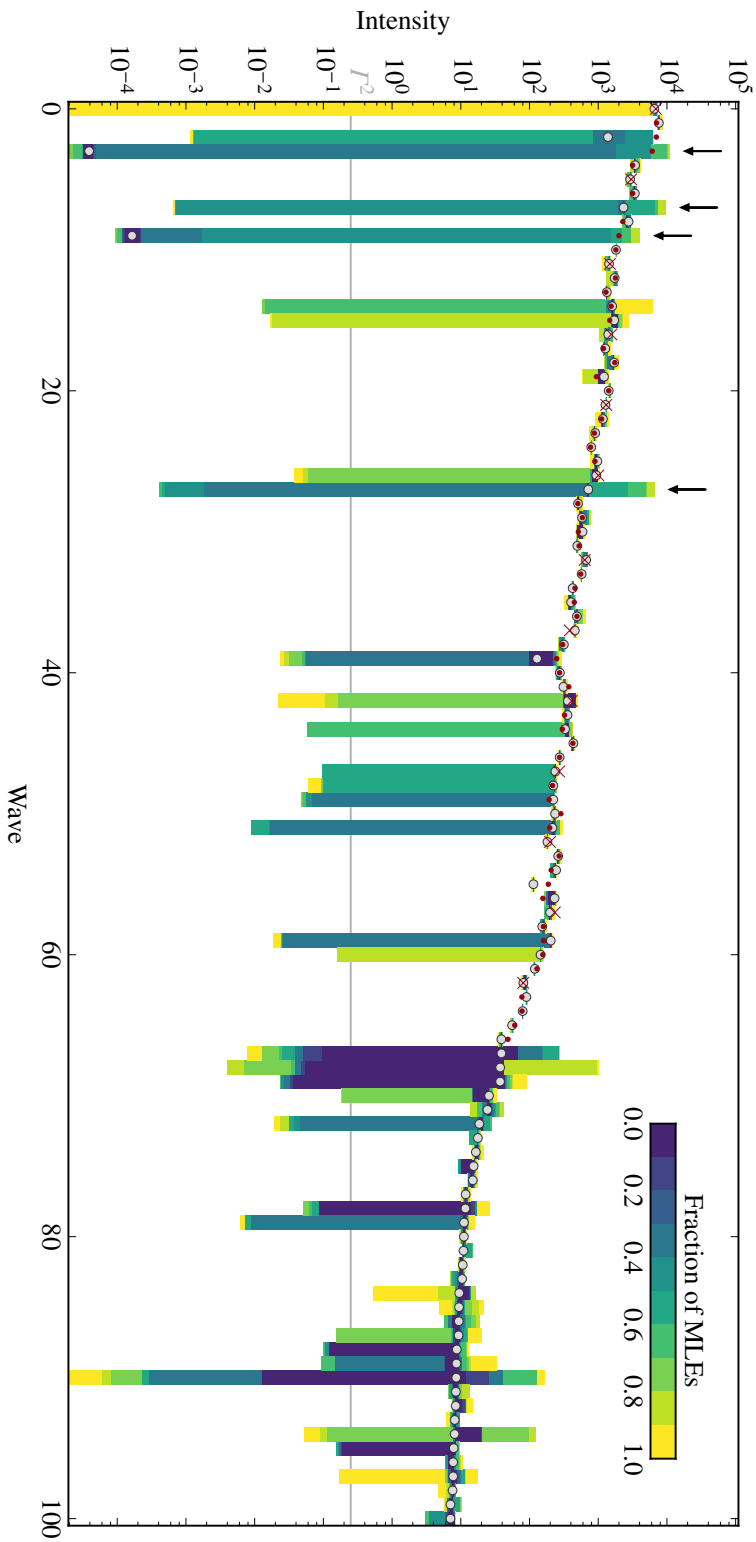


Figure 4.14: Partial-wave intensities for all MLEs on a Monte Carlo data sample with 42 334 events (zoomed version of figure 4.13). Gray points mark the intensities of the MLE with the highest likelihood, red points and crosses an MLE with the input model and the arrows point to misidentified waves.

assigned to the "noise" class by the algorithm, which means they are not classified as fitting with any of the clusters but do not form one of their own. The noise class contains 4.9 % of the total MLEs.

For every found cluster (except the one containing the overall best MLE), figure 4.16 shows in dark blue the intensities of the 15 waves under investigation from the best MLE within the respective cluster. In order to provide context to these intensities, in the background, the intensity variation plot from figure 4.13 is shown, where the colors have been desaturated. The reference result using the input model and the MLE with the overall best likelihood are drawn as faint red and gray lines, respectively. The arrows from figure 4.14 are also present in the background, pointing out misidentified waves. In the regions highlighted in gray, the x-axis has been contracted by a factor of 50 to better show the intensities of interest.

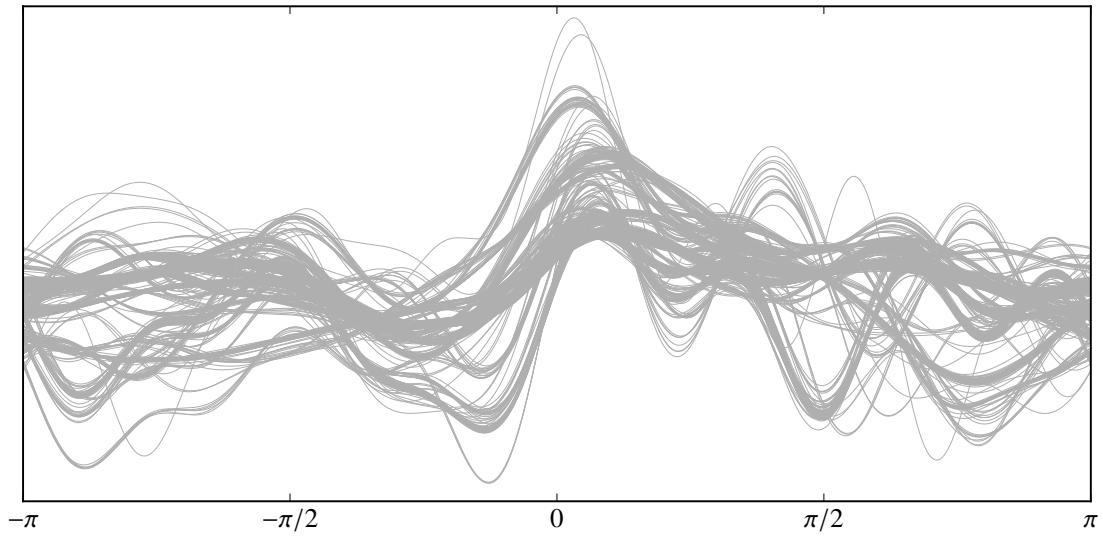
Even though figure 4.16 shows a wealth of information, the important message is that there is a complex pattern of intensities, where different combinations of waves with large intensities all describe the data reasonably well. This illustrates that the simple pattern found in the pairwise misidentification of waves is part of a more complex correlation structure which manifests itself as the observed intensity variations. While it is not directly evident from the figure, the third cluster, the best result of which is shown third from the top, correctly identifies all the waves from the input model without adding superfluous ones. However, the log-likelihood of the best result of this cluster is still 8.2 units worse, meaning the overall best result does describe the data better.

The likelihood distributions for the individual clusters, this time including the one with the best MLE, are shown in figure 4.17. The cluster containing the best MLE is shown on the top left, then left-to-right, top-to-bottom, the likelihood difference increases. At the bottom right, the likelihood distribution for the noise class is shown. The size of the cluster in percent of the total number of MLE is shown in the plots. One can see that also in likelihood, the clusters are rather well localized. It is noteworthy that the best cluster is also the biggest one. As expected, the noise class is distributed evenly across the whole range.

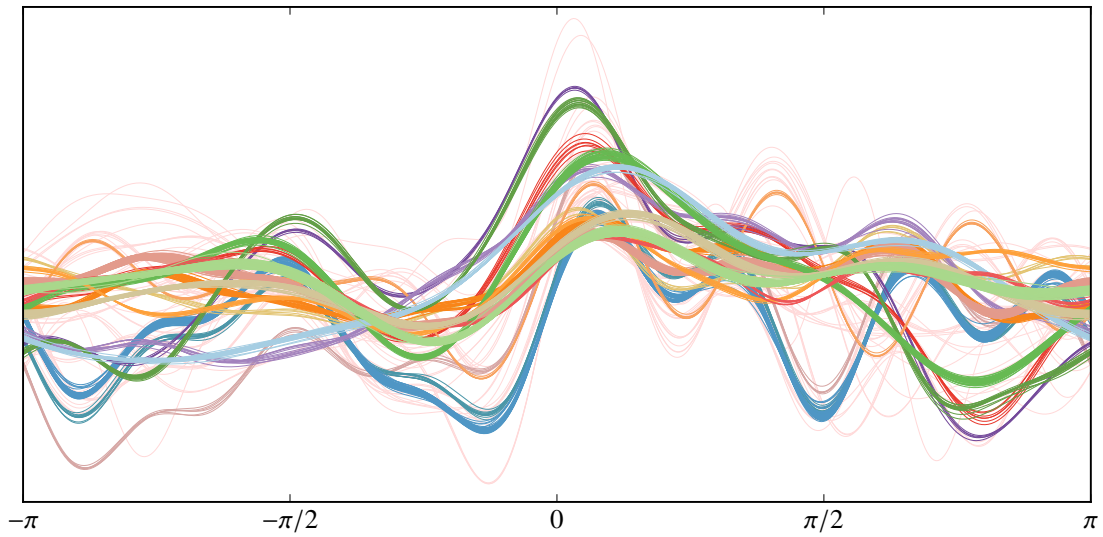
To study the effect that an imperfect partial-wave model has on the result, the same Monte Carlo data sample with 42 334 events was analyzed again, but instead of parameterizing the propagator of the isobars with relativistic Breit-Wigner amplitudes with mass-dependent widths as given by equation (3.43), they were parameterized with relativistic Breit-Wigner amplitudes with constant widths, as given by equation (3.56). The values for m_0 and Γ_0 remained the same. This means that the isobar parameterizations differ from those used to generate the event sample. A similar situation is expected when analyzing measured data, because many of the isobar parameters are not well known. The results are shown in figures 4.18 to 4.20.

The likelihood distribution in figure 4.18 shows much broader structures. The intensity variations in figure 4.19 are bigger, the two intensity steps are washed out and in the zoomed view in figure 4.20, one can see that a much larger number of waves are misidentified. However, the finding that waves which have a small intensity spread are reliably and correctly identified is still valid.

As a final study, 1000 MLEs without using the prior function were performed on the same Monte Carlo data sample containing 42 334 events. The results are shown in figures 4.21 and 4.22. As all the MLEs have likelihoods that lie within 1.2 units of each other, the likelihood distribution is not shown. The intensity variations also shrink drastically, but there is no step in the intensities in



(a)



(b)

Figure 4.15: Andrews plots for the partial-wave intensities of selected waves (a) without and (b) with coloring based on cluster classification.

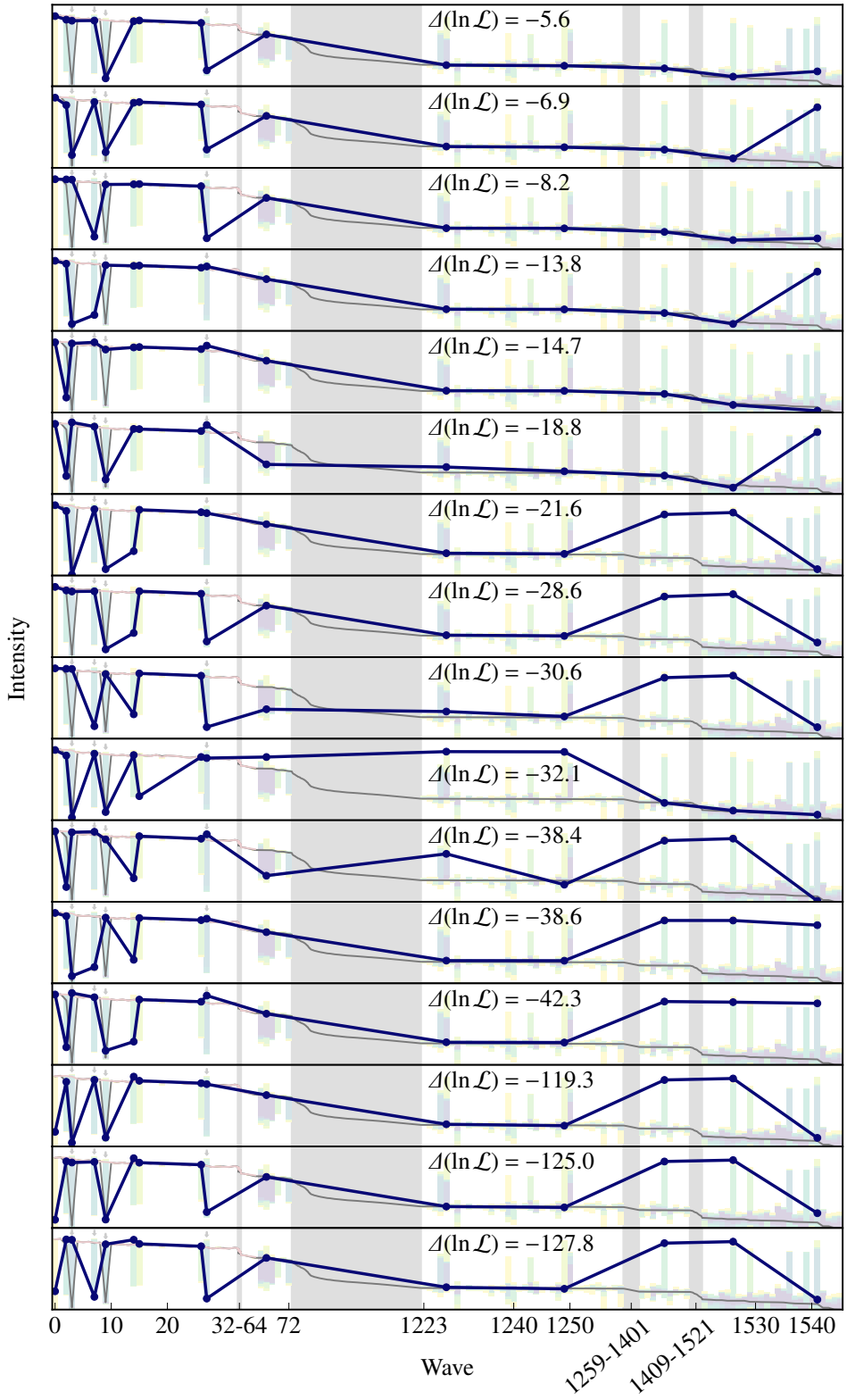


Figure 4.16: Partial-wave intensities of the best MLE from different clusters. See text for details.

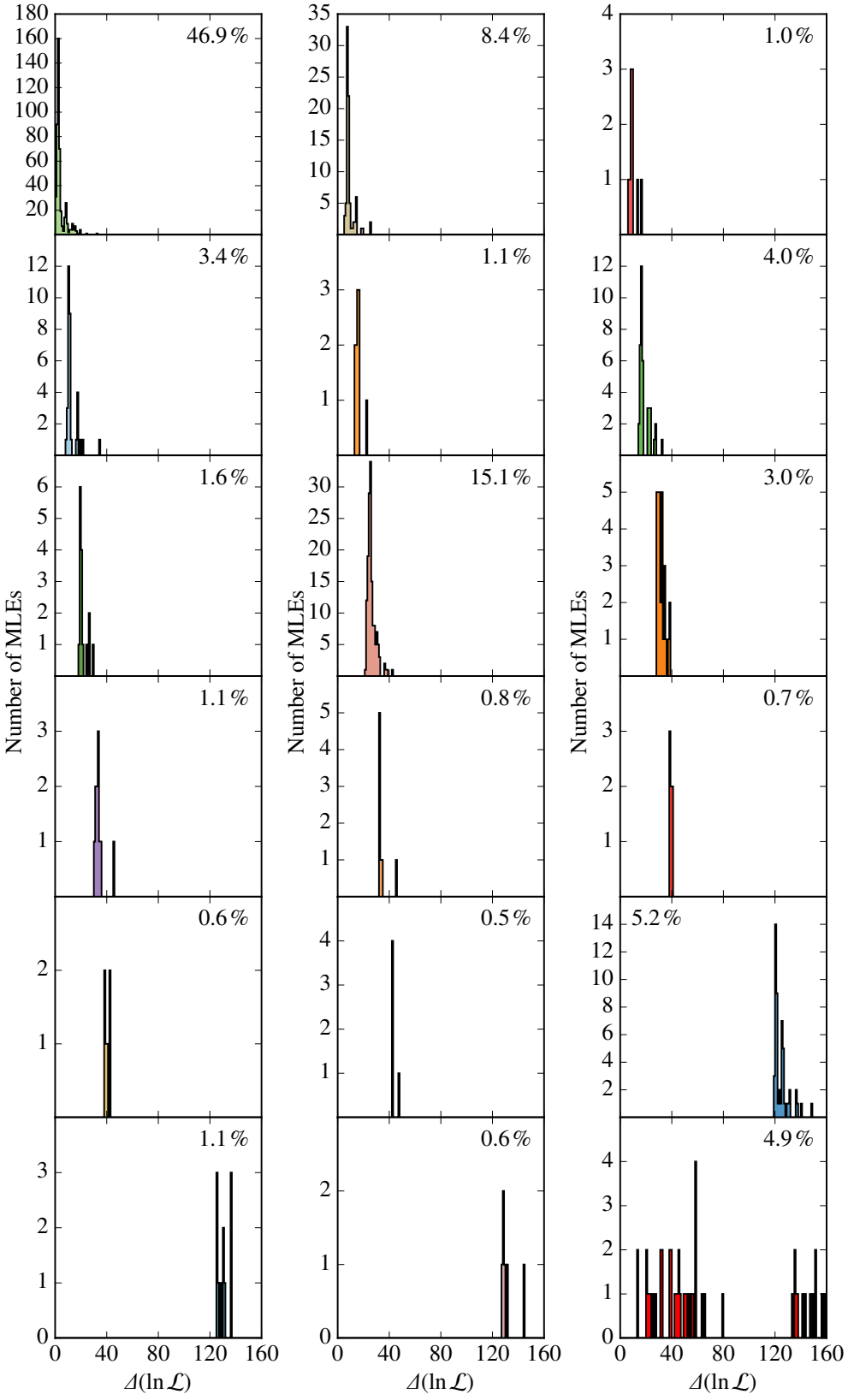


Figure 4.17: Likelihood distributions for all found clusters and the noise class on the bottom right.

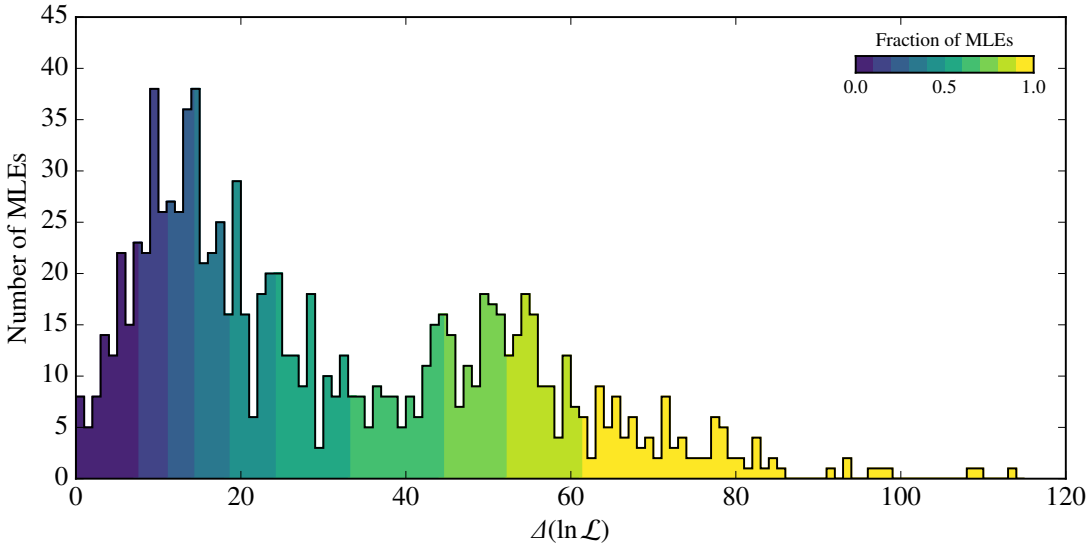


Figure 4.18: Likelihood distribution for all MLEs with constant-width Breit-Wigner isobars on a Monte Carlo data sample with 42 334 events.

figure 4.21. In the zoomed view, the arrows for misidentified waves are not shown because of their large number. It is evident that a great number of waves exhibit deceptively large intensities, leading to the conclusion that even though the likelihood distribution and the intensity variations look more stable, this approach is not suited for model selection.

To conclude, it was demonstrated that the biggest-conceivable-model method is able to identify many of the model components which are present in the Monte Carlo data sample. Furthermore, performing multiple runs of the MLE provides valuable information on the confidence to be placed in selected waves, with large intensity variations indicating that the corresponding wave is not unambiguously identified. The intensity variations and likelihood distributions are consequences of ambiguities in the description of the data and studies of the clustering of different results for the intensities of a subset of waves show that there are multiple models describing the data similarly well. However, despite these ambiguities, there are waves not exhibiting large intensity variations so their extraction is not affected. It is this kind of waves that will be in the focus of the discussion of the results of this method when it is applied to measured data.

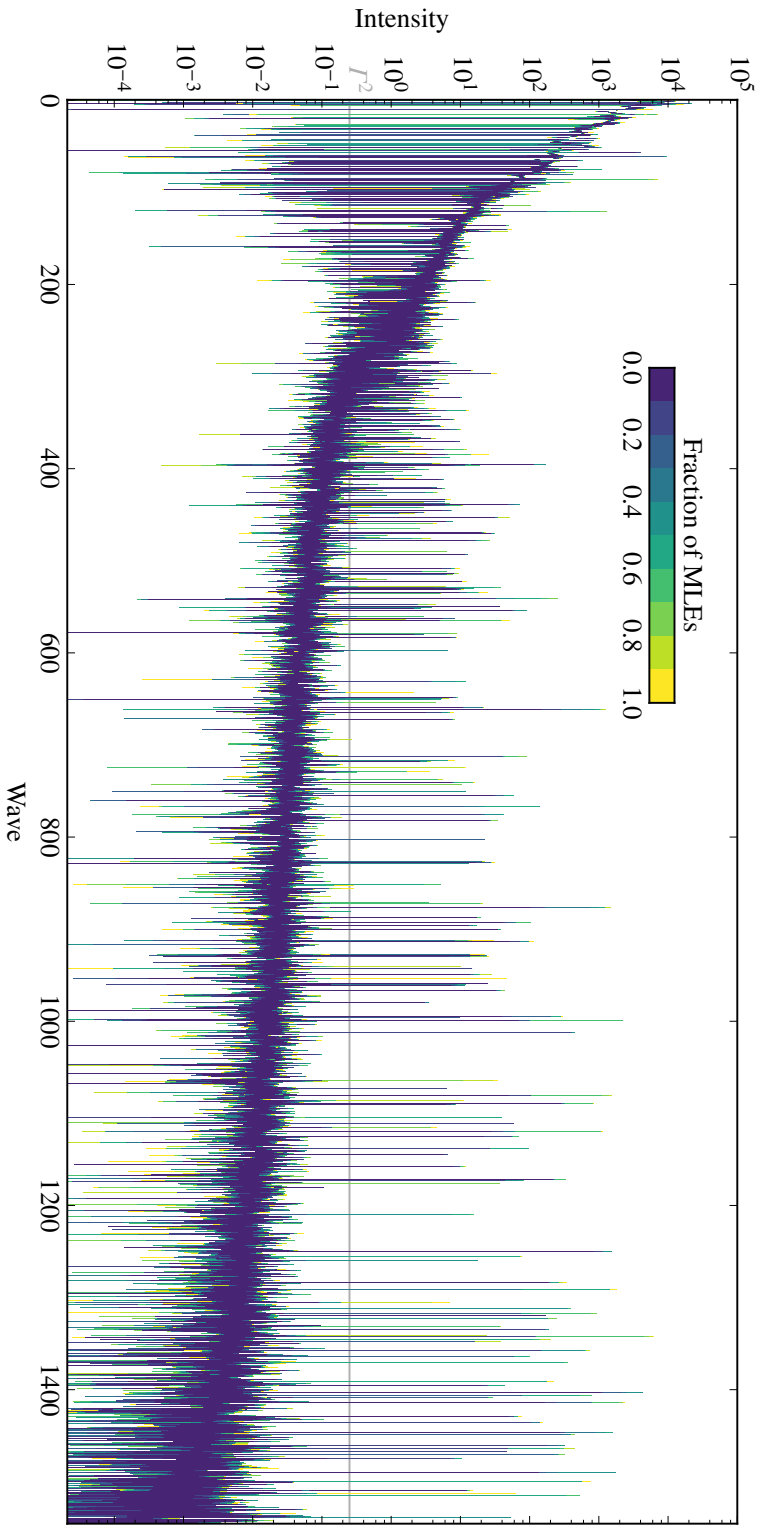


Figure 4.19: Partial-wave intensities for all MLEs with constant-width Breit-Wigner isobars on a Monte Carlo data sample with 42 334 events.

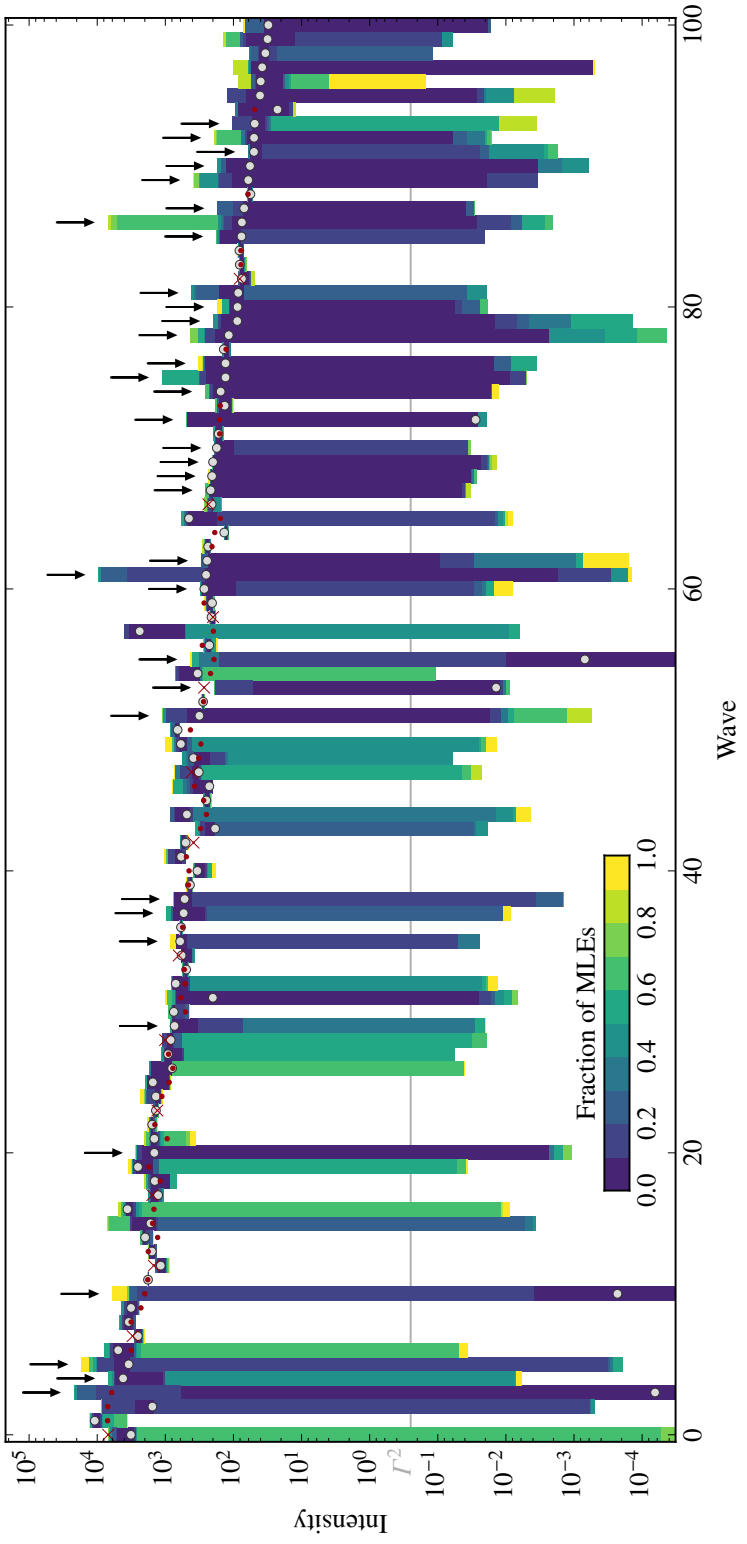


Figure 4.20: Partial-wave intensities for all MLEs with constant-width Breit-Wigner isobars on a Monte Carlo data sample with 42 334 events (zoomed version of figure 4.19). Gray points mark the intensities of the MLE with the highest likelihood, red points and crosses an MLE with the input model and the arrows point to misidentified waves.

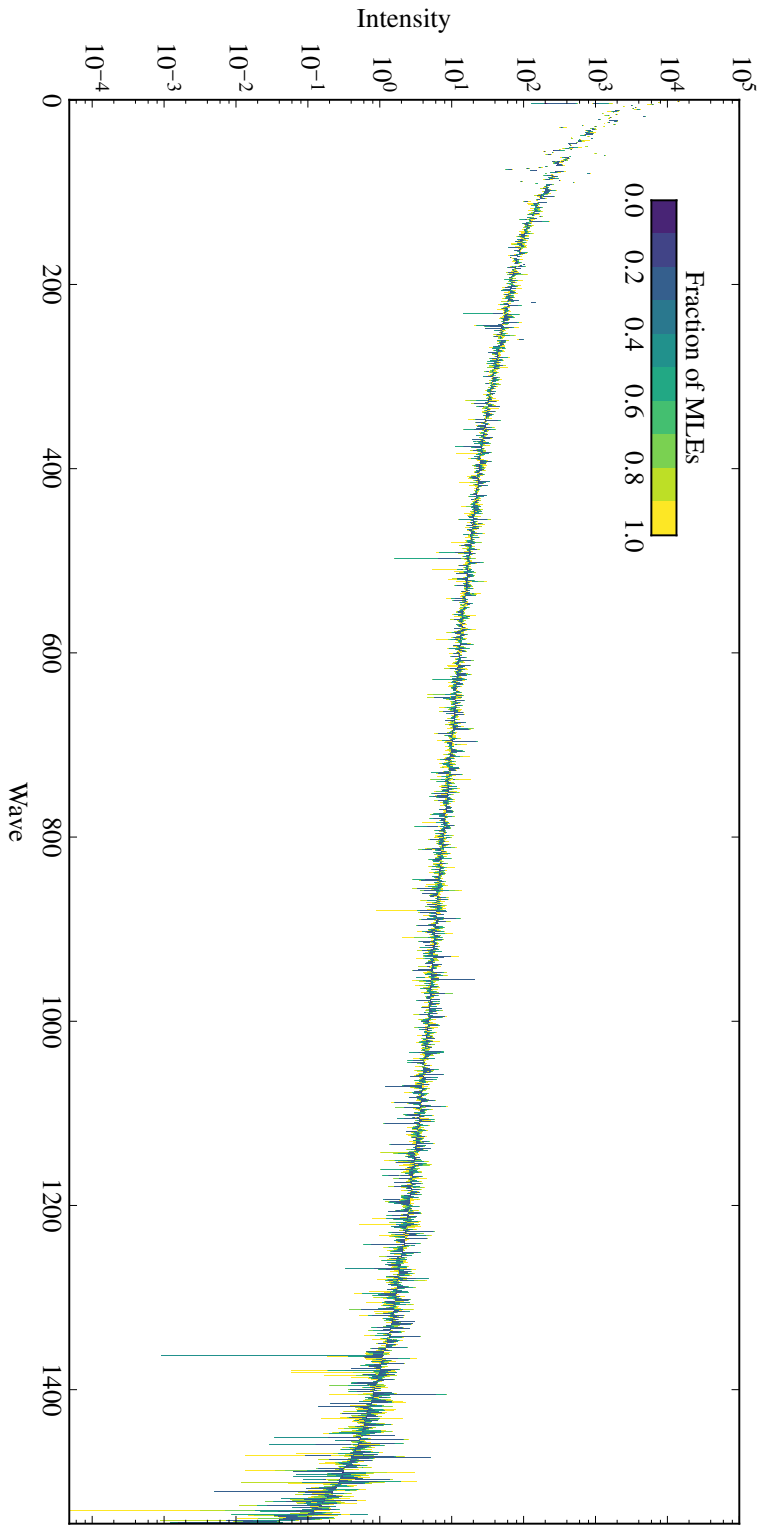


Figure 4.21: Partial-wave intensities for all MLEs without prior on a Monte Carlo data sample with 42 334 events.

