Bayesian Data Analysis in Baryon Spectroscopy

Stefanie Janneke Lewis

A Thesis presented for the degree of
Doctor of Philosophy

University of Glasgow

Nuclear Physics
School of Physics and Astronomy
University of Glasgow
Scotland

June 2014

Dedicated to

Mom, Dad, Granny, Grandpa, and Grandpa Jack.

Abstract

The strong interaction within a nucleon has been the focus of much theoretical and experimental work in nuclear and particle physics. Theorists have been improving lattice QCD calculations and developing quark models that define the inter-quark interactions, and experimentalists have spent years gathering data to support and improve these models. Finding nucleon resonance states provides essential information for the development of these theories and improves our understanding of the excited nucleon spectrum. There are a variety of quark models that have been proposed which each predict a unique resonance spectrum. Currently, these models predict resonances that have not been observed experimentally. It is important to experimentally determine which of these resonances exist. Historically, many of the existing measurements came as a result of nucleon-pion scattering experiments. It has been suggested, however, that some resonances may couple more strongly to other reaction channels, such as the $K\Lambda$ strangeness reaction channel analysed here. Pseudoscalar meson photoproduction experiments can be used to analyse such a reaction channel. In these experiments, a photon beam is incident on a stationary nucleon target and the reaction products are detected. The polarisation of the recoiling particle can often be determined or measured. In the $K\Lambda$ channel, the recoiling baryon is a Λ whose polarisation can be obtained without the use of any additional hardware through the self-analysing properties of the hyperon. These experiments can be completely described by four complex amplitudes, which can be accessed experimentally through sixteen polarisation observables. The polarisation observables are bilinear combinations of the amplitudes and as such have nontrivial correlations. They are dependent on the polarisations of the beam, target and recoiling particle. By selecting different polarisations of the beam or target, or by using a combination of polarisations, different observables can be measured. The amplitudes can be obtained once a sufficient selection of observables is determined. Currently, analyses of pseudoscalar meson photoproduction data is done using a binned fitting method. The use of binned fitting inevitably leads to some information from the data being lost. In this thesis, a new analysis method is presented, based on Bayesian statistics. The aim of such an approach is to maximise the

information yield from data. An event-by-event likelihood function reduces the information lost through histogram binning. It is shown that the use of a prior distribution in amplitude space can preserve the correlations between observables, also increasing the information yield. It is found that such an analysis programme leads to a significant extraction of information from existing data. It is also found that datasets from different experiments could be concatenated and analysed together using the programme presented in this work, and successfully extract observables. Information on observables to which the experiment is not directly sensitive can be found and visualised graphically. The development of this analysis programme is detailed in this thesis. Previously analysed data from two experiments are analysed using this analysis method, and the results are compared to those obtained in the past. It is shown that this Bayesian approach produces results that are consistent with accepted results and provides information on observables that are not directly measurable by a particular experiment. The data from two experiments is combined and analysed together, and it is shown that the results of the combined analysis are consistent with those obtained through separate analyses.

Declaration

The work in this thesis is based on research carried out at the Nuclear Physics Experimental Group, School of Physics and Astronomy University of Glasgow, Scotland. No part of this thesis has been submitted elsewhere for any other degree or qualification and is all my own work unless referenced to the contrary in the text.

Acknowledgements

There are many people that have helped me over the past three and a half years, and I would like to extend my gratitude for everything they have done.

First and foremost, I would like to thank my supervisors, Prof Dave Ireland and Dr Ken Livingston for the opportunity and for all their help during this project. They have not only helped in the preparation of my PhD work, but have also helped me develop my interest and enthusiasm. I am also owing to Dr Wim Vanderbauwhede for his help developing some of the more complicated computational aspects of my project. I would also like to thank Bryan McKinnon, Morgan Murray, Daria Sokhan and David Hamilton for their advice at many points during my PhD. Thanks also to Brian Vernarsky from CMU for his help with COBRA.

Life as a PhD student has been made much easier by the company of friendly and interesting colleagues, and for that I have the other PhD students to thank. Simon, Euan and Bruno have been excellent (never boring) officemates. Johan, Sian, Rachel, Gary, Mark and Jeff - thanks for making the group such a friendly environment. I would like to extend a special thank-you to two people who have shared their wisdom with me - Seian Al Jebali and Stuart Fegan. You both have helped me more than you know.

Towards the outskirts of the academic environment, I would like to express my appreciation for my good friends in the MCMP group who have kept me relatively sane, especially during the final stretch of my PhD. Thanks to Aaron, Francisco, Pablo (and Veronica!), and Rob for terrific evenings at the pub.

I would especially like to thank Ciaran, for all of his support and advice (and also for putting up with me during the writing-up stage). Thanks for your patience and understanding. Although this may have been a particularly stressful experience, it was made considerably less so by your company, and I appreciate everything that you have done for me.

Thanks to my family, for their many years of support and encouragement.

"Don't panic."

- Douglas Adams *The Hitchhiker's Guide to the Galaxy*

Contents

List of Figures

List of Tables

1

Introduction

The study of nuclear and particle physics was revolutionised in the mid-20th century when cosmic ray experiments such as those of Rochester and Butler [1] suggested the existence of unstable particles. These results are now credited as the discovery of what we know now as hyperons [2], particles containing at least one strange quark. This was the beginning of many new particle discoveries. In 1950, the Λ^0 baryon was found by Hopper and Biswas from the University of Melbourne [3]. Several years later, the first baryon resonance was discovered by Anderson, Fermi, Long and Nagle [4, 5] by comparing total cross-sections of pion-proton (πp) systems. The comparison of the $\pi^- p$ cross-section with that of $\pi^+ p$ suggested the presence of a resonance. These events sparked an increased interest in the study of elementary particles, and many more baryonic states and resonances were soon discovered [6]. It was not long before scientists attempted to sort these new states, and in 1961, Gell-Mann and Ne'eman introduced the Eightfold Way, their hadron

classification scheme [7] based on previous efforts by Gell-Mann and Nishijima [8]. Several years previously, the Sakata model had been developed [9] where it was postulated that protons, neutrons, and Λ particles were the fundamental particles in baryons and mesons. This model encouraged the symmetry-based approach and introduced the importance of the SU(3) symmetry group. This paradigm led Gell-Mann and Ne'eman to the conclusion that, assuming electric charge Q was included in the symmetry, only particles with an integer charge could be members of certain groupings, or families. It was determined that, in the simplest case, particles could either belong to 1-, 8-, or 10-member families. The baryon octet (shown in Fig. 1.1) was then introduced.

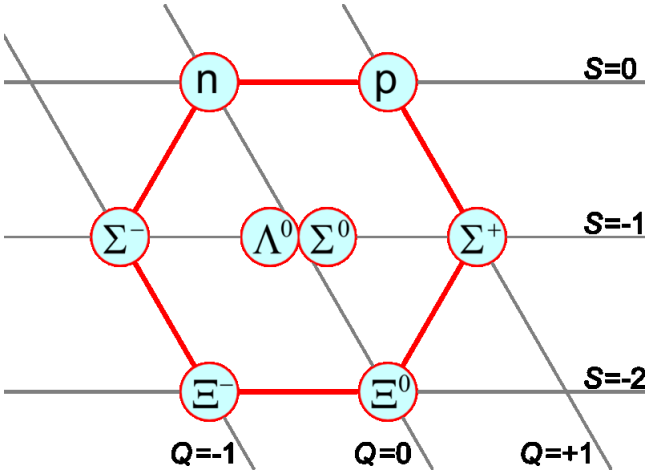


Figure 1.1: The baryon octet, a plot of strangeness against charge, Q [10].

His work on classifying hadrons led Gell-Mann to the concept of quarks, elementary particles that formed hadrons. Independently, Zweig arrived at the same conclusion [11]. These postulations incited the development of a variety of models and theories to explain the strong interaction and quarks.

1.1 Quark Models and QCD

The strong interaction is the fundamental force responsible for maintaining the structure of a nucleus. The force repelling the positively charged protons in the nucleus is overcome by the strong force, binding the nucleus. The strong interaction is

experienced by quark-based particles such as baryons and mesons, and is mediated by gluons through the colour charge. Non-quark-based particles, such as leptons, do not interact via the strong force.

During the 1970s, developments to the Standard Model [12] were made to include quarks, evidence of which had recently been found experimentally [13, 14], and the strong interaction, in the form of quantum chromodynamics (QCD) , the field theory associated with the strong force.

The discovery of the Δ^{++} hadronic state posed a significant challenge to the development of quark models. The Δ^{++}, a baryon made up of three 'up' quarks with parallel spin, contained fermionic components that appeared to occupy the same quantum state - a violation of the Pauli exclusion principle. In order to incorporate this phenomenon into the existing theories, a new "colour" quantum number was put forth, thereby allowing particles such as Δ^{++} to exist in accordance with Fermi statistics [15]. Three colours - red, green and blue - were introduced. It was suggested that only colour-neutral states could exist, under the rules that red + green + blue = white (for baryons), and any colour + its anticolour = white (for mesons). The introduction of this colour charge gave quantum chromodynamics its name.

QCD is a non-Abelian gauge theory associated with the strong interaction, and provides an insight into the properties and interactions of the six flavours of quarks (up, down, strange, charm, top and bottom) featured in the Standard Model through the inclusion of the 'colour' quantum number. One interesting feature of QCD is known as confinement - a phenomenon where free quarks cannot be experimentally observed [16]. This is presumed to be due to the strength of the interaction increasing as the separation distance increases. For this reason, it is impossible to remove a quark from a nucleon.

Whilst QCD is relatively successful in describing the strong interaction for quarks and gluons within a nucleon, complications arise when it is applied in a larger scale, such as the interactions and properties of nuclei. This is because the calculations involved in QCD become far more complex. Unlike other couplings, such as that of the weak interaction, the strong interaction becomes weaker as the distance between interacting particles decreases, a characteristic property of QCD known as asymp-

totic freedom. This is reflected in the behaviour of the strong interaction's running coupling constant, α_s, which decreases as energies increase toward the perturbative region, but increases towards unity as the energy decreases to the energy scale of nucleons [17–20]. This is a problem that is at the heart of much particle and nuclear physics research. Lattice-QCD, initially proposed by Wilson [21], takes a numerical approach to the complicated path integral calculations by discretising space-time. The spacing used in this discretisation determines both the accuracy and the computational cost of the calculation - the smaller the spacing, the more accurate and more computationally expensive the calculation will be. As developments in computing and parallelism are made, smaller lattice spacings can be applied and provide more accurate results. The current state of lattice QCD calculations predict, for example, a pion mass of around 300-400 MeV/c^2, compared to its mass of 134.5 MeV/c^2. Current developments in lattice QCD aim to exploit the recent advances in hardware and computing methods to handle the complex calculations. These calculations are computationally very expensive. Other methods of understanding the properties of nucleons involve the development of quark models, which can be used to predict spectra of baryonic states.

Quark models were introduced in an effort to explain the composite nature of mesons and baryons, as well as the newly discovered baryon states. As more and more hadronic states were discovered, the quark theories had to be adapted. Originally, it was thought that only three flavours of quarks existed - up, down and strange, but new discoveries led to the possibility of other, heavier flavours. The existence of these heavier quarks (charm, top and bottom) were confirmed in several experiments conducted between 1970 and 1995. Over time, many quark models were developed, such as the symmetric quark model and the diquark models, described below. These models each predicted a spectrum of nucleon resonance states. These predicted resonances greatly outnumber those that have been discovered experimentally, which has created a problem in hadronic physics known as the "missing resonance" problem. Experiments in a range of international facilities concentrate on learning more about the excited nucleon spectrum and identifying resonance states predicted by quark models. As each of these models predicts a unique spectrum of resonance

states, the experimental evidence of a resonance in one spectrum and not another would provide valuable insight to the properties and interactions of nucleons in the low-energy region.

Symmetric Quark Model

The symmetric quark models (SQMs) stemmed from the postulations of Gell-Mann and Ne'eman, who introduced the idea that hadrons were composed of smaller constituents, as illustrated in Figure 1.2. The model was based on the premise that hadrons were composite particles and that their constituents (quarks) were not point-like, as put forth in the earlier QCD-based theories. SQMs were developed on the principle that these non-point-like constituent quarks (known as "dressed" quarks) were symmetrically equivalent and had energy-dependent masses [22, 23]. The key characteristics of QCD - asymptotic freedom and quark confinement - were incorporated into the theory by introducing various potentials, as described in [24]. The nucleon resonance spectrum predicted by one quark model is shown in Figure 1.3. In this figure, it can be seen that the model predicts many resonances at the higher end of the mass scale - many more than have been found experimentally.

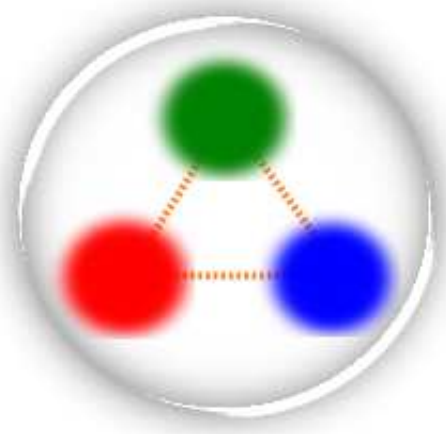


Figure 1.2: Symmetric quark model.

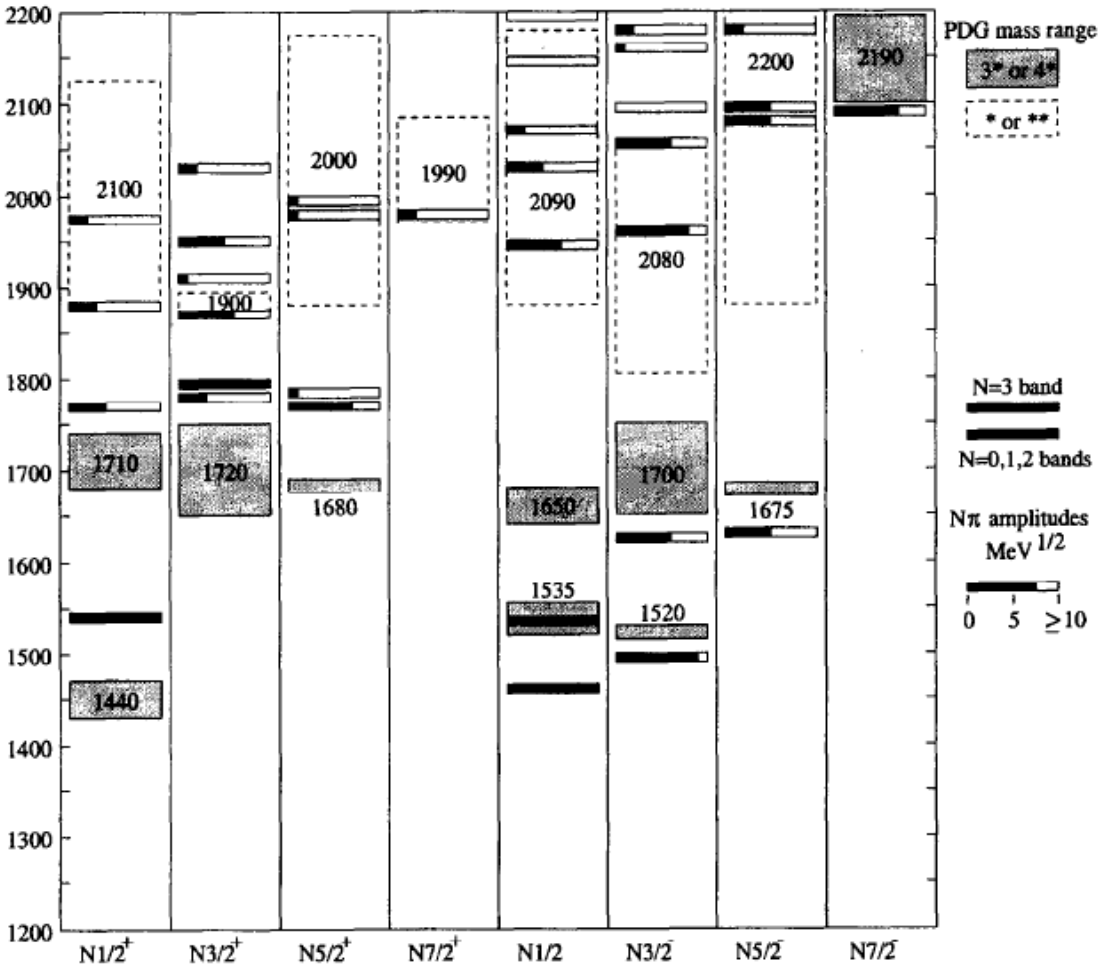


Figure 1.3: Nucleon mass predictions from a relativised quark model [25], where light-coloured boxes represent the one- or two-star resonant masses from the Particle Data Group (PDG) and darker coloured boxes depict states with three or four stars in the PDG.