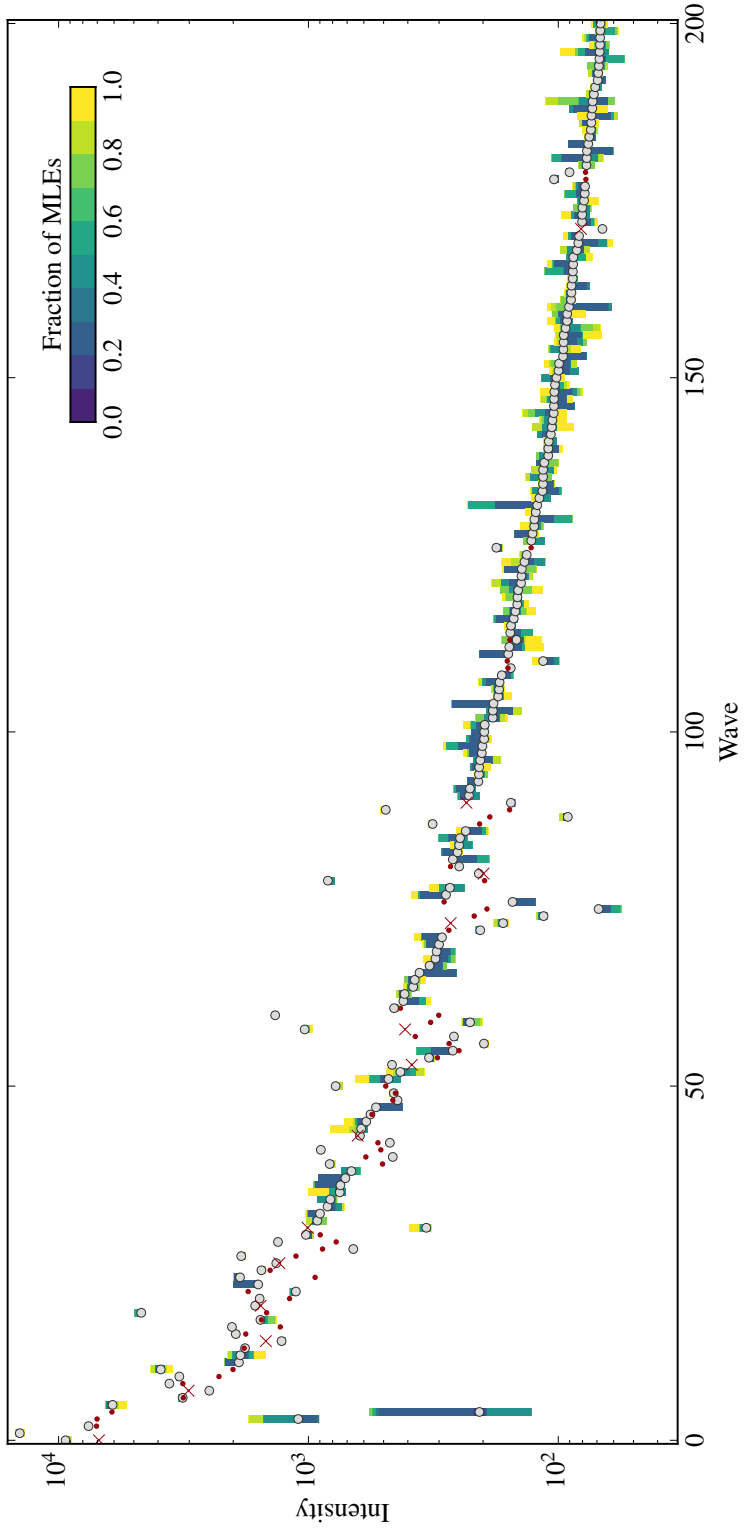


Figure 4.22: Partial-wave intensities for all MLEs without prior on a Monte Carlo data sample with 42 334 events (zoomed version of figure 4.21). Gray points mark the intensities of the MLE with the highest likelihood and red points and crosses an MLE with the input model.

4.3.2 Results on Measured Data (2008)

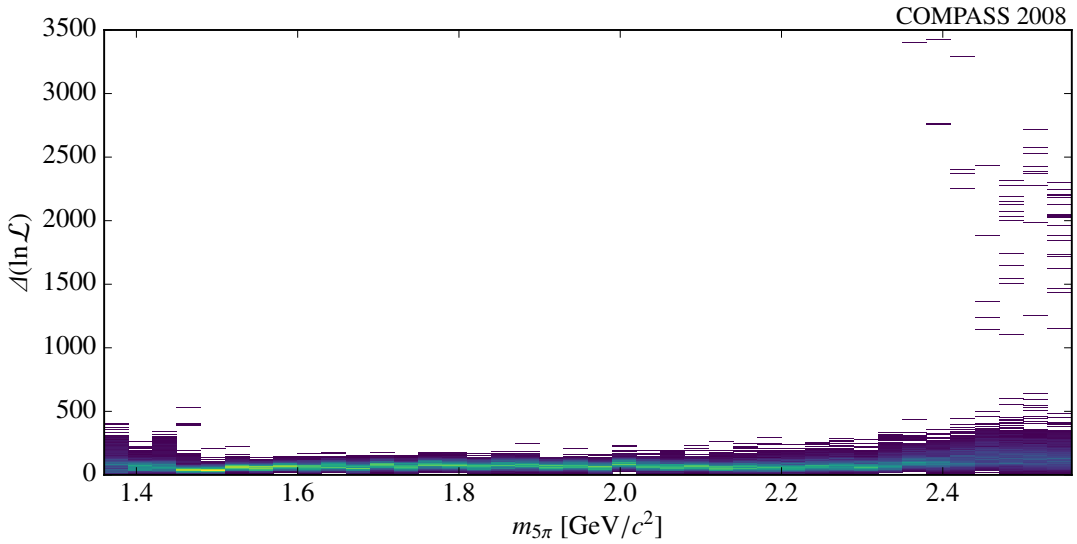
The BCM method has been applied to the measured event sample resulting from the selection described in section 2.2. The data have been subdivided into 40 bins with a width of $30\,\mathrm{MeV}/c^2$ ranging from 1.36 to $2.56\,\mathrm{GeV}/c^2$. Since this is a proof-of-principle study, only the t' region from 0.1 to $0.15\,(\mathrm{GeV}/c)^2$ has been used. For every mass bin independently, 1000 maximum-likelihood estimations have been performed.

Figure 4.23 shows the likelihoods for all MLE as a function of $m_{5\pi}$. The qualitative behavior from mass bin to mass bin is very similar. If we look at the likelihood distribution of all MLE for a single bin at around $2\,\mathrm{GeV}/c^2$ in figure 4.24, we can compare directly to the corresponding distributions obtained on a Monte Carlo generated event sample in the previous section (cf. figures 4.12 and 4.18). With the two samples being about equal in size, the likelihood structure should be similar, assuming that the Monte Carlo sample is indeed similar to the data. However, the distributions in figure 4.24 and figure 4.12 differ significantly in structure, with a very broad bump in the former versus several rather sharp features in the latter. Interestingly, figure 4.24 exhibits much more similarities with the likelihood distribution from the study using detuned amplitudes to parameterize the propagator of the isobars, shown in figure 4.18. This is not surprising, as the masses, widths and shapes of many of the isobars used in the model are not well known.

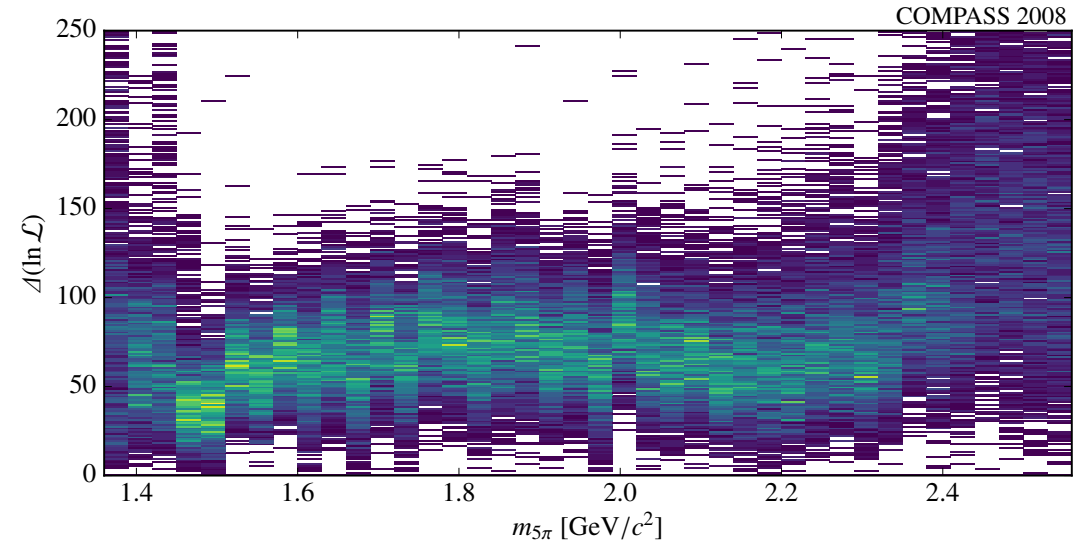
A feature of the likelihood distribution obtained from measured events, which is not reproduced in any of the studies on generated events, is the fact that the best MLE is an outlier. Possible explanations are that exclusive and non-exclusive backgrounds as well as imperfections in the detector simulation are absent in the Monte Carlo data set. It is entirely possible that the global maximum of the likelihood is not found. This might cause concern that this hypothetical undiscovered MLE might exhibit an entirely different set of production amplitudes. As we will see, there are a number of waves which have very similar amplitudes in a large fraction of the MLEs, so one can be reasonably certain that, for these waves, the results would not differ significantly in any undiscovered solution with higher likelihood.

When discussing the intensity distributions in the previous chapter, no prescription for actually selecting a model was given. The reason for this is that the intensity pattern observed when analyzing Monte Carlo data does not exhibit a very clear point where one could cut the wave set (see figures 4.10, 4.13 and 4.19). This problem is not present in measured data, which can be seen in the intensity distribution for the best MLE for one mass bin shown in figure 4.25. A clear step is visible with a drop in intensity of almost two orders of magnitude between two waves. Such a step is present in the intensity distributions of the best MLE for every mass bin in measured data, making the selection a wave set straightforward.

Figure 4.26 shows the size of the wave sets resulting from the model selection as a function of $m_{5\pi}$ together with the number of events in the corresponding mass bins. One can see that the wave-set size roughly follows the data-sample size. This is expected as an increasing number of small waves can be resolved when increasing the number of events. Since the MLE on different mass bins are independent from each other, no bin-to-bin continuity of the selected wave sets is enforced. Many waves are only present in the selected wave sets of a few bins, so that in the end, 1517 waves out of the 1546 waves in the pool appear in at least one selected set.



(a)



(b)

Figure 4.23: Likelihood distribution of all MLEs for all mass bins of the COMPASS 2008 event sample. In (b), the same plot is shown with a zoom on the y-axis.

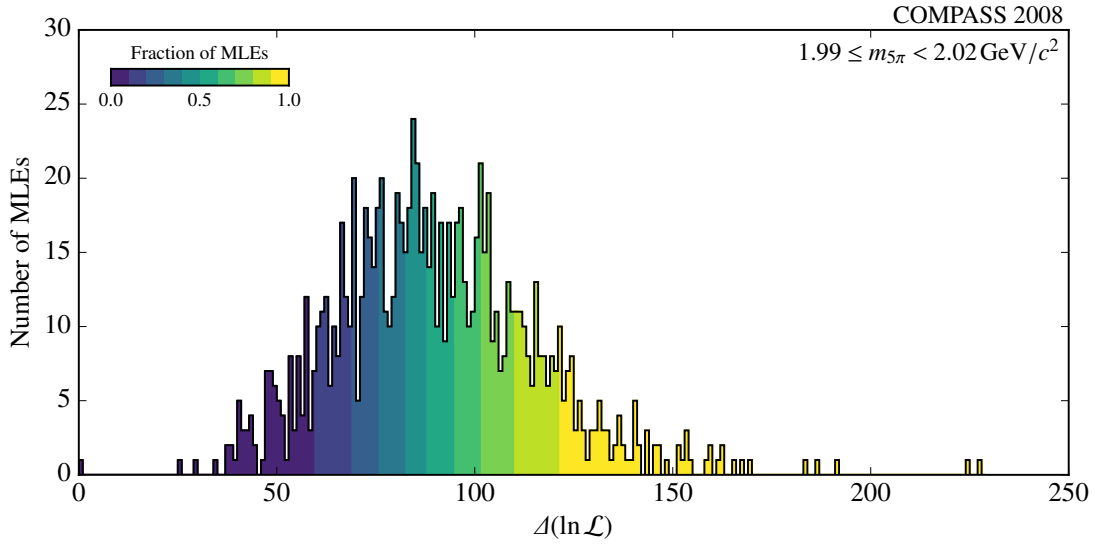


Figure 4.24: Likelihood distribution for all MLEs for the bin $1.99 \leq m_{5\pi} < 2.02\,\mathrm{GeV}/c^2$, containing 43043 events, of the COMPASS 2008 event sample.

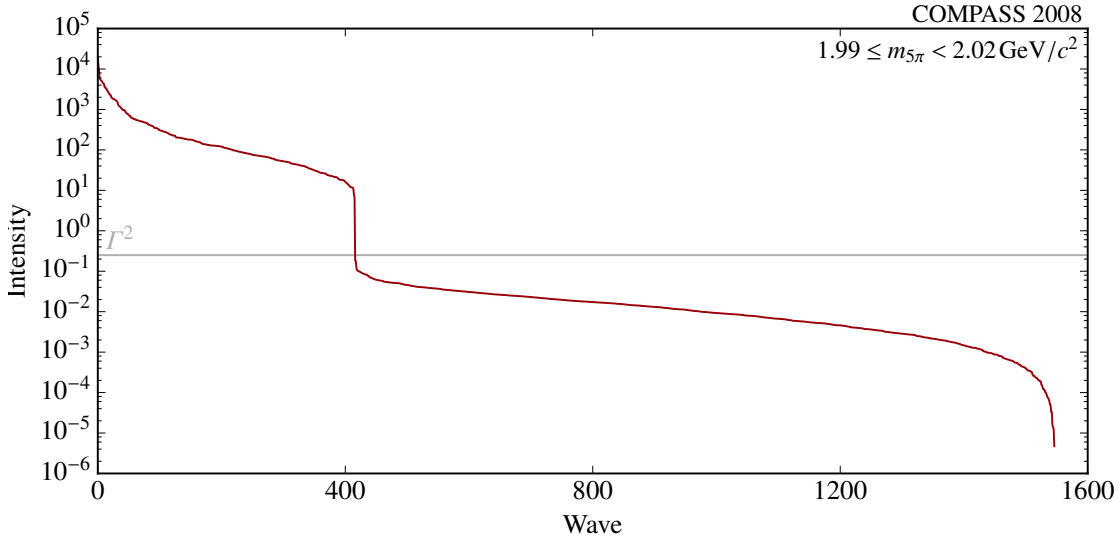


Figure 4.25: Ordered partial-wave intensity distribution for all waves for the best MLE for the bin $1.99 \leq m_{5\pi} < 2.02\,\mathrm{GeV}/c^2$, containing 43043 events, of the COMPASS 2008 event sample.

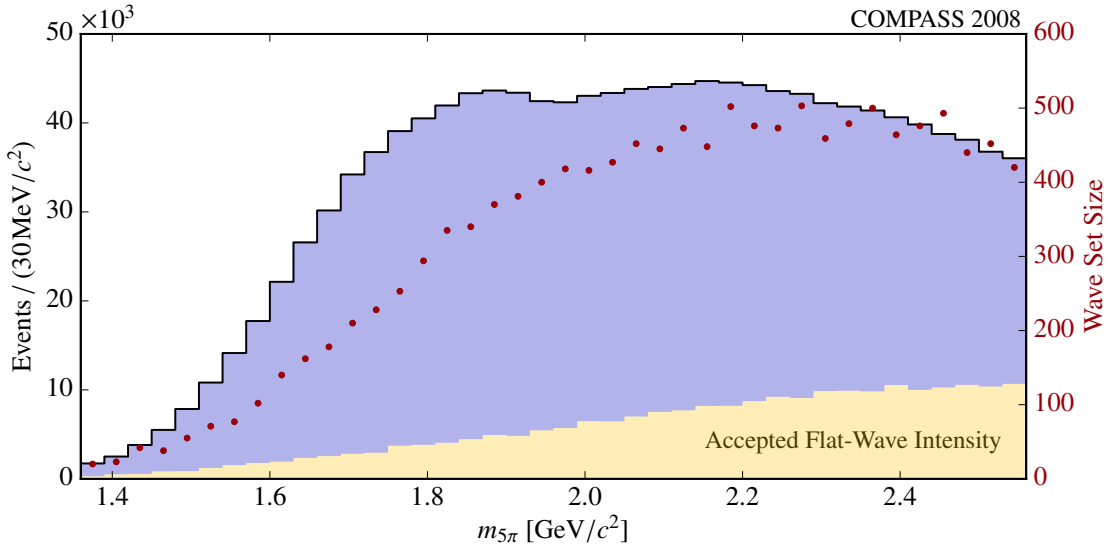
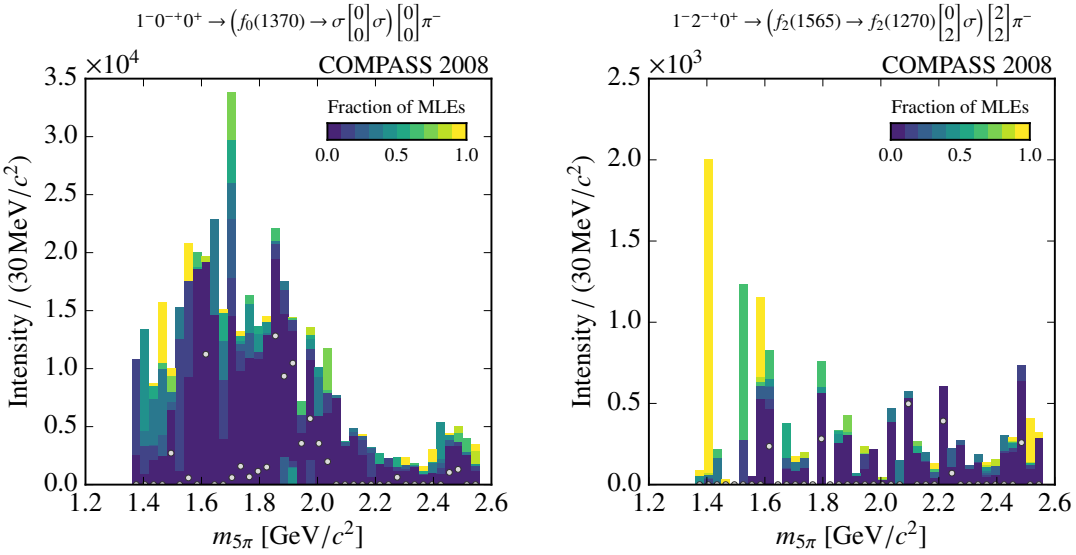


Figure 4.26: Selected wave set size (red), number of events (blue) and accepted flat-wave intensity (yellow) for the COMPASS 2008 event sample. The accepted flat-wave intensity is discussed in section 5.4.



(a) (b)

Figure 4.27: Partial-wave intensities for two unstable waves for all MLEs for the COMPASS 2008 event sample.

As discussed in the previous section, the intensity variations from all MLEs on a given event sample allow to determine which waves are reliably identified in the data. The visualization technique introduced for the Monte Carlo studies in section 4.3.1 is employed here in a modified form. While the gray points still show the intensity from the best MLE and the colored bars show the range of the intensities for the corresponding fraction of MLEs, the intensities are now shown as a function of $m_{5\pi}$ and for every wave separately. For practical reasons, not all 1517 waves can be shown. In the following, we focus only on those waves which we consider stable. A wave is considered stable if the intensity found in all 1000 MLEs is different from zero in more than one consecutive mass bin. With this criterion, one ends up with approximately 60 stable waves. As an example, two waves which do not meet the stability criterion are shown in figure 4.27. Starting from the $\sim$ 60 stable waves, a further selection was done by using combinations of the following conditions: (1) the intensity evolution over $m_{5\pi}$ is continuous, (2) the intensity variation over all 1000 MLE is comparably small, (3) the intensity is large, (4) the quantum numbers of the wave in question are those of well-known resonances, i.e. $J^{PC} \in \{0^{-+}, 1^{++}, 2^{-+}\}$ and (5) the narrow $f_1(1285)$ isobar is present in the decay. Note that conditions (1)-(3) are only defined qualitatively and for a wave to be selected, not all of the conditions above had to be met. As an example, the 4^{-+} wave shown in figure 4.33a does not fulfill conditions (4) and (5), but this is outweighed by conditions (1)-(3). Applying these conditions reduces the $\sim$ 60 waves from above to 18 waves, which are shown in figures 4.28 to 4.33. The existence of waves with such stable intensity is remarkable: even when taking into account all the MLEs, no matter how bad the likelihood, the fluctuation of the intensity of these waves is small and the variation with $m_{5\pi}$ is sensible. Throughout all Monte Carlo studies presented in section 4.3.1, waves with small intensity variations were correctly extracted, giving us good reason to trust such waves also here. However, it must be stressed that the stability is only useful to determine the reliability of the model-selection procedure within the context of the partial-wave decomposition formalism and the wave pool which was used. It is not useful to identify artifacts resulting from processes present in the event sample but not accounted for by the basic model assumptions, which include diffractive production and isobaric decay as discussed in chapter 3, or from too small a wave pool. This will be discussed in more detail in section 5.5.

It is important to note that the 18 waves which are presented here are not the wave sets which have been selected by the BCM method. They are just the waves which are reliably extracted, while all the other waves present in the selected models exhibit large intensity fluctuations and are therefore not considered to be reliable enough for a physical interpretation. However, they are still used in the next step, which is to perform a partial-wave decomposition without the prior using the selected wave sets. This is a shift from the usual point of view in this type of analysis: instead of compiling a model, used for all mass bins[j], with components interpreted in terms of their physical properties, here, most of the model components are just seen as pieces to describe the data and no physical interpretation is attributed to the extracted production amplitudes. Only waves which the model-selection procedure finds reliably are looked at further. It is hoped that by this, the signals in the stable waves are as free as possible from model leakage. However, signals which are affected by the ambiguities discussed in section 4.3.1 are lost, with their intensity distributed over two or more waves. With this approach, one does not get the complete physical picture which is present in the measured distributions, instead, one aims for maximal confidence in the extracted information.

[j] Actually, it is customary to remove certain waves from the wave set in mass bins where the the final-state mass m_X is too low for the wave to contribute, meaning waves are added to the model with increasing m_X.

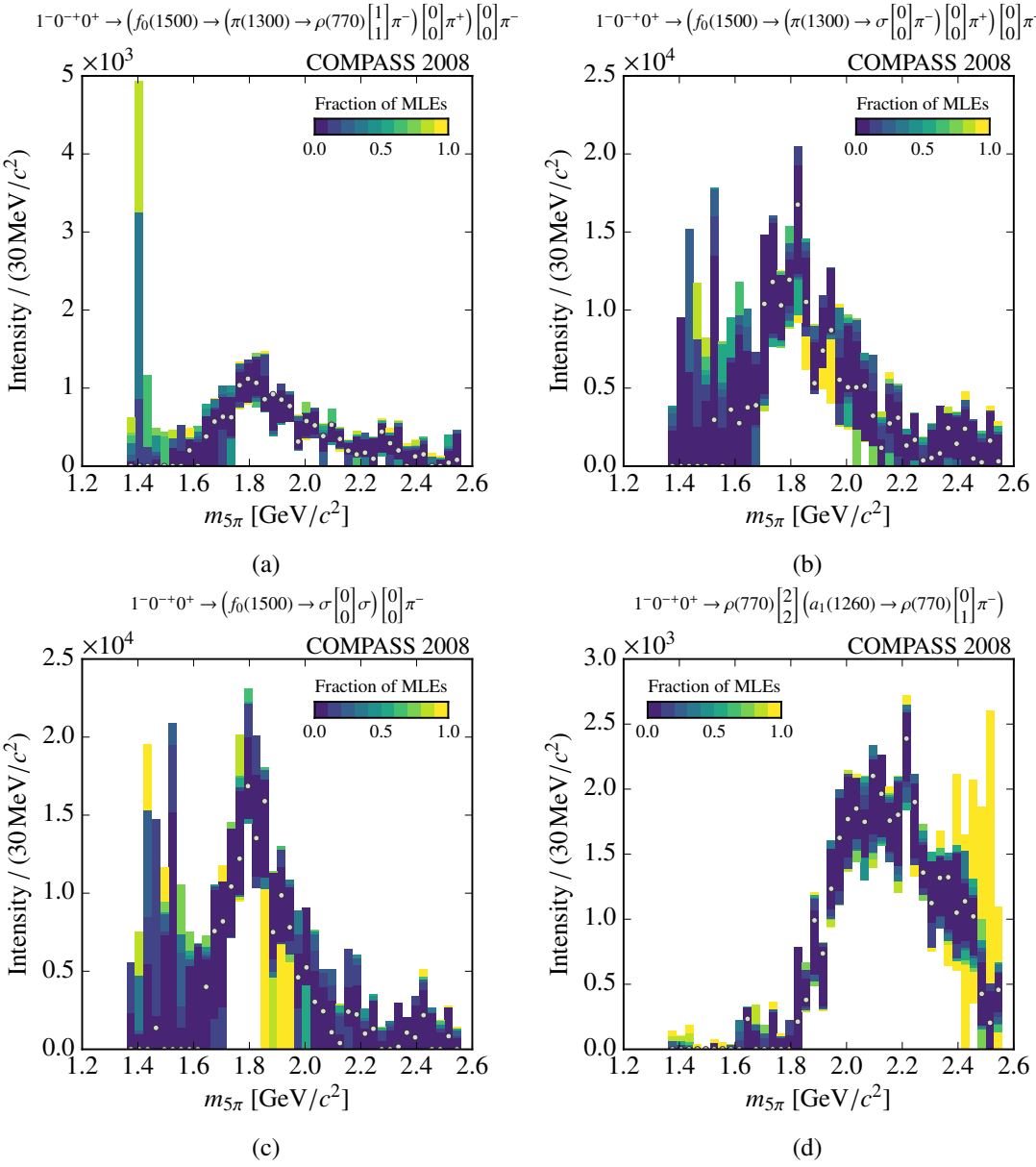


Figure 4.28: Partial-wave intensities for selected waves for all MLEs for the COMPASS 2008 event sample.

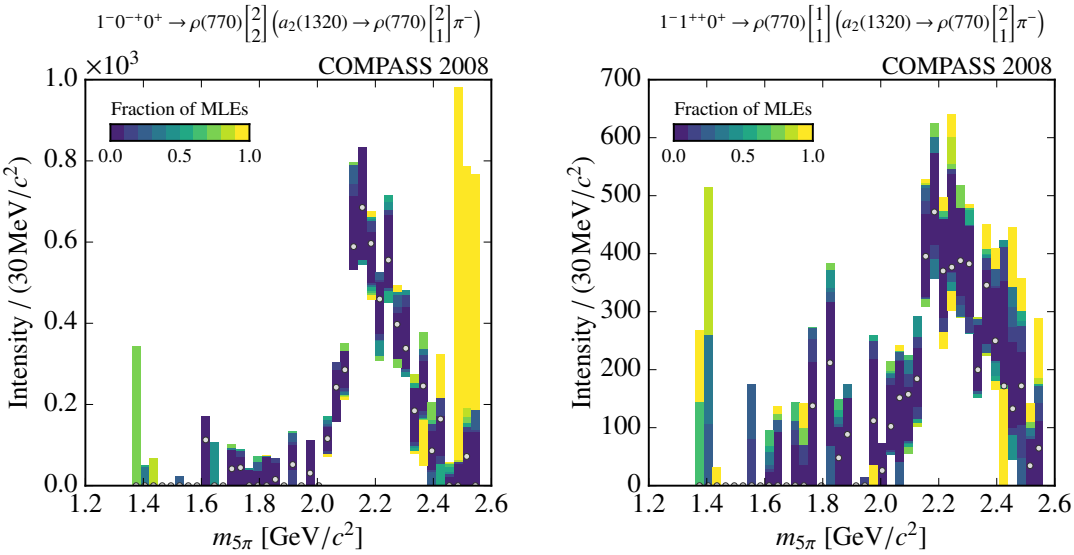


Figure 4.29: Partial-wave intensities for selected waves for all MLEs for the COMPASS 2008 event sample.

Due to the small width of the $f_1(1285)$ isobar, waves decaying into it can be extracted most reliably, even when their intensity is small. Out of the 16 waves containing the $f_1(1285)$, 6 have been found to be stable. They are shown in figures 4.31 and 4.32. The other 10 waves show either small or no intensity, with intensity spreads which are also comparably small.

Finally, figure 4.33b shows the flat-wave intensity and its fluctuation over all MLEs. It exhibits very little spread, indicating that the size of the background component is rather independent of the wave set. For higher masses, where the spread increases, the best MLE is generally on the lower end, meaning that lower background seems to yield a better description.

Having selected the wave sets and identified a number of waves which show promise, the next step is to use the selected waves sets in an MLE. The extracted production amplitudes of the interesting waves from this MLE will be discussed in chapter 5.

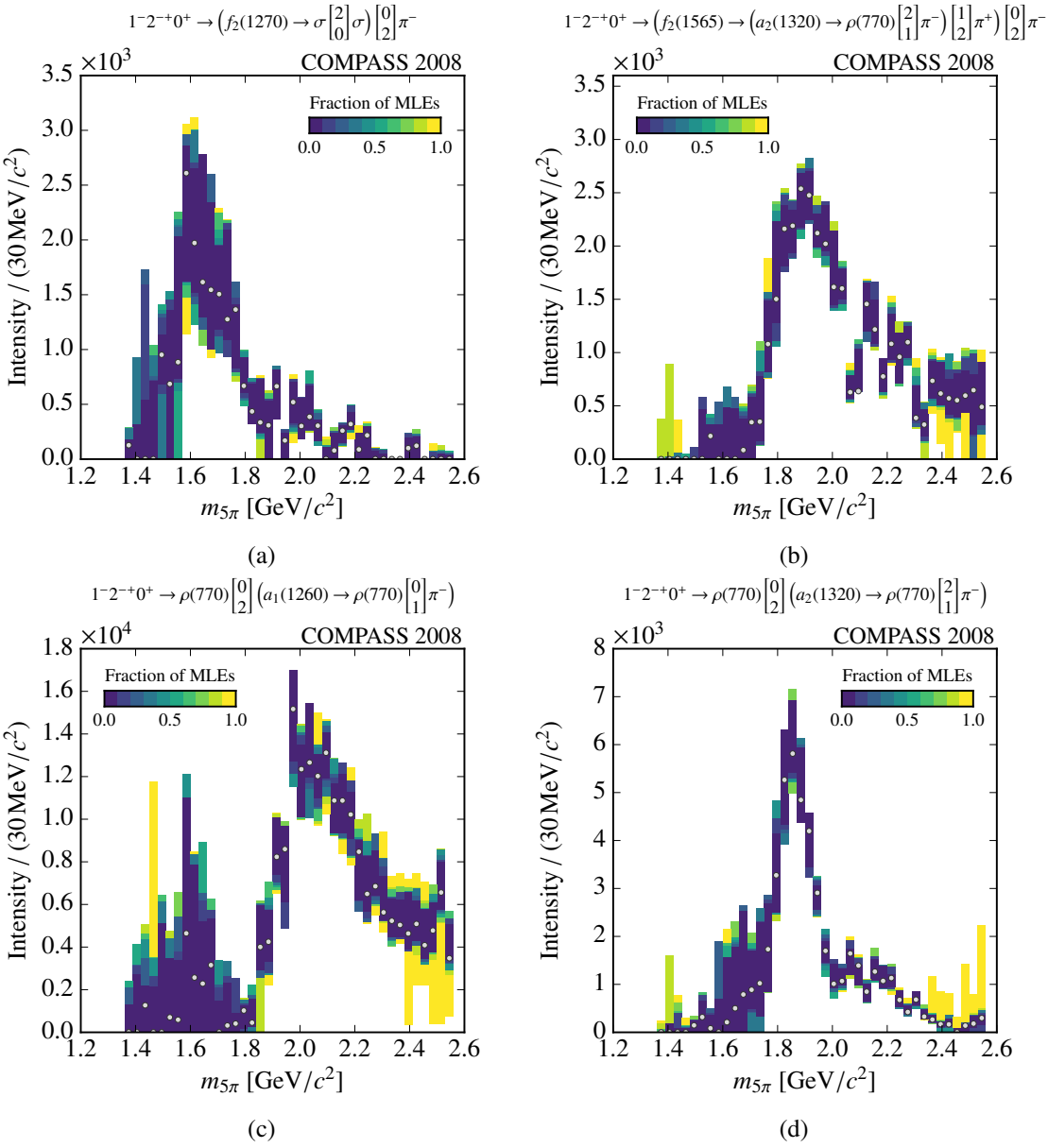


Figure 4.30: Partial-wave intensities for selected waves for all MLEs for the COMPASS 2008 event sample.