Quark-Diquark Models

There exist a large number of quark-diquark models, which all differ from the constituent quark model by one fundamental assumption. Unlike the SQM, the quark-diquark model predicts mass resonances by assuming that two of the three quarks within the hadron are tightly bound [26–29], as illustrated in Figure 1.4. This assumption results in a reduction in the number of degrees of freedom, leading to the prediction of fewer resonant states, shown in Figure 1.5. Although the diquark models predict fewer states than the SQM, there still remain numerous resonances

that have not been discovered experimentally.

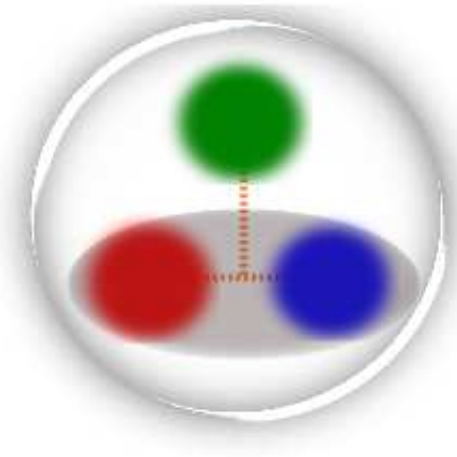


Figure 1.4: Diquark model.

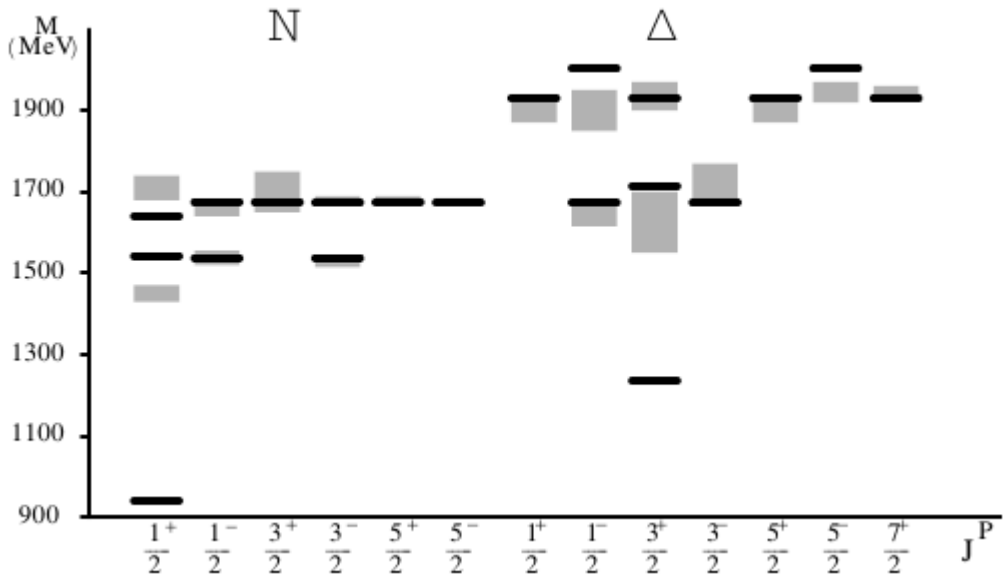


Figure 1.5: Mass resonances predicted by the quark-diquark model [29]. Grey-shaded boxes indicate experimental three- and four-star resonances listed in the PDG.

1.2 Experimental Efforts

The "missing resonance" problem is at the focus of many physics programmes around the world. The Particle Data Group (PDG) [30] is an international organisation that analyses experimental results and classifies resonant states using a star rating system. All predicted and experimentally observed resonant states are listed and assigned a star rating - four stars indicating those which have solid experimental evidence, are certain to exist and whose properties have been explored, ranging to one or no stars, for states who have not been observed or that have very poor evidence of existence. A list of baryon resonant states with their ratings is shown in Table 1.1.

Table 1.1: PDG Table of N* Resonant States [30]

N^*	Spin-parity J^P	Status
p	$1/2^+$	****
n	$1/2^+$	****
$N(1440)$	$1/2^+$	****
$N(1520)$	$3/2^-$	****
$N(1535)$	$1/2^-$	****
$N(1650)$	$1/2^-$	****
$N(1675)$	$5/2^-$	****
$N(1680)$	$5/2^+$	****
$N(1685)$		*
$N(1700)$	$3/2^-$	***
$N(1710)$	$1/2^+$	***
$N(1720)$	$3/2^+$	****
$N(1860)$	$5/2^+$	**
$N(1875)$	$3/2^-$	***
$N(1880)$	$1/2^+$	**
$N(1895)$	$1/2^-$	**
$N(1900)$	$3/2^+$	***
$N(1990)$	$7/2^+$	**
$N(2000)$	$5/2^+$	**
$N(2040)$	$3/2^+$	*
$N(2060)$	$5/2^-$	**
$N(2100)$	$1/2^+$	*
$N(2190)$	$7/2^-$	****
$N(2220)$	$9/2^+$	****
$N(2250)$	$9/2^-$	****
$N(2600)$	$11/2^-$	***
$N(2700)$	$13/2^+$	**

Historically, most of the world data collected on the existence of resonant states has been obtained through scattering experiments that result in πN final states [31]. The results of these πN scattering experiments, however, do not suggest the existence of many of these missing resonances predicted by various quark models. More sensitive measurement techniques were needed to find these states. Meson photoproduction experiments then emerged as a new, more sensitive approach. These experiments lead to information on resonances through the measurements of polarisation observables. Studies by Capstick and Roberts [25] revealed that some of these

missing resonances may strongly couple to certain reaction channels, in particular those channels where the final state includes a strange baryon. Figure 1.6 illustrates the total cross-section as a sum of the cross-sections of the various reaction channels.

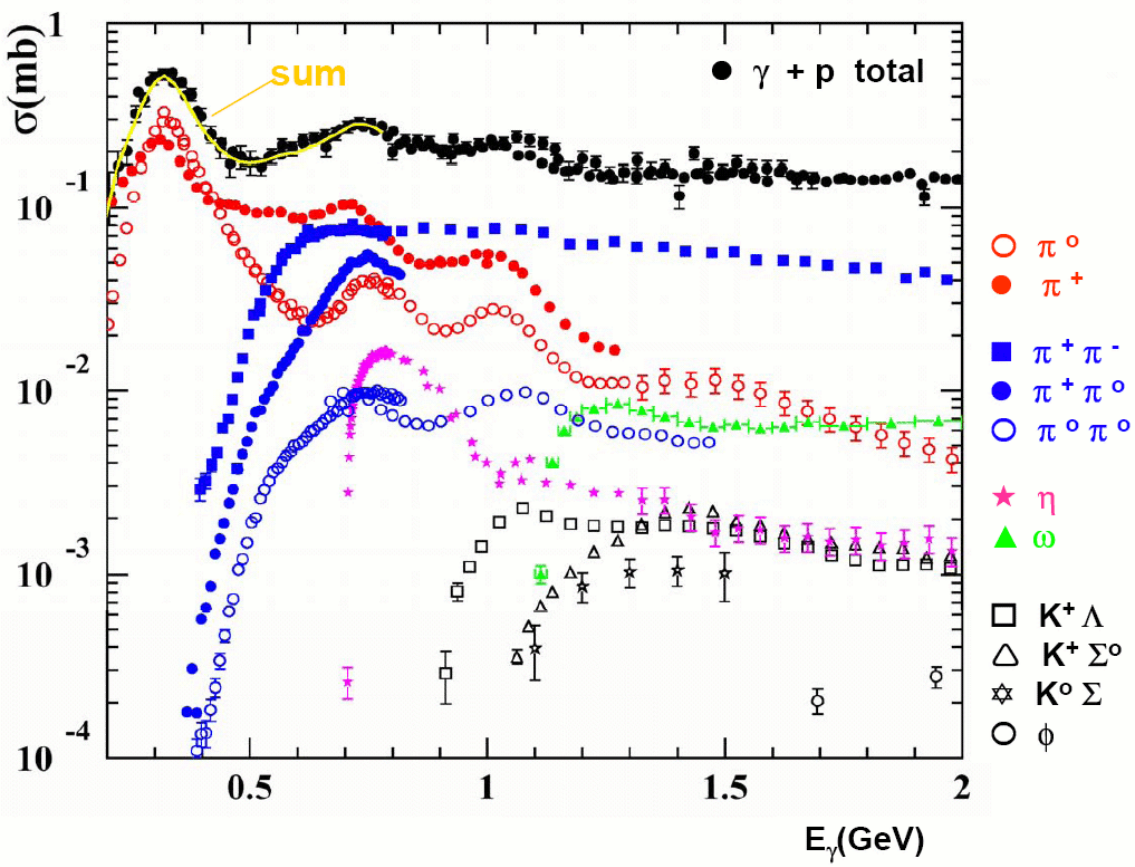


Figure 1.6: Cross-sections in photoproduction for the 0.2 - 2.0 GeV energy region. The separated reaction channels give insight into resonant states. [32]

The analysis presented in this thesis concentrates on meson photoproduction reactions. Such reactions can provide access to different couplings than pion production, and as such are being performed at numerous experimental facilities around the world, most notably including the Thomas Jefferson National Accelerator Facility (Jefferson Lab) . Jefferson Lab is home to CEBAF , the Continuous Electron Beam Accelerator Facility, which generates an electron beam with an energy of up to 5.7 GeV [33]. A diagram of the CEBAF accelerator is shown in Figure 1.7.

The CEBAF accelerator consists primarily of two superconducting linear accelerators and uses steering magnets to arc between them. The injected polarised

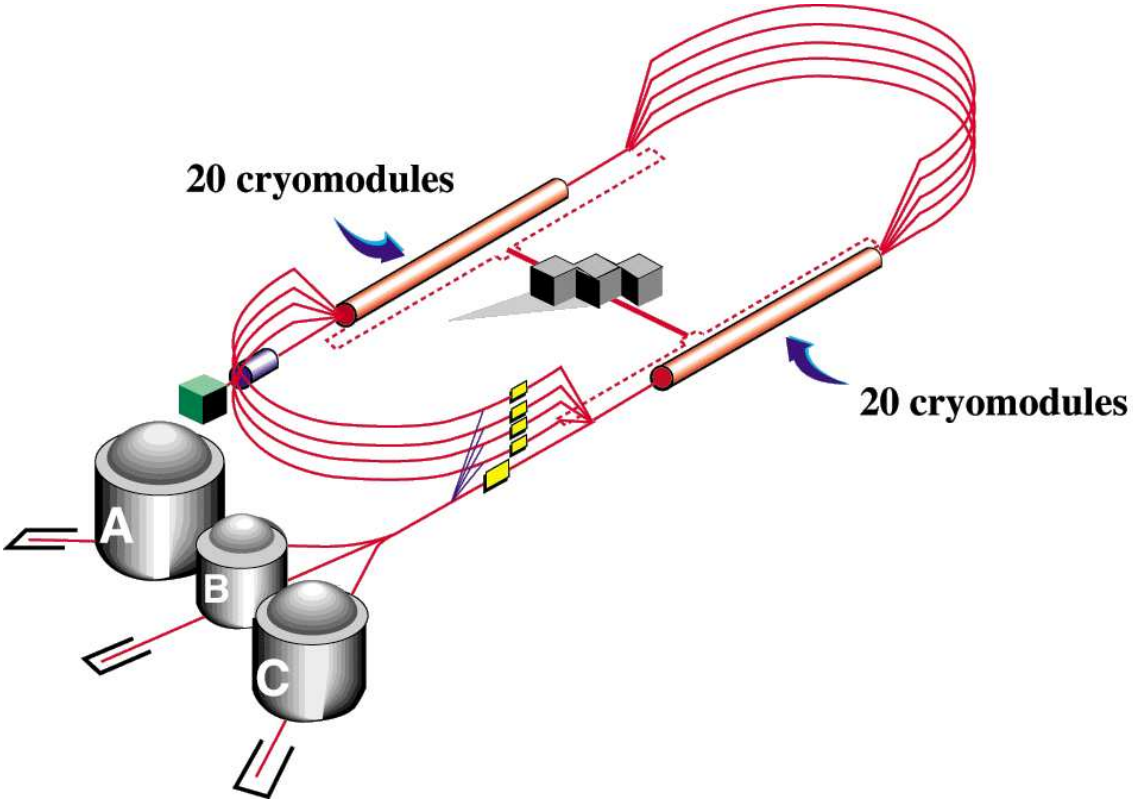


Figure 1.7: Diagram of the CEBAF accelerator used to generate electron beams at Jefferson Lab, where the end stations refer to the experimental halls [34, 35].

electrons are accelerated through five laps of the accelerator, until its desired energy is obtained (currently the maximum energy obtainable is 6 GeV). It is capable of simultaneously delivering the electron beam to three experimental halls - Hall A, Hall B and Hall C - with luminosities of the order of 10^{38} cm^{-2} s^{-1} in Halls A and C. The luminosity of the beam delivered to Hall B is significantly lower due to limits imposed by the detector itself. The collective aims of these three halls are to gain an understanding of QCD in the strong interaction regime and to increase what is known about nuclear structure. Hall B houses the CEBAF Large Acceptance Spectrometer (CLAS), a large acceptance detector with a high solid angle acceptance and high efficiency of detection of multiple charged particles in final states. The detector is used in experiments involving photon and electron beams, and is ideal for probing the nucleon resonance region. It is also the focus of the CLAS Collaboration, an international group of researchers involved in experiments using

the CLAS spectrometer in Hall B. The N* programme at CLAS consists of a wide range of experiments that aim to explore the excited nucleon spectrum and learn about nucleon resonant states [36–40]. Many experiments are based on pseudoscalar meson photoproduction reactions, described in Section 1.3. This type of experiment often requires one or more of the target or photon beam to be polarised in some way. Photon beams are either linearly polarised through the process of coherent bremsstrahlung [41], or circularly polarised through scattering of longitudinally polarised electrons. Coherent bremsstrahlung is a process in which a crystal radiator is used to coherently scatter electrons. Circularly polarised photon beams are produced through bremsstrahlung with an amorphous radiator. Polarisation of a target can also be done at Jefferson Lab by subjecting a target to cryogenic temperatures and a strong magnetic field, as discussed in [42, 43].
Currently, the CEBAF accelerator is undergoing an upgrade that will increase the maximum electron beam energy to 12 GeV, and as a consequence the CLAS detector is being rebuilt in order to accommodate the increased beam energies [44]. This work will enable the exploration of the nucleon spectrum in the region where the dominant degree of freedom is due to dressed quarks [45].

Although research from Jefferson Lab has made a significant contribution to the world data sets, other international facilities such as Mainz [46, 47] and Bonn [48] also pursue similar research goals and often perform complementary experiments.

1.3 Pseudoscalar Meson Photoproduction

A pseudoscalar meson is a meson with 0 spin and odd parity, such as the pion. The 0-spin property of pseudoscalar mesons makes them particularly useful in photoproduction reactions as it limits the degrees of freedom. The nonet of pseudoscalar mesons is shown in Figure 1.8.

Pseudoscalar meson photoproduction reactions can be a more sensitive alternative to pion-nucleon scattering experiments in the determination of nucleon resonances because of the different polarisation configurations available. Resonant states are found through the extraction of polarisation observables, which are related to

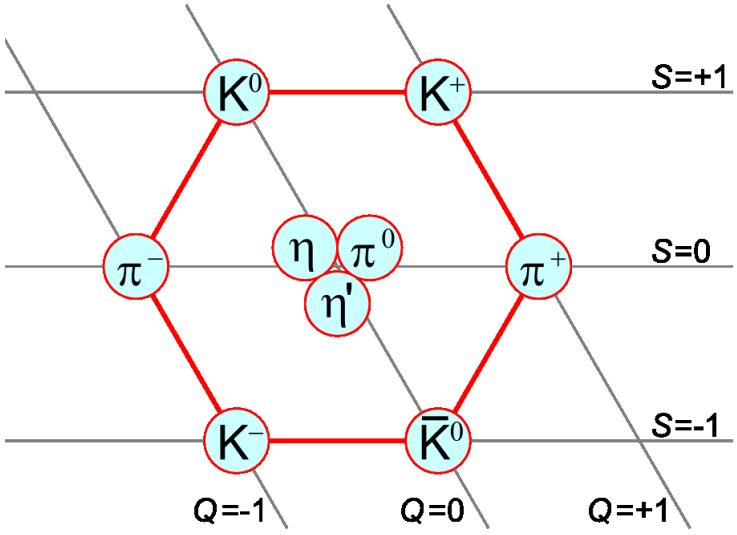


Figure 1.8: Nonet of pseudoscalar mesons.

the polarisations of the beam, target and recoiling particle in the experiment. Such reactions are often used to analyse strangeness reaction channels, where the recoiling particle is a hyperon - a baryon made up of three light quarks in which at least one quark is a strange quark. Hyperons, such as Λ and Σ^0, are intermediate states which decay to a pion and nucleon (and γ, in the case of the Σ^0). Reactions featuring a recoiling hyperon present one significant benefit. The hyperon is said to be "self-analysing" [49, 50] due to its parity-violating weak decay. This implies that the polarisation of the hyperon can be determined by considering the angular distributions of its decay products. A polarimeter device is therefore unnecessary in the detector system.

There are fifteen polarisation observables, referred to as Σ, P, T, E, F, G, H, C_x, C_z, O_x, O_z, T_x, T_z, L_x and L_z, and a differential cross-section term (σ) that are experimentally measurable, and are related to the polarisations of the photon beam, target, and recoiling hyperon. The photon beam can be either linearly polarised (P_L^γ), circularly polarised (P_C^γ), or unpolarised. The target can be longitudinally polarised (P_L^T), transversely polarised (P_T^T), or unpolarised, and information about the recoiling particle (P_x^R, P_y^R, P_z^R) may or may not be available. Different configurations of these polarisations can provide experimental access to different polarisation

observables. The differential cross-section equation of pseudoscalar meson photoproduction is given in Eqn. 1.1, where ϕ indicates the kaon azimuthal angle with respect to the photon polarisation. Once one configuration is chosen, this equation reduces significantly as many terms go to zero. This is discussed in more detail in Section 3.5.

$$\begin{aligned}
\sigma = \sigma_0\{1 &- P_L^\gamma P_T^T P_y^R \sin(\phi)\cos(2\phi) + \Sigma(-P_L^\gamma \cos(2\phi) + P_T^T P_y^R \sin(\phi)) \\
&+ T(P_T^T \sin(\phi) - P_L^\gamma P_y^R \cos(2\phi)) + P(P_y^R - P_L^\gamma P_T^T \sin(\phi)\cos(2\phi)) \\
&+ E(-P_C^\gamma P_L^T + P_L^\gamma P_T^T P_y^R \cos(\phi)\sin(2\phi)) + F(P_C^\gamma P_T^T \cos(\phi) + P_L^\gamma P_L^T P_y^R \sin(2\phi)) \\
&- G(P_L^\gamma P_L^T \sin(2\phi) + P_C^\gamma P_T^T P_y^R \cos(\phi)) - H(P_L^\gamma P_T^T \cos(\phi)\sin(2\phi) - P_C^\gamma P_L^T P_y^R) \\
&- C_x(P_C^\gamma P_x^R - P_L^\gamma P_T^T P_z^R \sin(\phi)\sin(2\phi)) - C_z(P_C^\gamma P_z^R + P_L^\gamma P_T^T P_x^R \sin(\phi)\sin(2\phi)) \\
&- O_x(P_L^\gamma P_x^R \sin(2\phi) + P_C^\gamma P_T^T P_z^R \sin(\phi)) - O_z(P_L^\gamma P_z^R \sin(2\phi) - P_C^\gamma P_T^T P_x^R \sin(\phi)) \\
&+ L_x(P_L^T P_x^R + P_L^\gamma P_T^T P_z^R \cos(\phi)\cos(2\phi)) + L_z(P_L^T P_z^R - P_L^\gamma P_T^T P_x^R \cos(\phi)\cos(2\phi)) \\
&+ T_x(P_T^T P_x^R \cos(\phi) - P_L^\gamma P_L^T P_z^R \cos(2\phi)) + T_z(P_T^T P_z^R \cos(\phi) + P_L^\gamma P_L^T P_x^R \cos(2\phi))\}
\end{aligned} \tag{1.1}$$

Pseudoscalar meson photoproduction can be completely described by four amplitudes. These amplitudes are dependent on variables such as energy and scattering angle and were developed by Chew, Goldberger, Low and Nambu (CGLN) [51]. Whilst CGLN amplitudes are useful in other analyses [52], more convenient representations such as helicity and transversity amplitudes are more common in strangeness photoproduction. Helicity and transversity amplitudes are not dependent on energy or scattering angle. Therefore, analyses that use these amplitudes are applied using data that fall into a relatively narrow range of these variables as it is assumed that the terms in Equation 1.1 do not vary much within this range. The available data is sorted into these ranges prior to analysis. These amplitude representations differ primarily in the choice of coordinate system. Amplitudes calculated in the helicity basis reflect the helicity states of the photon and nucleon [53], where components are defined in relation to the spin of each particle. Transversity amplitudes are defined in the basis describing the plane transverse to the scattering plane [54]. Although helicity amplitudes have, in the past, dominated most of the literature,

the work presented in this thesis will be based upon the transversity representation, shown below [54]. Transversity amplitudes can be seen to have an advantage over helicity amplitudes when single-polarisation data is more numerous than double-polarisation data, as the measurements of single-polarisation observables provides the moduli of the transversity amplitudes. The transversity amplitudes, b_1, b_2, b_3 and b_4 are defined in Eqns. 1.2 - 1.5.

$$b_1 = \langle +|M|+\rangle_{\perp} \tag{1.2}$$

$$b_2 = \langle -|M|+\rangle_{\perp} \tag{1.3}$$

$$b_3 = \langle +|M|-\rangle_{\parallel} \tag{1.4}$$

$$b_4 = \langle -|M|+\rangle_{\parallel} \tag{1.5}$$

where M is the transition matrix that describes the transition from the reaction's initial to final state.

Due to the behaviour of quantum states, only the moduli and relative phases are accessible, meaning that there are only seven degrees of freedom [55]. This can be further reduced to six degrees of freedom by normalising these amplitudes, as shown in Eqn. 1.6.

$$a_i = \frac{b_i}{\sqrt{|b_1|^2 + |b_2|^2 + |b_3|^2 + |b_4|^2}} \tag{1.6}$$

Table 1.2: Table of polarisation observables with their amplitude representations, where S denotes single-polarisation observables, BT refers to beam-target double polarisation observables, BR to beam-recoil double polarisation observabes and TR to target-recoil double polarisation observables [54–56].

Type	Observable	Transversity representation	Helicity representation
S	σ	$\|a_1\|^2 + \|a_2\|^2 + \|a_3\|^2 + \|a_4\|^2$	$\|h_1\|^2 + \|h_2\|^2 + \|h_3\|^2 + \|h_4\|^2$
	Σ	$\|a_1\|^2 + \|a_2\|^2 - \|a_3\|^2 - \|a_4\|^2$	$2\Re(h_1 h_4^* - h_2 h_3^*)$
	P	$\|a_1\|^2 - \|a_2\|^2 + \|a_3\|^2 - \|a_4\|^2$	$2\Im(h_1 h_3^* + h_2 h_4^*)$
	T	$\|a_1\|^2 - \|a_2\|^2 - \|a_3\|^2 + \|a_4\|^2$	$2\Im(h_1 h_3^* + h_2 h_4^*)$
BT	E	$2\Re(a_1 a_3^* + a_2 a_4^*)$	$\|h_1\|^2 - \|h_2\|^2 + \|h_3\|^2 - \|h_4\|^2$
	F	$2\Im(a_1 a_3^* - a_2 a_4^*)$	$2\Re(h_1 h_2^* + h_3 h_4^*)$
	G	$2\Im(a_1 a_3^* + a_2 a_4^*)$	$-2\Im(h_1 h_4^* + h_2 h_3^*)$
	H	$-2\Re(a_1 a_3^* - a_2 a_4^*)$	$-2\Im(h_1 h_3^* - h_2 h_4^*)$
BR	C_x	$-2\Im(a_1 a_4^* - a_2 a_3^*)$	$2\Re(h_1 h_3^* + h_2 h_4^*)$
	C_z	$2\Re(a_1 a_4^* + a_2 a_3^*)$	$\|h_1\|^2 + \|h_2\|^2 - \|h_3\|^2 - \|h_4\|^2$
	O_x	$2\Re(a_1 a_4^* - a_2 a_3^*)$	$-2\Im(h_1 h_2^* - h_3 h_4^*)$
	O_z	$2\Im(a_1 a_4^* + a_2 a_3^*)$	$2\Im(h_1 h_4^* - h_2 h_3^*)$
TR	T_x	$2\Re(a_1 a_2^* - a_3 a_4^*)$	$-2\Re(h_1 h_4^* + h_2 h_3^*)$
	T_z	$2\Im(a_1 a_2^* - a_3 a_4^*)$	$-2\Re(h_1 h_2^* - h_3 h_4^*)$
	L_x	$-2\Im(a_1 a_2^* + a_3 a_4^*)$	$2\Re(h_1 h_3^* - h_2 h_4^*)$
	L_z	$2\Re(a_1 a_2^* + a_3 a_4^*)$	$\|h_1\|^2 - \|h_2\|^2 - \|h_3\|^2 + \|h_4\|^2$

The amplitudes can be accessed experimentally through fifteen polarisation observables and a cross-section. These observables are bilinear combinations of the amplitudes and are therefore correlated. Two examples of such relations are listed below in Equations 1.7 and 1.8 [56, 57].

$$C_x^2 + C_z^2 + O_x^2 + O_z^2 = 1 + T^2 - P^2 - \Sigma^2 \tag{1.7}$$

$$FG - EH = P - \Sigma T \tag{1.8}$$

By polarising the beam and target in different ways, it is possible to obtain information on different polarisation observables. In order to determine the underlying

amplitudes, a "complete measurement" is required [56, 58]. Such a measurement requires that at least eight observables, including one from each double-polarisation category, are measured. It has been argued, however, that this definition of a complete measurement assumes perfectly precise measurements, and practically, additional measurements are required to resolve ambiguities arising from imperfect experimentation [55].

1.3.1 The $\gamma p \rightarrow K^+\Lambda$ reaction

The reaction of interest in this work involves a polarised photon beam and an unpolarised stationary proton target, as illustrated in Fig. 1.9, where a Λ hyperon is an intermediate state.

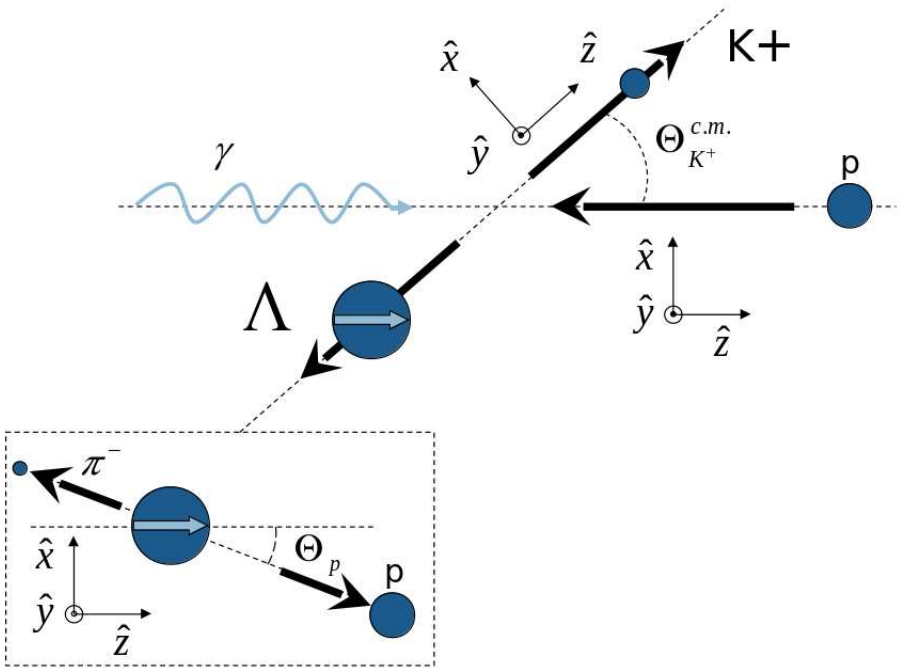


Figure 1.9: Diagram of a pseudoscalar meson photoproduction reaction where a photon beam is incident on a stationary proton target. A kaon and Λ are produced. The Λ then decays and a pion and proton are produced [59].

The Λ hyperon has a mass of $1115.68\,MeV/c^2$ and mean lifetime of $2.6 \times 10^{-10}s$. It decays preferentially to a proton and pion ($\Lambda \rightarrow p\pi^-$) with a branching ratio of 63.8%, and to a neutron and pion ($\Lambda \rightarrow n\pi^0$) with a branching ratio of 35.8%.

As described earlier, the self-analysing characteristic of the Λ is advantageous and enables the use of recoil information in analyses without the need for additional hardware. The recoil information is determined through the detection of the decay products, the proton and pion.
In this thesis, data from two experimental configurations will be analysed, both of which involve a polarised photon beam and an unpolarised target. One set-up (known as g8b) involves a linearly polarised beam, whereas the other (known as g1c) uses circular polarisation of the beam. Both of these experiments were run at Jefferson Lab and will be discussed further in Chapter 4.

1.4 Outline

The work in this thesis is concentrated on the development and testing of a new analysis method. A Bayesian approach, called *nested sampling* [60], is applied to pseudoscalar meson photoproduction data taken by the CLAS Collaboration at Jefferson Lab. This analysis method is based on an event-by-event likelihood function. Chapter 2 will discuss the underlying statistical formalism involved in nested sampling. The nested sampling algorithm will be featured, along with a description of Markov chain Monte Carlo, which also features in the analysis method. The third chapter will focus on the development of the software programme. The current methods of analysis will be discussed, followed by the derivations of likelihood functions. Testing of the analysis programme using simulated data from an event generator will then be shown, and results are compared to those obtained through other analysis methods. Chapter 4 features the analysis of real data, taken from the g8b and g1c experiments at CLAS. Results from the nested sampling analysis are compared to published results, and comparisons between both experiments are performed. One significant benefit of a nested sampling analysis programme is illustrated in Chapter 5, where data from both experiments is combined and analysed simultaneously. The results from this combined analysis are compared to those from the separate analyses and plots showing new information are shown. Additional plots from these analyses are included in Appendix B. The summary and outlook are given in Chap-

ter 6. Appendix A features the computing theory behind data parallelism and GPU programming, a tool which was employed in Chapter 3.

2

Inference

Data analysis in the field of baryon spectroscopy usually stems from the *frequentist* branch of statistics. This approach assumes that there exist underlying fixed statistical parameters and that the measured data are sampled randomly from all independent and identically distributed values. Bayesian statistics is founded on the assumption that the data are fixed values, and parameters such as the mean are variable and are updated as more information is observed. These parameters are not considered to be fixed, rather they are expressed in terms of probabilities. There are benefits of using a Bayesian statistics framework. Thomas Bayes' mathematical theorem, Bayes' Theorem, was posthumously put forth to the Royal Society in 1763 [61], but was largely ignored until it was used to develop Laplace's famous works [62]. The introduction of Bayesian statistics to a frequentist community was slow and fraught with controversy. It meant a complete paradigm shift for statisticians, one which proved to take time. In recent years, however, the theorem has

found its place at the heart of many statistical and analytical applications in modern science, mathematics and computing.
Both statistical approaches have their advantages and disadvantages - one approach cannot be said to be "better" than the other. They are simply two different interpretations of a statistical problem. In some cases, the conclusions reached through one method may be more useful than the other, but this depends on the suitability of each approach on the given problem and the desired information.

In this chapter, the fundamental concepts of statistics and inference relevant to this work will be discussed. In Section 2.1, the ideas that form the backbone of Bayesian statistics will be explored and compared to frequentism. Section 2.2 will focus on the use of Markov Chain Monte Carlo methods and their role in data analysis. A detailed description of the nested sampling algorithm that is integral to this work will be given in Section 2.3.

2.1 Bayesian Formalism

Many current analysis tools rely on frequentist statistics. Frequentism was heavily influenced by the likes of Neyman, who introduced the idea of a confidence interval in the 1930s [63], and Fisher, who developed the maximum likelihood estimation method and was heavily critical of Bayesian statistics [64]. In frequentist statistics, it is assumed that the statistical properties of a distribution, such as the mean and variance, are fixed values, and the data can be repeatedly and randomly sampled [65]. High numbers of data ensure values that vary only very slightly, supporting the idea that the mean of a distribution is constant. The underlying assumption is that the randomly sampled data accurately reflects reality, an assumption that is often reliable [66]. There are, however, some criticisms. It has been argued that repeating measurements and trials under identical conditions is impossible. If all conditions were exactly the same from one measurement to another, there would be no difference in the outcome. This leads to scepticism in the interpretation of statistical parameters, such as errors. The mean squared error is the basis on which different

statistical methods are compared, and is considered to represent the average squared error expected in repeated measurements taken under identical circumstances. It is argued that as replicating experiments or measurements under completely unchanged conditions is impossible, this interpretation is unrealistic [66].
Under the frequentist paradigm, the probability of an event A occurring is based on its relative frequency in large numbers of trials. For example, rolling an unweighted 6-sided die 100 times would result in a '4' approximately 17 times. The frequency of rolling a '4' would then be 17/100, or 0.17 ($\approx \frac{1}{6}$). The probability of then rolling a '4' on a subsequent roll would be described as $\Pr(4) = \frac{1}{6}$. This approach does not require any knowledge or expectation of the results, and this is the key difference between frequentist and Bayesian statistics [67].