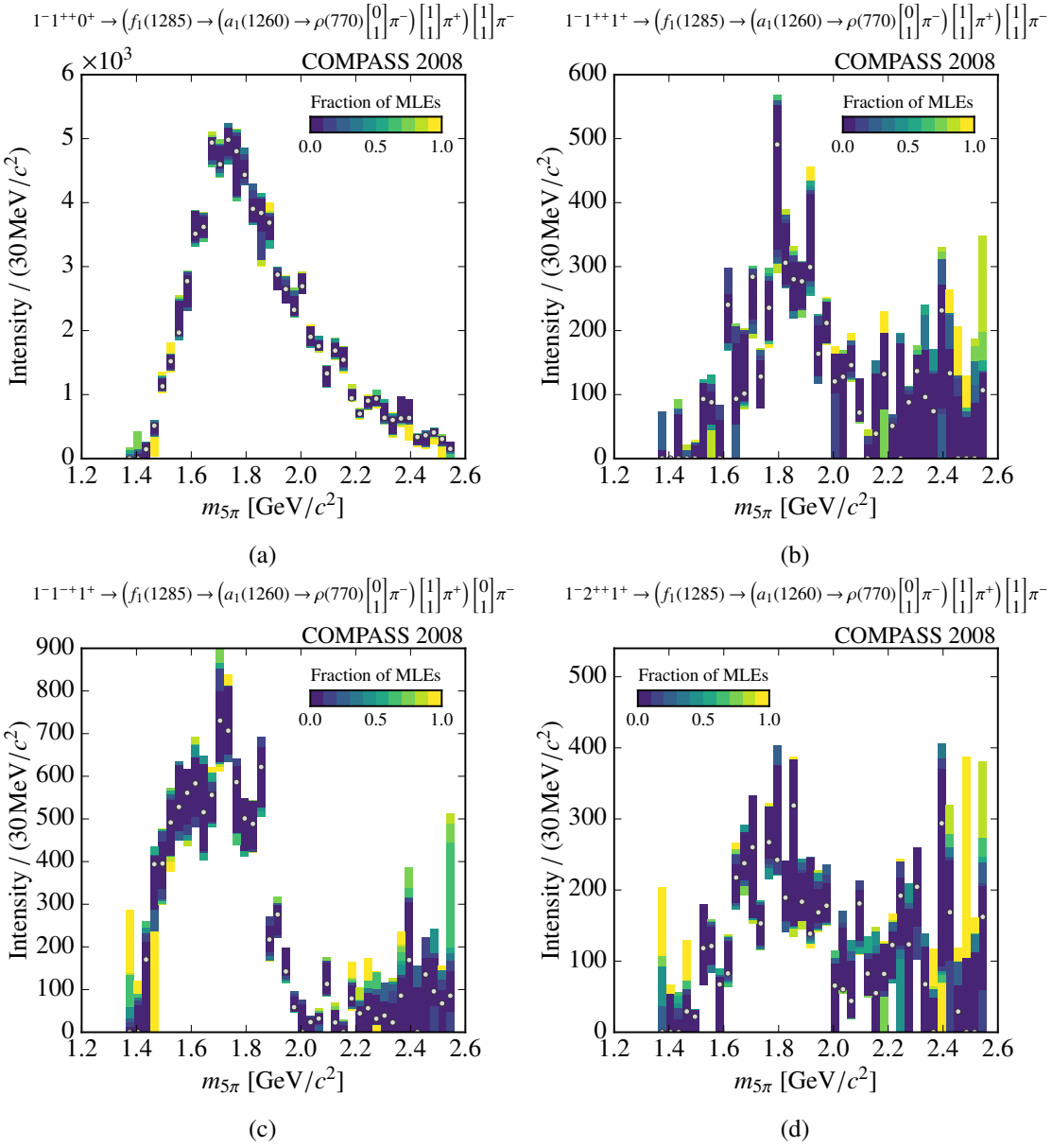


Figure 4.31: Partial-wave intensities for selected waves containing the $f_1(1285)$ isobar for all MLEs for the COMPASS 2008 event sample.

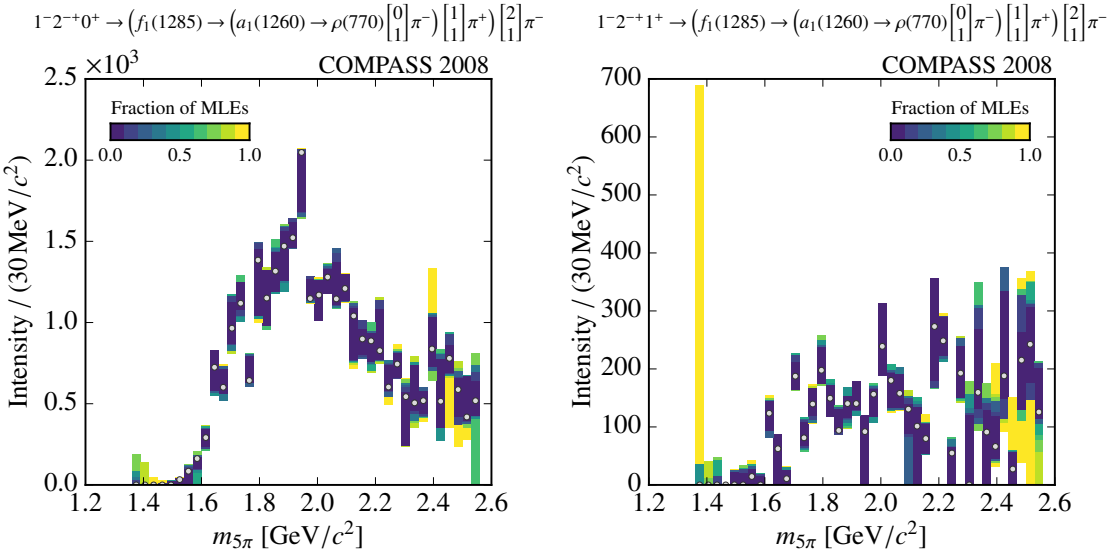


Figure 4.32: Partial-wave intensities for selected waves containing the $f_1(1285)$ isobar for all MLEs for the COMPASS 2008 event sample.

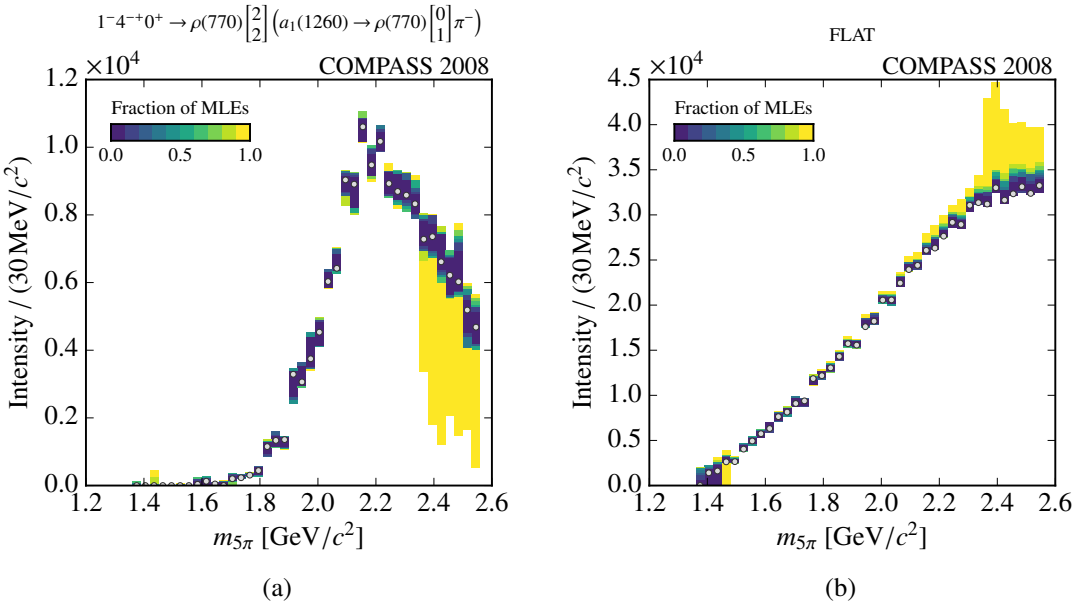
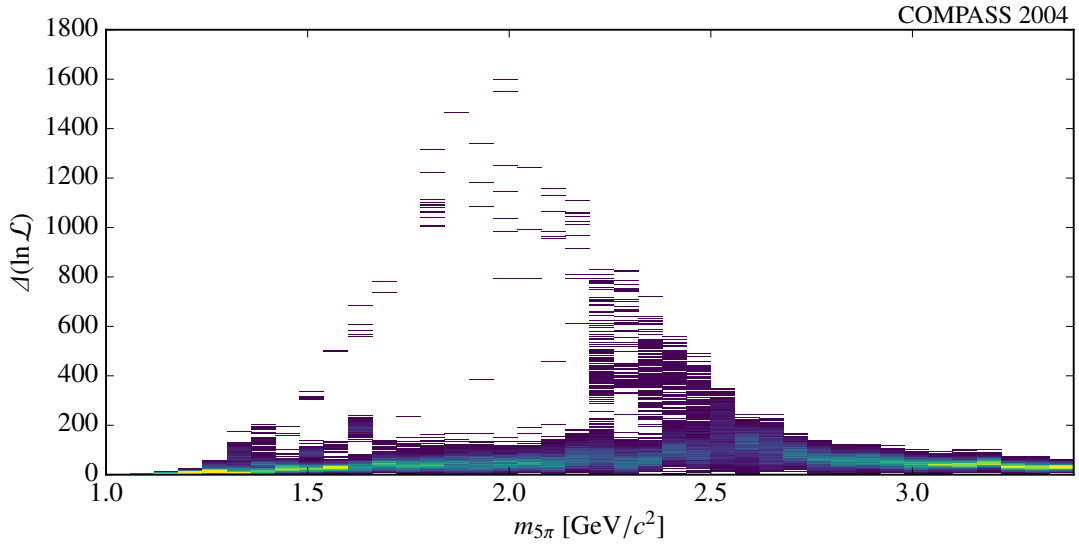
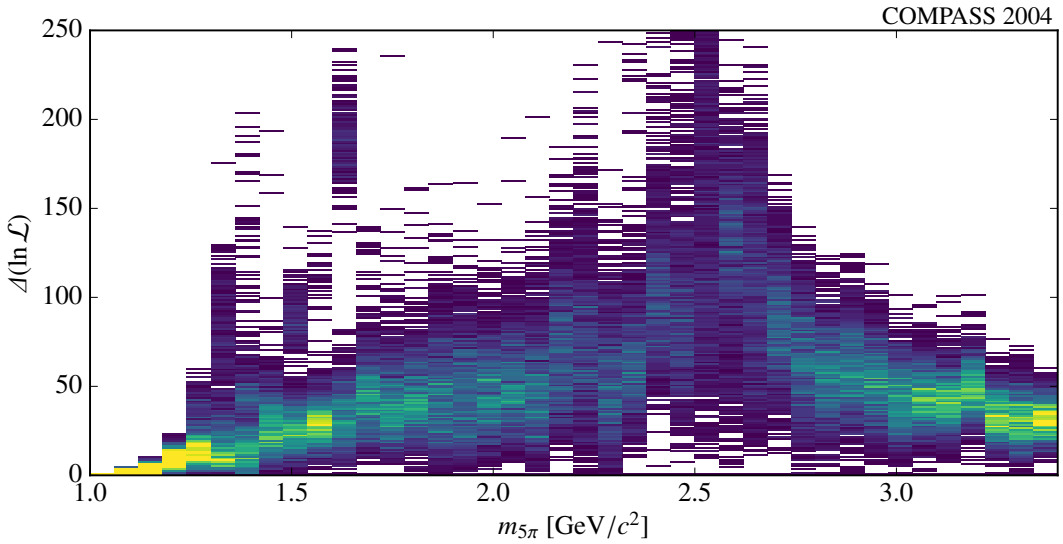


Figure 4.33: Partial-wave intensities for selected waves for all MLEs for the COMPASS 2008 event sample.



(a)



(b)

Figure 4.34: Likelihood distribution of all MLEs for all mass bins of the COMPASS 2004 event sample. In (b), the same plot is shown with a zoom on the y-axis.

4.3.3 Results on Measured Data (2004)

The BCM method was also applied to the data set from the COMPASS 2004 campaign, which was the basis for the analysis in [10]. It contains data from the reaction $\pi^- + \text{Pb} \to 5\pi + \text{Pb}$ and covers the t' region from 0 to $5 \times 10^{-3}\,(\text{GeV}/c)^2$. The $m_{5\pi}$ region from 1.0 to 3.4 GeV/c^2 was analyzed in 60 MeV/c^2-wide five-pion mass bins. The event selection and the Monte Carlo event samples used for the normalization integrals were taken over from [10] without changes. These events were then analyzed in the same way as the 2008 data.

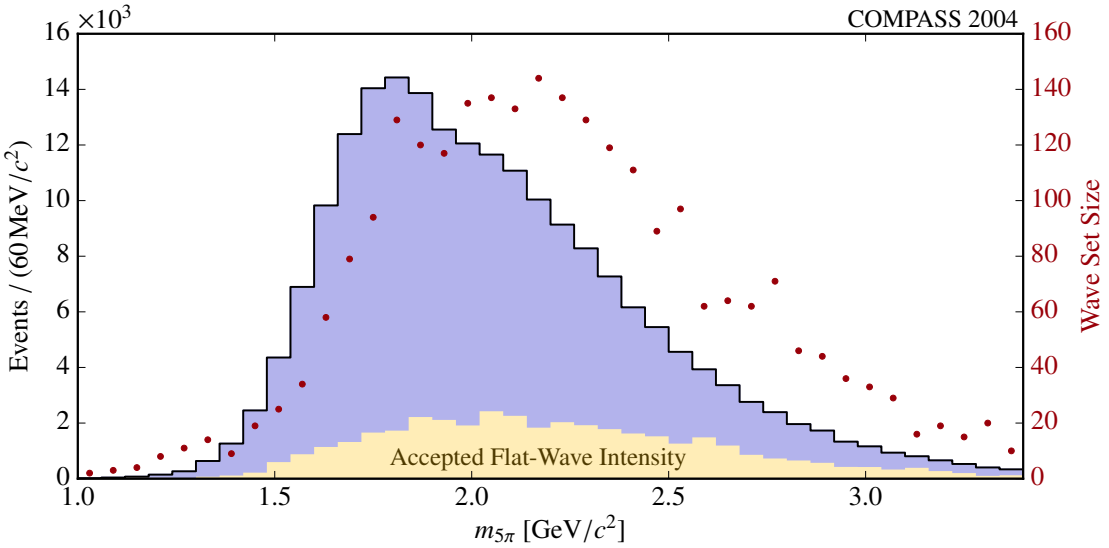


Figure 4.35: Selected wave set size (red), number of events (blue) and accepted flat-wave intensity (yellow) for the COMPASS 2004 event sample.

The same wave pool was used as biggest conceivable model. However, when investigating the results, it was found that the waves with $M = 1$ did not contribute significant intensity. This is expected because of the small t' of the COMPASS 2004 sample. Therefore, $M = 1$ waves were removed from the pool to reduce intensity variations. The pool with only $M = 0$ waves with has a size of 626. All results shown here have been obtained with the smaller pool.

Figure 4.34 shows the likelihood distribution for all mass bins. The picture is qualitatively very similar to the one shown in figure 4.23, with the exception of the low $m_{5\pi}$ region, where the fluctuation is very small. The probable cause for this is the small number of events there, with each of the first four bins having less than 200 entries.

In figure 4.35, the wave set size as a function of $m_{5\pi}$ is shown together with the number of events in the corresponding bins. Due to the smaller data set, less waves are needed to describe the data compared to 2008 (cf. figure 4.35). The wave set size follows the number of events also here, but at higher masses, wave sets are generally bigger.

Naturally, because of the smaller data set, the criteria for stable waves had to be loosened. This is true even for the considerably smaller pool with only $M = 0$ waves. Instead of demanding that the intensity found in all 1000 MLEs is different from zero, only the intensities from the best 10 % of MLEs had to be different from zero to select a wave. In the end, with the relaxed condition, 10 stable waves remain, which are shown in figures 4.36 to 4.38.

All stable waves found in the 2004 data sample are also among the 18 stable waves from the 2008 sample, except the $0^{-+} \to \rho(770)\ a_1(1260)$ wave shown in figure 4.36c. This wave also exhibits similar intensity rather stably in 2008, but did not meet the stricter stability criterion. When comparing the intensities from 2004 and 2008, note the different range of the x-axis. As far as the large intensity fluctuations allow to tell, the evolution of the intensity with $m_{5\pi}$ is compatible between the two data

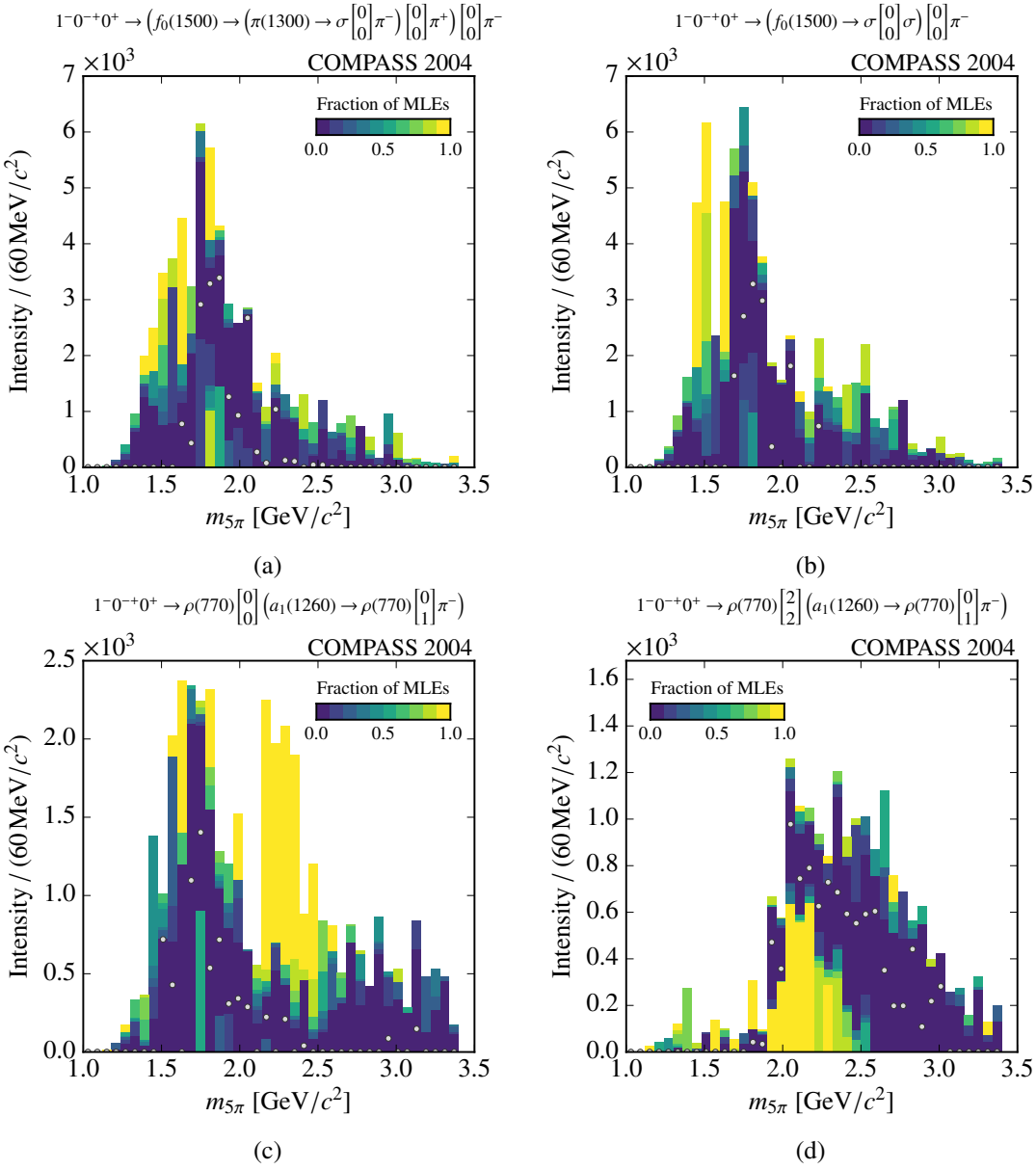


Figure 4.36: Partial-wave intensities for selected waves for all MLEs for the COMPASS 2004 event sample.

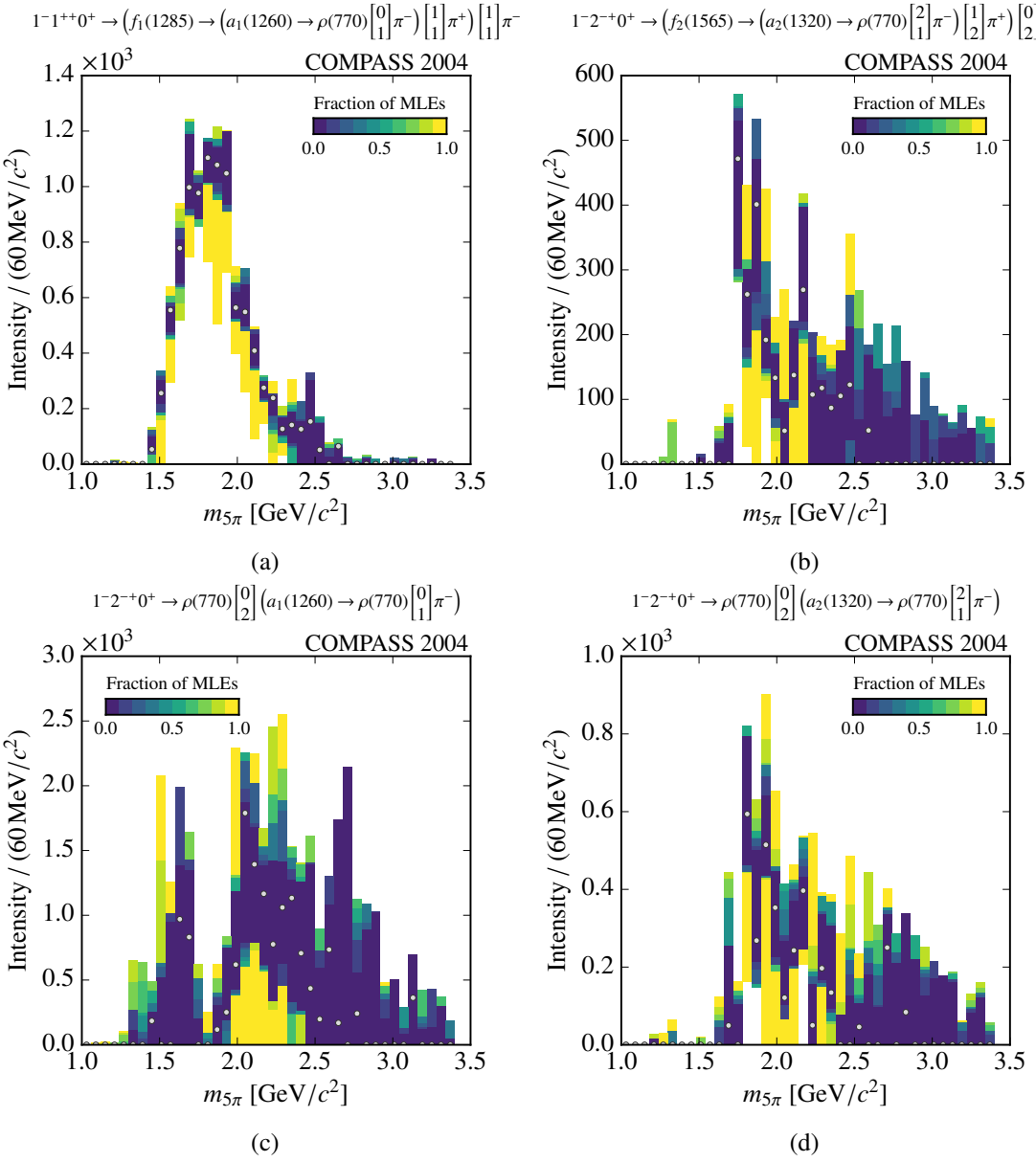
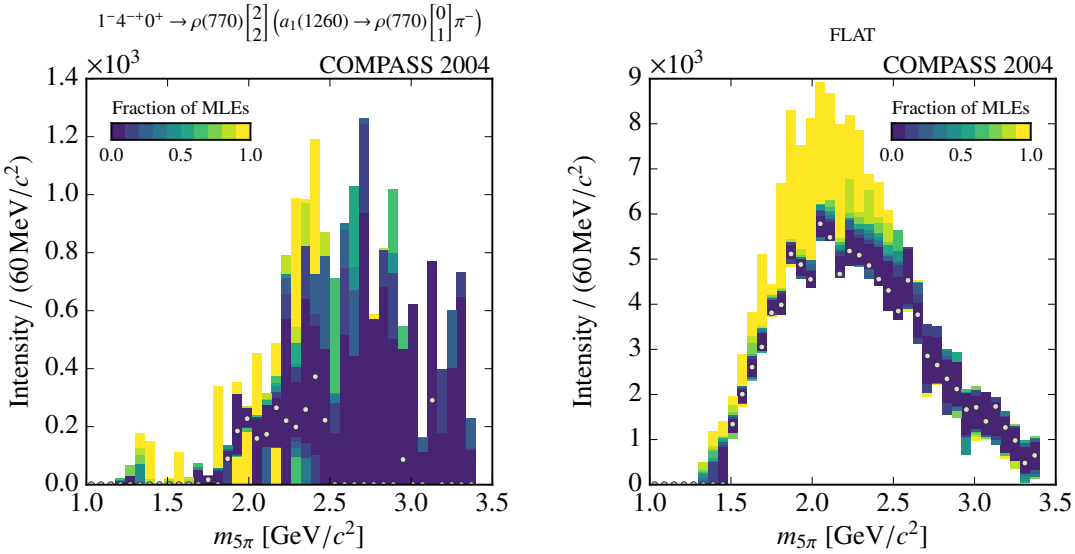


Figure 4.37: Partial-wave intensities for selected waves for all MLEs for the COMPASS 2004 event sample.



(a) (b)

Figure 4.38: Partial-wave intensities for selected waves for all MLEs for the COMPASS 2004 event sample.

sets. The $1^{++}0^+$ wave decaying into $f_1(1285)\,\pi^-$ shown in figure 4.37a exhibits a remarkable stability compared to the other 9 waves. Similar behavior was already observed in the 2008 data.

As before, the flat wave in figure 4.38b shows reliable extraction with the best MLEs typically exhibiting intensities on the lower end of the spread. In relation to the number of events, the flat wave has similar intensity in 2004 and 2008.

Comparing these results to the ones in [10], which were obtained using the genetic algorithm, one finds that the wave set size is generally larger, 20-140 versus 31, while the number of stable waves is smaller when using the BCM method. While there are several waves exhibiting similar intensities, the overall picture could not be reproduced. In light of the conclusion of section 4.2.3, which found that the genetic algorithm is prone to local convergence behavior, we may hypothesize that the solution found in [10] is a local one and that the model-selection uncertainties are severely underestimated. This means one cannot distinguish between reliable signals and unreliable ones. A nice example of this is the $2^{-+} \to \rho(770)\ a_1(1260)$ wave shown in figure 4.37c. While the wave set found by the genetic algorithm finds two peaks there, the BCM method reveals that the peak at lower masses is actually not stable and its interpretation as a physical signal is potentially doubtful.

Due to the reduced stability because of the smaller 2004 event sample and because there are no features showing up which are not present in the 2008 data set already, the 2004 data set has not been considered for further analysis.

CHAPTER 5

Results and Discussion

"That's one fine-looking barbecue pit."

Homer Jay Simpson

In the biggest-conceivable-model method, wave sets are selected by performing an MLE with an additional prior term and the full wave pool as model. The effect of the prior on waves which exhibit significant intensity is small by construction, so in principle, one could discuss the extracted amplitudes from the model selection directly. However, to exclude any bias on the results, the selected models were used in a maximum-likelihood estimate with the unmodified likelihood function from equation (3.32).

As mentioned at the end of the previous chapter, we will only discuss the results from the 2008 data set, which has been described at the beginning of section 4.3.2. In total, fifty maximum-likelihood estimations with random start values were performed for every mass bin, with all finding the same maximum likelihood estimate. This means that the partial-wave decomposition is remarkably stable despite the large number of waves in the selected model, if compared to e.g. the wave pool or the three-pion analysis, where a maximum of 88 waves is used [6].

5.1 Acceptance

The acceptance $\eta(\tau)$ was introduced when the partial-wave decomposition procedure was derived in section 3.2, more specifically in equation (3.14) on page 20. It is an important factor when judging the quality of the description of the data by the model, especially the comparison between the model prediction and the data. While the acceptance is included in the model in the form of normalization integrals, it is useful to project it to the set of variables which are used by the likelihood function. As described in section 3.3.1, these are eight angles in the Gottfried-Jackson and helicity frames and the three masses of the isobars. As there are three topologies for the isobaric decay into five pions, as shown in figure 3.2, there are three different possibilities to construct the reference frames and therefore three sets of eight angles to examine. Obviously, theses three sets of angles are not independent from each other, because the number of independent variables is still eleven. But, as

we are looking at one-dimensional projections of an eleven-dimensional distribution, there might be features visible in one projections which are hidden in another. Some angles are shared between different topologies and the isobar mass distributions are the same in every topology as well. When the redundancies are removed, twenty-three different projections remain, twenty angles and three masses.

The acceptance is determined by generating Monte Carlo events which are distributed isotropically in the five-pion phase space and then processing the events with a detector simulation and performing the event selection like on measured data. Because all the angular distributions are flat in phase space, the events accepted after the detector simulation directly show the shape of the acceptance. The dependence of the acceptance on the various angles is shown in figures 5.1 to 5.5. The distributions are arranged vertically for all bins in five-pion mass, with the mass increasing from bottom to top. The diagram on the top gives the topology, where the subsystem corresponding to the angle shown is highlighted in orange. Due to the symmetrizations discussed in section 3.3.2, there are multiple entries per event. As only those permutations are applied which lead to a change in the respective variable, this number is different for every angle. It is labeled "⊗" and given at the bottom of the distributions. Because the generated events are not uniformly distributed in the three isobar masses, the accepted events cannot be used directly to show the acceptance. Instead, in figure 5.5, the number of accepted events has been divided by the number of generated events for every isobar-mass bin. If there are only a small number of generated events, the statistical fluctuations get very large. For this reason, isobar-mass bins which have less than 50 events are not shown at all.

In general, the acceptance is rather flat in the projections onto the angles. Some of the distributions in $\cos(\vartheta)$ show a decreasing acceptance at the extremal values. Another departure from the flat distribution occurs in one of the φ_{12} angles shown in figure 5.4 on the far left, where there is a dip at $\varphi_{12} = 0$. Correlating the reference frames, which are different in every event, with the laboratory frame is very difficult, so that it was not possible to identify the origin of these features.

It should be pointed out that the flatness of the one-dimensional projections of the acceptance distribution does not necessarily mean that the eleven-dimensional acceptance distribution is flat, as modulations can be averaged out by projecting. However, as such acceptance modulations are taken into account by the ansatz for the likelihood function, see equation (3.14), they are not expected to distort the results.

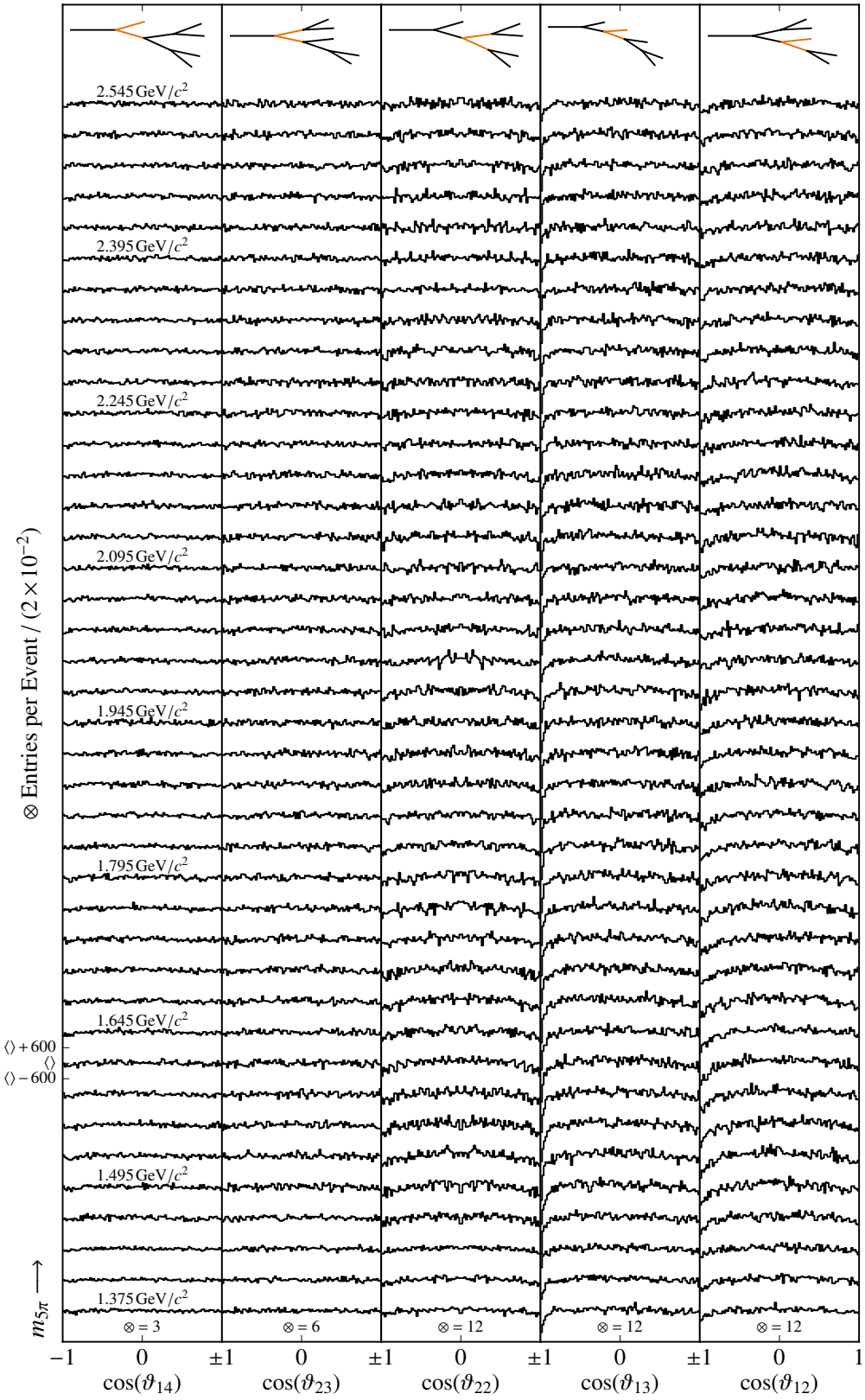


Figure 5.1: Acceptance distributions for all $m_{5\pi}$ bins. See text for details.