In the Bayesian approach, properties like the mean or standard deviation are seen as variables that describe some distribution, and the data are treated as fixed values. An initial 'guess', regarded as the *prior* distribution, is required and the choice of prior can change the nature of the process of estimation [65]. It can be considered as an estimate or expectation of the results prior to any calculation or statistical analysis. It can also be used to introduce physical constraints. The use of a prior is one of the reasons that Bayesian analysis has been scrutinised for so many years, and why it continues to cause controversy. The choice of prior does indeed affect the estimation process, but a poor choice of prior that still includes all physically allowed values for the variable(s) being estimated will still arrive at the same conclusion, although it may take significantly longer to reach a convergent result [62], which is known as the *posterior* distribution, or simply the posterior . The statistical quantities that define the posterior are then used to describe the results of the analysis [62].

It would be helpful at this point to introduce an example where the frequentist and Bayesian approaches can be compared. The Λ hyperon has a well-measured mass of 1.1157 GeV/c^2. In this example, we want to determine the mass of the Λ hyperon from a simple distribution. Let us assume that an experiment was performed that

measured the mass of the hyperon. Usually, the mass distribution is Lorentzian, but for the sake of simplicity it will be approximated by a Gaussian. Also, for the sake of simplicity, this will be treated as a one-dimensional problem by assuming that the width of the distribution is known. Using the more traditional frequentist analysis, the measured data points are sorted into bins and a frequency histogram is produced, as shown in Figure 2.1. A pre-defined fit function (given in Equation 2.1.1) is applied, using the χ^2 fitting method. The fit is defined over a specified range. The range is explored by calculating the derivatives of the fit function and moving towards a minimum.

$$y = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(\frac{(x-\mu)}{\sigma})^2\right) \tag{2.1.1}$$

where μ is the mean of the distribution, σ is the width and c is a constant.

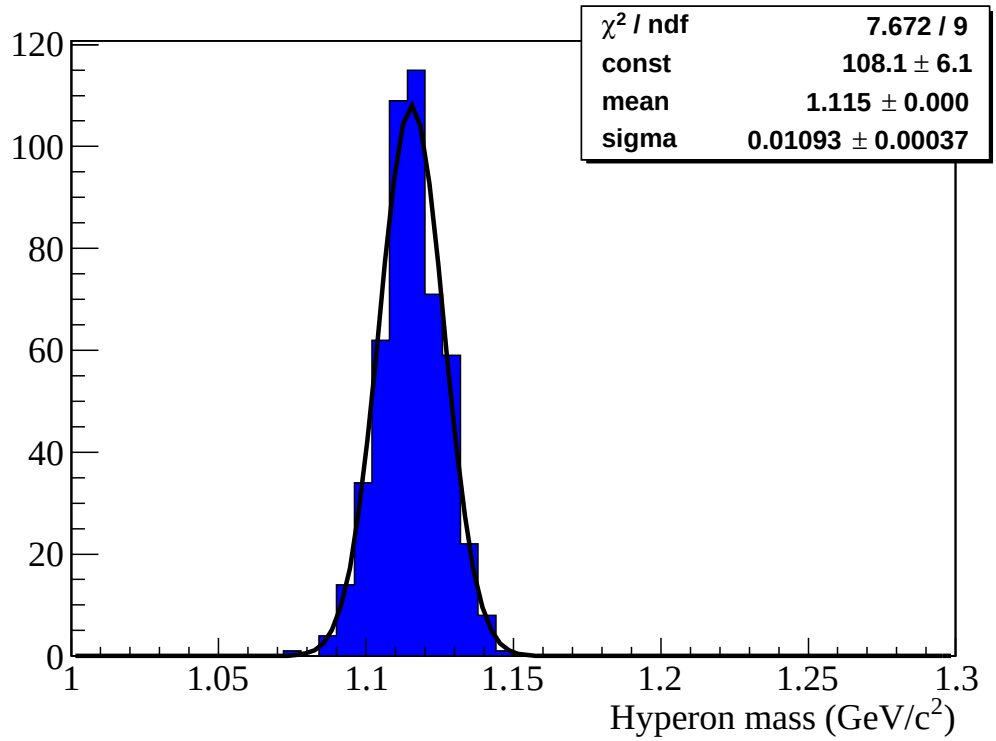


Figure 2.1: Histogram of hyperon mass with a χ^2 fit.

The resulting mass is then expressed as the mean mass with an error, $M \pm \delta M$.

Using a Bayesian approach, there are several things that need to be defined. First, the prior distribution must be stated. In this problem, it has been assumed that the mass must be within the region [1.0, 1.3]. The most intuitive choice of prior would then be a uniform distribution on this region. This type of prior is known as an *uninformed* prior. The likelihood function used in this case is based on the probability density function of a normal distribution, shown in Equation 2.1.2.

$$\mathcal{L}(\mu) = \sum_i \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{(2\sigma^2)}} \tag{2.1.2}$$

where μ is the mass, σ is the width (constant in this case) and x_i denote the data points.

For each datum measured, the posterior probability distribution function (pdf) is updated, as illustrated in Figure 2.2. As more data points are measured, the distribution converges on a mean mass value (in this case, a value of 1.1157).

In contrast to the result obtained using a binned fitting approach where the mean mass and an associated error are found ($M \pm \delta M$), the hyperon mass found using this method is expressed in the form of a probability distribution.

The posterior is related to the prior and the data via Bayes' Theorem, shown in Equation 2.1.3 [62].

$$\Pr(X|Y,I) = \frac{\Pr(Y|X,I) \times \Pr(X|I)}{\Pr(Y|I)} \tag{2.1.3}$$

where I denotes any background information, X represents the hypothesis and Y is the data. $\Pr(X|Y,I)$ refers to the posterior, or the probability of the hypothesis given the data and other information. $\Pr(X|I)$ describes the prior distribution, and $\Pr(Y|X,I)$ is the likelihood, or the probability of the data given the hypothesis. The term $\Pr(Y|I)$ denotes the probability of the data given the background information. It can generally be considered as a constant of proportionality or normalisation constant.

The basic premise of Bayesian statistics can be expressed in the form of a simple expression [60]:

$$\textit{Prior} \times \textit{Likelihood} \Rightarrow \textit{Evidence} \times \textit{Posterior} \tag{2.1.4}$$

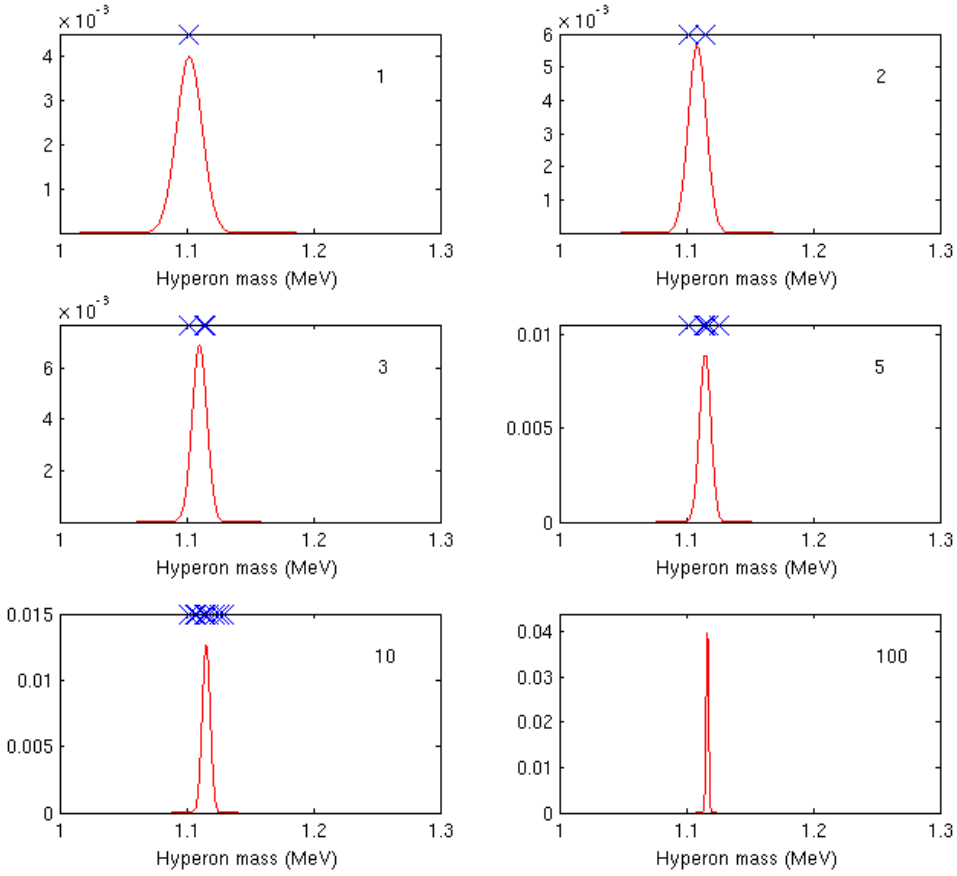


Figure 2.2: Evolution of the posterior pdf in the experiment of measuring Λ hyperon mass. In each plot, the number of events used is indicated in the top right corner. The values of these data points are shown as crosses along the top border of the plot (omitted in the last plot).

or, in mathematical terms,

$$Prior\,distribution = \pi(\theta)d\theta \tag{2.1.5}$$

$$Likelihood\,function = L(\theta) \tag{2.1.6}$$

$$Evidence = Z = \int L(\theta)\,\pi(\theta)d\theta \tag{2.1.7}$$

$$Posterior\,distribution = p(\theta)d\theta \tag{2.1.8}$$

where θ represents the variable(s) being obtained through the analysis.

The output of a Bayesian calculation contains two parts: the evidence and the posterior. The evidence, Z, is useful in comparing different statistical models. This evidence term is not accessible in many Bayesian analysis methods. It is one of the main distinguishing features of the nested sampling algorithm, described in detail in Section 2.3. Bayes factors (K), or ratios of evidence, are used to compare any two models X_1 and X_2 at any time without the need for any recalculations [60].

$$K = \frac{Z_1}{Z_2} = \frac{\int \Pr(X_1)\ \Pr(Y|X_1 dX_1)}{\int \Pr(X_2)\ \Pr(Y|X_2 dX_2)} \tag{2.1.9}$$

A K value of less than one indicates that the data supports model 2. A K value greater than one suggests that model 1 is more likely, and the strength of this support is related to the value of K. A K value significantly greater than 1 indicates that model 1 is substantially more likely.

The evidence can be described as a term that relates to the region where the posterior is most populated. This region is situated within the prior distribution, and therefore comprises some fraction of the prior. This fraction is defined as $\exp^{-H}$, where H denotes the *information*, otherwise known as the negative entropy [60].

$$H = \int log(\frac{p(\theta)}{\pi(\theta)})p(\theta)d\theta \tag{2.1.10}$$

It is essentially a measure of the proportion of the prior distribution that contains the majority of the posterior [60]. This number can range over several orders

of magnitude. The value of H can be in the thousands in cases where the likelihood is dominant in an unusual area of the prior, such as in the region of boundaries.

2.2 Markov Chain Monte Carlo

In many frequentist analysis methods, a region is explored by calculating the derivative of a fit function and moving towards a minimum. Often, however, it is useful to use an alternative method of exploration, especially in cases where a function cannot be easily differentiated. A popular alternative (especially in multivariate problems) is the use of a random walk such as Markov chain Monte Carlo, or MCMC. This is a method usually used to generate random samples from a probability distribution. In recent years, there has been a significant increase in the use of MCMC, especially in Bayesian applications [68]. Most often, it is used to sample a posterior distribution to determine its statistical properties. MCMC is particularly useful in multi-dimensional problems as it is straightforward to explore a multi-dimensional space.

A Markov chain is a sequence of random states where the value of the next state is dependent solely on the value of the current state (and not on the previous values in the chain) [62, 68]. Let $X = X_0, X_1, X_2 ...$ be a sequence of random variables, sampled from a distribution $\pi(\cdot)$. Then X is a Markov chain if the value X_{n+1} is determined from sampling a distribution only dependent on X_n. An example of such a walk is shown in Figure 2.3. From a computing standpoint, it is generally straightforward to construct a Markov chain algorithm. There are many such algorithms, the most commonly used being the Metropolis-Hastings algorithm.

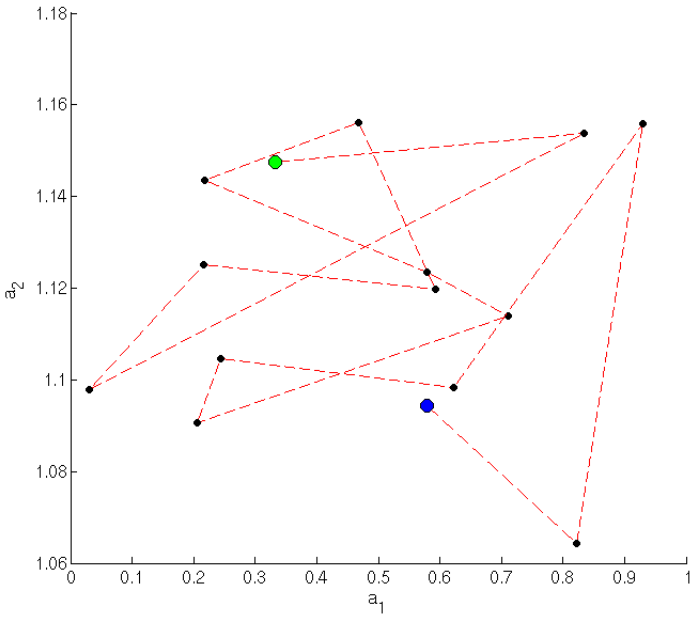


Figure 2.3: Diagram of a random walk involving two parameters, a_1 and a_2. The initial point is indicated with a green marker, and the final point is shown with a blue marker.

The Metropolis-Hastings is an extension of the Metropolis algorithm [69], and was developed by Nicholas Metropolis *et al* in 1953 [70]. It is a random-walk Markov chain method used extensively in physics. The algorithm itself is easy to implement and adheres to the following basic structure [69, 71]:

Select a sample θ_0 randomly from the starting distribution, $\pi(\theta)$.

For a given number of iterations ($t = 1, ..., m$), repeat:

Select a sample randomly from the proposal distribution, $q(\theta^{(t)} \mid \theta^{(t-1)})$.

Calculate the likelihood of this new sample.

Determine the joint posterior distribution, $h(\theta^*) = \mathcal{L}(I|\theta)\pi(\theta)$,

where $\mathcal{L}(I|\theta)$ is the likelihood of the data (I) given observable(s), θ.

Calculate the acceptance ratio, $r = \frac{h(\theta^*)q(\theta^{(t-1)}|\theta^*)}{h(\theta^{(t-1)})q(\theta^*|\theta^{(t-1)})}$.

Accept the new value of θ if $r \geq 1$.

If $r < 1$, the new value of θ is set as θ^* with a probability of r, and $\theta^{(t-1)}$ with a probability of $(1 - r)$.

The use of m steps in the algorithm ensures that the final value of the sample will be sufficiently different from its initial value.

MCMC is frequently used to infer statistical quantities of an observable (or observables) by drawing samples from a distribution [68,72]. These samples are selected through the use of a Markov chain. Nested sampling is a special case, where MCMC is applied to the prior distribution, rather than to the posterior [62]. In the nested sampling algorithm (discussed in Section 2.3), a prior distribution is explored by selecting a sample to use as an initial state for a Markov chain and evolving it using the Metropolis-Hastings algorithm [60]. Here, MCMC is being used in a capacity almost reverse to its more common purpose.

2.3 Nested Sampling

The nested sampling algorithm is a recently developed method in statistical Bayesian analysis. It was developed by Skilling in 2004 [73] and has since been introduced into a wide range of scientific and mathematical disciplines. The algorithm is unique in that it provides access to both pieces of the Bayesian output: the posterior, which is usually obtained through Bayesian analysis, and the evidence, Z (described in Section 2.1). It is commonly used as a modern model comparison technique. Whilst the determination of evidence is the key feature of this method, the posterior distribution has been the outcome of significance in this project.
The programme itself can be broken up into several functional blocks, as shown in Figure 2.4.