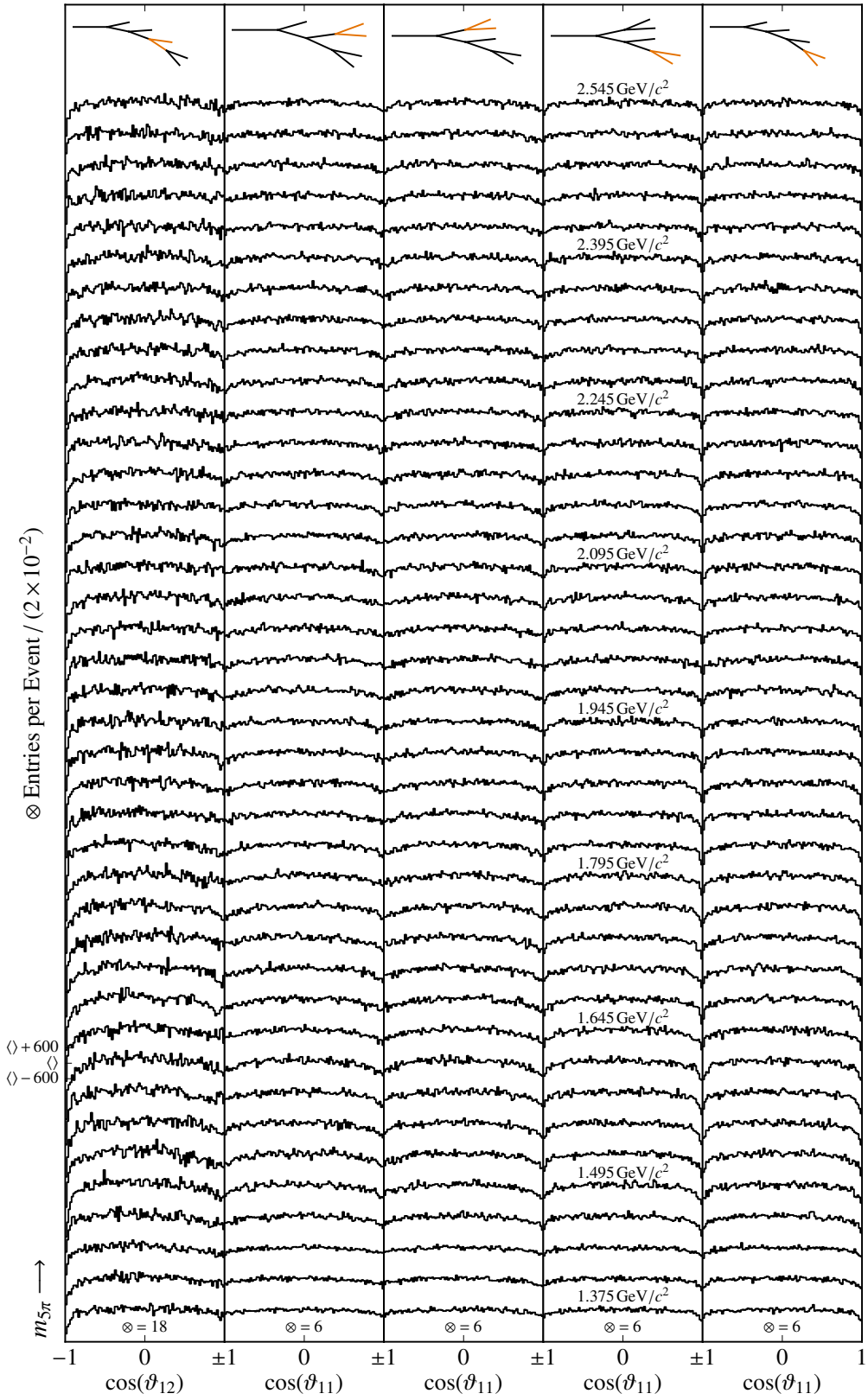


Figure 5.2: Acceptance distributions for all $m_{5\pi}$ bins. See text for details.

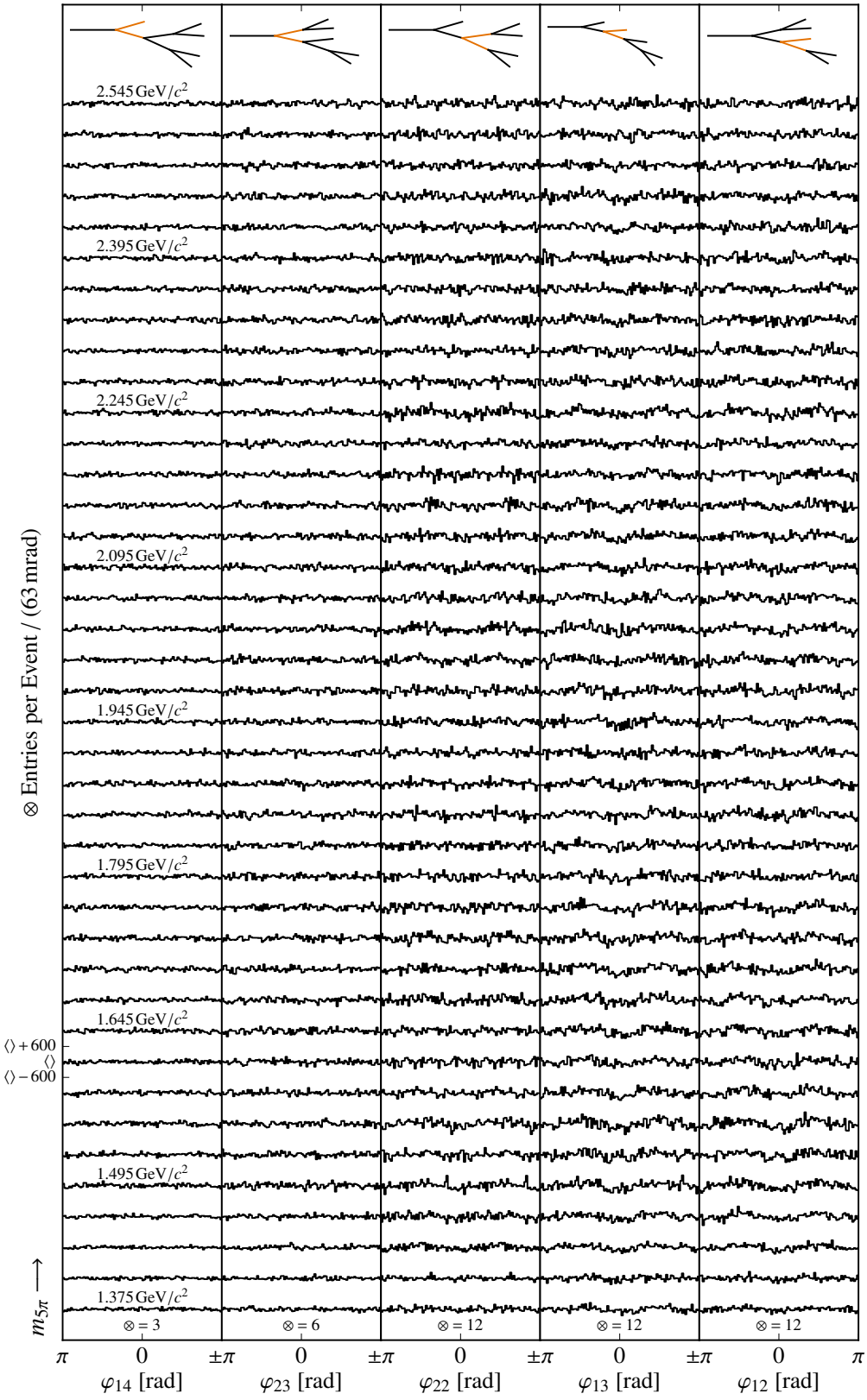


Figure 5.3: Acceptance distributions for all $m_{5\pi}$ bins. See text for details.

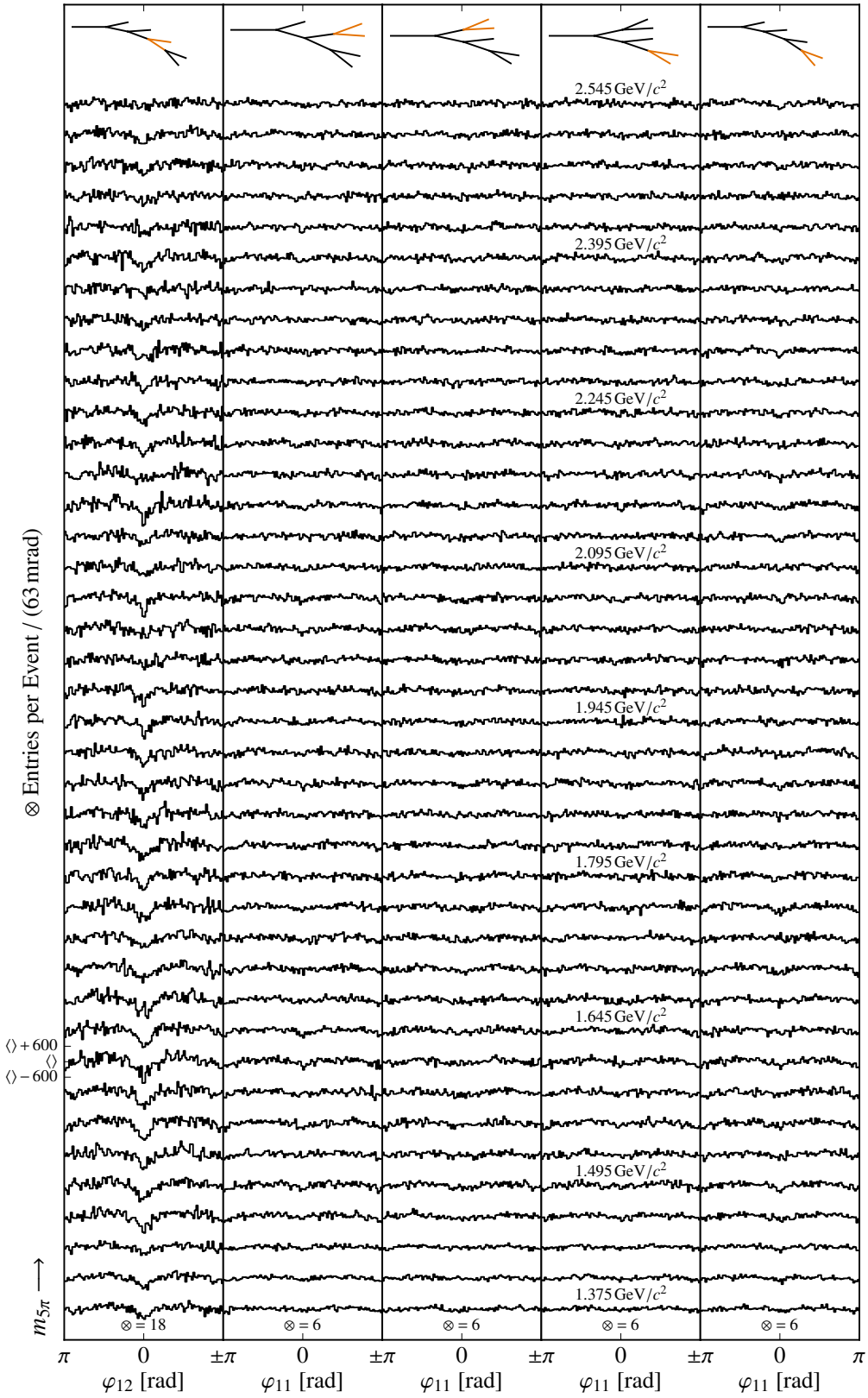


Figure 5.4: Acceptance distributions for all $m_{5\pi}$ bins. See text for details.

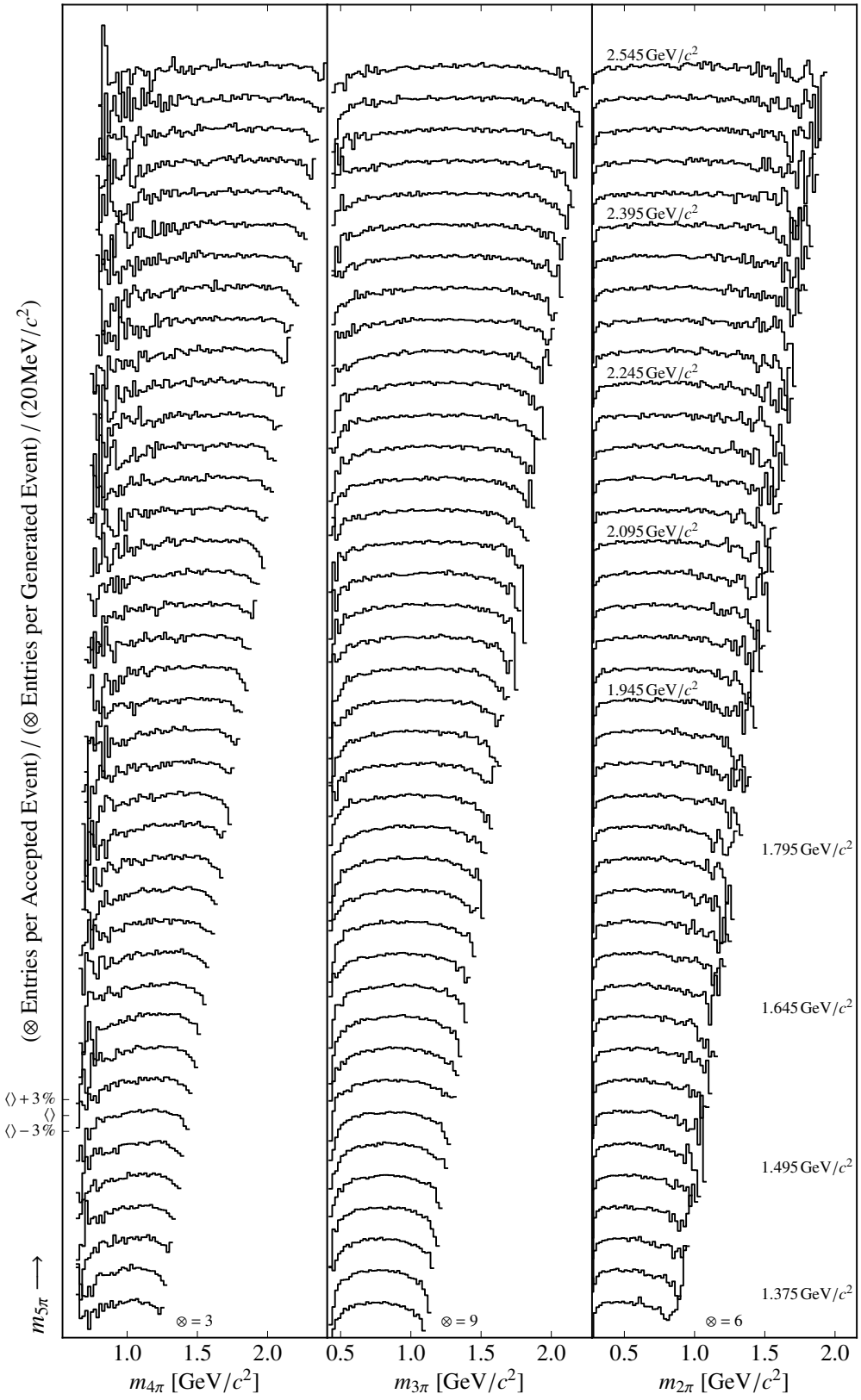


Figure 5.5: Acceptance distributions for all $m_{5\pi}$ bins. See text for details.

5.2 Comparison of the MLE Result with Data

The Monte Carlo event sample used to study the acceptance distributions can also be used to compare the MLE result with the measured data set. Approximating the integral in equation (3.16), that yields the expected number of events $\overline{N}$, with a sum, as shown in equation (3.24), taking the intensity from equation (3.20) and the transformation from equation (3.27), one obtains

$$\overline{N} = \frac{V_\Omega}{N^{\mathrm{MC}}} \sum_{\mathfrak{k}=1}^{N^{\mathrm{Acc}}} \left[\sum_{\epsilon=\pm1} \sum_{r=1}^{N_r^\epsilon} \left| \sum_i \frac{{}^r T_i^\epsilon}{\sqrt{\mathcal{P}_{ii}}} \Psi_i^\epsilon(\tau_\mathfrak{k}^{\mathrm{Acc}}) \right|^2 + \frac{T_{\mathrm{flat}}^2}{V_\Omega} \right] \tag{5.1}$$

We can cancel V_Ω and interpret every term in the sum over $\mathfrak{k}$ as the weight for the corresponding event

$$w_\mathfrak{k} = \frac{1}{N^{\mathrm{MC}}} \left[\sum_{\epsilon=\pm1} \sum_{r=1}^{N_r^\epsilon} \left| \sum_i \frac{{}^r T_i^\epsilon}{\sqrt{V_\Omega^{-1} \mathcal{P}_{ii}}} \Psi_i^\epsilon(\tau_\mathfrak{k}^{\mathrm{Acc}}) \right|^2 + T_{\mathrm{flat}}^2 \right] \tag{5.2}$$

Note that by the definition of $\mathcal{P}_{ij}$ in equation (3.28), V_Ω cancels also there. The distribution of the accepted Monte Carlo events weighed by $w_\mathfrak{k}$ can then be directly compared with the measured distributions.

Figures 5.6 to 5.10 show such a comparison for the angular and isobar-mass distributions in the same fashion as the acceptance plots explained in the previous section. Generally, the agreement is very good, however, there are several instances where at large polar angles and high five-pion masses, MLE result and data diverge. In the two-pion mass distributions, the peak of the $\rho(770)$ is not described well at higher $m_{5\pi}$. Considering the large pool of waves which was used in the model selection procedure, the appearance of such deviations is rather surprising. It leads to the conclusion that one or more waves needed to describe the given data set are missing from the wave pool. One possible explanation is that rules used for constructing the wave pool, given in section 4.1, where too restrictive. In particular the limit $L \leq 2$ might be too small, but due to the combinatorics of the decay topologies, increasing it would increase the wave pool dramatically. With the other construction rules staying the same, $L \leq 3$ or $L \leq 4$ would yield a wave pool with 3635 and 5711 waves, respectively. Also, even if all the required waves were present in the wave pool, imperfect isobar parameterizations might lead to discrepancies. Another possible reason for the deviations are features present in the data which are not compatible with the model assumptions altogether. We will explore this issue in section 5.5.

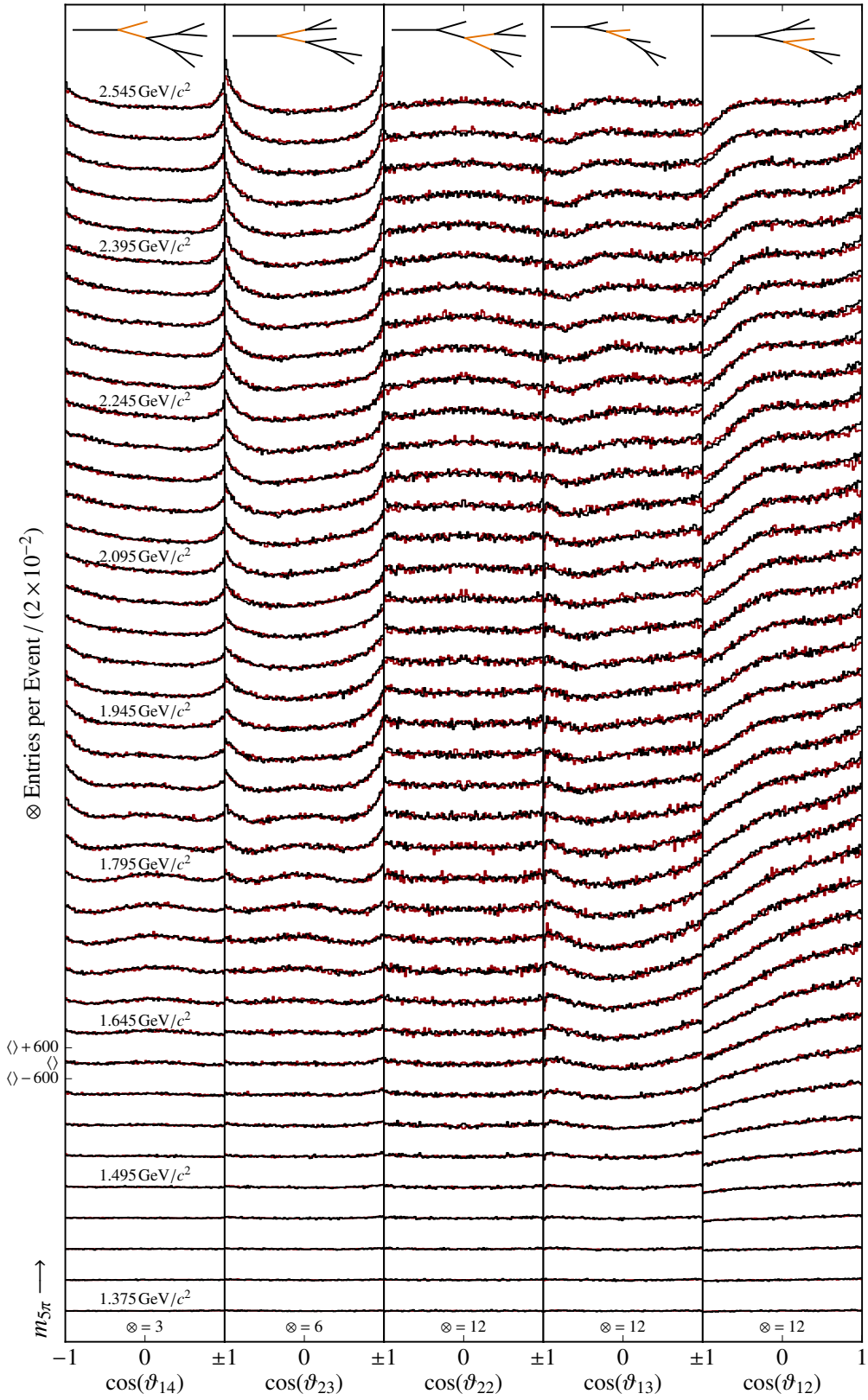


Figure 5.6: Comparison between model (red) and measured data (black). See text for details.

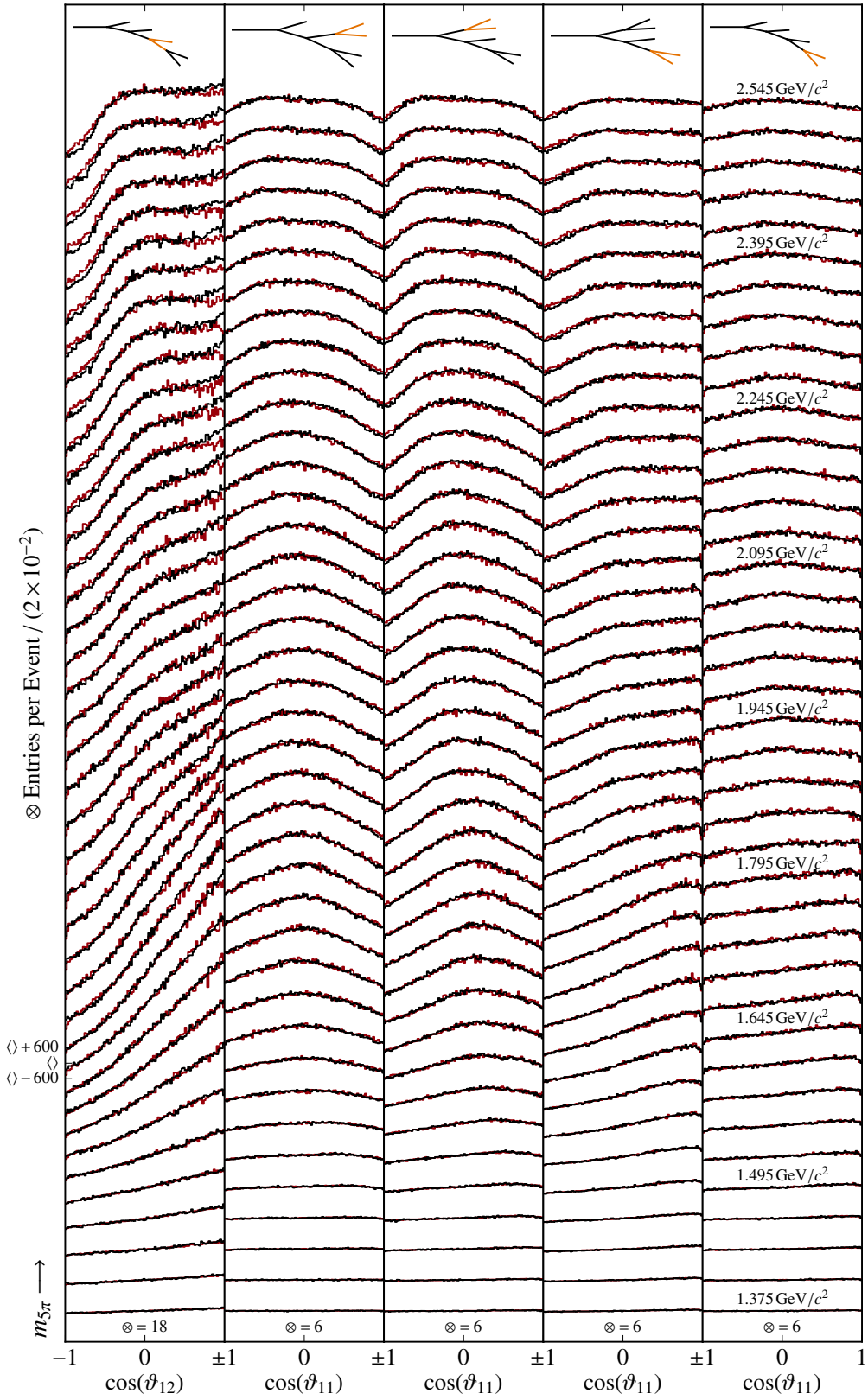


Figure 5.7: Comparison between model (red) and measured data (black). See text for details.

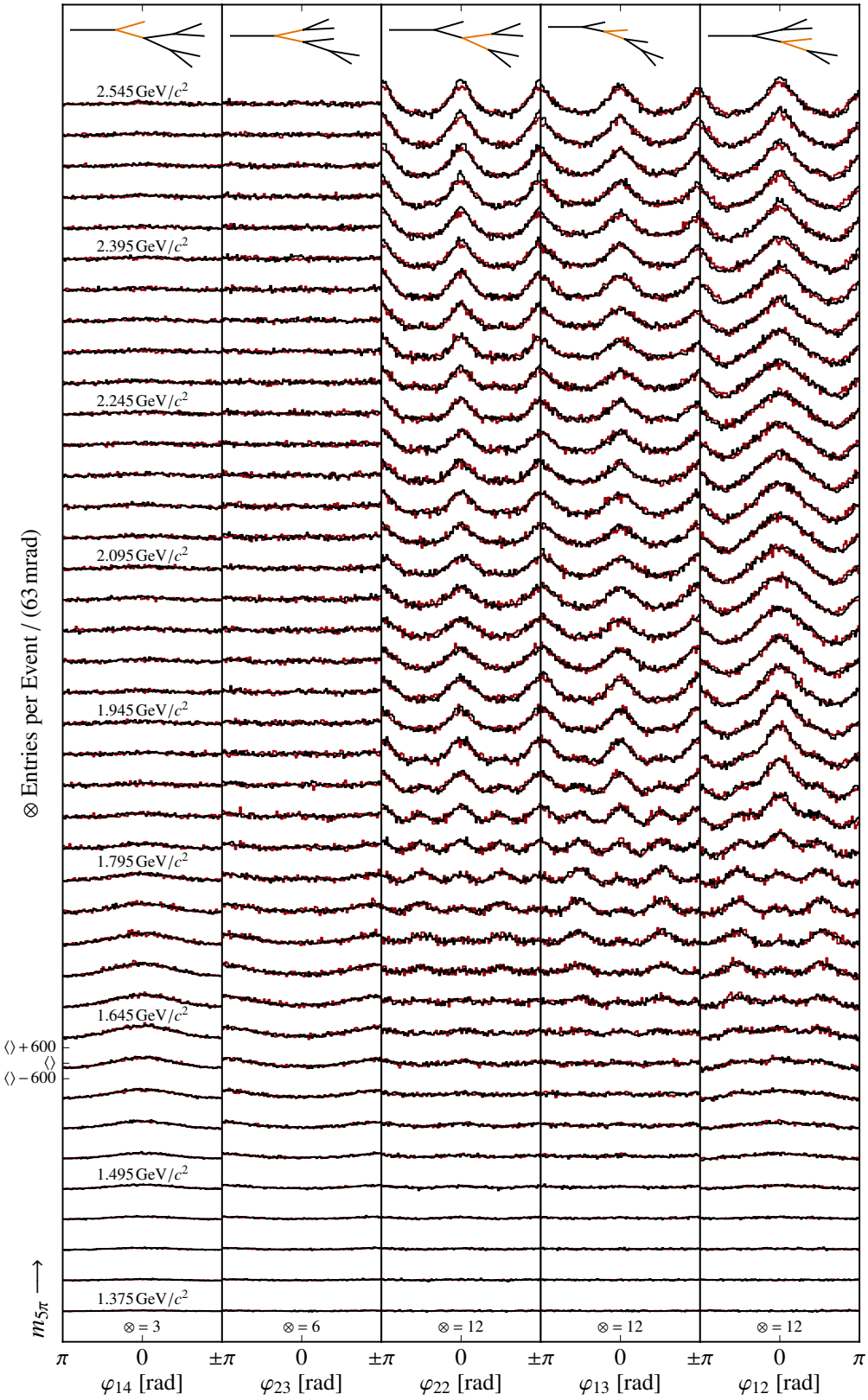


Figure 5.8: Comparison between model (red) and measured data (black). See text for details.

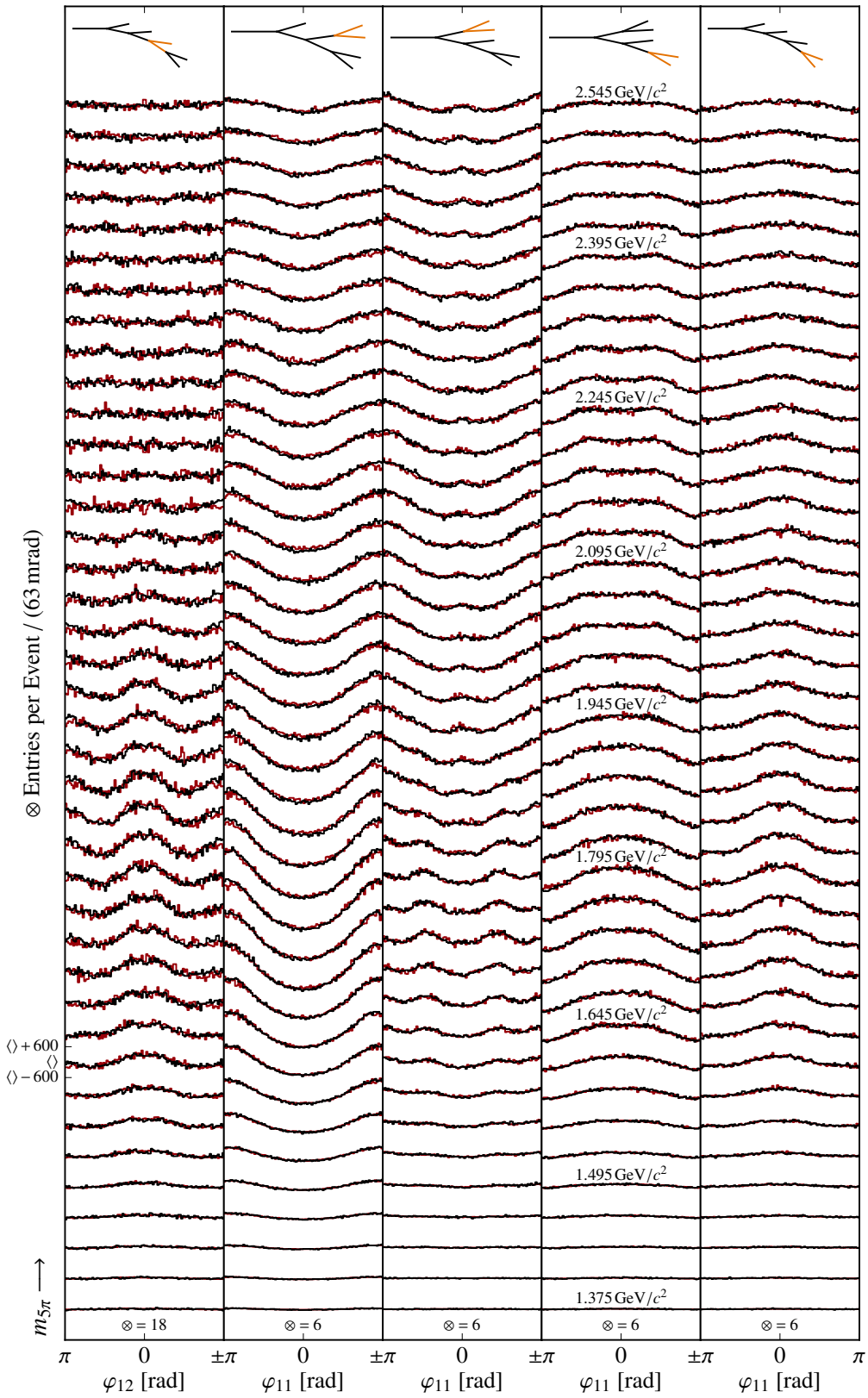


Figure 5.9: Comparison between model (red) and measured data (black). See text for details.