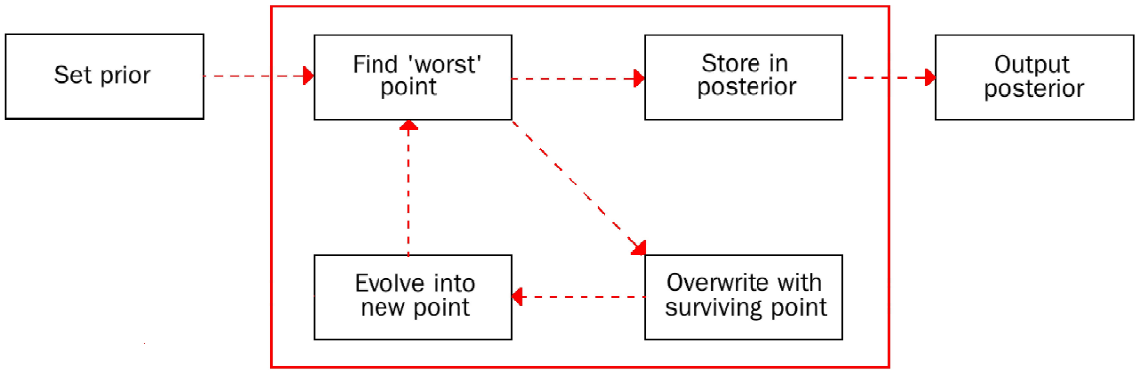


Figure 2.4: The nested sampling algorithm. The loop continues to find the least likely (worst) point until some termination condition is met, at which point the resulting posterior is output.

The algorithm is generic and can be applied to a wide range of problems, but some of the components are problem-specific [60]. The prior is one such component. For a particular problem, a prior distribution that covers the full range of possible values must be established. When nothing is known about the variables in question, a uniform prior distribution is often used. The likelihood function is the heart of the analysis method. The experimental data and the function itself are used to calculate the likelihood of any point in the prior. Once the prior has been defined, a set of sample points are selected and their associated likelihood values are determined. At this point, the main loop of the nested sampling algorithm is entered and more generic aspects of the programme are used [60, 62].

Once in the main event loop, the sample point with the lowest likelihood is found and stored in the posterior distribution. The sample point is then overwritten with a copy of a randomly selected surviving sample and is evolved into a new, distinct point. The evidence, Z, and information, H, are updated based on the value of the overwritten point. The loop continues until some termination condition is met. In the simplest case, the termination condition can be a pre-set number of iterations. For a more robust rule, it is possible to use the information, H, to determine when the programme should end.

Such a generic termination condition has been derived by Skilling [60]. The principle is to end the programme when the current likelihood no longer increases by more than some small amount. It has been shown [60] that an equivalent termination condition could be determined:

"continue iterating until the count i significantly exceeds NH" [60]

where N represents the number of iterations that have so far been performed. This termination condition is useful, but is still a matter for the user's judgement. The expression "significantly exceeds" is qualitative and requires an interpretation by the user.

In many likelihood-based analysis methods, problems can arise in cases where the likelihood function contains discontinuities, plateaus or other such features. This is not the case in nested sampling - the algorithm is immune to such problems. There is no need to differentiate the likelihood function, and therefore any features of the function that make differentiation difficult or impossible are completely avoided by this approach [60].

It may be helpful to explore the use of nested sampling in a simple example. Here, the problem is to find the x-coordinate of an object. In order to keep the example as trivial as possible, the object's position is defined by a function, which has been translated into a likelihood function (Equation 2.3.11). As such, there is no need to introduce data points. For this problem, it has been assumed that the object's position along the x-axis lies between $x = 0$ and $x = 5$. An intuitive choice of prior be a uniform distribution covering all possible x-coordinates.

$$L(x) = 6x - x^2 \tag{2.3.11}$$

From the prior distribution, a set number of samples are selected (in this case, six samples were chosen). Each sample's value is then used to calculate its associated likelihood value using the likelihood function given (and shown in Figure 2.5).

The likelihood value associated with each sample describes the probability of that sample point being the true x-coordinate of the object. The sample with the

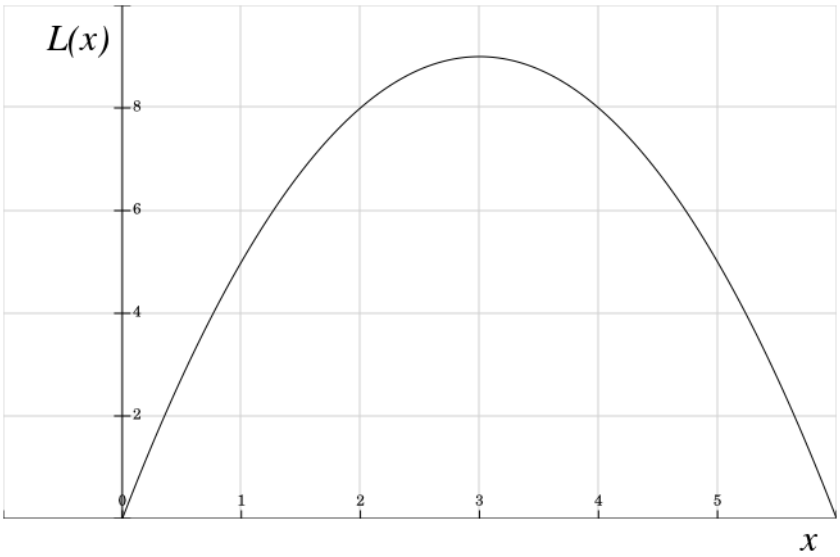


Figure 2.5: Plot of quadratic likelihood function, with a maximum at $x = 3$.

lowest likelihood is found (i.e. the worst sample), and is stored in the posterior. All values associated with the sample, including the value of the sample point itself (its x-coordinate) and its associated likelihood, are all stored in the algorithm's output. The worst sample is then overwritten with a copy of a randomly selected surviving sample point. This copy is evolved in order to obtain a distinct new point, usually through the use of a random walk, such as a Markov chain Monte Carlo. The likelihood of this new point is calculated and compared to that of the overwritten point. It must be greater than the likelihood of the worst sample, and this condition is ensured by the random walk algorithm.

The diagram in Figure 2.6 shows the process of the nested sampling algorithm visually. After the first iteration, the sample point with the lowest likelihood value is, in essence, moved to a region of the prior that is more likely. The next iteration moves the next-least-likely point to a region of higher likelihood, and so on, until a cluster forms around one value. This value is then deemed to be the most likely x-coordinate of the object.

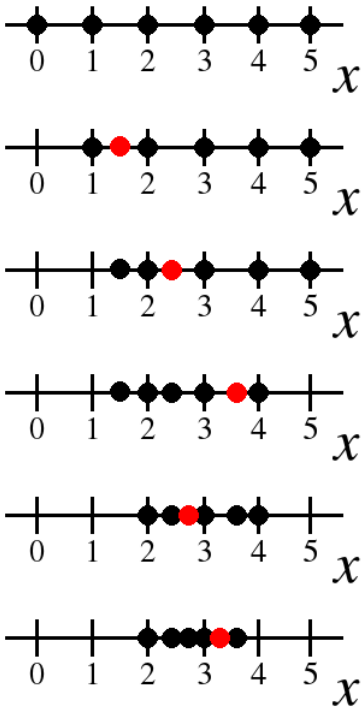


Figure 2.6: Diagram of the iterative nature of the nested sampling algorithm. Red points indicate the new point at each iteration. After each iteration, the values approach $x = 3$ - the position of the object.

2.4 Summary

Bayesian inference is an alternative paradigm to the more conventional frequentist statistics. Despite its slow and gradual emergence into mainstream statistics, the Bayesian approach has become widely used in a range of applications, including physics. The main difference between Bayesianism and frequentism is the way in which data and statistical properties are regarded. In the more familiar frequentist approach, values such as the mean and standard deviation of a distribution are treated as fixed values, which can be found by randomly sampling data. The fundamental principle is that there exists one true value of the mean (for example), which can be found by repeatedly sampling data. Bayesianism views the data as fixed, and statistical properties are expressed as probabilities. Unlike frequentist statistics, where no previous knowledge or expectation is required, Bayesian methods make use of a prior distribution. The prior distribution is an initial distribution which covers the range of possible values for the variable in question. The prior is often regarded as an initial estimate of the expected result(s). In recent years, there have been significant developments into Bayesian analysis methods, including the nested sampling algorithm introduced by Skilling, and is the focus of this work. Nested sampling works by associating samples from the prior with a likelihood value, determined by a likelihood function. The sample with the lowest likelihood value is overwritten with a copy of another point, chosen at random, and moved in order to obtain a distinctly new point. This movement, or exploration, is usually done through the use of a random walk algorithm, and in this work, a Markov chain Monte Carlo was used. The overwritten point is saved to the posterior distribution, and after many iterations, the posterior shows the region of highest likelihood for the variable being determined. The algorithm is immune to features of the likelihood function that can occasionally result in poor outcomes, such as discontinuities, nondifferentiable functions and plateaus. It is also unique in that it provides access to the evidence term, which can be used to compare statistical models.

3

Development of Method with Simulated Data

Within the CLAS Collaboration and in many other international hadron spectroscopy groups, there is a significant effort in analysing the large amounts of data acquired from pseudoscalar meson photoproduction experiments. The development of new mathematical analysis methods that can maximise the information yield from these experiments has been increasingly explored. Bayesian analysis has become more common across many fields of physics, but had not been applied to the experiments performed at CLAS until very recently. This chapter illustrates some of the more common current analysis methods and fully details the development of a Bayesian analysis programme called *nested sampling* that can be used to increase the amount of information obtained from the same data. Also included in this chapter is the methodology for deriving likelihood functions, which can, in prin-

ciple, also be applied to some non-Bayesian methods of analysis. Work in combining datasets from different experiment types is presented, where observables from multiple experiments can be extracted simultaneously. This is possible by exploiting the fundamental characteristics of the nested sampling algorithm.

3.1 Current methods of analysis

There are several different approaches currently being applied to data from pseudoscalar meson photoproduction experiments in the field of hadron spectroscopy. Initially, polarisation observables were extracted through a binned fitting method. This sort of analysis is quick and simple, and effective for single polarisation observable extraction. There are, however, several disadvantages inherent to binned fitting, and complications arise in cases where double polarisation observables are extracted. More recently, a maximum likelihood (ML) method was explored, which features an event-by-event likelihood function [74]. Both of these analysis techniques, however, share a common problem, in that they both treat polarisation observables as independent parameters and do not inherently include the constraints imposed by observable correlations. As such, both methods result in a loss of the information contained in the observable correlations. There is no guarantee that, when multiple observables are extracted from the same data set, the results will be consistent. Despite this issue, these two methods are the most common currently in use in data analysis of pseudoscalar meson photoproduction. They are both described in this section.

3.1.1 Binned Fitting Approach

Most data analysis used to extract polarisation observables from pseudoscalar meson photoproduction experiments is still performed using a binned χ^2 fit. In this method, and in the simplest case, data from two polarisation settings are used to fill two histograms of equal binning. These histograms are then combined to obtain an asymmetry histogram, using Equation 3.1.1, and a sinusoidal function is fitted. The amplitude of the function then provides access to the beam asymmetry polarisation

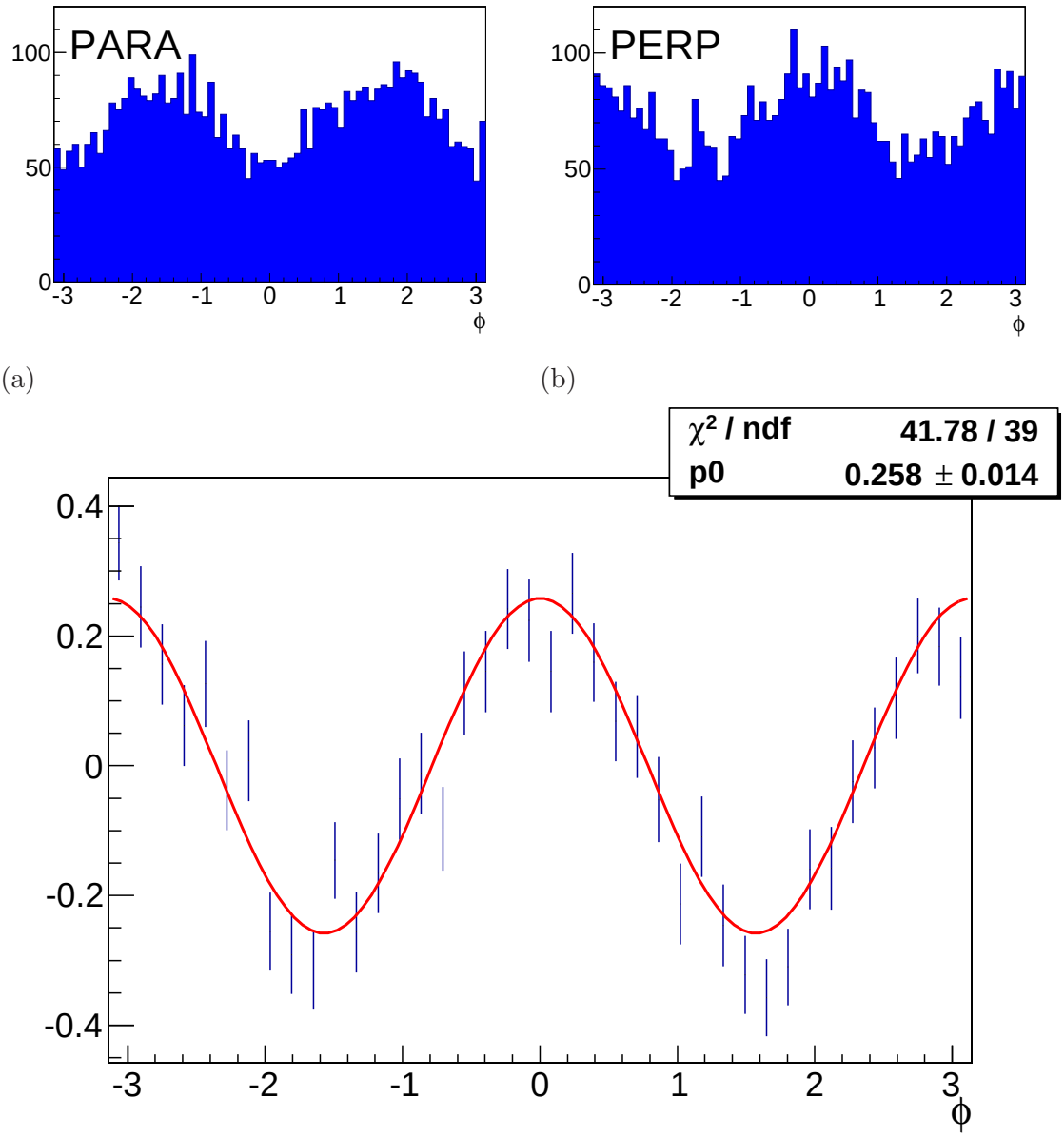
observable, Σ.

$$A_\sigma = \frac{\sigma^\perp - \sigma^\parallel}{\sigma^\perp + \sigma^\parallel} = P_\gamma \Sigma \cos(2\phi) \tag{3.1.1}$$

Figure 3.1 depicts the process used to extract the beam asymmetry, Σ, from simulated data with high statistics. Σ is a single polarisation observable, and as such can be extracted using a simple fit. For double polarisation observables, this method becomes more complicated.

Once Σ has been determined, its value can then be used to find the other observables. For the extraction of O_i, where $i = x$, z, two-dimensional histograms are filled with ϕ, the kaon azimuthal angle, and $\cos\theta_i$, the cosine of polar angle defined by the proton and hyperon rest frame (x and z axes, respectively) for each of the two beam settings [75]. Two-dimensional asymmetry histograms are created (shown in Figure 3.2) and "sliced". Each slice corresponds to a projection of the 2-D asymmetry onto the ϕ-axis, resulting in a 1-D histogram. The recoil polarisation of the hyperon modulates the phase shift of the $\cos(2\phi)$ distribution, and from this shift, the beam-recoil double polarisation observables can be determined. The fit shown in Equation 3.1.2 can then be applied to each slice.

$$A = P_\gamma(\Sigma \cos(2\phi) \; + \; \alpha x(z) O_{x(z)} \sin(2\phi)) \tag{3.1.2}$$



(a) (b)

(c)

Figure 3.1: (a) Distribution of ϕ angles when the beam has parallel polarisation. (b) similarly shows the ϕ angles when the beam is polarised perpendicularly. (c) displays the resulting asymmetry of these two histograms. A $\cos(2\phi)$ curve is fitted (shown in red).