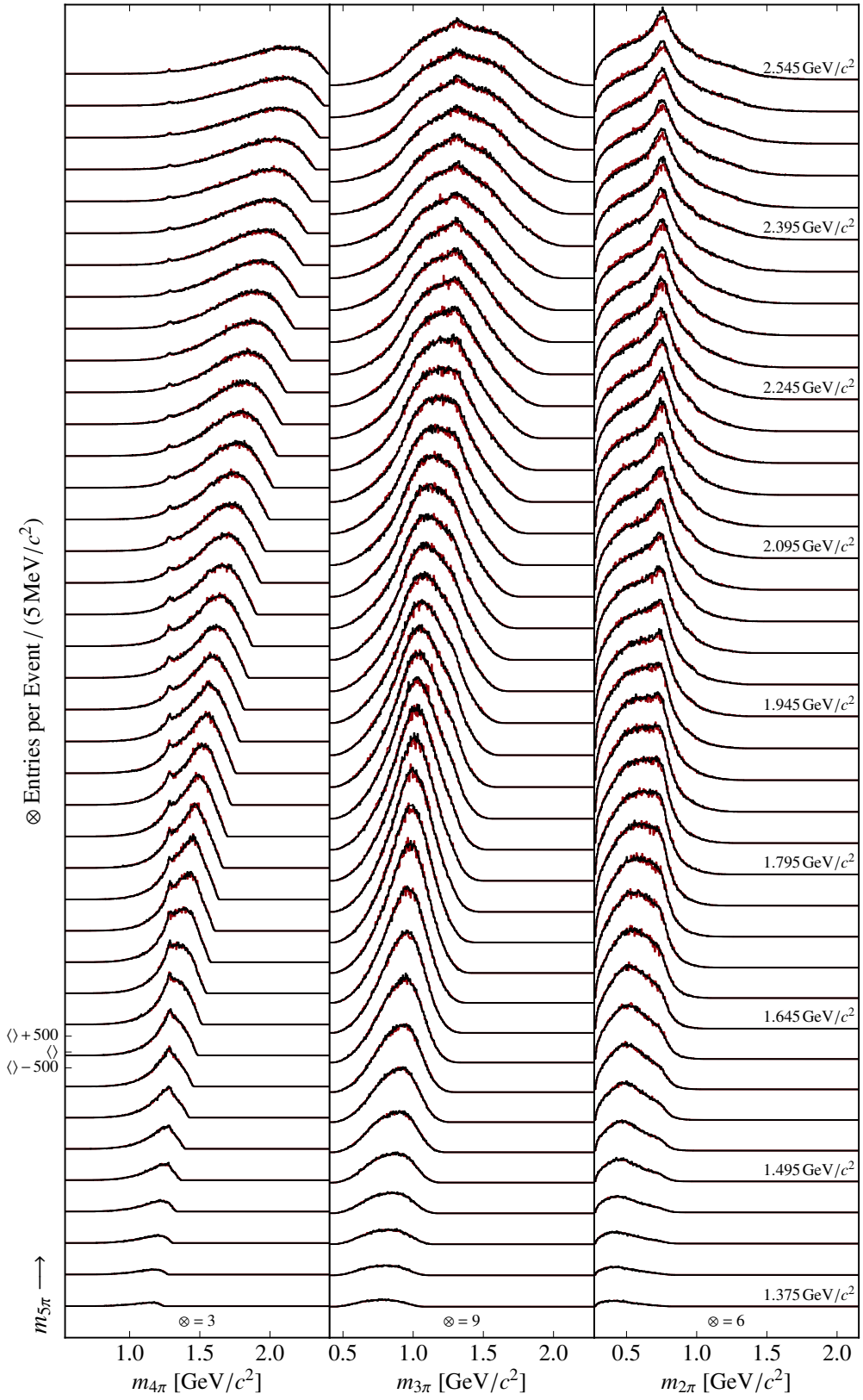


Figure 5.10: Comparison between model (red) and measured data (black). See text for details.

5.3 Quantity of Interest and Comparison with other Channels

The extracted parameters in the partial-wave decomposition described in section 3.2 are the production amplitudes $T_i = \mathcal{T}_i \cdot \sqrt{\mathcal{P}_{ii}}$. For this discussion, we introduce

$$^{\dagger}\mathcal{P}_{ij} = \frac{1}{N^{\mathrm{MC}}} \sum_{\mathrm{t}=1}^{N^{\mathrm{MC}}} \Psi_i^*(\tau_{\mathrm{t}}^{\mathrm{MC}})\, \Psi_j(\tau_{\mathrm{t}}^{\mathrm{MC}}) \tag{5.3}$$

and use this definition in equation (3.28) to write

$$\mathcal{P}_{ij} = V_{\Omega}\, ^{\dagger}\mathcal{P}_{ij} \tag{5.4}$$

For simplicity, the index ϵ for the reflectivity has been omitted. In section 3.2, the quantity V_{Ω} has been introduced as $V_{\Omega} = \int \mathrm{d}\phi_n(\tau)$, which is the total volume of the n-body, or in this case five-body, phase space. By construction, V_{Ω} does not explicitly appear in the partial-wave decomposition, but to interpret the five-pion mass dependence of T_i, the five-pion mass dependence of V_{Ω} has to be understood. This can best be seen in equation (3.54), which gives the parameterization of the spin-density matrix used in the resonance extraction described in section 3.4 and which contains $\mathcal{P}_{ii}$ and thereby also V_{Ω}. As $\int \mathrm{d}\phi_n(\tau)$ is a strictly increasing function of the n-body mass, one or more additional factors are required to dampen the growth of V_{Ω} so that the parameterization reproduces the measured $m_{5\pi}$ spectrum. Strictly speaking, these are separate from V_{Ω} and belong to the parameterization of $\mathcal{T}_i$, however, it simplifies the notation to change the definition of V_{Ω} so that it includes all additional factors and the remaining $\mathcal{T}_i$ can be parameterized by a Breit-Wigner function as given in equations (3.43) and (3.56). Therefore, we redefine

$$V_{\Omega}(m_{5\pi}) = V_{5\pi}(m_{5\pi})\, \xi(m_{5\pi}) \tag{5.5}$$

now writing the dependence on the five-pion mass explicitly. The term $V_{5\pi}(m_{5\pi}) = \int \mathrm{d}\phi_5(\tau; m_{5\pi})$ is the total volume of the five-pion phase space and the term $\xi(m_{5\pi})$ represents the damping factors.

To interpret the mass dependence of the extracted T_i, we need both $V_{5\pi}(m_{5\pi})$ and $\xi(m_{5\pi})$. The quantity $V_{5\pi}$ was directly extracted from binned Monte Carlo events. For $\xi(m_{5\pi})$, the situation is more complicated. In [10], a single exponential

$$\xi_{\mathrm{ff}}(m_{5\pi}) = \exp\left[-c_{\mathrm{ff}}\,(m_{5\pi} - m_{\mathrm{ts}})^2\right] \tag{5.6}$$

was used in a form-factor approach[a], assuming that all produced resonances have the same form factor. The slope parameter c_{ff} was a free parameter in the resonance extraction and the value obtained was $c_{\mathrm{ff}} = 1.742\,(\mathrm{GeV}/c^2)^{-2}$. For the 2008 data set, the same procedure yielded a similar value of $1.635\,(\mathrm{GeV}/c^2)^{-2}$. In the analysis of the $\pi^- + p \to \pi^-\pi^+\pi^- + p$ and $\pi^- + p \to \pi^-\pi^0\pi^0 + p$ reactions, point-like resonances are assumed so that $\xi_{\mathrm{ff}} = 1$. However, the distortion of the production phase

[a] In [10], the mass term in the exponential is linear in $m_{5\pi}$, however, the source code of the framework from that time uses a squared mass term as given here.

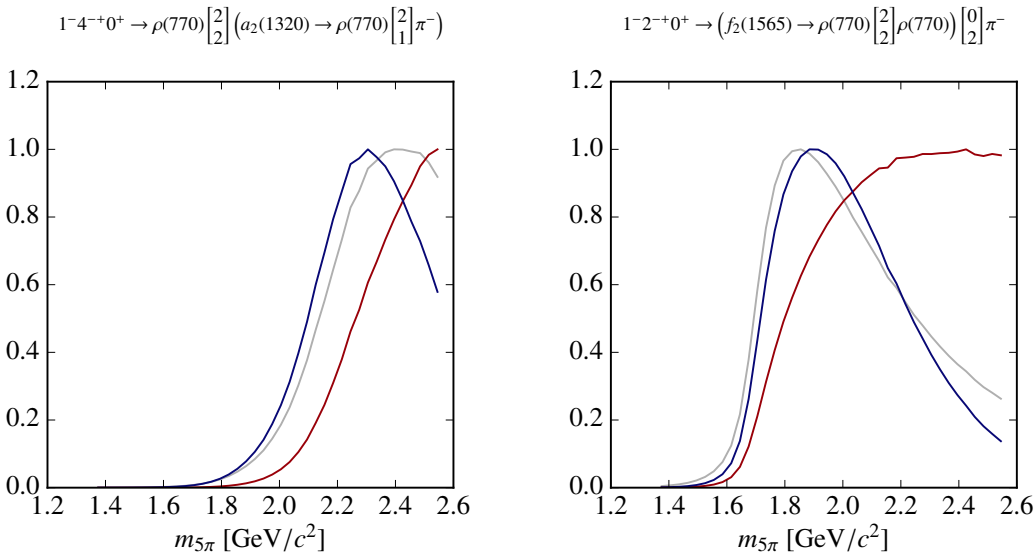


Figure 5.11: The five-pion mass dependence of ${}^{\dagger}\mathcal{P}_{ii}$ (gray), $[V_{5\pi}\,\xi_{\text{pps}}\,{}^{\dagger}\mathcal{P}_{ii}]$ (red) and $[V_{5\pi}\,\xi_{\text{ff}}\,{}^{\dagger}\mathcal{P}_{ii}]$ (blue) for two waves. All three curves have been normalized such that their maximum is one.

space by Pomeron exchange is taken into account [8, 9][b] by using

$$\xi_{\text{pps}}(m_{5\pi}) = \left| \frac{250720}{m_{5\pi}^{2.4749}} - \frac{172190}{m_{5\pi}^{2.8420}} \right| \frac{1}{m_{5\pi}} \tag{5.7}$$

As ξ does not depend on the decay process, at least approximately, it should not depend on the final state under study. Because the form factor and the production-phase-space factor are not mutually exclusive, it is possible that neither parameterization given above is sufficient and a new approach combining both needs to be found. The consequences of choosing one or the other parameterization for $\xi(m_{5\pi})$ will be discussed below. For now, let us note that the quantity in which one would expect to see resonances in the form of the familiar Breit-Wigner amplitude is

$$\begin{aligned} \mathcal{T}_i(m_{5\pi}) &= T_i(m_{5\pi}) \,/\, \sqrt{V_{\Omega}(m_{5\pi})\,{}^{\dagger}\mathcal{P}_{ii}(m_{5\pi})} \\ &= T_i(m_{5\pi}) \,/\, \sqrt{V_{5\pi}(m_{5\pi})\,\xi(m_{5\pi})\,{}^{\dagger}\mathcal{P}_{ii}(m_{5\pi})} \end{aligned} \tag{5.8}$$

It is worth pointing out that the square root in equation (5.8) is real-valued, meaning the phase differences between waves do not depend on it.

To illustrate the effect of the different parameterizations for ξ, figure 5.11 shows the five-pion mass dependence of ${}^{\dagger}\mathcal{P}_{ii}$, $[V_{5\pi}\,\xi_{\text{pps}}\,{}^{\dagger}\mathcal{P}_{ii}]$ and $[V_{5\pi}\,\xi_{\text{ff}}\,{}^{\dagger}\mathcal{P}_{ii}]$ for two different waves. The extent of the effect is more visible when the extracted intensities are actually divided by these terms. A particularly instructive example is the $J^{PC} = 4^{-+}$ wave shown in figure 5.12, which shows a clear peak in intensity. In $|\mathcal{T}|^2$ however, the peak is shifted but still present if ξ_{pss} is used, while it all but vanishes when

[b] Note that [8] and [9] disagree on the factor $1/m_{5\pi}$, which originates in different conventions regarding the determination of $V_{5\pi}$. The parameterization of ξ_{pps} as given by equation (5.7) is the one which is consistent with $V_{5\pi}$ used here.

using ξ_{ff}. Note that in figure 5.12, the vertical bars indicate the bin width and the horizontal bars the statistical uncertainties. Even though this is a worst-case example, it illustrates how dramatically the interpretation of the results may change with the choice of ξ. As stated above, it is not obvious which parameterization is more appropriate and it might very well be that neither is applicable in this case. Therefore, a different approach is chosen here: the quantities which will be shown are

$$^{\dagger}\mathcal{T}_i(m_{5\pi}) = T_i(m_{5\pi}) \,/\, \sqrt{^{\dagger}\mathcal{P}_{ii}(m_{5\pi})} \tag{5.9}$$

or rather, the intensities

$$\left|^{\dagger}\mathcal{T}_i(m_{5\pi})\right|^2 = \left|T_i(m_{5\pi})\right|^2 \,/\, {}^{\dagger}\mathcal{P}_{ii}(m_{5\pi}) \tag{5.10}$$

While they are very difficult to interpret by eye, they are as model-independent as possible, making comparison with future research straightforward.

While the model-selection procedure already provides a measure for the stability of the extraction of the production amplitudes, there is another independent way to cross-check some of the results. Several of the isobars used in the analysis of the five-pion channel have been observed to decay not only into charged pions, but also into a combination of η and pions. Two notable examples are

$$f_1(1285) \to 2\pi^+ + 2\pi^- \quad \text{and} \quad f_1(1285) \to \eta + \pi^+ + \pi^-$$

as well as

$$a_2(1320) \to 2\pi^{\pm} + \pi^{\mp} \quad \text{and} \quad a_2(1320) \to \eta + \pi^{\pm}$$

This means that a partial-wave decomposition of the reaction $\pi^- + p \to \eta\,\pi^-\pi^+\pi^- + p$ will have a number of waves in common with the five-pion channel, i.e. they only differ in the decay of one of the isobars. For simplicity, we use the abbreviations "$\eta(3\pi)$" and "5π" from here on. Similarly, there are related waves also in the $\pi^- + p \to \pi^-\pi^+\pi^- + p$ channel, abbreviated to "3π" from here on, namely

$$f_0(1500) \to 2\pi^+ + 2\pi^- \quad \text{and} \quad f_0(1500) \to \pi^+ + \pi^-$$

as well as

$$f_2(1270) \to 2\pi^+ + 2\pi^- \quad \text{and} \quad f_2(1270) \to \pi^+ + \pi^-$$

For related waves, the production amplitudes $\mathcal{T}_i$ should be identical between the two channels. If we assume this to be true, we can deduce the quantities which can be directly compared to each other

$$\begin{aligned}
\left|\mathcal{T}_i^{5\pi}\right|^2 &= \left|\mathcal{T}_i^{\eta(3\pi)}\right|^2 \\
\Leftrightarrow \frac{\left|T_i^{5\pi}\right|^2}{V_{5\pi} \cdot {}^{\dagger}\mathcal{P}_{ii}^{5\pi} \cdot \xi} &= \frac{\left|T_i^{\eta(3\pi)}\right|^2}{V_{\eta(3\pi)} \cdot {}^{\dagger}\mathcal{P}_{ii}^{\eta(3\pi)} \cdot \xi} \\
\Leftrightarrow \frac{\left|T_i^{5\pi}\right|^2}{{}^{\dagger}\mathcal{P}_{ii}^{5\pi}} &= \frac{\left|T_i^{\eta(3\pi)}\right|^2}{{}^{\dagger}\mathcal{P}_{ii}^{\eta(3\pi)}} \cdot \frac{V_{5\pi}}{V_{\eta(3\pi)}} \\
\Leftrightarrow \left|{}^{\dagger}\mathcal{T}_i^{5\pi}\right|^2 &= \left|{}^{\dagger}\mathcal{T}_i^{\eta(3\pi)}\right|^2 \cdot \frac{V_{5\pi}}{V_{\eta(3\pi)}}
\end{aligned} \tag{5.11}$$

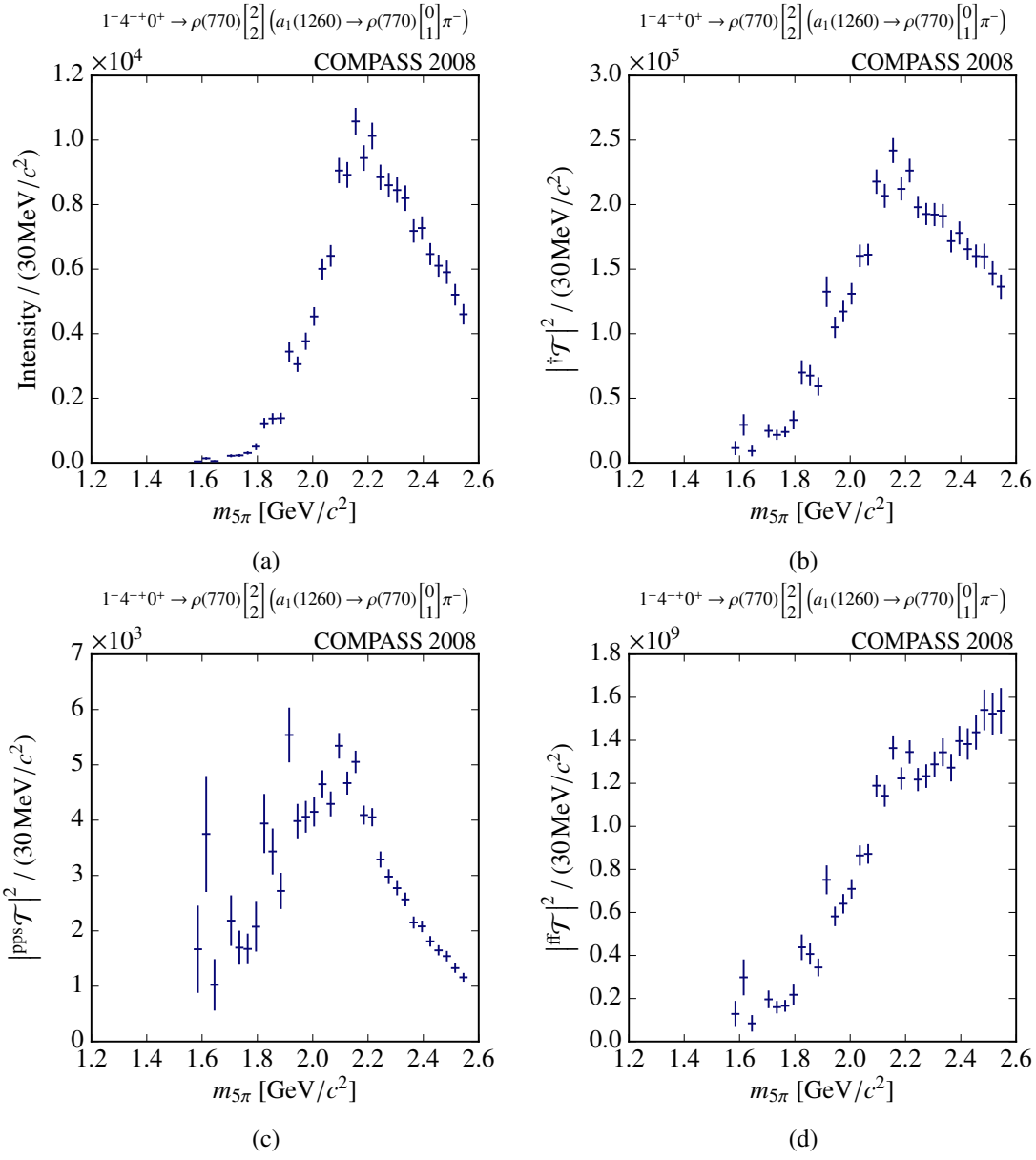


Figure 5.12: The extracted partial-wave intensity (a), divided by ${}^{\dagger}\mathcal{P}_{ii}$ (b), by $[V_{5\pi}\,\xi_{\mathrm{pps}}\,{}^{\dagger}\mathcal{P}_{ii}]$ (c) and by $[V_{5\pi}\,\xi_{\mathrm{ff}}\,{}^{\dagger}\mathcal{P}_{ii}]$ (d) for one wave.

where we assumed that ξ is indeed channel-independent. The same argument can be applied analogously to the 3π channel. However, as discussed after equation (3.34), in the partial-wave decomposition formalism used here, the couplings of isobars to their decay modes are included in the production amplitude. Therefore, a scaling factor should appear when comparing results from the two channels, which corresponds to the branching-fraction ratio of the different isobar decays.

Unfortunately, in the analysis of the $\eta(3\pi)$ and 3π channels, the factors $[{}^{\dagger}\mathcal{P}^{x}_{ii}\, V_x]$, with $x \in \{\, \eta(3\pi),\ 3\pi \,\}$, are not normalized from wave to wave. To keep the scale meaningful despite this property, the normalizations

$$\begin{aligned} \sum_{\mathfrak{k}=1}^{N^{5\pi}_{\text{bins}}} \frac{1}{{}^{\dagger}\mathcal{P}^{5\pi}_{ii}(m_{5\pi,\mathfrak{k}})} &\stackrel{!}{=} N^{5\pi}_{\text{bins}} \\ \sum_{\mathfrak{k}=1}^{N^{x}_{\text{bins}}} \frac{V_{5\pi}(m_{5\pi,\mathfrak{k}})}{{}^{\dagger}\mathcal{P}^{x}_{ii}(m_{x,\mathfrak{k}})\, V_x(m_{x,\mathfrak{k}})} &\stackrel{!}{=} N^{x}_{\text{bins}} \quad x \in \{\, \eta(3\pi),\ 3\pi \,\} \end{aligned} \tag{5.12}$$

were applied for every wave. The number of final-state mass bins N_{bins} is used to compensate for different bin widths in the different channels. The introduction of this normalization ensures that the scale of $|{}^{\dagger}\mathcal{T}_i|^2$ is on average the same as the scale of $|T_i|^2$, which, by construction, is measured in number of events. The different shapes of the corresponding ${}^{\dagger}\mathcal{P}^{x}_{ii}$ and V_x with $x \in \{\, 5\pi,\ \eta(3\pi),\ 3\pi \,\}$ are corrected for, while the ratios of their total integrals over m_{FS} are lost. As a consequence, the shapes of intensity distributions should be comparable but it should not be possible to extract branching-fraction ratios from the results as they are presented here. Because at small m_{FS}, ${}^{\dagger}\mathcal{P}_{ii}$ can become very small, the normalization in equation (5.12) is very sensitive to small differences in the decay-amplitude calculation in different channels and fluctuations due to the numerical precision of the Monte Carlo integration. Therefore, the summation is only performed in the m_{FS} interval from 1.60 to 2.56 GeV/c^2 with N^{x}_{bins} set accordingly.

5.4 Results

In section 4.3.2, eighteen 5π waves where identified as stably and reliably extracted. For these waves, the intensities $|{}^{\dagger}\mathcal{T}_i|^2$ as defined in equation (5.10) and the corresponding phase differences are shown in figures 5.13 to 5.15 and 5.17 to 5.19. For the $\eta(3\pi)$ channel, a partial-wave decomposition based on the same 2008 data set and the same t' bin is available [30], using a hand-picked model consisting of 88 waves. The acceptance correction which was applied only takes into account the geometrical acceptance without using the full detector simulation [49]. The partial-wave decomposition of the 3π channel presented in [6] has been repeated for the t' bin also used in the five-pion analysis, including a full acceptance correction [30]. Where related waves from the $\eta(3\pi)$ or 3π channels exist, their $|{}^{\dagger}\mathcal{T}_i|^2$ have been included in the corresponding figures, with the scaling given by equation (5.11) applied.

As discussed in the previous section, the phase differences are not affected by the transformations given by equations (5.10) and (5.11). Where applicable, the corresponding phase differences for the $\eta(3\pi)$ channel are also included in in the figures. The wave with $J^{PC} = 1^{++}$ decaying into $f_1(1285)\,\pi^-$, which is used as a reference for all phase differences, is not present in the 3π channel. Therefore,

the phase differences for that channel cannot be included in the same way and will be shown and discussed separately.

Towards the lower and higher mass limits, ${}^{\dagger}\mathcal{P}_{ii}$ can get very small, so that points with small intensity and errors compatible with zero are scaled up to dominate the intensity plots. For the 5π channel, such points have been removed from the plots if they lay far above all other points and their position has been marked with an arrow at the top edge. For the $\eta(3\pi)$ and 3π channels, points have been removed for which the scaling factor due to the application of equation (5.11) was exceptionally large, which removes also low-lying points with very large error bars. To emphasize the subtly different procedures, the position of points removed due to scaling have been marked with a gray cross at the top edge.

To make the comparison of the intensity shapes easier, the remaining points for both $\eta(3\pi)$ and 3π have been scaled such that their integral over m_X is equal to the integral of the points from the 5π channel. This scaling factor is given in the plots.

As stated in the previous section, the shown intensity $|{}^{\dagger}T_i|^2$ cannot be easily interpreted without dividing it by $V_{5\pi}\,\xi$. For this reason, we refrain from discussing the shapes of the intensity distributions shown here, focusing instead on the comparison with the $\eta(3\pi)$ and 3π channels and the phase differences.

Figures 5.13a to 5.13c show three different $0^{-+}0^{+} \to f_0(1500)\,\pi^-$, $L = S = 0$ waves with different decays of the $f_0(1500)$. All three 5π waves are compared with the same wave from the 3π channel, where the $f_0(1500)$ decays only into $\pi^+\pi^-$. As per the discussion above, the three 5π waves should look the same, which is the case on a qualitative level. However, it is evident that the two waves with $f_0(1500) \to \pi(1300)\,\pi^-$ exhibit a broader distribution than the $f_0(1500) \to \sigma\,\sigma$ mode. The mass and width of the $\pi(1300)$ are not very well known [4], so that an inadequate isobar parameterization is one possible explanation for this discrepancy. In 3π, the structure is slightly narrower than in the $f_0(1500) \to \sigma\,\sigma$ wave, although the agreement in figure 5.13c is generally good. For all three 5π waves, the phase differences with respect to the $1^{++} \to f_1(1285)\,\pi^-$ wave are rather flat and therefore markedly different from what one would expect from a Breit-Wigner resonance (see figure 3.6b). For the reasons discussed at the end of the previous section, branching fractions cannot be extracted from the scaling factors shown in the plots. Nevertheless, one can attempt to verify that the order of magnitude of the scaling factors are plausible by comparing them to previously measured branching fractions. As only the branching fraction $\mathcal{B}[f_0(1500) \to 4\pi]$ is listed in [4], the three $0^{-+}0^{+} \to f_0(1500)\,\pi^-$ waves need to be combined for a meaningful comparison. This has been implemented using the formalism discussed in appendix D and is shown in figure 5.13d, again together with the 3π results. It is worth to note that the phase for the combined intensity is not defined. One can see that the agreement with the 3π channel is better than for any of the single waves, at least in the peak region. The large error bars can be attributed to the correlations between the production amplitudes, a residual feature of the ambiguities already observed in the model selection and discussed in section 4.3.2. The branching-fraction ratio $\mathcal{B}[f_0(1500) \to 4\pi]/\mathcal{B}[f_0(1500) \to 2\pi]$ is 1.42 ± 0.13, so that the scaling factor of 6 is compatible, considering the systematic effects: apart from the effect of the normalization mentioned above, the branching-fraction ratio from [4] includes all charge combinations in the decay final states, whereas we are only investigating the fully-charged decay modes. Also, there are other $0^{-+}0^{+} \to f_0(1500)\,\pi^-$ waves in the wave set which are not looked at because they are not considered stable after the model-selection procedure, but might contain a significant amount of intensity, skewing the scaling factor.

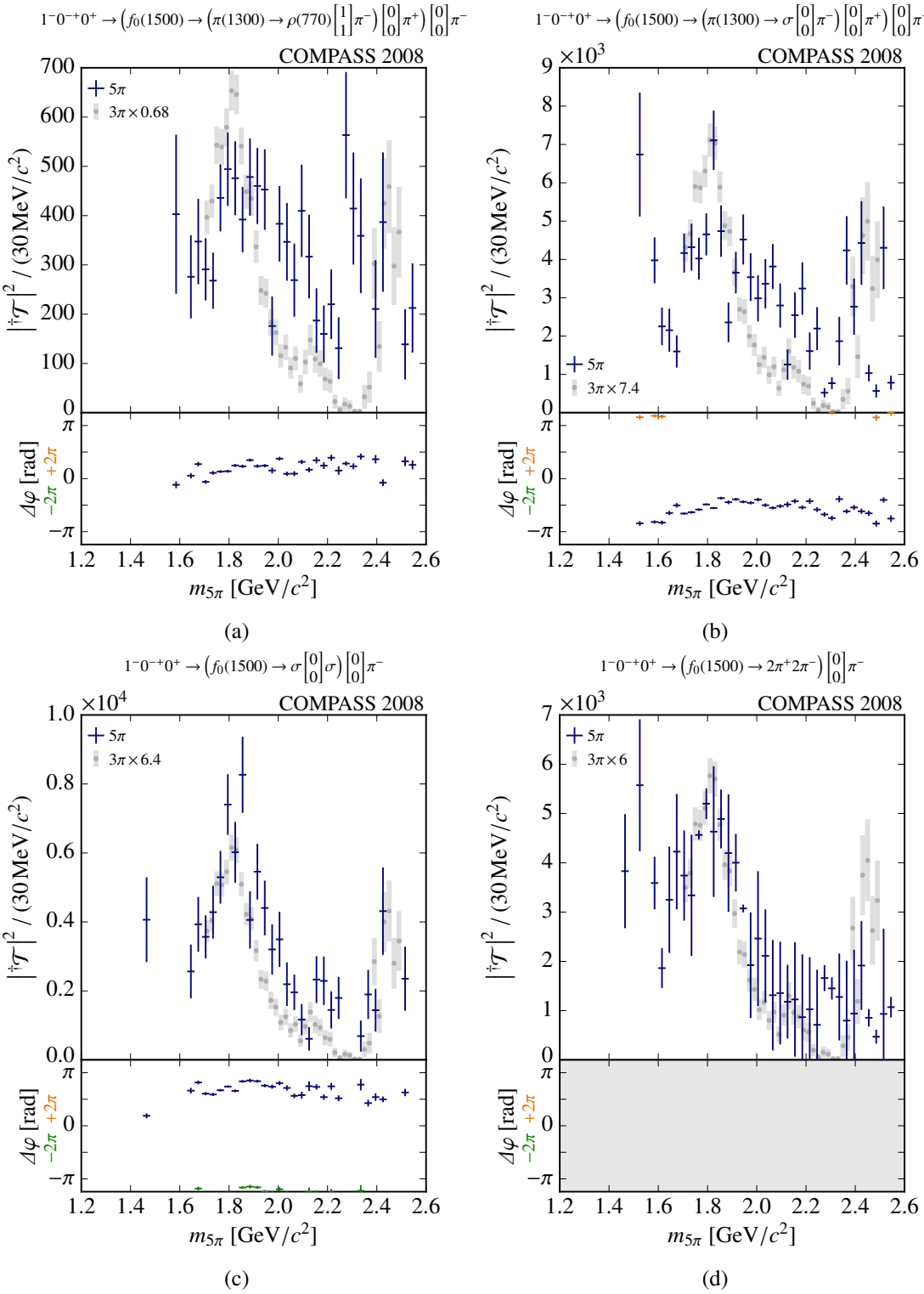


Figure 5.13: Intensities and phase differences for $0^{-+}0^+ \to f_0(1500)\,\pi^-$ waves (a)-(c) and their combined intensity (d). The phase difference is calculated by subtracting the phase of the $1^-1^{++}0^+ \to f_1(1285)\,\pi^-$ wave shown in figure 5.17a.

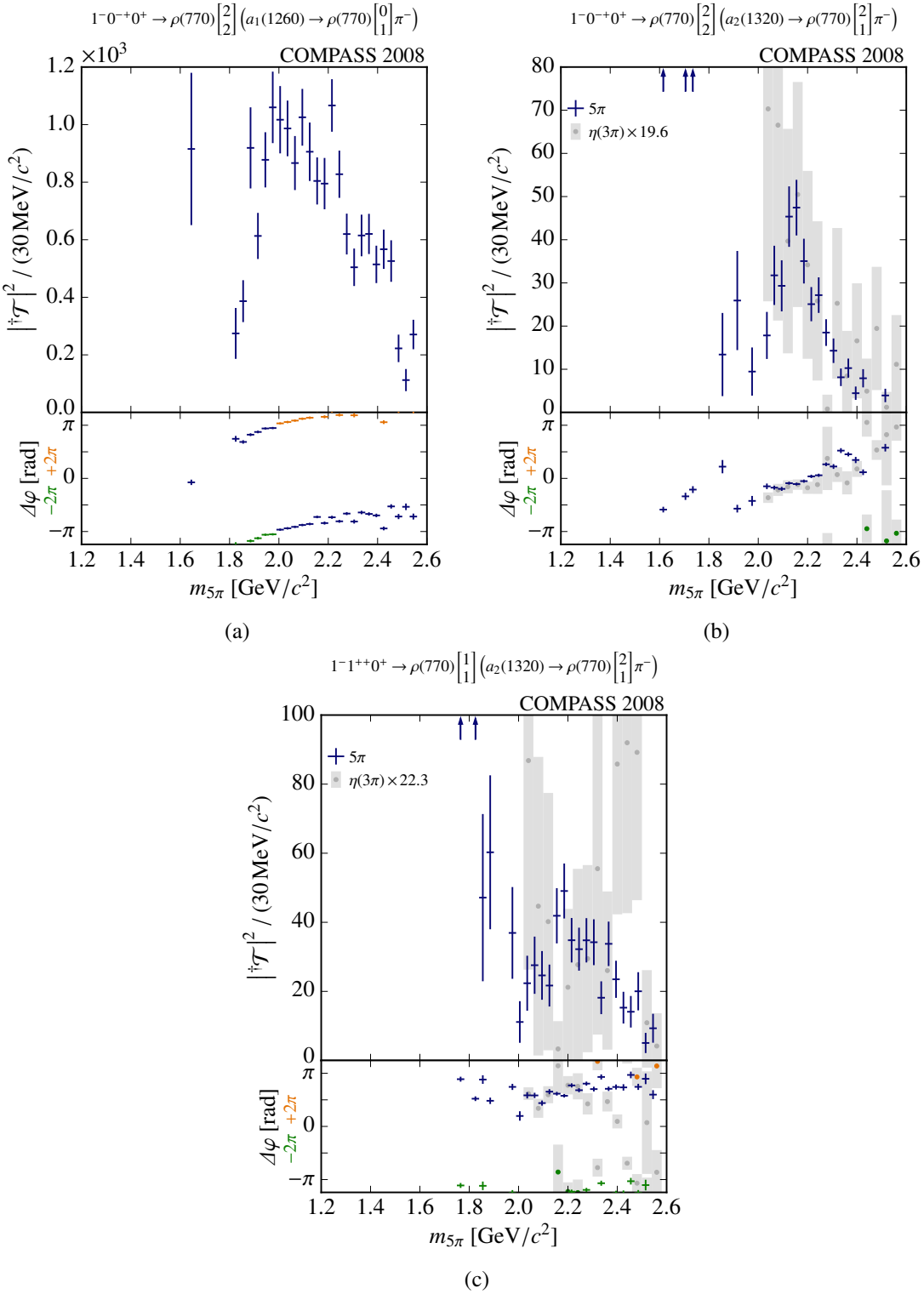


Figure 5.14: Intensities and phase differences for selected waves of the partial-wave decomposition. The phase difference is calculated by subtracting the phase of the $1^-1^{++}0^+ \to f_1(1285)\pi^-$ wave shown in figure 5.17a.