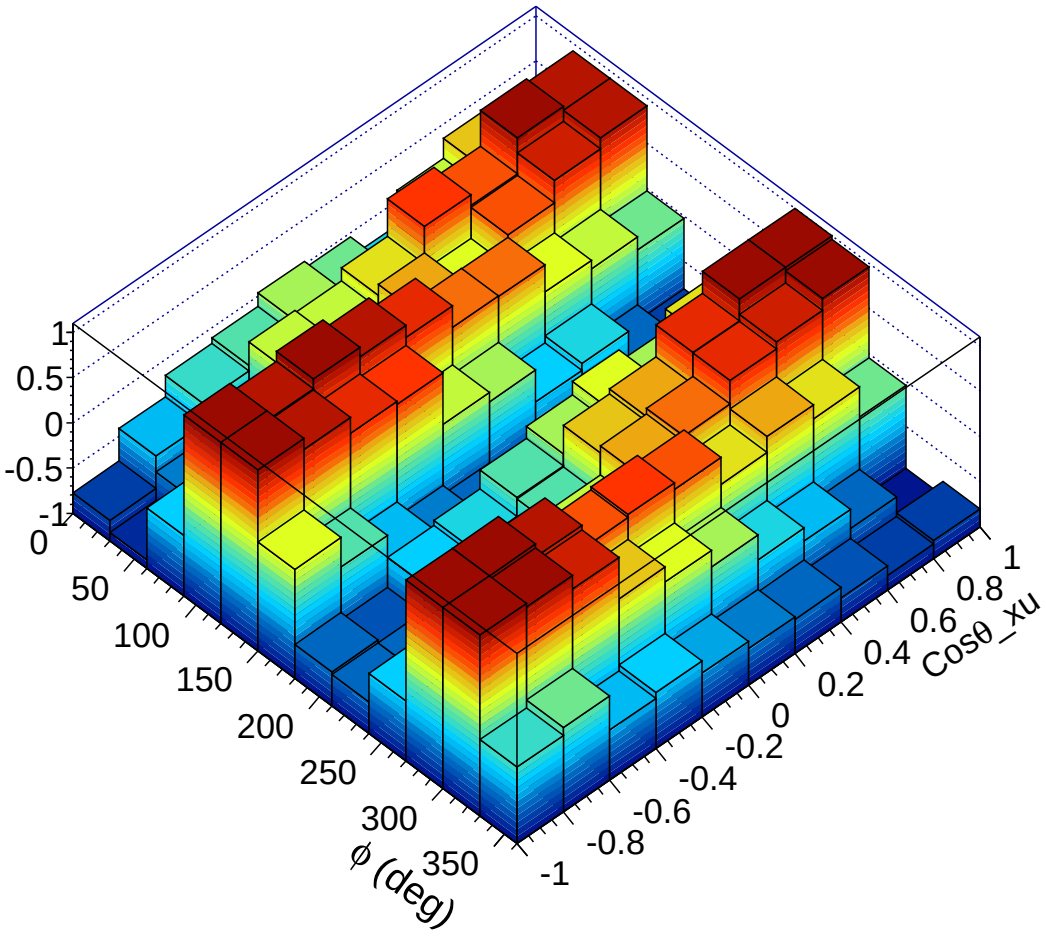


Figure 3.2: Two-dimensional histogram of ϕ and $\cos(\theta_x)$ [76]

The values of parameter p_1 are then used to fill a new histogram, a simple linear fit is applied and the gradient of the fitted line provides the value of the observable.

There are several aspects of this analysis technique that must be mentioned. The involvement of binned histograms inherently results in the loss of information from the data. This is unavoidable when histograms are used, and can be subjective based on the binning choices that are made. Often, a fixed value of a previously measured observable is required to extract others, such as in this case when the extraction of O_x and O_z is dependent on the value of Σ. This can lead to compounded errors and less accurate results. Also, as stated previously, there is no consideration for the correlations between different observables.

3.1.2 Maximum Likelihood

Despite its prevalence in data analysis in many related fields, maximum likelihood techniques in hadron spectroscopy have only recently been introduced. Maximum likelihood estimation relies on an event-by-event likelihood function and therefore overcomes several of the shortcomings of the binned fit.

Maximum likelihood is an estimation technique that determines which observable values make the data most likely [77, 78]. The concept is most clearly explained for the simplest case of a univariable problem. Let $x_1, \ldots, x_N$ be a set of data points, and a be the relevant observable. The maximum likelihood technique then finds the value of a for which Equation 3.1.3 is a maximum [78]:

$$\mathcal{L}(x_1, \ldots, x_N; a) = \prod \Pr(x_i, a) \tag{3.1.3}$$

The concept can easily be extended to cases with many observables. The set of observable values that maximises the likelihood of the data is found, usually through the use of simultaneous equations.

Practically, it is often more convenient to find the maximum of the logarithm of the likelihood. It can also be beneficial to find the position of the maximum by calculating the derivative of the likelihood function, although this is not always possible. Maximum likelihood is a simple and straightforward method of estimation, with invariance under parameter transformation, no information loss from binning, and a suitability for problems with multiple observables. Despite these benefits, however, there are some disadvantages to the maximum likelihood approach. For small datasets, there can be a bias imposed on the results which can cause the outcome to be consistently above or below the true value (although this can be remedied by normalisation) [78]. This bias is reduced as the size of the dataset increases.

In nuclear and particle physics, a minimisation package included in the ROOT framework called TMinuit [79] is often used to minimise the negative of the likelihood (or log(likelihood)) function. It is commonly used in χ^2 fitting as well. The software package is widely used, but can exacerbate some undesirable features of the maximum likelihood method. In most uses of TMinuit, there is a heavy dependence on the first derivatives. This can, in some cases, prove to be a problem when the

function cannot easily be differentiated algebraically. It is also possible that using TMinuit will result in the position of a local maximum, which can be problematic for multimodal likelihood functions.

3.2 Nested Sampling Algorithm

In previous methods of analysis, observables have been treated as free and independent variables in fit functions or maximum likelihood programmes. This approach therefore forfeits a large amount of information contained in the observable correlations. In contrast, the nested sampling programme developed here explores *amplitude space* - that is, the prior distribution consists of sets of amplitude probability density functions rather than polarisation observables. Polarisation observable values are then calculated from these complex amplitudes in a completely consistent manner. This leads to a number of significant advantages: The information obtained from a dataset is maximised by preserving all correlations between polarisation observables, the results of a nested sampling analysis are always consistent, and it is possible to obtain information on observables to which a particular experiment is not directly sensitive. The use of an event-by-event likelihood function also provides certain benefits over a binned fitting method. Whereas in previous analysis programmes, information was inherently lost through the binning of histograms, an event-by-event treatment of data bypasses this issue. More interestingly, the Bayesian nature of the algorithm and the fact that amplitude space is being explored makes possible the analysis of concatenated datasets - that is, data from different (although consistent) experiments can be combined and a Bayesian programme can be used to extract observables from all included experiments in a completely consistent manner. It should be noted that this method of combining datasets is possible using other methods of analysis (such as χ^2 and maximum likelihood), although this would be significantly more challenging to implement. The binning of histograms would need to be consistent across both datasets, and this can result in a substantial loss of information.

The analysis was written in C++, using ROOT. Although the heart of the algorithm was kept completely generic, there were several functions that were problem-specific. The generation of a prior distribution, for example, the likelihood function and the function responsible for evolving points. These will be discussed below.

3.2.1 Prior Generation

The prior distribution used in this programme consisted of the four normalised complex amplitudes introduced in Section 1.3, from which all polarisation observables are calculated. Eight independent values (corresponding to the complex amplitudes' real and imaginary components) were randomly generated from the surface of a unit 7-sphere. They were then used to define the four complex amplitudes, from which the relevant polarisation observables were calculated. The sample size used in this code was 3000 - that is, there were 3000 sets of four complex amplitudes and polarisation observables. Tests were performed to ensure that this construction of a prior distribution uniformly covered the physical regions of the observables. A plot of one of the complex amplitudes is given in Figure 3.3 and the prior distributions for each of the five observables are shown in Figure 3.4.

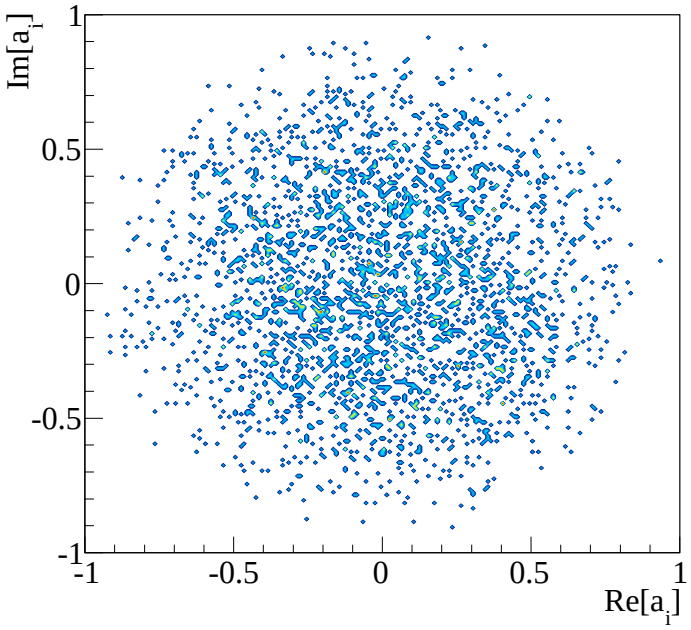


Figure 3.3: Argand plot showing a 3000-sample prior distribution of a complex amplitude. As all four amplitudes are generated in a similar way, their distributions are similar.

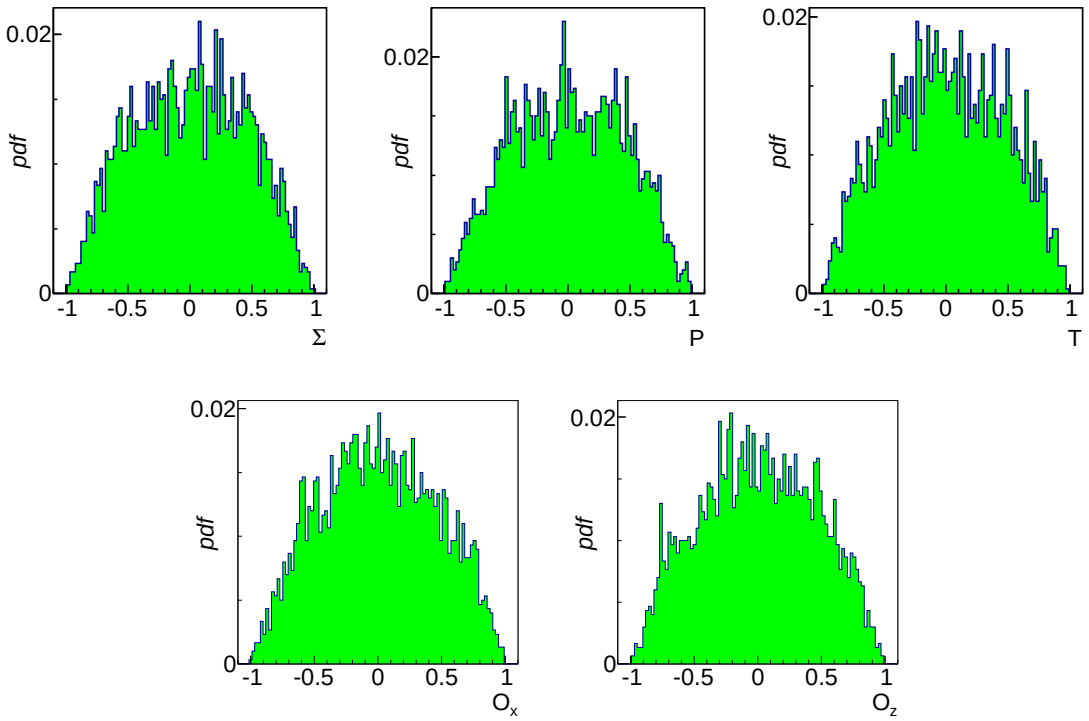


Figure 3.4: Prior distributions for the five polarisation observables Σ, P, T, O_x, and O_z, respectively. These observable values were calculated from the 3000 sets of four complex amplitudes, one of which is shown in Figure 3.3.

3.2.2 Likelihood function

Once a prior distribution has been generated, each sample point is assigned a likelihood value which is dependent on the data and the values associated with the relevant observable quantities. During the initial development of the nested sampling program, a simplified likelihood function was adopted. The final likelihood functions will be discussed in Section 3.5 in detail. The likelihood function is problem-specific, although in the case of pseudoscalar meson photoproduction data analysis, the form is usually fairly similar for different experiments. The likelihood function simplified for a specific configuration is given in Equation 3.2.1.

$$\mathcal{L} = \prod_{i=1}^{N} \Pr(i) \tag{3.2.1}$$

where

$$\Pr(i) = \frac{1}{2}\left\{1 \pm \frac{P_{\gamma i}[(\Sigma + \alpha T \cos\theta_{yi})\cos(2\phi_i) \; + \; \alpha(O_x \cos\theta_{xi} + O_z \cos\theta_{zi})\sin(2\phi_i)]}{1 + \alpha P \cos\theta_{yi}}\right\} \tag{3.2.2}$$

There are several terms to define in Equation 3.2.2 that describe an event. $P_{\gamma i}$ is the degree of polarisation of the beam and is bounded by $[0, 1]$. The $\cos\theta_{x,y,zi}$ terms represent the direction cosines defined by the polar angles of the proton in the hyperon rest frame, and ϕ_i is the kaon azimuthal angle. The addition or subtraction in Equation 3.2.2 depends on the beam polarisation state. The other variables are not event-dependent. α is the weak decay constant of the decaying hyperon, and Σ, P, T, O_x and O_z are the polarisation observables that serve as an input to the likelihood function and which define the sample point. Due to the typical sizes of the likelihood values, the function returns the logarithm of the likelihood value.

Each call to the likelihood function involves looping over all events in the data file. In cases of moderate statistics, this is typically of the order of 5000 - 10000 events. During each loop of the nested sampling algorithm, the likelihood function is called twenty times, and the nested sampling algorithm itself loops around 50000 times. The iterative nature of the likelihood function resulted in a run-time significantly slower than other commonly applied methods of analysis. Because of

this, multithreading and GPU programming was investigated. The details of this investigation are discussed in Section 3.4.

3.2.3 Exploration

When each sample in the prior distribution has been given a likelihood value, the nested sampling algorithm identifies the point with the lowest likelihood, or the least likely sample. It then overwrites that point with a copy of a surviving point, and evolves the value slightly. The function ensures that the likelihood of the new point is higher than that of the original through the use of a Markov Chain Monte Carlo. This exploration or evolution is performed by the Explore function.

The duplicate sample point is adjusted slightly by adding small random numbers to the real and imaginary components of the four complex amplitudes. The observables are then re-calculated and an associated likelihood value is calculated. A comparison between the sample's new likelihood value and the value it originally held determines whether the magnitude of the small random numbers should be increased or decreased. This enables a well-distributed exploration of the sample space.

3.2.4 Results on Perfect Simulated Data

Initial tests of the nested sampling programme were performed using data from a simulation of an ideal set-up. The data were created with an event generator whose input was a set of values of the relevant polarisation observables. A consistent set of observable values were input into the generator, and a text file containing the relevant data members was created. The generated data represented a completely pure signal and contained no noise or background contributions. The degree of polarisation was set to be constant for the purposes of testing. The simplified likelihood function was used and the results obtained were consistent with the values input into the event generator.

A simple Monte Carlo event generator was used to create simulated data for testing purposes. The main input for the event generator was a configuration file which contains a value for each relevant polarisation observable, the photon energy used and the degree of photon polarisation. On the command line, the number of events to be produced and the polarisation state are specified. Phase space distributions of K^+ and Λ particles were then produced. The Λ particles then decayed into their decay products (a proton and π^-). The relevant kinematic properties (kaon azimuthal (ϕ) angle, direction cosines) are generated for each event. Approximately 70% of the events are rejected based on the value of the cross-section determined by the kinematic variables. For the testing phase, an equal number of events in each polarisation state were generated for simplicity. The event generator outputs a text file containing the kaon azimuthal angle (ϕ), flag indicating the polarisation state, and the three direction cosines for each event.

The effects of low, moderate and high statistics were studied by varying the size of the generated dataset. Sets of data consisting of 1000, 5000 and 10 000 events were tested and the results are shown in Figures 3.5 to 3.7.