In figure 5.14a, a $0^{-+} \to \rho(770)\, a_1(1260)$ wave is shown which does not have any corresponding waves in the $\eta(3\pi)$ or 3π channels. It exhibits intensities at higher masses and does feature a phase motion, albeit a rather shallow one. Figures 5.14b and 5.14c show a 0^{-+} and 1^{++} wave, respectively, both decaying to $\rho(770)\, a_2(1320)$. Both waves are in agreement with the results from $\eta(3\pi)$, although the statistical fluctuations are large in the $\eta(3\pi)$ channel because of the relatively lower branching fraction of $a_2(1320) \to \eta\pi^{\pm}$. The scaling factors for the $\eta(3\pi)$ waves will be discussed separately in the next section. The phase motion of the 0^{-+} wave in figure 5.14b is interesting, rising by $\sim \pi$ over a mass range of about $300\,\mathrm{MeV}/c^2$, which could be indicative of a resonant structure in this wave.

Figure 5.15 shows all the stable 5π $J^{PC} = 2^{-+}$ waves without the $f_1(1285)$ in the decay. The $2^{-+}0^+ \to f_2(1270)\,\pi^-$ wave in figure 5.15a has a related wave in the 3π channel. Using the isospin Clebsch-Gordan coefficients to account for the fact that the branching fraction ratio given by [4] includes both the $\pi^-\pi^+$ and $\pi^0\pi^0$ decay modes, one arrives at a value of $\mathcal{B}[f_2(1270) \to 2\pi^+2\pi^-]/\mathcal{B}[f_2(1270) \to \pi^+\pi^-] = 0.05 \pm 0.007$. Indeed, the scaling factor found is very small, although also here, a quantitative comparison is not possible due to the normalization from equation (5.12). As far as one can tell given the large statistical fluctuations, the two channels are in agreement. The phase difference in the 5π channel is constant and does not show a resonance signature. For the $2^{-+}0^+ \to f_2(1565)\,\pi^-$ wave shown in figure 5.15b, the agreement between the five-pion and $\eta(3\pi)$ channels is very good, both in intensity and in phase. Also here, the phase difference are does not show any motion and is not what would be expected from a resonance.

For the $2^{-+}0^+ \to \rho(770)\ a_1(1260)$ wave shown in figure 5.15c, there is a very interesting phase motion of about π. Unfortunately, this motion occurs at a place where the intensity is very small, making it difficult to ascertain whether this is actually a physical signature or an artifact from the ill-determined phase at this point. For the $\eta(3\pi)$ channel, the $2^{-+}0^+ \to \rho(770)\ a_2(1320)$ wave shown in figure 5.15d seems to have only very small intensity and does not follow the clear and substantial peak seen in five pions. The scaling factor is unusually large. One hypothesis to explain this behavior is the presence of model artifacts affecting the $\rho(770)\ a_2(1320)$ waves of the $\eta(3\pi)$ channel. While the model-selection procedure shows very stable extraction of this wave in the five-pion channel (see figure 4.30d on page page 81), this does not necessarily have to be the case for $\eta(3\pi)$. In fact, a number of $\rho(770)\ a_2(1320)$ waves have been added to the $\eta(3\pi)$ partial-wave model because of their stability in the model selection of the five-pion channel, which might have introduced ambiguities in the MLE, leading to the observed discrepancies. Another possible explanation is that there are differences in the parameterizations of the decay amplitudes in the 5π and $\eta(3\pi)$ channels.

Having discussed now two waves which have related partners in the 3π channel, it is possible to compare phase differences for these waves between the 5π and 3π results. Figure 5.16 shows the phase difference of the three $0^{-+}0^+ \to f_0(1500)\,\pi^-$ waves discussed above and shown in figure 5.13, with respect to to the $2^{-+}0^+ \to f_2(1270)\,\pi^-$ wave shown in figure 5.15a. Because the couplings of the isobar decays can be complex-valued and are inseparable from the production amplitudes, different isobar decays can have a phase shift, which should be constant over m_{FS}. For this reason, the 5π phase differences have been shifted so that they match best. The phase shift applied is given in the plots. One can see that the phases of the three waves in the 5π channel have a very similar shape, which is expected. However, while the initial rise of the phase at lower masses is in good agreement between the 5π and 3π channels, the phase motion levels off at about $1.9\,\mathrm{GeV}/c^2$ in 5π, while it nicely continues to rise to complete almost 180° in 3π, indicating the presence of the $\pi(1800)$ in the $0^{-+}0^+ \to f_0(1500)\,\pi^-$ wave in the 3π channel. The reason for this is unclear, but possible

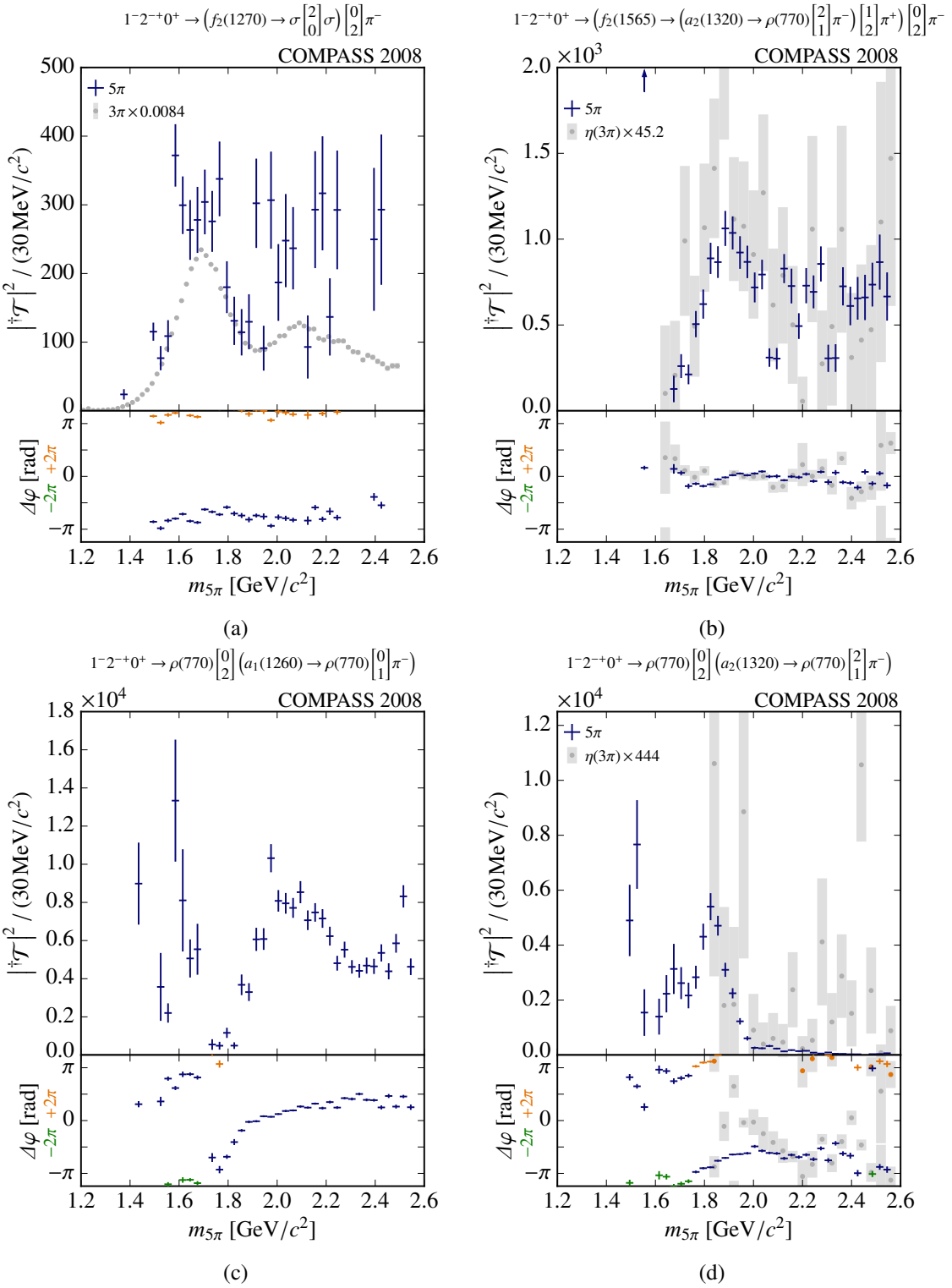


Figure 5.15: Intensities and phase differences for selected waves of the partial-wave decomposition. The phase difference is calculated by subtracting the phase of the $1^{++}0^+ \to f_1(1285)\,\pi^-$ wave shown in figure 5.17a.

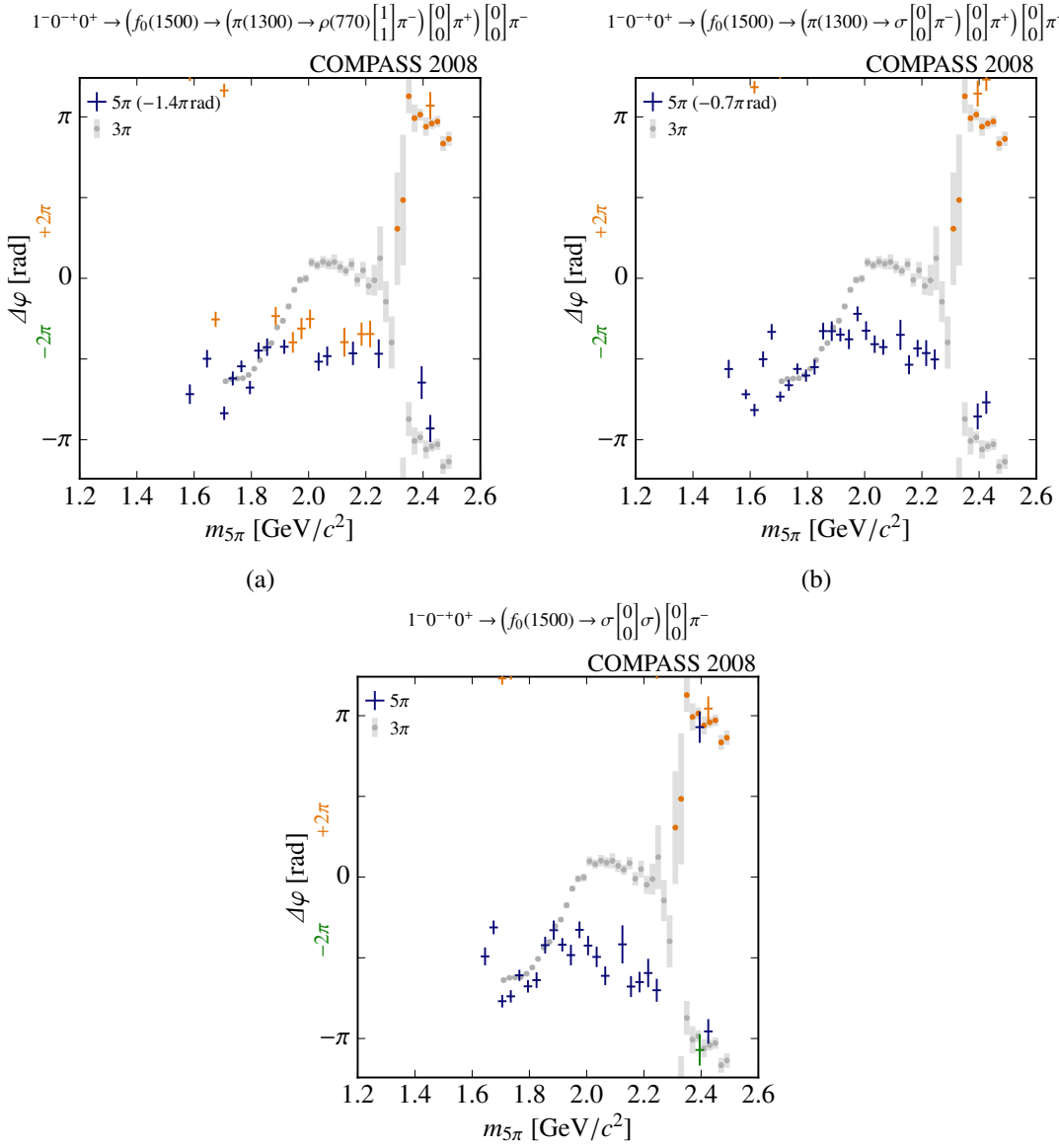


Figure 5.16: Phase differences of the $0^{-+}0^+ \to f_0(1500)\,\pi^-$ waves with respect to the $2^{-+}0^+ \to f_2(1270)\,\pi^-$ wave shown in figure 5.15a, with the latter subtracted.

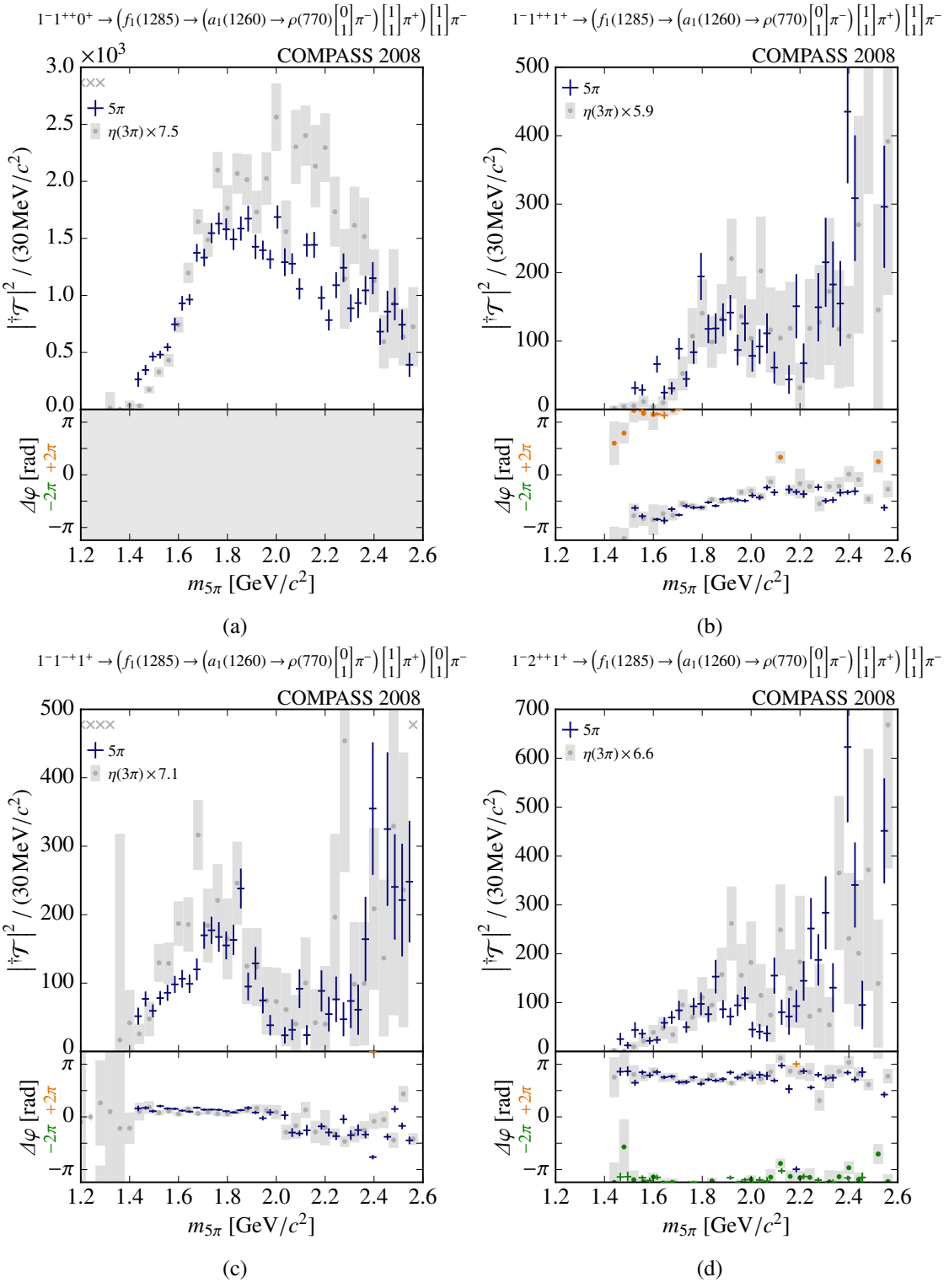


Figure 5.17: Intensities and phase differences for selected waves containing $f_1(1285)$ isobar of the partial-wave decomposition. The phase difference is calculated by subtracting the phase of the $1^-1^{++}0^+ \to f_1(1285)\,\pi^-$ wave shown in (a).

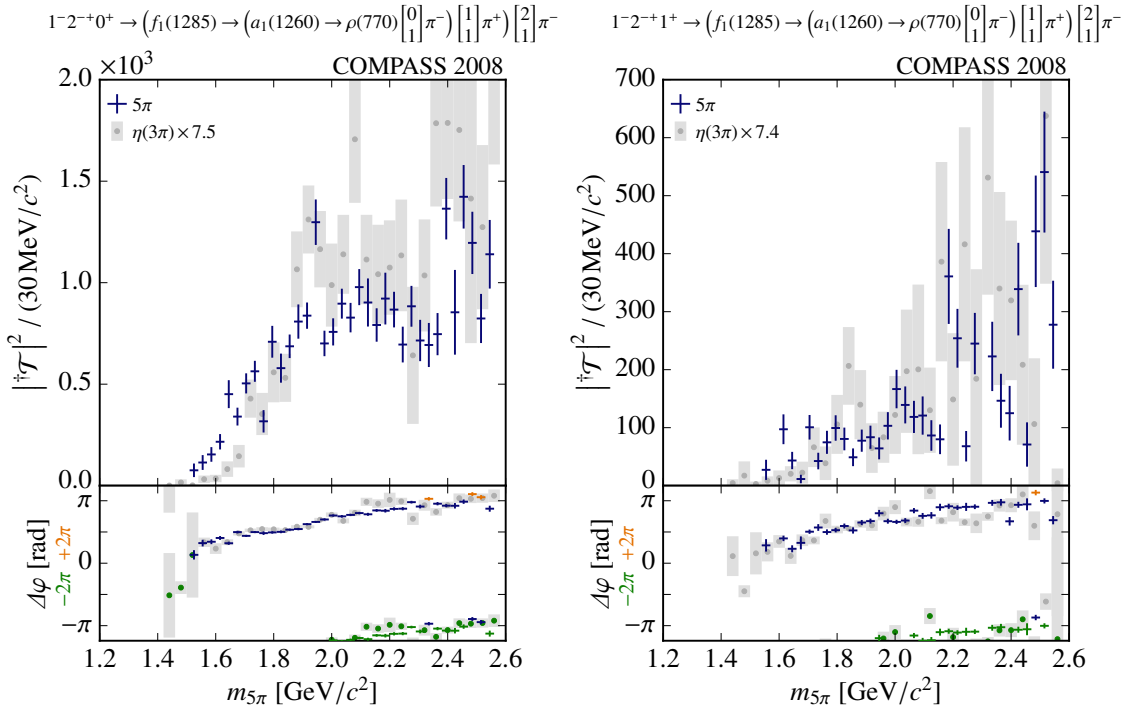
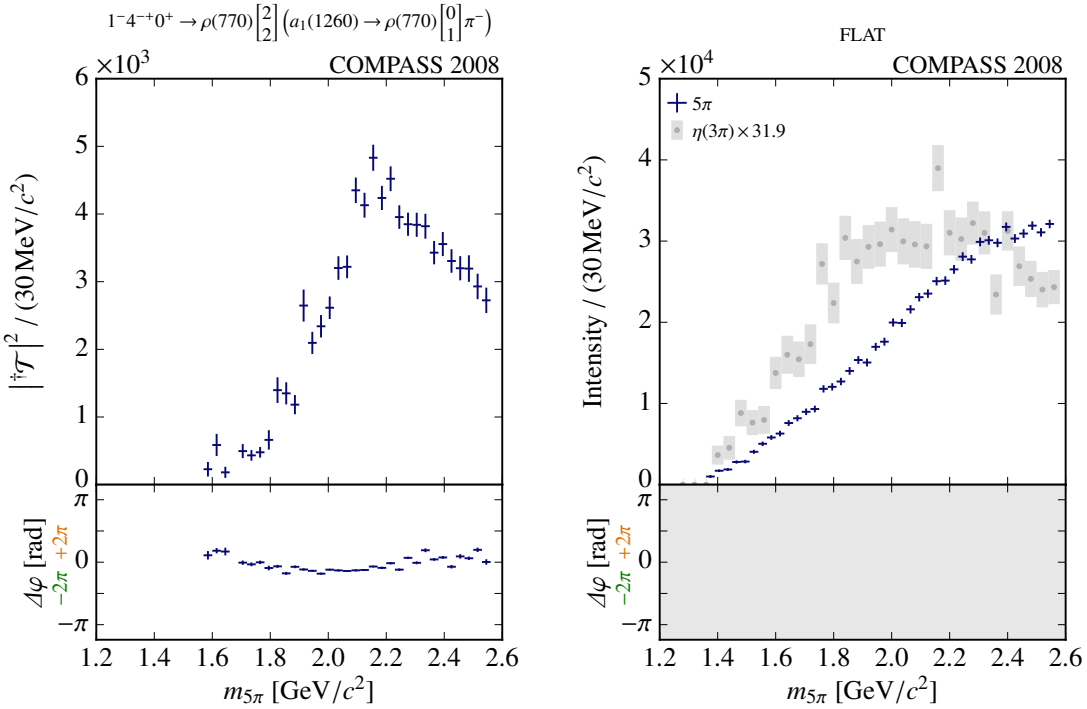


Figure 5.18: Intensities and phase differences for selected waves containing $f_1(1285)$ isobar of the partial-wave decomposition. The phase difference is calculated by subtracting the phase of the $1^-1^{++}0^+ \to f_1(1285)\,\pi^-$ wave shown in figure 5.17a.

explanations will be discussed in the next section. It is worth pointing out that for the $\eta(3\pi)$ channel, shifting the phase is not necessary to come to an agreement with the phases of the 5π channel. Also here, such phase shifts are only needed for the two waves with a $\pi(1300)$ in the decay, lending further support to the hypothesis that the isobar parameterization of this resonance is not adequate.

The stable 5π waves decaying into $f_1(1285)\,\pi^-$ are shown in figures 5.17 and 5.18. Figure 5.17a shows the $J^{PC} = 1^{++}1^+$ wave. The agreement is not as good between the 5π and $\eta(3\pi)$ channels, where there is an enhancement in the low-mass region in the 5π channel. Around 2.1 GeV/c^2, the $\eta(3\pi)$ channel exhibits higher intensity. For the $1^{++}1^+$, $1^{-+}1^+$ and $2^{++}1^+$ waves shown in figures 5.17b to 5.17d, the agreement both in intensity and in the phase difference are very good. In the $1^{++}1^+$ wave, there is a shallow phase motion. This is somewhat surprising as it is measured against the $M = 0$ variant of the same wave, so that any resonance showing up in one of the waves should also appear in the other and the phase difference between them is expected to be flat. For the $1^{-+}1^+$ and $2^{++}1^+$ waves, the phase difference is flat in m_{FS}, both for the 5π and $\eta(3\pi)$ channel. In figures 5.18a and 5.18b, the $2^{-+} \to f_1(1285)\,\pi^-$ waves are shown for $M = 0$ and $M = 1$, respectively. In the $M = 1$ wave, the 5π channel shows somewhat higher intensity at low masses, similar to what has been observed for the $1^{++}0^+$ wave. Apart from that, the agreement between the two channels is very good both in intensity and phase. The phase motions are similar for the two waves, with a slow rise over a very large m_{FS} range, which does not appear to conform to an interpretation as a resonance.



(a) (b)

Figure 5.19: Intensities and phase differences for selected waves of the partial-wave decomposition. The phase difference is calculated by subtracting the phase of the $1^-1^{++}0^+ \to f_1(1285)\pi^-$ wave shown in figure 5.17a.

Finally, figure 5.19a shows a $J^{PC} = 4^{-+}$ wave, which exhibits considerable intensity, but with little phase motion. The flat wave in figure 5.19b grows with mass. The behavior in $\eta(3\pi)$ differs somewhat, however, considering the different background processes and different wave sets in the two channels, this is expected. When the flat-wave intensity is multiplied by the acceptance bin by bin, it can be compared directly to the measured $m_{5\pi}$ spectrum. This "accepted flat-wave intensity" is shown in figure 4.26 on page page 77. One can see that despite the large number of waves present in the wave pool during the model selection, the intensity of the flat wave accounts for a substantial fraction of the measured events. This means that a significant portion of the data are compatible with an isotropic background.

5.5 Discussion

When looking at the results from the five-pion channel presented in the previous section, the most striking feature is the absence of clear, resonance-like phase motions. One benign reason for the lack of such phase motions is the phenomenon called phase locking, which appears when there is the same resonance appearing in two different waves. When two waves are phase-locked, their phase motions are the same and the phase difference between the two does not show any motion. The situation in the 5π data, however, is different as the majority of waves do not exhibit significant phase motion. This could only be explained by phase-locking if the resonances in the different waves would have very similar masses and widths. This seems rather unlikely.

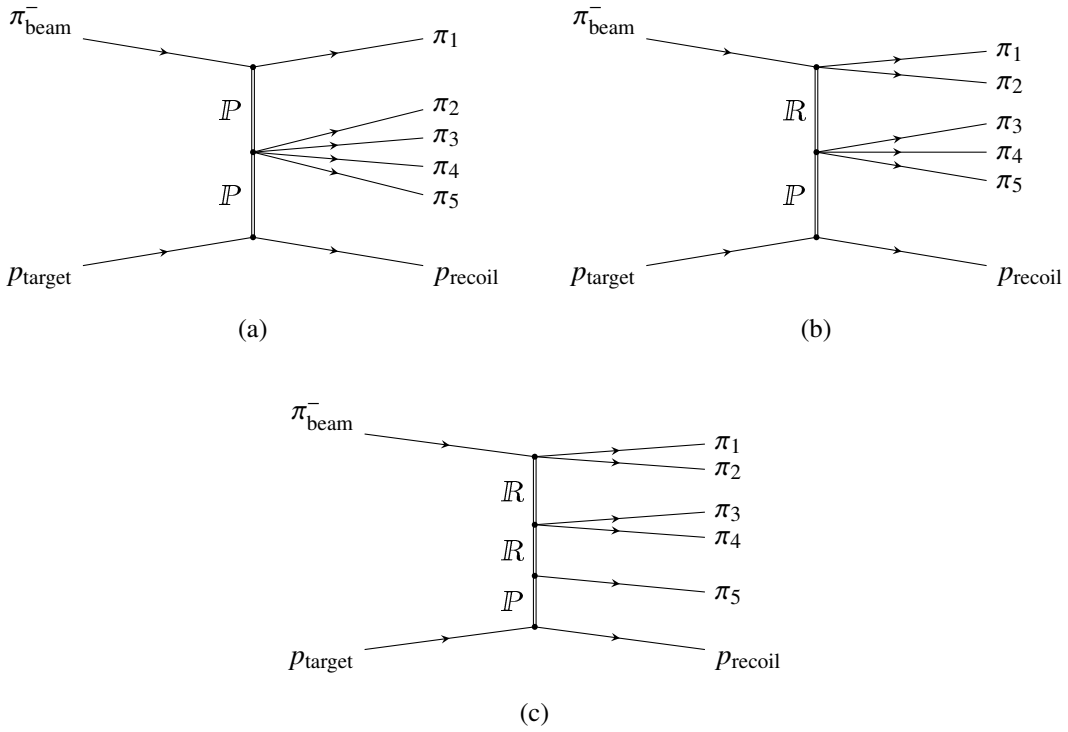


Figure 5.20: Exclusive five-pion production by multiperipheral processes.

Another reason for the lack of phase motions are background processes which produce an exclusive five-pion system without the formation of an intermediate five-pion state X. A prominent candidate for such a production mechanism is multiperipheral production, three examples of which are shown in figure 5.20. As one can see, there is no five-pion resonance in these diagrams, which is why these kinds of backgrounds are also called non-resonant backgrounds. Even though there is no five-pion resonant present, resonances will still appear in the four-, three- and two-pion subsystems and the angular distribution will not be flat. If multiperipheral production is present in the measured data set but not in the model, it will be decomposed into partial waves in the MLE so that the data is described and therefore distort the intensities and phases. The fraction of multiperipheral events present in the data set presented here is unknown. However, there are several indications that it may be substantial. The wave sets required to describe the data are very big, and most of the waves do not show stable behavior in the model selection. The reason for this could be that a combination of many waves is required to describe the multiperipheral processes. In this situation, the optimal description can change drastically in between $m_{5\pi}$ bins, leading to the unstable waves seen in the model selection[c]. Even though the wave sets are large, the flat wave is still very big and, as was shown in section 5.2, the comparison of model and and measured data show discrepancies in several places. Taken together, these findings suggest that there is something missing from the model, and multiperipheral processes are an obvious candidate.

To interpret the results of the partial-wave decomposition in terms of resonances, the mass dependence of the spin-density matrix needs to be parameterized by a model, as described in section 3.4. However,

[c] An alternative explanation for the large wave sets is that the form of the prior given by equation (4.24) is not restrictive enough for the problem at hand. The unstable waves might be caused by ambiguities or the uncertainties in the isobar mass shape parameterizations.

there are two prerequisites for this procedure to be meaningful: (1) the functions with which the intensities and phases should be described have to be known and (2) the $m_{5\pi}$ dependence of $\mathcal{P}_{ii}$ needs to be known for all waves which should be parameterized. Concerning (1), as discussed in section 3.4, Breit-Wigner amplitudes are usually employed together with a function accounting for background events. However, the lack of Breit-Wigner-like signatures in the phases presented here makes it difficult to motivate such an approach. In section 5.3, a number of issues with (2) have been discussed. In principle, it would be possible to learn about the $m_{5\pi}$ dependence of $\mathcal{P}_{ii}$ from the data during the resonance extraction, as was attempted in [10]. However, with the lack of clear resonance signals, it would be very difficult to interpret the results of such a study. Conversely, if there was exact knowledge about (2), it might be possible to better understand the phase motions and intensity signatures observed in the data. With the unclear situation in both areas, resonance extraction does not yield reliable and trustworthy results.

The comparison with the $\eta(3\pi)$ and 3π channels, however, is very encouraging. The intensities and phases for most of the waves in common with the $\eta(3\pi)$ channel are in very good agreement. This means that the lack of phase motions is not an artifact of the model selection or any systematic effects in the measurement of 5π, but is a feature of the measured production process. With both partial-wave decompositions using the same model assumptions, multiperipheral processes present in the data seem to affect both channels similarly. The comparison with the 3π channel shows good agreement in the intensities, but the phase motions seem to be damped in 5π. This implies that the effect causing the lack of phase motions affects the 5π and $\eta(3\pi)$ channels, but is not present or significantly reduced in 3π. This might provide some clues as to its origin.

Despite the normalization of the phase-space correction factors discussed at the end of section 5.3, it is worth checking whether the intensity scaling factors are plausible for the $\eta(3\pi)$ channel. All the waves with the same decay should end up with the same scaling factor, which is realized to a reasonable degree in the waves with the $f_1(1285)\,\pi^-$ decay. Averaging over these scaling factors gives $S[X \to f_1(1285)\,\pi^-] = 7 \pm 0.6$. Because the acceptance correction in $\eta(3\pi)$ only takes into account geometrical acceptance, an overall correction factor r is needed. Using the branching-fraction ratio for the decay of the $f_1(1285)$ [4] and the average scaling factor S for that wave as calculated above, one obtains

$$r = \frac{\mathcal{B}\left[f_1(1285) \to \eta\pi^+\pi^-\right]}{\mathcal{B}\left[f_1(1285) \to 2\pi^+\,2\pi^-\right]} \cdot S\left[X \to f_1(1285)\,\pi^-\right] = \frac{0.35 \pm 0.15}{0.11} \cdot 7 \pm 0.6 = 22 \pm 10 \tag{5.13}$$

Only the uncertainties on $\mathcal{B}[f_1(1285) \to \eta\pi^+\pi^-]$ and $S[X \to f_1(1285)\,\pi^-)]$ have been included in the error propagation, as they are far bigger than the uncertainty on $\mathcal{B}[f_1(1285) \to 2\pi^+\,2\pi^-]$. Now, we can use this to predict the scale factor for other waves, in particular

$$\begin{aligned} S\left[2^{-+}0^+ f_2(1565)\,\pi^-\right] &= r \cdot \frac{\mathcal{B}\left[a_2(1320) \to \rho(770)\,\pi^\pm\right]}{\mathcal{B}\left[a_2(1320) \to \eta\,\pi^\pm\right]} \\ &= 22 \pm 10 \cdot \frac{0.35}{0.145} \\ &= 53 \pm 24 \end{aligned} \tag{5.14}$$

Where we have taken into account the isospin Clebsch-Gordan coefficient for the decay of $a_2(1320) \to \rho(770)^0\pi^\pm$. Compared to the value from figure 5.15b, which is 45.2, this matches reasonably well and indicates that the results are consistent. As only the branching ratio of the $a_2(1320)$ enters in this

calculation, the same scaling factor should appear also in the waves decaying to $\rho(770)$ $a_2(1320)$. The unusually large scaling factor shown in figure 5.15d was already pointed out and discussed in the previous section. For the two waves in figure 5.14, the scaling factor lies somewhat lower at $S \approx 20$, which is still consistent with the result above. It is surprising that this calculation agrees this well with the data, as an effect of the normalization of the phase-space correction factors is expected.

CHAPTER 6

Conclusions and Outlook

> "No lang, a sälbem Sunntig, het är gwartet a der Schtell..."
>
> Dr. Hans-Peter "Mani" Matter

The reaction $\pi^- + p \to \pi^-\pi^+\pi^-\pi^+\pi^- + p$ is very challenging to analyze in the context of a partial-wave decomposition, which makes it an ideal channel to enhance and refine the analysis techniques. The high number of final-state particles gives rise to a large number of possible contributing partial waves without much hierarchy, making an automated model-selection procedure a necessity.

An exclusive sample of five-pion events has been selected from the data set measured by the COMPASS experiment during the 2008 data-taking campaign. An automated model-selection procedure, developed by [10] in the form of a genetic algorithm, was subjected to a series of tests before finding it unsuitable. Therefore, a novel method, the biggest-conceivable-model (BCM) method, was developed and studied extensively on Monte Carlo and measured data. During these studies, it was found that this method not only allows to find a partial-wave model, but is also capable of identifying correlation patterns of the different waves in the model. This was used to judge the stability of the extracted waves. While the BCM method was studied in depth, there is room for refinement in the choice of the prior function and its parameters. Also the correlations between the waves are an interesting avenue for further research. Finding a possibility to resolve the correlations and thereby increase the number of stable waves has the potential of greatly improving the understanding of the five-pion data set.

With only a minimal change required in the likelihood function, the BCM method can easily be implemented in partial-wave decompositions of other final states. Indeed, it was already applied successfully to the $\pi^- + p \to \pi^-\eta\eta + p$ channel [9] and, in a pilot study, to the $\pi^- + p \to \pi^-\pi^+\pi^- + p$ channel [50]. Together with the techniques to investigate and visualize the stability of waves, the BCM method has the potential to introduce an easy-to-use automated model-selection procedure to partial-wave decompositions in general, which can reduce bias, improve the reliability of the results and the understanding of the structure of the data and the possible models.

The BCM model-selection method was applied to the measured five-pion data, where it provided a wave set and identified 18 stable waves. With a model selected, a partial-wave decomposition

was performed and the stable waves were studied in more detail. Unfortunately, due to the way the phase-space and damping factors are included in the formalism, it was not possible to interpret the intensity distributions resulting from the partial-wave decomposition. The phases, which are not affected by this problem, do not show any clear resonance signatures. Nevertheless, where possible, the results of the partial-wave decomposition were compared to similar analyses of the $\pi^- + p \to \pi^-\pi^+\pi^- + p$ and $\pi^- + p \to \eta\pi^-\pi^+\pi^- + p$ reactions and the stable five-pion waves were found to be in good agreement. Due to the lack of resonance signatures and the unclear situation concerning the parameterization of the damping factors ξ, an extraction of resonance parameters was not attempted.

The two problems preventing an extraction of resonance parameters at this point also pose an compelling opportunity. The exact nature of the modulation of the phase space with the final-state mass is not well understood, especially the way the phase space is damped towards higher masses. The five-pion channel is an interesting place to study this, because the modulation is very strong there. Similarly, there are indications that multiperipheral background plays an important role in the measured data set. There is progress in understanding such backgrounds and implementing them into partial-wave decompositions for channels with three-particle final states [30]. If such an approach can be generalized and applied to the five-pion channel, it would be possible to test and refine it in a largely independent environment.

APPENDIX A

Calculation of det($\mathcal{N}$)

Any complex number $\hat{z} \in \mathbb{C}$ can be written in matrix form

$$z = \begin{pmatrix} \mathrm{Re}(\hat{z}) & -\mathrm{Im}(\hat{z}) \\ \mathrm{Im}(\hat{z}) & \mathrm{Re}(\hat{z}) \end{pmatrix}, \quad \text{with} \quad z \in \mathbb{R}^{2\times 2} \tag{A.1}$$

It is easy to see that this representation is isomorph to the complex numbers. Now, consider a positive-definite, Hermitian, complex matrix $\hat{M} \in \mathbb{C}^{n\times n}$. We write the real matrix

$$\mathbb{R}^{2n\times 2n} \ni N = \begin{pmatrix} \mathrm{Re}(\hat{M}) & -\mathrm{Im}(\hat{M}) \\ \mathrm{Im}(\hat{M}) & \mathrm{Re}(\hat{M}) \end{pmatrix} \tag{A.2}$$

with the $n \times n$ block matrices $\mathrm{Re}(\hat{M})$ and $\mathrm{Im}(\hat{M})$ obtained by taking the real and imaginary part of every element of $\hat{M}$, respectively. We are now interested in the relation between $\det(\hat{M})$ and $\det(N)$. First, we transition the elements of $\hat{M}$ from complex numbers to matrices using equation (A.1) and get a matrix of matrices $M \in \left(\mathbb{R}^{2\times 2}\right)^{n\times n}$, for which all the usual matrix operations work as expected. Obviously, $\det(M)$ and $\det(\hat{M})$ are different notations for the same value, with $\det(M) \in \mathbb{C}$ and $\det(\hat{M}) \in \mathbb{R}^{2\times 2}$. Then, let us note that, without changing $\det(N)$, we can reorder the rows and columns of N so that it has the same structure as M, only without having the entries grouped in submatrices:

$$\begin{aligned} M &= \begin{pmatrix} \begin{pmatrix} \mathrm{Re}\,(M_{11}) & -\mathrm{Im}\,(M_{11}) \\ \mathrm{Im}\,(M_{11}) & \mathrm{Re}\,(M_{11}) \end{pmatrix} & \cdots & \begin{pmatrix} \mathrm{Re}\,(M_{1n}) & -\mathrm{Im}\,(M_{1n}) \\ \mathrm{Im}\,(M_{1n}) & \mathrm{Re}\,(M_{1n}) \end{pmatrix} \\ \vdots & & \vdots \\ \begin{pmatrix} \mathrm{Re}\,(M_{n1}) & -\mathrm{Im}\,(M_{n1}) \\ \mathrm{Im}\,(M_{n1}) & \mathrm{Re}\,(M_{n1}) \end{pmatrix} & \cdots & \begin{pmatrix} \mathrm{Re}\,(M_{nn}) & -\mathrm{Im}\,(M_{nn}) \\ \mathrm{Im}\,(M_{nn}) & \mathrm{Re}\,(M_{nn}) \end{pmatrix} \end{pmatrix} \\ N &= \begin{pmatrix} \mathrm{Re}\,(M_{11}) & -\mathrm{Im}\,(M_{11}) & \cdots & \mathrm{Re}\,(M_{1n}) & -\mathrm{Im}\,(M_{1n}) \\ \mathrm{Im}\,(M_{11}) & \mathrm{Re}\,(M_{11}) & & \mathrm{Im}\,(M_{1n}) & \mathrm{Re}\,(M_{1n}) \\ \vdots & & & \vdots & \\ \mathrm{Re}\,(M_{n1}) & -\mathrm{Im}\,(M_{n1}) & \cdots & \mathrm{Re}\,(M_{nn}) & -\mathrm{Im}\,(M_{nn}) \\ \mathrm{Im}\,(M_{n1}) & \mathrm{Re}\,(M_{n1}) & & \mathrm{Im}\,(M_{nn}) & \mathrm{Re}\,(M_{nn}) \end{pmatrix} \end{aligned} \tag{A.3}$$

We diagonalize M, which is possible because $\hat{M}$ is positive definite and Hermitian. The eigenvalues λ_i of a Hermitian matrix are real, which translates to a diagonal matrix when written in the notation defined in equation (A.1). Therefore the result of the diagonalization can be written as

$$D_M = \operatorname{diag}\left(\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_1\end{pmatrix}, \cdots, \begin{pmatrix}\lambda_n & 0\\ 0 & \lambda_n\end{pmatrix}\right) \in \left(\mathbb{R}^{2\times 2}\right)^{n\times n} \tag{A.4}$$

and we get for

$$\det(M) = \det(D_M) = \begin{pmatrix}\prod_{i=1}^{n}\lambda_i & 0\\ 0 & \prod_{i=1}^{n}\lambda_i\end{pmatrix} \Leftrightarrow \det(\hat{M}) = \prod_{i=1}^{n}\lambda_i \tag{A.5}$$

which is not surprising.

To calculate the determinant of N, we first note that block matrices can be multiplied blockwise, i.e. the blocks are treated as entries in the multiplication and then resulting products of blocks are expanded. This means, we can use the same diagonalization we used for M, but treating the submatrices as blocks, which yields,

$$D_N = \operatorname{diag}(\lambda_1, \lambda_1, \lambda_2, \lambda_2, \ldots, \lambda_n, \lambda_n) \in \mathbb{R}^{2n\times 2n} \tag{A.6}$$

and therefore

$$\det(N) = \det(D_N) = \prod_{i=1}^{n}\lambda_i^2 = \det(\hat{M})^2 \tag{A.7}$$

APPENDIX B

Test-Model Parameters

Table B.1 gives the model parameters used to generate events for model selection testing, discussed in sections 4.2.3 and 4.3.1. The production amplitudes are shown in the complex plane in figure B.1. As event samples of different size have been generated with this model, its parameters are scaled such that the total intensity equals 1000 events. In the table, the "Percent" column gives the fraction of the intensity of a single wave to the (incoherent) sum of intensities, which, in this case, is 1462.576.

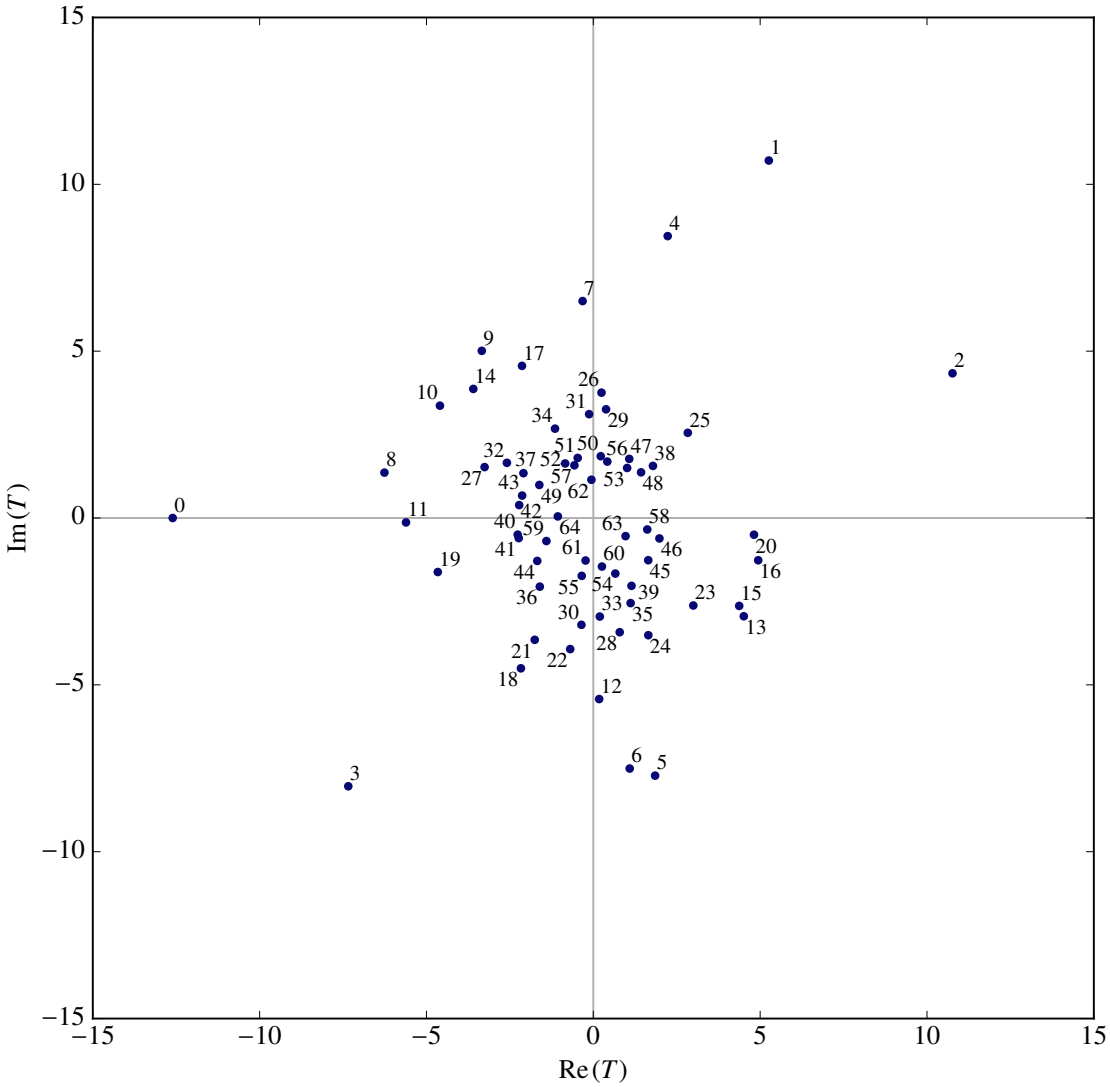


Figure B.1: Production amplitudes used for model-selection testing. The numbers give the wave index as defined in table B.1.

Table B.1: Wave set and production amplitudes used for model-selection testing.

Wave Index	Wave	Re(T)	Im(T)	Intensity	Percent
0	$1^-1^{++}0^+ \to \sigma \begin{bmatrix} 0 \\ 1 \end{bmatrix} \left(a_1(1260) \to \rho(770) \begin{bmatrix} 0 \\ 1 \end{bmatrix} \pi^- \right)$	−12.606	0.0	158.906	10.86
1	$1^-1^{++}0^+ \to \left(\rho(1700) \to \left(a_1(1260) \to \rho(770) \begin{bmatrix} 0 \\ 1 \end{bmatrix} \pi^- \right) \begin{bmatrix} 0 \\ 1 \end{bmatrix} \pi^+ \right) \begin{bmatrix} 0 \\ 1 \end{bmatrix} \pi^-$	5.26	10.713	142.435	9.74
2	$1^-0^{-+}0^+ \to \sigma \begin{bmatrix} 0 \\ 0 \end{bmatrix} \left(\pi(1800) \to \sigma \begin{bmatrix} 0 \\ 0 \end{bmatrix} \pi^- \right)$	10.765	4.336	134.696	9.21
3	$1^-0^{-+}0^+ \to f_0(980) \begin{bmatrix} 0 \\ 0 \end{bmatrix} \left(\pi(1800) \to \sigma \begin{bmatrix} 0 \\ 0 \end{bmatrix} \pi^- \right)$	−7.343	−8.039	118.545	8.11
4	$1^-1^{++}0^+ \to \left(\rho(1700) \to \rho(770) \begin{bmatrix} 0 \\ 1 \end{bmatrix} \sigma \right) \begin{bmatrix} 0 \\ 1 \end{bmatrix} \pi^-$	2.229	8.448	76.342	5.22
5	$1^-1^{++}0^+ \to \sigma \begin{bmatrix} 2 \\ 2 \end{bmatrix} \left(a_2(1320) \to \rho(770) \begin{bmatrix} 2 \\ 1 \end{bmatrix} \pi^- \right)$	1.852	−7.72	63.034	4.31
6	$1^-1^{++}0^+ \to \left(f_0(1500) \to \left(\pi(1300) \to \sigma \begin{bmatrix} 0 \\ 0 \end{bmatrix} \pi^- \right) \begin{bmatrix} 0 \\ 0 \end{bmatrix} \pi^+ \right) \begin{bmatrix} 1 \\ 0 \end{bmatrix} \pi^-$	1.091	−7.508	57.559	3.94
7	$1^-1^{++}0^+ \to \left(f_0(1500) \to \sigma \begin{bmatrix} 0 \\ 0 \end{bmatrix} \sigma \right) \begin{bmatrix} 1 \\ 0 \end{bmatrix} \pi^-$	−0.321	6.501	42.369	2.9
8	$1^-0^{-+}0^+ \to \sigma \begin{bmatrix} 0 \\ 0 \end{bmatrix} \left(\pi(1300) \to \sigma \begin{bmatrix} 0 \\ 0 \end{bmatrix} \pi^- \right)$	−6.26	1.361	41.038	2.81
9	$1^-1^{++}0^+ \to \left(f_1(1285) \to \left(a_1(1260) \to \rho(770) \begin{bmatrix} 0 \\ 1 \end{bmatrix} \pi^- \right) \begin{bmatrix} 1 \\ 1 \end{bmatrix} \pi^+ \right) \begin{bmatrix} 1 \\ 1 \end{bmatrix} \pi^-$	−3.342	5.01	36.273	2.48
10	$1^-1^{++}0^+ \to f_0(980) \begin{bmatrix} 2 \\ 2 \end{bmatrix} \left(a_2(1320) \to \rho(770) \begin{bmatrix} 2 \\ 1 \end{bmatrix} \pi^- \right)$	−4.596	3.366	32.458	2.22
11	$1^-0^{-+}0^+ \to f_0(980) \begin{bmatrix} 1 \\ 1 \end{bmatrix} \left(a_1(1260) \to \sigma \begin{bmatrix} 1 \\ 0 \end{bmatrix} \pi^- \right)$	−5.613	−0.132	31.52	2.16

Continued on next page

Table B.1: Wave set and production amplitudes used for model-selection testing (continued).

Wave Index	Wave	Re(T)	Im(T)	Intensity	Percent
12	$1^-2^{-+}0^+ \to \left(f_2(1565) \to \left(a_1(1260) \to \rho(770)\begin{bmatrix}0\\1\end{bmatrix}\pi^-\right)\begin{bmatrix}1\\1\end{bmatrix}\pi^+\right)\begin{bmatrix}0\\2\end{bmatrix}\pi^-$	0.175	−5.425	29.461	2.01
13	$1^-1^{++}0^+ \to f_0(980)\begin{bmatrix}0\\1\end{bmatrix}\left(a_1(1260) \to \rho(770)\begin{bmatrix}0\\1\end{bmatrix}\pi^-\right)$	4.51	−2.945	29.014	1.98
14	$1^-2^{-+}0^+ \to f_2(1270)\begin{bmatrix}0\\2\end{bmatrix}\left(\pi(1800) \to \sigma\begin{bmatrix}0\\0\end{bmatrix}\pi^-\right)$	−3.597	3.868	27.902	1.91
15	$1^-0^{-+}0^+ \to \sigma\begin{bmatrix}1\\1\end{bmatrix}\left(a_1(1260) \to \sigma\begin{bmatrix}1\\0\end{bmatrix}\pi^-\right)$	4.372	−2.634	26.055	1.78
16	$1^-1^{++}0^+ \to \rho(770)\begin{bmatrix}0\\1\end{bmatrix}\left(\pi_2(1670) \to \rho(770)\begin{bmatrix}1\\1\end{bmatrix}\pi^-\right)$	4.942	−1.267	26.029	1.78
17	$1^-1^{++}0^+ \to \left(\rho(1450) \to \left(a_1(1260) \to \rho(770)\begin{bmatrix}0\\1\end{bmatrix}\pi^-\right)\begin{bmatrix}0\\1\end{bmatrix}\pi^+\right)\begin{bmatrix}0\\1\end{bmatrix}\pi^-$	−2.137	4.559	25.357	1.73
18	$1^-2^{-+}0^+ \to f_2(1270)\begin{bmatrix}0\\2\end{bmatrix}\left(\pi(1300) \to \sigma\begin{bmatrix}0\\0\end{bmatrix}\pi^-\right)$	−2.17	−4.504	24.997	1.71
19	$1^-1^{++}0^+ \to \left(f_0(1500) \to \rho(770)\begin{bmatrix}0\\0\end{bmatrix}\rho(770)\right)\begin{bmatrix}1\\0\end{bmatrix}\pi^-$	−4.661	−1.619	24.342	1.66
20	$1^-2^{-+}0^+ \to \rho(770)\begin{bmatrix}0\\2\end{bmatrix}\left(a_2(1320) \to \rho(770)\begin{bmatrix}2\\1\end{bmatrix}\pi^-\right)$	4.815	−0.501	23.437	1.6
21	$1^-1^{++}0^+ \to \left(\rho(1700) \to \left(a_1(1260) \to \sigma\begin{bmatrix}1\\0\end{bmatrix}\pi^-\right)\begin{bmatrix}0\\1\end{bmatrix}\pi^+\right)\begin{bmatrix}0\\1\end{bmatrix}\pi^-$	−1.756	−3.65	16.407	1.12
22	$1^-1^{++}0^+ \to \rho(770)\begin{bmatrix}2\\1\end{bmatrix}\left(\pi(1300) \to \sigma\begin{bmatrix}0\\0\end{bmatrix}\pi^-\right)$	−0.693	−3.928	15.908	1.09
23	$1^-0^{-+}0^+ \to \left(f_0(1500) \to \left(\pi(1300) \to \sigma\begin{bmatrix}0\\0\end{bmatrix}\pi^-\right)\begin{bmatrix}0\\0\end{bmatrix}\pi^+\right)\begin{bmatrix}0\\0\end{bmatrix}\pi^-$	2.996	−2.622	15.855	1.08

Continued on next page

Table B.1: Wave set and production amplitudes used for model-selection testing (continued).

Wave Index	Wave	Re(T)	Im(T)	Intensity	Percent
24	$1^-1^{++}0^+ \to \sigma \begin{bmatrix}1\\0\end{bmatrix} \left(\pi(1300) \to \rho(770) \begin{bmatrix}1\\1\end{bmatrix} \pi^-\right)$	1.648	−3.512	15.055	1.03
25	$1^-2^{-+}0^+ \to f_0(980) \begin{bmatrix}0\\2\end{bmatrix} \left(\pi_2(1670) \to \rho(770) \begin{bmatrix}1\\1\end{bmatrix} \pi^-\right)$	2.832	2.554	14.542	0.99
26	$1^-2^{-+}0^+ \to \left(f_2(1565) \to \rho(770) \begin{bmatrix}0\\2\end{bmatrix} \rho(770)\right) \begin{bmatrix}0\\2\end{bmatrix} \pi^-$	0.246	3.757	14.173	0.97
27	$1^-1^{++}0^+ \to \left(\rho(1700) \to \left(\pi(1300) \to \sigma \begin{bmatrix}0\\0\end{bmatrix} \pi^-\right) \begin{bmatrix}1\\0\end{bmatrix} \pi^+\right) \begin{bmatrix}0\\1\end{bmatrix} \pi^-$	−3.258	1.526	12.945	0.89
28	$1^-1^{++}0^+ \to \left(\rho(1700) \to \rho(770) \begin{bmatrix}0\\1\end{bmatrix} \sigma\right) \begin{bmatrix}2\\1\end{bmatrix} \pi^-$	0.791	−3.422	12.339	0.84
29	$1^-2^{-+}0^+ \to \left(f_2(1565) \to \left(a_2(1320) \to \rho(770) \begin{bmatrix}2\\1\end{bmatrix} \pi^-\right) \begin{bmatrix}1\\2\end{bmatrix} \pi^+\right) \begin{bmatrix}0\\2\end{bmatrix} \pi^-$	0.382	3.258	10.759	0.74
30	$1^-0^{-+}0^+ \to \left(f_0(1500) \to \left(a_1(1260) \to \sigma \begin{bmatrix}1\\0\end{bmatrix} \pi^-\right) \begin{bmatrix}1\\1\end{bmatrix} \pi^+\right) \begin{bmatrix}0\\0\end{bmatrix} \pi^-$	−0.356	−3.201	10.374	0.71
31	$1^-1^{++}0^+ \to \left(f_1(1420) \to \left(a_1(1260) \to \rho(770) \begin{bmatrix}0\\1\end{bmatrix} \pi^-\right) \begin{bmatrix}1\\1\end{bmatrix} \pi^+\right) \begin{bmatrix}1\\1\end{bmatrix} \pi^-$	−0.126	3.113	9.704	0.66
32	$1^-2^{-+}0^+ \to \rho(770) \begin{bmatrix}2\\0\end{bmatrix} \left(a_1(1260) \to \rho(770) \begin{bmatrix}0\\1\end{bmatrix} \pi^-\right)$	−2.59	1.654	9.442	0.65
33	$1^-1^{++}0^+ \to \rho(770) \begin{bmatrix}1\\1\end{bmatrix} \left(a_1(1260) \to \rho(770) \begin{bmatrix}0\\1\end{bmatrix} \pi^-\right)$	0.195	−2.953	8.761	0.6
34	$1^-2^{-+}0^+ \to \rho(770) \begin{bmatrix}0\\2\end{bmatrix} \left(a_1(1260) \to \rho(770) \begin{bmatrix}0\\1\end{bmatrix} \pi^-\right)$	−1.148	2.678	8.487	0.58
35	$1^-0^{-+}0^+ \to \rho(770) \begin{bmatrix}2\\2\end{bmatrix} \left(a_1(1260) \to \rho(770) \begin{bmatrix}0\\1\end{bmatrix} \pi^-\right)$	1.118	−2.55	7.751	0.53

Continued on next page

Table B.1: Wave set and production amplitudes used for model-selection testing (continued).

Wave Index	Wave	Re(T)	Im(T)	Intensity	Percent
36	$1^-0^{-+}0^+ \to \left(\rho(1700) \to \left(a_1(1260) \to \rho(770)\begin{bmatrix}0\\1\end{bmatrix}\pi^-\right)\begin{bmatrix}0\\1\end{bmatrix}\pi^+\right)\begin{bmatrix}1\\1\end{bmatrix}\pi^-$	−1.604	−2.057	6.803	0.47
37	$1^-1^{++}0^+ \to \sigma\begin{bmatrix}1\\0\end{bmatrix}\left(\pi(1800) \to \sigma\begin{bmatrix}0\\0\end{bmatrix}\pi^-\right)$	−2.094	1.343	6.19	0.42
38	$1^-1^{++}0^+ \to f_2(1270)\begin{bmatrix}2\\1\end{bmatrix}\left(a_1(1260) \to \sigma\begin{bmatrix}1\\0\end{bmatrix}\pi^-\right)$	1.787	1.561	5.628	0.38
39	$1^-1^{++}0^+ \to \left(f_0(1370) \to \left(\pi(1300) \to \rho(770)\begin{bmatrix}1\\1\end{bmatrix}\pi^-\right)\begin{bmatrix}0\\0\end{bmatrix}\pi^+\right)\begin{bmatrix}1\\0\end{bmatrix}\pi^-$	1.144	−2.033	5.441	0.37
40	$1^-2^{-+}0^+ \to \sigma\begin{bmatrix}1\\1\end{bmatrix}\left(a_1(1260) \to \rho(770)\begin{bmatrix}0\\1\end{bmatrix}\pi^-\right)$	−2.266	−0.498	5.382	0.37
41	$1^-0^{-+}0^+ \to \rho(770)\begin{bmatrix}0\\0\end{bmatrix}\left(a_1(1260) \to \sigma\begin{bmatrix}1\\0\end{bmatrix}\pi^-\right)$	−2.235	−0.603	5.357	0.37
42	$1^-2^{-+}0^+ \to \left(f_0(1370) \to \left(\pi(1300) \to \sigma\begin{bmatrix}0\\0\end{bmatrix}\pi^-\right)\begin{bmatrix}0\\0\end{bmatrix}\pi^+\right)\begin{bmatrix}2\\0\end{bmatrix}\pi^-$	−2.222	0.387	5.086	0.35
43	$1^-1^{++}0^+ \to \left(f_0(1500) \to \left(\pi(1300) \to \rho(770)\begin{bmatrix}1\\1\end{bmatrix}\pi^-\right)\begin{bmatrix}0\\0\end{bmatrix}\pi^+\right)\begin{bmatrix}1\\0\end{bmatrix}\pi^-$	−2.134	0.672	5.005	0.34
44	$1^-2^{-+}0^+ \to \left(f_1(1285) \to \left(a_1(1260) \to \rho(770)\begin{bmatrix}0\\1\end{bmatrix}\pi^-\right)\begin{bmatrix}1\\1\end{bmatrix}\pi^+\right)\begin{bmatrix}2\\1\end{bmatrix}\pi^-$	−1.68	−1.286	4.477	0.31
45	$1^-2^{-+}0^+ \to \left(f_2(1270) \to \rho(770)\begin{bmatrix}0\\2\end{bmatrix}\rho(770)\right)\begin{bmatrix}2\\2\end{bmatrix}\pi^-$	1.646	−1.265	4.311	0.29
46	$1^-1^{++}0^+ \to \left(f_2(1270) \to \left(a_1(1260) \to \rho(770)\begin{bmatrix}0\\1\end{bmatrix}\pi^-\right)\begin{bmatrix}1\\1\end{bmatrix}\pi^+\right)\begin{bmatrix}1\\2\end{bmatrix}\pi^-$	1.981	−0.612	4.301	0.29
47	$1^-0^{-+}0^+ \to \left(f_0(1370) \to \left(a_1(1260) \to \rho(770)\begin{bmatrix}0\\1\end{bmatrix}\pi^-\right)\begin{bmatrix}1\\1\end{bmatrix}\pi^+\right)\begin{bmatrix}0\\0\end{bmatrix}\pi^-$	1.077	1.772	4.3	0.29

Continued on next page

Table B.1: Wave set and production amplitudes used for model-selection testing (continued).

Wave Index	Wave	Re(T)	Im(T)	Intensity	Percent
48	$1^-0^{-+}0^+ \to f_2(1270)\begin{bmatrix}0\\0\end{bmatrix}\left(\pi_2(1670) \to \rho(770)\begin{bmatrix}1\\1\end{bmatrix}\pi^-\right)$	1.431	1.369	3.921	0.27
49	$1^-1^{++}0^+ \to \left(f_0(1370) \to \left(\pi(1300) \to \sigma\begin{bmatrix}0\\0\end{bmatrix}\pi^-\right)\begin{bmatrix}0\\0\end{bmatrix}\pi^+\right)\begin{bmatrix}1\\0\end{bmatrix}\pi^-$	−1.618	0.991	3.598	0.25
50	$1^-2^{-+}0^+ \to \left(\eta_2(1645) \to \left(a_2(1320) \to \rho(770)\begin{bmatrix}2\\1\end{bmatrix}\pi^-\right)\begin{bmatrix}0\\2\end{bmatrix}\pi^+\right)\begin{bmatrix}1\\2\end{bmatrix}\pi^-$	0.223	1.852	3.48	0.24
51	$1^-0^{-+}0^+ \to \left(f_0(1500) \to \left(\pi(1300) \to \rho(770)\begin{bmatrix}1\\1\end{bmatrix}\pi^-\right)\begin{bmatrix}0\\0\end{bmatrix}\pi^+\right)\begin{bmatrix}0\\0\end{bmatrix}\pi^-$	−0.468	1.8	3.459	0.24
52	$1^-1^{++}0^+ \to \rho(770)\begin{bmatrix}2\\1\end{bmatrix}\left(\pi(1300) \to \rho(770)\begin{bmatrix}1\\1\end{bmatrix}\pi^-\right)$	−0.844	1.634	3.382	0.23
53	$1^-4^{-+}0^+ \to \rho(770)\begin{bmatrix}2\\2\end{bmatrix}\left(a_1(1260) \to \rho(770)\begin{bmatrix}0\\1\end{bmatrix}\pi^-\right)$	1.013	1.5	3.276	0.22
54	$1^-1^{++}0^+ \to \left(\rho(1450) \to \left(\pi(1300) \to \sigma\begin{bmatrix}0\\0\end{bmatrix}\pi^-\right)\begin{bmatrix}1\\0\end{bmatrix}\pi^+\right)\begin{bmatrix}2\\1\end{bmatrix}\pi^-$	0.66	−1.665	3.209	0.22
55	$1^-2^{-+}0^+ \to \left(f_0(1500) \to \left(a_1(1260) \to \sigma\begin{bmatrix}1\\0\end{bmatrix}\pi^-\right)\begin{bmatrix}1\\1\end{bmatrix}\pi^+\right)\begin{bmatrix}2\\0\end{bmatrix}\pi^-$	−0.347	−1.733	3.125	0.21
56	$1^-1^{++}0^+ \to \left(f_2(1565) \to \left(\pi(1300) \to \rho(770)\begin{bmatrix}1\\1\end{bmatrix}\pi^-\right)\begin{bmatrix}2\\0\end{bmatrix}\pi^+\right)\begin{bmatrix}1\\2\end{bmatrix}\pi^-$	0.419	1.691	3.035	0.21
57	$1^-3^{++}0^+ \to \left(\rho(1450) \to \rho(770)\begin{bmatrix}0\\1\end{bmatrix}\sigma\right)\begin{bmatrix}2\\1\end{bmatrix}\pi^-$	−0.567	1.579	2.814	0.19
58	$1^-3^{++}0^+ \to \sigma\begin{bmatrix}2\\1\end{bmatrix}\left(a_1(1260) \to \rho(770)\begin{bmatrix}0\\1\end{bmatrix}\pi^-\right)$	1.617	−0.343	2.734	0.19
59	$1^-3^{++}0^+ \to \rho(770)\begin{bmatrix}2\\3\end{bmatrix}\left(\pi_2(1670) \to f_2(1270)\begin{bmatrix}0\\2\end{bmatrix}\pi^-\right)$	−1.406	−0.69	2.452	0.17

Continued on next page

Table B.1: Wave set and production amplitudes used for model-selection testing (continued).