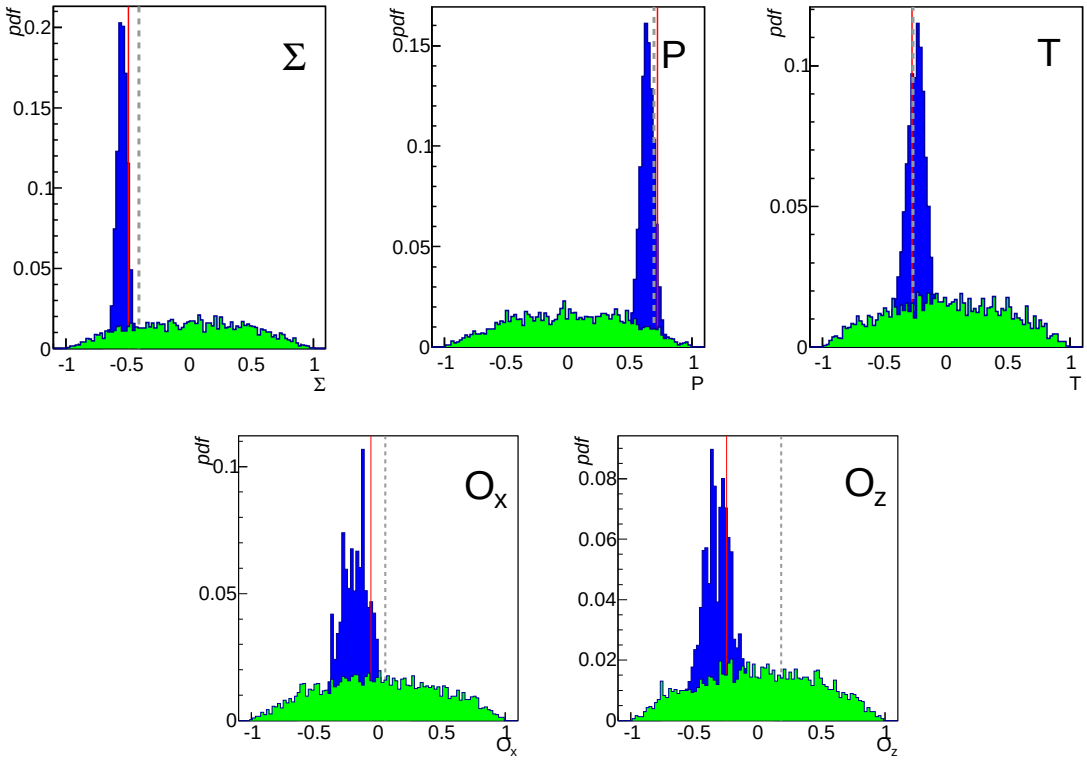


Figure 3.5: Results of nested sampling analysis on dataset of 1000 events. The prior distributions are illustrated by the green histograms, the resulting posterior distributions (pdfs) by the blue peaks and the true value (input into the event generator) of each observable is indicated by the vertical red line. Results obtained from a χ^2 fitting method are shown with a dashed grey line. The observables shown are (a) Σ, (b) P, (c) T, (d) O_x and (e) O_z.

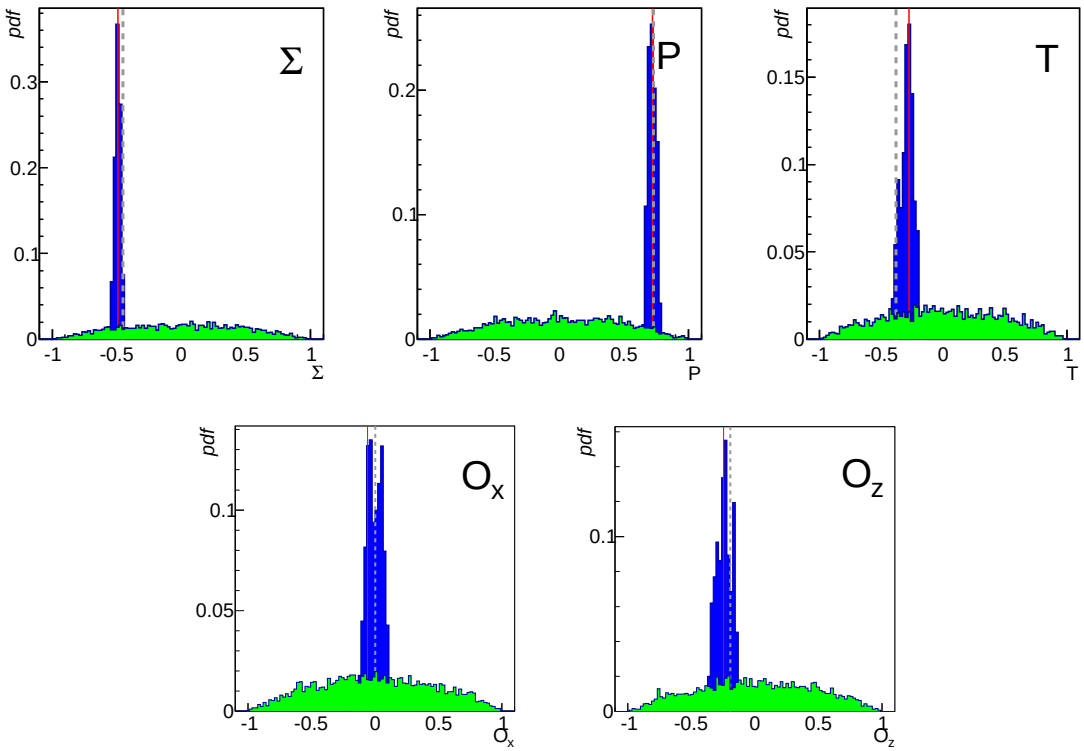


Figure 3.6: Results of nested sampling analysis on dataset of 5000 events. The prior distributions are illustrated by the green histograms, the resulting posterior distributions by the blue peaks and the true value (input into the event generator) of each observable is indicated by the vertical red line. Results obtained from a χ^2 fitting method are shown with a dashed grey line. The observables shown are (a) Σ, (b) P, (c) T, (d) O_x and (e) O_z.

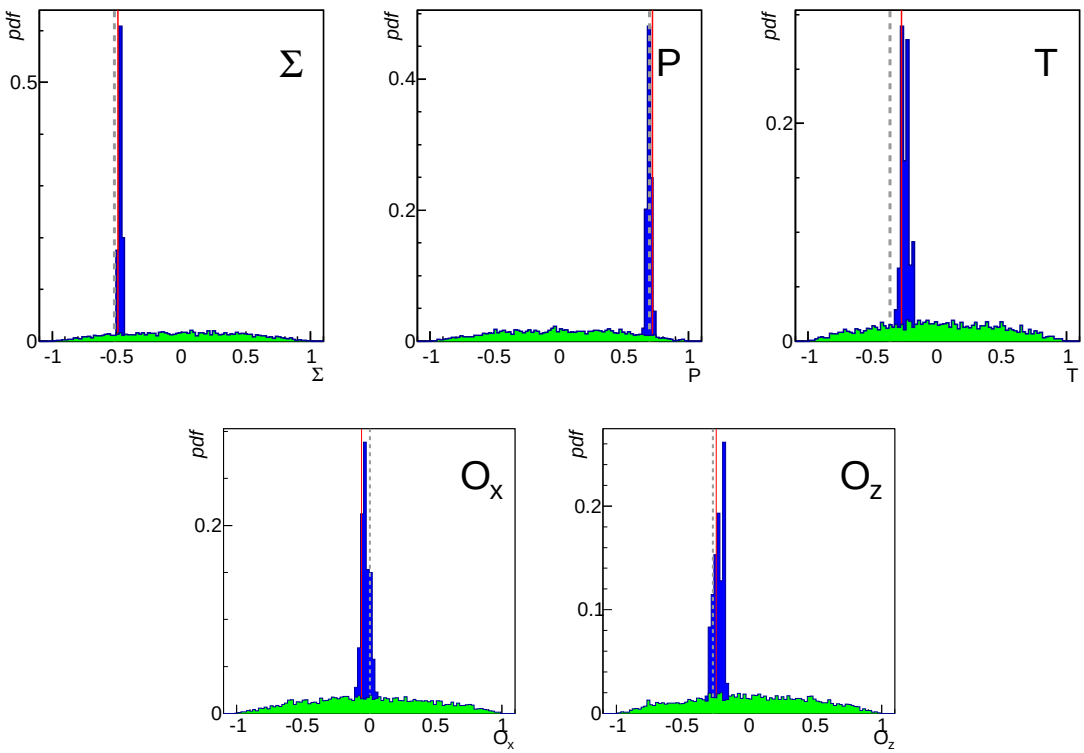


Figure 3.7: Results of nested sampling analysis on dataset of 10000 events. The prior distributions are illustrated by the green histograms, the resulting posterior distributions by the blue peaks and the true value (input into the event generator) of each observable is indicated by the vertical red line. Results obtained from a χ^2 fitting method are shown with a dashed grey line. The observables shown are (a) Σ, (b) P, (c) T, (d) O_x and (e) O_z.

It was found that even at low statistics, the results of a nested sampling analysis were accurate and it was determined that the algorithm and simplified likelihood function were performing as expected.

3.2.5 Systematic Errors

The normalised amplitude values that make up the prior distribution are used to calculate values of the polarisation observables that are inherently bound by [-1, 1]. This implies that under no circumstances can the values of these polarisation observables be above +1 or below -1. This is generally an advantage, however there are some cases where this restriction can be problematic. The normalisation of the amplitudes assumes a precise knowledge of the polarisations and luminosity, and in most cases any error in measurement is negligible in comparison to other sources of systematic errors. It is possible, however, that systematic errors in the measurement of photon polarisation (for example) can be present. In cases where the values of some polarisation observables are close to a boundary, the effects of this systematic error would be difficult to determine, as the observable values would be suppressed in order to keep them within the allowed region. In a binned fit method, the resulting observable values would, in this case, be outwith the physical region, but may provide some information on the scale of the systematic errors present.

3.3 Benchmarking

The accuracy of the analysis methods described above were compared quantitatively by benchmarking the obtained results against the input values used to generate the simulated data. Ten data sets were simulated by an event generator using consistent sets of polarisation observable values as inputs. These data sets were then analysed using a binned χ^2 fit and nested sampling. The values obtained by each analysis method were compared to the input parameters through two values, the distance squared and the residuals squared, given by the following two equations [74]:

$$D^2 = \sum_i (\mathcal{O}_i^{true} - \mathcal{O}_i^{calc})^2 \tag{3.3.1}$$

$$R^2 = \sum_i (\frac{\mathcal{O}_i^{true} - \mathcal{O}_i^{calc}}{\sigma_i^{calc}})^2 \tag{3.3.2}$$

where $\mathcal{O}$ denotes a generic observable and σ^{calc} indicates the associated error. The distance squared (Equation 3.3.1) describes, in essence, the difference between the observable values extracted through analysis and their corresponding input values. The residuals squared (Equation 3.3.2) provides a quantitative description of how well the uncertainties are determined. It is expected that the sum of the squared residuals would be approximately equal to the number of observables being extracted, with a significantly higher value indicating an underestimation of the associated uncertainties [74].

The results from each of the ten data sets are shown in the plots in Figure 3.8. These plots indicate that not only does an unbinned nested sampling extraction provide more accurate results, but does so more consistently than a binned χ^2 fit. The estimation of uncertainties has also been shown to be more accurate, with the sum of the squares of residuals clustering close to five, the number of observables being extracted in this analysis. The binned fit yields a significantly higher value and suggests that the uncertainties are frequently being underestimated. The plots also show more consistent behaviour as the number of events used in the nested sampling analysis is increased.

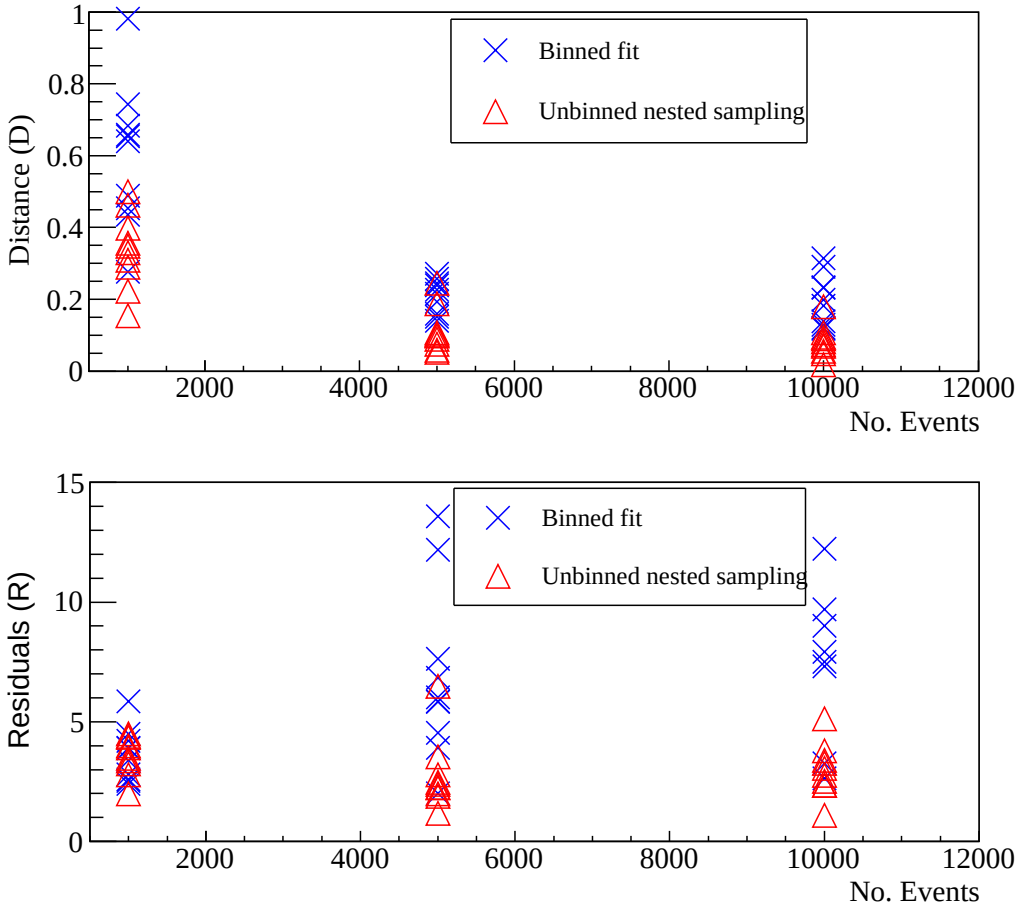


Figure 3.8: Benchmarking results comparing accuracy of nested sampling to the commonly used χ^2 fitting technique.

3.4 Data Parallelism

Throughout the development and testing of the nested sampling programme, it was found that the runtime was significantly longer than previous analysis programmes. This made the development a slow, tedious process and it was decided that this needed to be addressed. The use of data parallelism and GPU programming was explored in an effort to speed up the analysis programme. There exist multiple tools for parallelising code, two of which were investigated and tested. The work described in this section has also been published in [80]. The simplest tool explored was OpenMP [81], an application programming interface (API) that can be used to multithread code on a CPU. It works through the use of compiler directives that specify the loop or region of code that will be parallelised [82]. The implementation into existing code is relatively quick and uncomplicated, but this approach does not allow the programme to run on the GPU. The second tool explored was the OpenCL (Open Computing Language) framework, designed to parallelise code on multicore devices (including GPUs and multicore CPUs) [83]. This approach required a significant amount of work and its implementation was nontrivial. Collaboration with the University of Glasgow's School of Computing Science resulted in the development of a wrapper class [84] that simplified the considerable initialisations required to use OpenCL. The OpenCL implementation allowed the programme to be parallelised on either a multicore CPU or on a GPU, enabling the user to test the programme on more than one device. As such, both devices were tested with the OpenCL implementation of the code. A more complete description of the theory behind data parallelism and GPU programming can be found in Appendix A.

Several methods of multithreading and data parallelisation were attempted using data generated from a simplified linearly polarised beam with recoiling proton simulation. Three implementations were made for comparison purposes:

- Multithreading on the CPU with OpenMP
- Data parallelisation on the CPU/GPU with OpenCL
- Optimised unthreaded on the CPU (standard)

The correlations between dataset size and number of threads over which to parallelise the data were explored through a variety of tests. The implementations were all run on a high-end computer system with the following specifications:

- NVidia®Tesla™C2075 Companion Processor
- 448 CUDA Cores
- 1.15 GHz core frequency
- 6GB GDDR5 RAM
- 144 GB/s memory bandwidth
- Intel i7 2700K 3.5GHz Quad Core 8MB cache processor

Software development kits for both the Intel and NVidia implementations of OpenCL were installed on the system. The OpenCL version of the nested sampling programme involved a kernel function containing the likelihood calculation. The kernel function was called from the host programme and the necessary data was transferred during the kernel call. The kernel programme was written in a language based on C-99. A rudimentary programme called CLCC [85] tested the kernel code compilation and provided primitive error messages. It was used to debug the kernel code before implementing it into the programme. A wrapper class called OclWrapper was developed by Dr Vanderbauwhede of the University of Glasgow's School of Computing Science to reduce the amount of initialisation coding required. This class allowed the user to specify the number of threads to use and the device on which the programme was to be run, either the CPU or the GPU. The OclWrapper class was used in the coding of the OpenCL implementation of the nested sampling programme.