Wave Index	Wave	Re(T)	Im(T)	Intensity	Percent
60	$1^-2^{-+}0^+ \to \left(f_2(1565) \to \left(\pi(1300) \to \rho(770)\begin{bmatrix}1\\1\end{bmatrix}\pi^-\right)\begin{bmatrix}2\\0\end{bmatrix}\pi^+\right)\begin{bmatrix}0\\2\end{bmatrix}\pi^-$	0.258	−1.455	2.184	0.15
61	$1^-2^{-+}0^+ \to \left(\rho_3(1690) \to \left(a_2(1320) \to \rho(770)\begin{bmatrix}2\\1\end{bmatrix}\pi^-\right)\begin{bmatrix}2\\2\end{bmatrix}\pi^+\right)\begin{bmatrix}1\\3\end{bmatrix}\pi^-$	−0.232	−1.274	1.677	0.11
62	$1^-2^{-+}1^+ \to \left(f_1(1285) \to \left(a_1(1260) \to \rho(770)\begin{bmatrix}0\\1\end{bmatrix}\pi^-\right)\begin{bmatrix}1\\1\end{bmatrix}\pi^+\right)\begin{bmatrix}2\\1\end{bmatrix}\pi^-$	−0.057	1.147	1.318	0.09
63	$1^-3^{++}1^+ \to \left(\rho(1700) \to \left(a_1(1260) \to \rho(770)\begin{bmatrix}0\\1\end{bmatrix}\pi^-\right)\begin{bmatrix}0\\1\end{bmatrix}\pi^+\right)\begin{bmatrix}2\\1\end{bmatrix}\pi^-$	0.967	−0.543	1.229	0.08
64	$1^-1^{++}1^+ \to \left(\rho(1450) \to \left(\pi(1300) \to \sigma\begin{bmatrix}0\\0\end{bmatrix}\pi^-\right)\begin{bmatrix}1\\0\end{bmatrix}\pi^+\right)\begin{bmatrix}0\\1\end{bmatrix}\pi^-$	−1.063	0.047	1.132	0.08

APPENDIX C

Additional Plots from the Genetic Algorithm

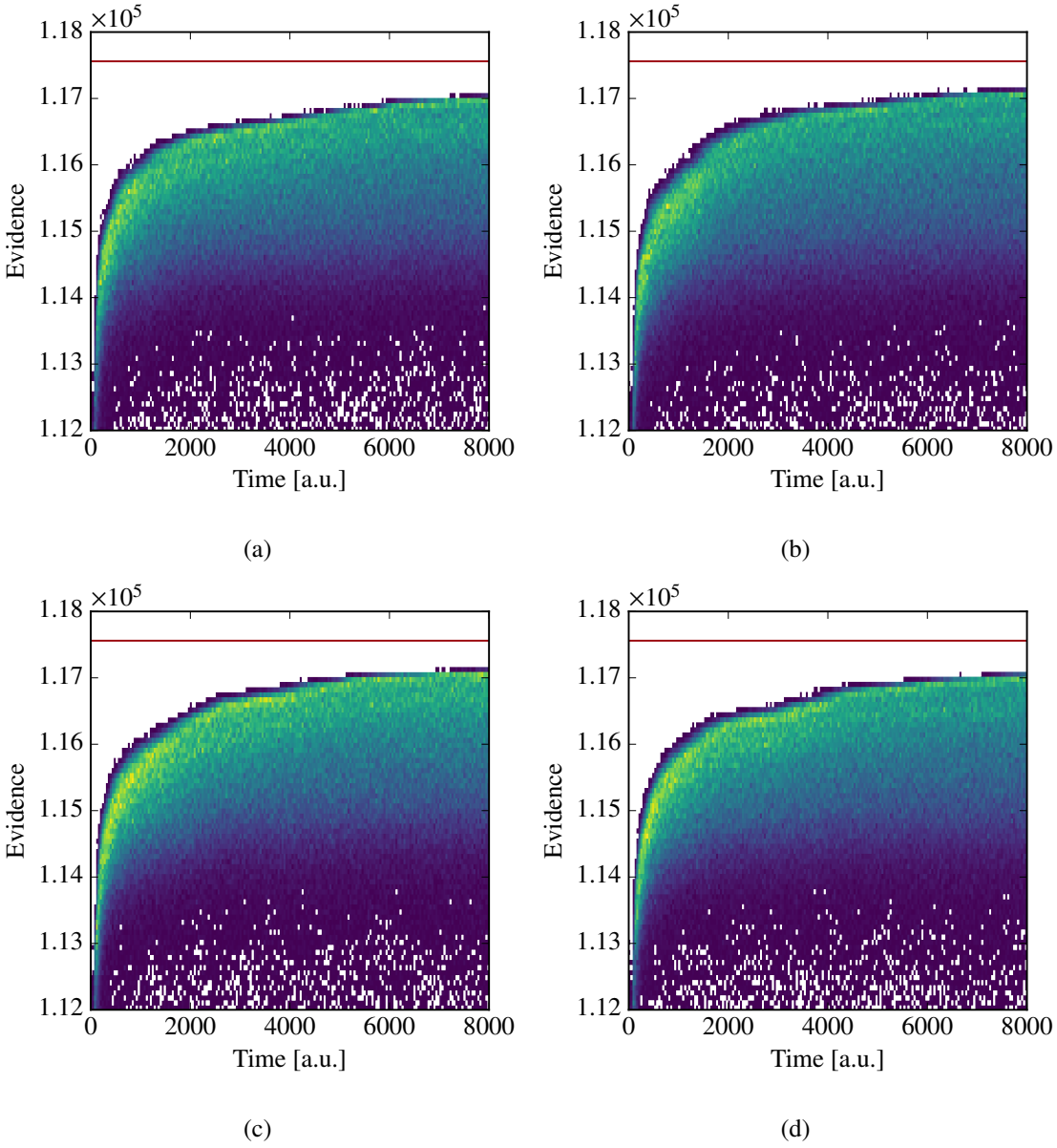


Figure C.1: Comparison of the evidence evolution comparison for four different runs of the genetic algorithm on the same generated data set.

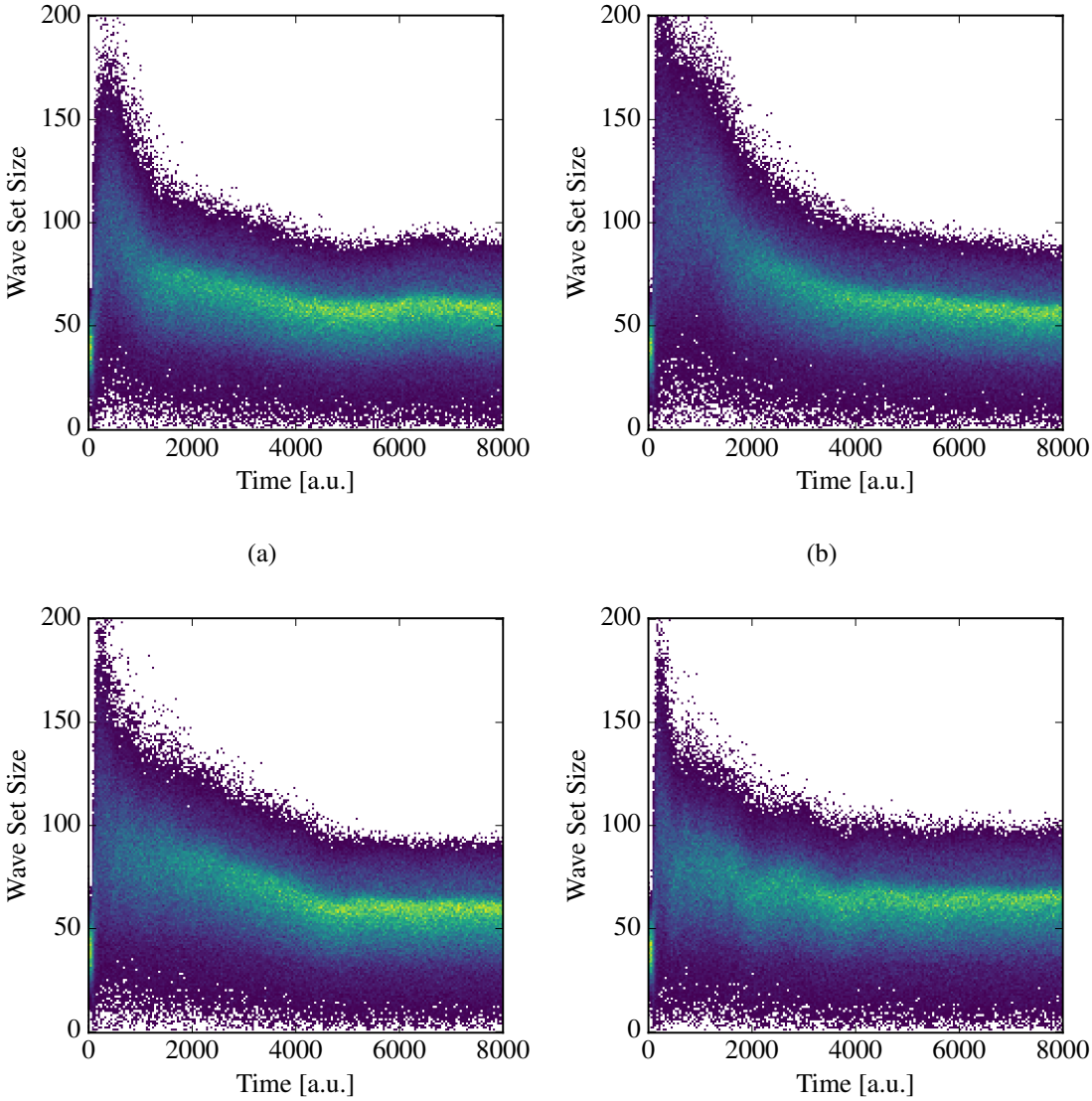


Figure C.2: Comparison of the wave set size evolution comparison for four different runs of the genetic algorithm on the same generated data set.

APPENDIX D

Intensity of a Subset of Waves

To calculate the intensity for a subset $\mathcal{W}$ of waves in the partial-wave decomposition, where we will exclude the flat wave for simplicity, we start from equation (3.16). As we are interested in the acceptance-corrected intensity, we do not include the acceptance in the integral and write

$$\mathcal{I}_{\mathcal{W}} \equiv \overline{N}_{\mathcal{W}} = \int \mathrm{d}\phi_n(\tau)\, \mathcal{I}_{\mathcal{W}}(\tau) \tag{D.1}$$

$$= \int \mathrm{d}\phi_n(\tau) \sum_{\epsilon=\pm1} \sum_{r=1}^{N_r^\epsilon} \left| \sum_{i \in \mathcal{W}^\epsilon} {}^r\mathcal{T}_i^\epsilon\, \Psi_i^\epsilon(\tau) \right|^2 \tag{D.2}$$

$$\approx \frac{V_\Omega}{N^{\mathrm{MC}}} \sum_{\mathrm{t}=1}^{N^{\mathrm{MC}}} \sum_{\epsilon=\pm1} \sum_{r=1}^{N_r^\epsilon} \left| \sum_{i \in \mathcal{W}^\epsilon} {}^r\mathcal{T}_i^\epsilon\, \Psi_i^\epsilon(\tau_{\mathrm{t}}) \right|^2 \tag{D.3}$$

$$= \frac{V_\Omega}{N^{\mathrm{MC}}} \sum_{\mathrm{t}=1}^{N^{\mathrm{MC}}} \sum_{\epsilon=\pm1} \sum_{r=1}^{N_r^\epsilon} \sum_{i,j \in \mathcal{W}^\epsilon} \frac{{}^rT_i^\epsilon}{\sqrt{\mathcal{P}_{ii}^\epsilon}} \frac{{}^rT_j^{\epsilon *}}{\sqrt{\mathcal{P}_{jj}^\epsilon}}\, \Psi_i^\epsilon(\tau_{\mathrm{t}})\, \Psi_j^{\epsilon *}(\tau_{\mathrm{t}}) \tag{D.4}$$

$$= \sum_{\epsilon=\pm1} \sum_{r=1}^{N_r^\epsilon} \sum_{i,j \in \mathcal{W}^\epsilon} \frac{{}^rT_i^\epsilon}{\sqrt{\mathcal{P}_{ii}^\epsilon}} \frac{{}^rT_j^{\epsilon *}}{\sqrt{\mathcal{P}_{jj}^\epsilon}} \frac{V_\Omega}{N^{\mathrm{MC}}} \sum_{\mathrm{t}=1}^{N^{\mathrm{MC}}} \Psi_i^\epsilon(\tau_{\mathrm{t}})\, \Psi_j^{\epsilon *}(\tau_{\mathrm{t}}) \tag{D.5}$$

$$= \sum_{\epsilon=\pm1} \sum_{r=1}^{N_r^\epsilon} \sum_{i,j \in \mathcal{W}^\epsilon} {}^rT_i^\epsilon\, {}^rT_j^{\epsilon *} \frac{\mathcal{P}_{ij}^\epsilon}{\sqrt{\mathcal{P}_{ii}^\epsilon\, \mathcal{P}_{jj}^\epsilon}} \tag{D.6}$$

We used the definition of the intensity from equation (3.13), approximated the integral with a Monte Carlo sum as shown in equation (3.24), used the definition of $\mathcal{P}_{ij}^\epsilon$ from equation (3.28) and the normalization given by equation (3.27). The indices i and j run over all the waves in the subset $\mathcal{W}^{\epsilon=\pm1}$.

By construction, the intensity $\mathcal{I}_{\mathcal{W}}$ is measured in number of acceptance-corrected events. To compare with other channels and to arrive at a less model-dependent formulation, it is useful to construct the analogue quantity to $|{}^\dagger\mathcal{T}_i|^2$ given by equation (5.10). For simplicity and to be consistent with the formulas in section 5.3, we are assuming $\epsilon = +1$ and $N_r = 1$ for this discussion and are therefore

omitting the indices for the reflectivity ϵ and the rank r. Let us first define an effective amplitude $\mathcal{T}_{\mathcal{W}}$ and a set of coefficients $C_i \in \mathbb{C}$ such that

$$\mathcal{T}_i = C_i \mathcal{T}_{\mathcal{W}} \quad \forall\, i \in \mathcal{W} \tag{D.7}$$

Using this in equation (D.3) yields

$$\begin{aligned}
\mathcal{I}_{\mathcal{W}} &= \frac{V_\Omega}{N^{\mathrm{MC}}} \sum_{\mathfrak{k}=1}^{N^{\mathrm{MC}}} \left| \sum_{i \in \mathcal{W}} \mathcal{T}_i \Psi_i(\tau_{\mathfrak{k}}) \right|^2 \\
&= \frac{V_\Omega}{N^{\mathrm{MC}}} \sum_{\mathfrak{k}=1}^{N^{\mathrm{MC}}} \left| \sum_{i \in \mathcal{W}} C_i \mathcal{T}_{\mathcal{W}} \Psi_i(\tau_{\mathfrak{k}}) \right|^2 \\
&= \mathcal{T}_{\mathcal{W}} \mathcal{T}_{\mathcal{W}}^* \sum_{i,j \in \mathcal{W}} C_i \left(\frac{V_\Omega}{N^{\mathrm{MC}}} \sum_{\mathfrak{k}=1}^{N^{\mathrm{MC}}} \Psi_i(\tau_{\mathfrak{k}}) \Psi_j{}^*(\tau_{\mathfrak{k}}) \right) C_j{}^* \\
&= \mathcal{T}_{\mathcal{W}} \mathcal{T}_{\mathcal{W}}^* \sum_{i,j \in \mathcal{W}} C_i \mathcal{P}_{ij} C_j{}^*
\end{aligned} \tag{D.8}$$

If we now define

$$\mathcal{P}_{\mathcal{W}} \equiv \sum_{i,j \in \mathcal{W}} C_i \mathcal{P}_{ij} C_j{}^* \quad \text{and} \quad T_{\mathcal{W}} \equiv \mathcal{T}_{\mathcal{W}} \sqrt{\mathcal{P}_{\mathcal{W}}} \tag{D.9}$$

we can see that

$$\mathcal{I}_{\mathcal{W}} = \left| T_{\mathcal{W}} \right|^2 \tag{D.10}$$

as for every individual wave. Analog to the definition of ${}^{\dagger}\mathcal{P}_{ij}$ given by equation (5.3) we can define

$${}^{\dagger}\mathcal{P}_{\mathcal{W}} = \frac{\mathcal{P}_{\mathcal{W}}}{V_\Omega} = \sum_{i,j \in \mathcal{W}} C_i \, {}^{\dagger}\mathcal{P}_{ij} C_j{}^* \tag{D.11}$$

Now, we have all the ingredients to write the quantity of interest

$$\left| {}^{\dagger}\mathcal{T}_{\mathcal{W}} \right|^2 = \left| T_{\mathcal{W}} \right|^2 / \, {}^{\dagger}\mathcal{P}_{\mathcal{W}} \tag{D.12}$$

provided we know the coefficients C_i. By construction, $\mathcal{P}_{\mathcal{W}} \in \mathbb{R}$ does not change if all C_i are shifted by an arbitrary phase φ. Therefore, we can assume $C_1 \in \mathbb{R}$. If we now rewrite the conditions in equation (D.7) using the definition in equation (D.9), we obtain

$$\frac{T_i}{\sqrt{\mathcal{P}_{ii}}} = C_i \frac{T_{\mathcal{W}}}{\sqrt{\mathcal{P}_{\mathcal{W}}}} \quad \forall\, i \in \mathcal{W} \tag{D.13}$$

Because we know that $|T_{\mathcal{W}}|^2 = \mathcal{I}_{\mathcal{W}}$, which can be calculated using equation (D.6), we square both sides and in addition, multiply them by V_Ω and are left with

$$\frac{\left| T_i \right|^2}{{}^{\dagger}\mathcal{P}_{ii} \mathcal{I}_{\mathcal{W}}} = \frac{\left| C_i \right|^2}{{}^{\dagger}\mathcal{P}_{\mathcal{W}}} \quad \forall\, i \in \mathcal{W} \tag{D.14}$$

One can see that multiplying all C_i by an arbitrary real factor, one can scale ${}^{\dagger}\mathcal{P}_{\mathcal{W}}$ at will while equation (D.14) is still valid. This means for every ${}^{\dagger}\mathcal{P}_{\mathcal{W}} > 0$, a set of C_i can be found such that all the conditions above are fulfilled. However, to be able to investigate the m_{FS} dependence of ${}^{\dagger}\mathcal{T}_{\mathcal{W}}$, the m_{FS} dependence of ${}^{\dagger}\mathcal{P}_{\mathcal{W}}$ needs to be known, at least up to a common real factor for all m_{FS} bins. It is unclear how to obtain this m_{FS} dependence, however one approach which proved to yield sensible results is to define another set of coefficients ${}^{\ddagger}C_i \in \mathbb{C}$ such that

$$T_i = {}^{\ddagger}C_i \, T_{\mathcal{W}} \quad \forall\, i \in \mathcal{W} \tag{D.15}$$

With ${}^{\ddagger}C_1 \in \mathbb{R}$ as before and the fact that $|T_{\mathcal{W}}|^2 = \mathcal{I}_{\mathcal{W}}$, the ${}^{\ddagger}C_i$ are fully determined. We can use them to write

$${}^{\dagger}\mathcal{P}_{\mathcal{W}} = \sum_{i,j \in \mathcal{W}} {}^{\ddagger}C_i \, {}^{\dagger}\mathcal{P}_{ij} \, {}^{\ddagger}C_j^{\,*} \tag{D.16}$$

With equation (D.12), it is now possible to calculate $|{}^{\dagger}\mathcal{T}_{\mathcal{W}}|^2$. Such an intensity distribution is shown in figure 5.13d. While the ${}^{\dagger}\mathcal{P}_{\mathcal{W}}$ and $|{}^{\dagger}\mathcal{T}_{\mathcal{W}}|^2$ obtained by introducing equations (D.15) and (D.16) are valid and consistent, the motivation for its use is based on the empirical observation that the resulting m_{FS} dependence of ${}^{\dagger}\mathcal{T}_{\mathcal{W}}$ looks reasonable. Therefore, due caution is appropriate when interpreting the results of this approach.

For illustration, we determine the relation between C_i and ${}^{\ddagger}C_i$ by using equation (D.13), and obtain

$$\frac{{}^{\ddagger}C_i}{C_i} = \sqrt{\frac{{}^{\dagger}\mathcal{P}_{ii}}{{}^{\dagger}\mathcal{P}_{\mathcal{W}}}} \quad \forall\, i \in \mathcal{W} \tag{D.17}$$

Bibliography

[1] J. J. Dudek et al., "Toward the excited isoscalar meson spectrum from lattice QCD," Phys. Rev. **D88** (2013) 094505, arXiv: 1309.2608 [hep-lat]. (cited on page 1)

[2] G. E. Moore, "Cramming more components onto integrated circuits," Electronics (1965). (cited on page 1)

[3] M. Roos, "Tables of Elementary Particles and Resonant States," Rev. Mod. Phys. **35** (1963) 314–323. (cited on page 1)

[4] K. A. Olive et al., "Review of Particle Physics," Chin. Phys. **C38** (2014) 090001, and 2015 update. (cited on pages 1, 107, 110, 117)

[5] C. Adolph et al., "Observation of a New Narrow Axial-Vector Meson $a_1(1420)$," Phys. Rev. Lett. **115** (2015) 082001, arXiv: 1501.05732 [hep-ex]. (cited on page 2)

[6] C. Adolph et al., "Resonance Production and $\pi\pi$ S-wave in $\pi^- + p \to \pi^-\pi^-\pi^+ + p_{\text{recoil}}$ at 190 GeV/c," (2016), arXiv: 1509.00992 [hep-ex], submitted to Phys. Rev. D. (cited on pages 2, 15, 18, 32, 89, 106)

[7] S. Schmeing, "Resonance Extraction in Diffractive 3π Production using 190 GeV/c π^- at the COMPASS experiment (CERN)," master's thesis: Technische Universität München, 2014, URL: http://wwwcompass.cern.ch/compass/publications/theses/2014_dpl_schmeing.pdf. (cited on page 2)

[8] S. Wallner, "Extraction of Resonance Parameters of Light Meson Resonances in the Charged Three-Pion Final State at the COMPASS Experiment (CERN)," master's thesis: Technische Universität München, 2015, URL: http://wwwcompass.cern.ch/compass/publications/theses/2015_dpl_wallner.pdf. (cited on pages 2, 103)

[9] S. Uhl, PhD thesis: Technische Universität München, 2016, in preparation. (cited on pages 2, 103, 119)

[10] S. Neubert, "First Amplitude Analysis of Resonant Structures in the 5-Pion Continuum at COMPASS," PhD Thesis: Technische Universität München, 2012, URL: https://wwwcompass.cern.ch/compass/publications/theses/2012_phd_neubert.pdf. (cited on pages 2, 29, 40, 42, 46, 47, 49, 51, 84, 88, 102, 117, 119)

[11] P. Abbon et al., "The COMPASS experiment at CERN," Nucl. Instrum. Meth. **A577** (2007) 455–518, arXiv: hep-ex/0703049 [hep-ex]. (cited on page 3)

[12] P. Abbon et al., "The COMPASS Setup for Physics with Hadron Beams," Nucl. Instrum. Meth. **A779** (2015) 69–115, arXiv: 1410.1797 [physics.ins-det]. (cited on page 3)

[13] Q. Weitzel, "Precision Meson Spectroscopy: Diffractive Production at COMPASS and Development of a GEM-based TPC for PANDA," PhD thesis: Technische Universität München, 2008, URL: https://wwwcompass.cern.ch/compass/publications/theses/2008_phd_weitzel.pdf. (cited on page 8)

[14] G. Ascoli et al., "Partial Wave Analysis of the 3π Decay of the A_2," Phys. Rev. Lett. **25** (1970) 962. (cited on page 15)

[15] J. Hansen et al., "Formalism and assumptions involved in partial-wave analysis of three-meson systems," Nuclear Physics B **81** (1974) 403–430, URL: http://www.sciencedirect.com/science/article/pii/0550321374902417. (cited on page 15)

[16] C. Daum et al., "Diffractive Production of 3π States at 63 and 94 GeV," Nucl. Phys. **B182** (1981) 269. (cited on pages 15, 18)

[17] S. U. Chung and T. L. Trueman, "Positivity Conditions on the Spin Density Matrix: A Simple Parametrization," Phys. Rev. **D11** (1975) 633. (cited on pages 15, 17)

[18] C. W. Salgado and D. P. Weygand, "On the Partial-Wave Analysis of Mesonic Resonances Decaying to Multiparticle Final States Produced by Polarized Photons," Phys. Rept. **537** (2014) 1–58, arXiv: 1310.7498 [nucl-ex]. (cited on pages 15, 17, 32)

[19] S. Mandelstam, "Determination of the Pion-Nucleon Scattering Amplitude from Dispersion Relations and Unitarity. General theory," Phys. Rev. **112** (1958) 1344–1360. (cited on page 16)

[20] P. A. M. Dirac, "Quantum theory of emission and absorption of radiation," Proc. Roy. Soc. Lond. **A114** (1927) 243–265. (cited on page 18)

[21] M. Alekseev et al., "Observation of a $J^{PC} = 1^{-+}$ exotic resonance in diffractive dissociation of 190-GeV/c π^- into $\pi^-\pi^-\pi^+$," Phys. Rev. Lett. **104** (2010) 241803, arXiv: 0910.5842 [hep-ex]. (cited on page 18)

[22] R. H. Landau, M. J. Páez, and C. C. Bordeianu, *A Survey of Computational Physics*, Princeton University Press, 2008, ISBN: 978-0-691-13137-5. (cited on page 22)

[23] E. Wigner, "Einige Folgerungen aus der Schrödingerschen Theorie für die Termstrukturen," Z. Phys. **43** (1927) 624–652. (cited on page 26)

[24] J. Blatt and V. Weisskopf, *Theoretical Nuclear Physics*, New York: John Wiley & Sons, 1952. (cited on page 26)

[25] F. von Hippel and C. Quigg, "Centrifugal-barrier effects in resonance partial decay widths, shapes, and production amplitudes," Phys. Rev. **D5** (1972) 624–638. (cited on page 26)

[26] G. Breit and E. Wigner, "Capture of Slow Neutrons," Phys. Rev. **49** (1936) 519–531. (cited on page 28)

[27] J.-H. Lee, "The BNL Four-Body Partial Wave Analysis System," 2008, not published. (cited on page 29)

[28] K. L. Au, D. Morgan, and M. R. Pennington, "Meson Dynamics Beyond the Quark Model: A Study of Final State Interactions," Phys. Rev. **D35** (1987) 1633. (cited on page 32)

[29] F. Krinner (Technische Universität München), personal communication, 2016. (cited on page 36)

[30] D. Ryabchikov (IHEP, Protvino), personal communication, 2016. (cited on pages 36, 106, 120)

[31] R. E. Kass and A. E. Raftery, "Bayes Factors," J. Am. Stat. Assoc. **90** (1995) 773–795, URL: http://www.jstor.org/stable/2291091. (cited on page 41)

[32] F. Feroz and M. P. Hobson, "Multimodal nested sampling: an efficient and robust alternative to MCMC methods for astronomical data analysis," Mon. Not. Roy. Astron. Soc. **384** (2008) 449, arXiv: 0704.3704 [astro-ph]. (cited on page 41)

[33] F. Feroz, M. P. Hobson, and M. Bridges, "MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics," Mon. Not. Roy. Astron. Soc. **398** (2009) 1601–1614, arXiv: 0809.3437 [astro-ph]. (cited on page 41)

[34] F. Feroz et al., "Importance Nested Sampling and the MultiNest Algorithm," (2013), arXiv: 1306.2144 [astro-ph.IM]. (cited on page 41)

[35] J. Skilling, "Nested sampling for general Bayesian computation," Bayesian Anal. **1** (2006) 833–859. (cited on page 41)

[36] D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms*, Version 7.2, Cambridge Univ. Press, 2003. (cited on page 42)

[37] F. Beaujean (Exzellenzcluster Universe, Garching), personal communication, 2015. (cited on page 56)

[38] *ROOT a Data analysis Framework*, URL: http://root.cern.ch/ (visited on 2016-02-26). (cited on page 56)

[39] F. James et al., *Minuit*, URL: http://seal.web.cern.ch/seal/snapshot/work-packages/mathlibs/minuit/ (visited on 2016-02-26). (cited on page 56)

[40] S. G. Johnson, *The NLopt nonlinear-optimization package*, URL: http://ab-initio.mit.edu/nlopt (visited on 2016-02-26). (cited on page 56)

[41] J. Nocedal, "Updating quasi-Newton matrices with limited storage," Math. Comp. **35** (1980) 773–782. (cited on page 56)

[42] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," Math. Program. **45** (1989) 503–528. (cited on page 56)

[43] L. Luksan, *Ladislav Luksan - subroutines*, URL: http://www.uivt.cas.cz/~luksan/subroutines.html (visited on 2016-02-26). (cited on page 56)

[44] M. Betancourt (University of Warwick, Coventry), personal communication, 2015. (cited on page 56)

[45] D. F. Andrews, "Plots of High-Dimensional Data," Biometrics **28** (1972) 125–136, URL: http://www.jstor.org/stable/2528964. (cited on page 62)

[46] M. Ester et al., "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise," *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), Portland, Oregon, USA*, 1996 226–231, URL: http://www.aaai.org/Library/KDD/1996/kdd96-037.php. (cited on page 62)

[47] *scikit-learn: machine learning in Python*, URL: http://scikit-learn.org (visited on 2016-02-26). (cited on page 62)

[48] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," J. Mach. Learn. Res. **12** (2011) 2825–2830. (cited on page 62)

[49] M. N. Diefenbach, "Monte Carlo Analysis Of The Acceptances In The $\pi^-\eta'$ And $\pi^-\eta$ Final States At COMPASS (CERN)," bachelor's thesis: Ludwig-Maximilians-Universität München, 2011, URL: https://wwwcompass.cern.ch/compass/publications/theses/2011_bac_diefenbach.pdf. (cited on page 106)

[50] O. J. Drotleff, "Model Selection for Partial-Wave Analysis of $\pi^- + p \to \pi^-\pi^+\pi^- + p$ at the COMPASS Experiment at CERN," diploma thesis: Technische Universität München, 2015, URL: http://wwwcompass.cern.ch/compass/publications/theses/2015_dpl_drotleff.pdf. (cited on page 119)

List of Figures

Additional Plots from the Genetic Algorithm

List of Tables

Own Contributions

Using the data set measured by the COMPASS collaboration in 2008, I selected an event sample of exclusive five-pion events. For the preselection, the software framework of the COMPASS collaboration was used and the software code was based on the work of Florian Haas. For the second part of the event selection, a custom software was developed from scratch. For partial-wave decompositions, the ROOTPWA software package was used, which is continuously developed and maintained by our group. Notable contributors are Oliver Drotleff, Boris Grube, Sebastian Neubert and Sebastian Uhl. In ROOTPWA, I implemented the automatic isospin symmetrization and the phase space calculation for the relativistic Breit-Wigner functions, the latter with the help of Jan Friedrich. The genetic algorithm had been implemented into ROOTPWA by Sebastian Neubert and has been rewritten from scratch by me. Additionally, in ROOTPWA, I implemented the biggest-conceivable-model method and integrated the NLopt optimization package. In collaboration with the other ROOTPWA developers, the framework was restructured and new, self-describing data formats were developed and implemented. Much of the methodology of verifying the model-selection procedures was based on earlier work and developed further and refined by me. Under my supervision Oliver Drotleff wrote his diploma thesis using the BCM method. All figures in this thesis have been created by me using matplotlib, tikz and povray, with the exception of the background image of the COMPASS spectrometer. The work was supervised by Boris Grube and in parts by Jan Friedrich.

Several fields of activity did not make it into this document. In the 2011 and 2012 data-taking campaigns, I was responsible for the operation of the silicon detectors. For Monte Carlo simulations, I developed a new method of simulating incoming beam particles such that they have the same properties as the physical beam. In collaboration with Jan Friedrich, I developed a likelihood-based method of selecting events using the Recoil Proton Detector of COMPASS.

Acknowledgements

First and foremost, I would like to thank Prof. Stephan Paul, for his guidance and support, for giving me the opportunity to work at E18 and for the very interesting and challenging thesis topic. I greatly appreciate the confidence which he placed in me by sending me to CERN.

Gerhard Mallot, who supervised me during the three years I spent at CERN, deserves thanks for always supporting me and my work.

Both supervisors made it possible for me to visit several conferences, where I learned a great deal and was able to show some of my work, for which I am very grateful.

Boris Grube not only guided my work, patiently explained physics to me and proofread this whole document, but also shared an office with me at CERN and, inexplicably, agreed to do it again after my return to Munich. He suffered through a number of my quirks, in particular through my “taste” in music. Thank you for your help and your support.

Jan Friedrich always took the time to listen to my problems and offered great advice in several challenging situations. I learned a lot from the many interesting and fruitful discussions with him. Thank you for supporting me and for never stopping to challenge me.

I thank Dima Ryabchikov for supporting this analysis and patiently answering and discussing my many questions. He spent a lot of time to make it possible to compare our results and I owe him for that.

Suh-Urk Chung was kind enough to share his experience with me, for which I would like to thank him.

Felix Böhmer reliably brought me to tears with laughter. For this, and for his unique perspective on life, I thank him very much. No matter where the universe will take us, we will always have the magical time with the Borsäure.

The collaboration with Sebastian Uhl on ROOTPWA was not only very productive, but also great fun. Thank you for this and for pointing out and fixing many of my bugs.

Fabian Krinner was kind enough to help me out when my math skills were at an end. Thank you for all the valuable discussions and I hope you did not suffer too much from all the pull request comments.

I thank Alexander Austregesilo for sharing many laughs and some very interesting discussions with me and Daniel Greenwald for being himself.

I thank Oliver Drotleff for enduring me during his time as a diploma student and for the excellent contributions to ROOTPWA.

Christian Dreisbach, with whom I spent great times discussing software development philosophy, agreed to take on the silicon project and made it possible for me to concentrate on writing this thesis. If not for him, it would probably have taken another year. Thank you!

To all the current and past members of E18, in particular Sverre Dørheim, Steffi Grabmüller, Florian Haas, Stefan Huber, Igor Konorov, Alexander Mann and Stefan Wallner: thank you for the discussions and the fun and exciting times together. Karin Frank deserves special thanks for keeping E18 running.

All the colleagues from the COMPASS collaboration deserve thanks for taking the data this work is based on, for many interesting discussions and a lot of fun on and off work. I thank Sergei Gerassimov for the many coffees together and the fun times talking shop about C++, variables names and programming in general.

I thank Marcus “Mess” Barth for drastically increasing the amount of derailed thought trains since my diploma thesis, Stefan “Lichti” Lichtmannecker for sharing many virtual beers and Fabian Heidelberg for sharing many virtual flights.

My parents Romy and Charly and my brother Raphael were always there for me and I owe them much. For shaping me and for allowing me to pursue this path, I thank them from the bottom of my heart.

I cannot begin to thank my wife, Teresa Negrini. Without her, I probably would have starved in front of a computer screen at some point. Thank you for the superhuman things you had to do to make it possible for me to finish this thesis. I will be forever in your debt. Sophia, you are the light of my life, thank you for teaching me to see the world through your eyes.