The OpenMP version of the code was much simpler to implement and required only a few additional lines of code, in the form of pragmas, or compiler directives. These OpenMP pragmas were placed around the section of code that was to be multithreaded (the event loop in the likelihood function), and the number of threads

over which to parallelise the data was an argument that could be set at the command line.

For each dataset size and for each implementation of the nested sampling program, the optimal number of threads was found by performing timing tests using a range of numbers of threads. The results of these tests are shown in Figures 3.9-3.11.

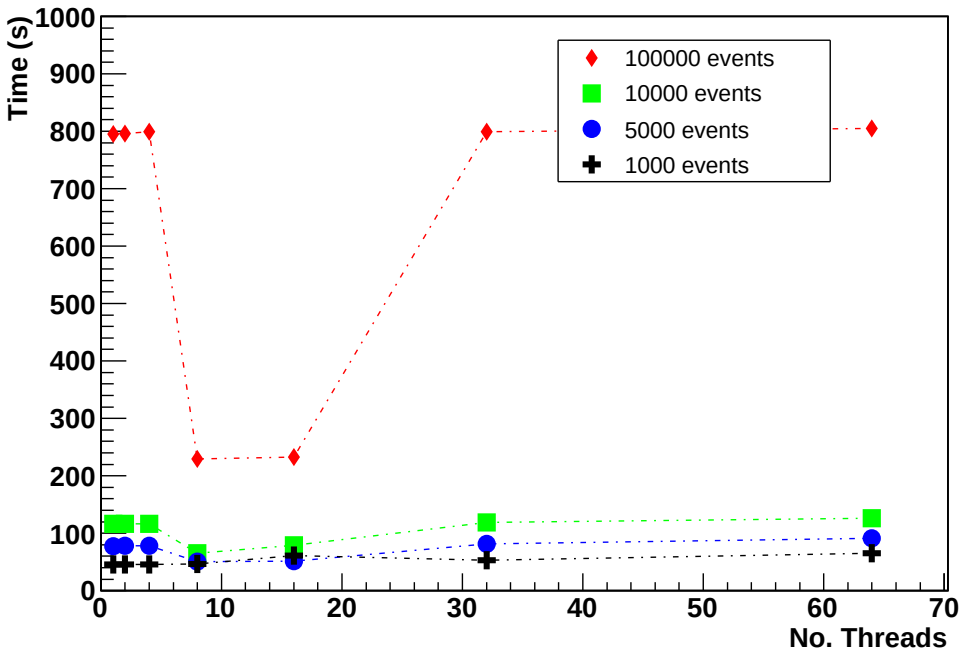


Figure 3.9: Initial thread test of OpenCL implementation run on the CPU.

From these tests, it was found that the optimal number of threads was dependent on the size of the dataset as well as the implementation selected, and that the impact of running the programme using the optimal number of threads was significantly more prominent for larger datasets. Using more threads than the optimal number results in a slower runtime, as more time is required in combining results from the many threads. This behaviour can be seen by the clear minimum shown in Figure 3.9.

Once the optimal number of threads was determined for given initial conditions, the runtimes of the implementations were compared for each dataset size.

The results in Figure 3.12 indicate that, contrary to the initial expectations of the GPU outperforming all other implementations, the most significant speed-ups

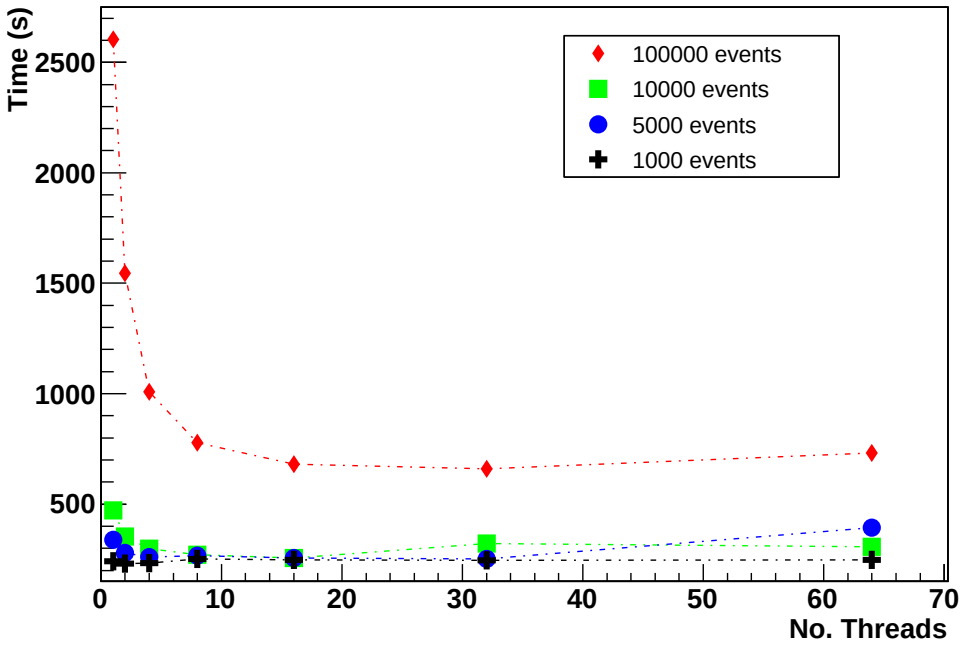


Figure 3.10: Initial thread test of OpenCL implementation run on the GPU.

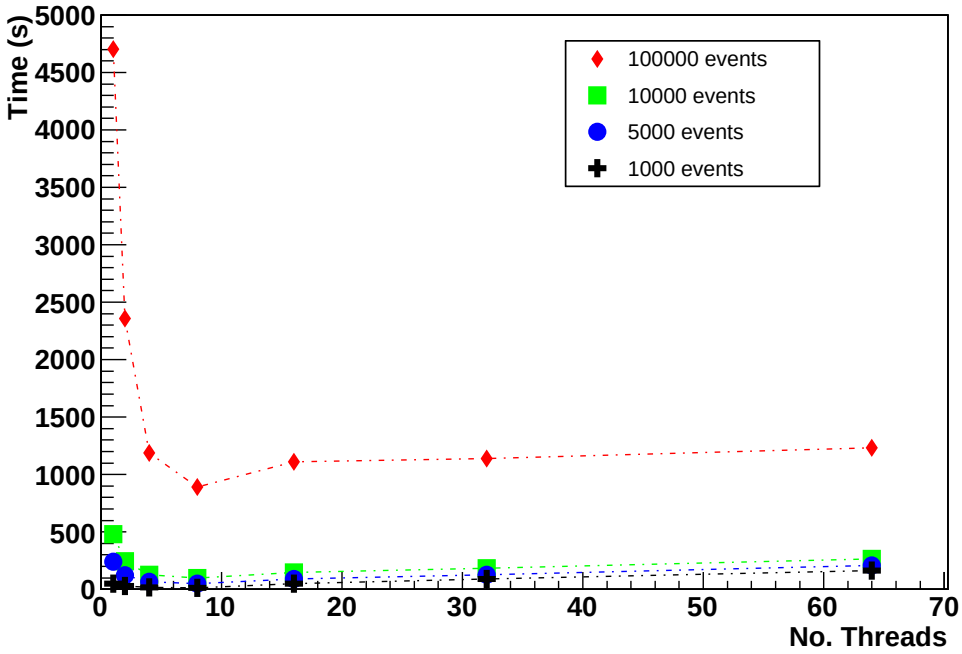


Figure 3.11: Initial thread test of OpenMP implementation run on the CPU.

were obtained using OpenMP or OpenCL on the CPU. For smaller datasets, in the region of 0-7000 events, OpenMP multithreading provided the most effective speed-up. As the size of the datasets was increased, OpenCL on the CPU started to show the best results, and there are hints that if the dataset sizes were to increase into millions or billions of events, the OpenCL implementation using the GPU would be the fastest.

This unexpected result was attributed to the overhead caused by transferring data to the specified compute unit. To quantify this, data transfer time tests were performed using OpenCL. For these tests, the OpenCL implementation of the nested sampling programme was run using an empty kernel function, where no calculations of any kind were being performed on the compute device. The time per iteration was then solely attributed to the time to transfer data from the main memory to the compute unit's local memory. The results of these tests are displayed in Tables 3.1 and 3.2.

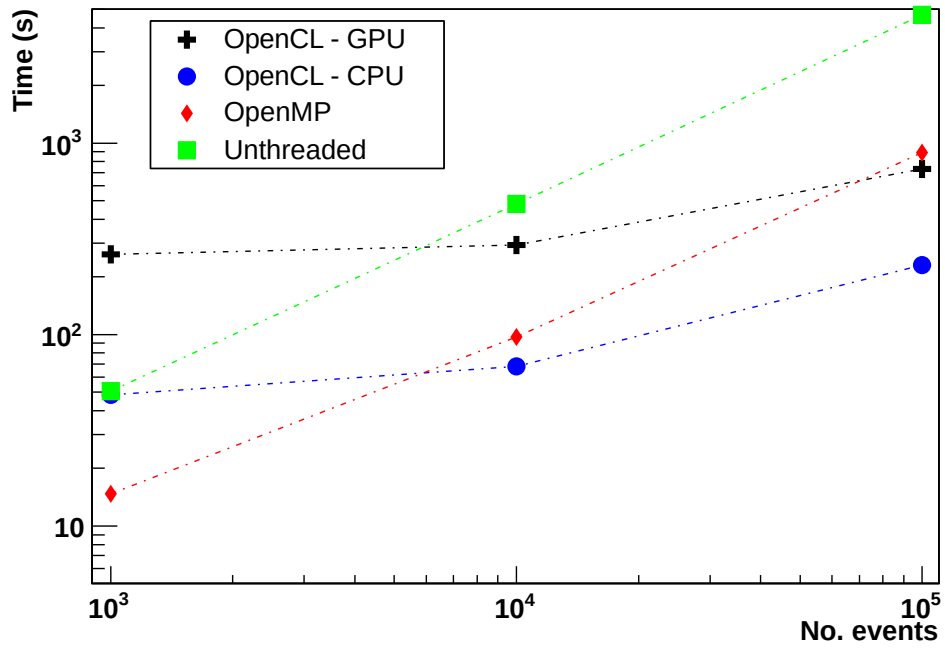


Figure 3.12: Results showing runtimes of each implementation with increasing dataset size. Each implementation was run with its optimal number of threads.

Table 3.1: Data transfer results per iteration for OpenCL on the GPU.

No. events	Data transfer time (ms)	Total time (ms)	% Data Transfer
1000	0.227 ± 0.008	0.277 ± 0.0023	82
5000	0.231 ± 0.010	0.328 ± 0.0027	70
10000	0.232 ± 0.004	0.3997 ± 0.0041	58
100000	0.229 ± 0.005	1.560 ± 0.0010	15

Table 3.2: Data transfer results per iteration for OpenCL on the CPU.

No. events	Data transfer time (ms)	Total time (ms)	% Data Transfer
1000	0.021 ± 0.0007	0.0457 ± 0.001	46
5000	0.021 ± 0.0022	0.0776 ± 0.001	27
10000	0.021 ± 0.0007	0.114 ± 0.0006	18
100000	0.021 ± 0.0011	0.790 ± 0.0074	3

It can be seen that much more time is required to transfer data to the GPU than to the CPU, and that a sufficient number of calculations must be performed in order to overcome this overhead. It was found that for the expected sizes of data sets that are likely to be encountered in such pseudoscalar meson photoproduction data analysis at CLAS, the OpenMP implementation was able to provide the best speed-up with the simplest code modifications to the nested sampling programme.

3.5 Derivation of Likelihood Functions

The cross section equation of pseudoscalar meson photoproduction has 32 terms in total. Many of these terms, however, go to zero when the reaction configuration is specified. Below is the expression of the total cross section of pseudoscalar meson photoproduction [86].

$$\begin{aligned}
\sigma = \sigma_0\{1 &- P_L^\gamma P_T^T P_y^R \sin(\phi)\cos(2\phi) + \Sigma(-P_L^\gamma \cos(2\phi) + P_T^T P_y^R \sin(\phi)) \\
&+ T(P_T^T \sin(\phi) - P_L^\gamma P_y^R \cos(2\phi)) + P(P_y^R - P_L^\gamma P_T^T \sin(\phi)\cos(2\phi)) \\
&+ E(-P_C^\gamma P_L^T + P_L^\gamma P_T^T P_y^R \cos(\phi)\sin(2\phi)) + F(P_C^\gamma P_T^T \cos(\phi) + P_L^\gamma P_L^T P_y^R \sin(2\phi)) \\
&- G(P_L^\gamma P_L^T \sin(2\phi) + P_C^\gamma P_T^T P_y^R \cos(\phi)) - H(P_L^\gamma P_T^T \cos(\phi)\sin(2\phi) - P_C^\gamma P_L^T P_y^R) \\
&- C_x(P_C^\gamma P_x^R - P_L^\gamma P_T^T P_z^R \sin(\phi)\sin(2\phi)) - C_z(P_C^\gamma P_z^R + P_L^\gamma P_T^T P_x^R \sin(\phi)\sin(2\phi)) \\
&- O_x(P_L^\gamma P_x^R \sin(2\phi) + P_C^\gamma P_T^T P_z^R \sin(\phi)) - O_z(P_L^\gamma P_z^R \sin(2\phi) - P_C^\gamma P_T^T P_x^R \sin(\phi)) \\
&+ L_x(P_L^T P_x^R + P_L^\gamma P_T^T P_z^R \cos(\phi)\cos(2\phi)) + L_z(P_L^T P_z^R - P_L^\gamma P_T^T P_x^R \cos(\phi)\cos(2\phi)) \\
&+ T_x(P_T^T P_x^R \cos(\phi) - P_L^\gamma P_L^T P_z^R \cos(2\phi)) + T_z(P_T^T P_z^R \cos(\phi) + P_L^\gamma P_L^T P_x^R \cos(2\phi))\}
\end{aligned} \tag{3.5.1}$$

where σ_0 is the reduced cross-section,

Σ is beam polarisation observable,

P is recoil polarisation observable,

T is target polarisation observable,

E, F, G, H are the beam-target double polarisation observables,

C_x, C_z, O_x, O_z are the beam-recoil double polarisation observables,

T_x, T_z, L_x, L_z are the target-recoil double polarisation observables,

P_L^γ, P_C^γ are the linear and circular photon polarisations, respectively,

$P_{L,T}^T$ are the longitudinal and transverse target polarisation components, respectively, and

$P_{x,y,z}^R$ are the recoil polarisation components in the x, y and z directions.

The type of reaction is determined by the polarisations of the beam and target, and whether the recoil polarisation information was obtained. The different polarisation settings for the beam and target result in a number of terms in Equation 3.5.1 going to zero.

Table 3.3: Types of Reactions

	Polarisation	Values going to zero
Beam	Unpolarised	P_L^γ, P_C^γ
	Linear	P_C^γ
	Circular	P_L^γ
Target	Unpolarised	P_L^T, P_T^T
	Longitudinal	P_T^T
	Tranverse	P_L^T
Recoil	Detected	n/a
	Not detected	P_x^R, P_y^R, P_z^R

For each combination of polarisations, a reduced cross section equation was determined.

In Tables 3.3 and 3.4, subscripts are used to denote the reaction types.

$B_{U,L,C}$ specifies beam polarisation - unpolarised, linearly polarised and circularly polarised, respectively.

$T_{U,L,T}$ specifies target polarisation - unpolarised, longitudinally polarised and transversely polarised, respectively, and

$R_{Y,N}$ specifies whether the recoil polarisation was detected or not detected, respectively.

Table 3.4: Reduced Cross Section Equation, by Configuration

Configuration	σ_{Red}/σ_0
$B_U\,T_U\,R_N$	$= 1$
$B_U\,T_U\,R_Y$	$= 1 + P\,P_y^R$
$B_U\,T_L\,R_N$	$= 1$
$B_U\,T_L\,R_Y$	$= 1 + P\,P_y^R + L_x\,P_x^R\,P_L^T$
$B_U\,T_T\,R_N$	$= 1 + T\,P_T^T \sin(\phi)$
$B_U\,T_T\,R_Y$	$= 1 + P\,P_y^R + (\Sigma\,P_y^R + T)\,P_T^T \sin(\phi) + (T_x\,P_x^R + T_z\,P_y^R)\,P_T^T \cos(\phi)$
$B_C\,T_U\,R_N$	$= 1$
$B_C\,T_U\,R_Y$	$= 1 + P\,P_y^R - C_x\,P_C^\gamma\,P_x^R - C_z\,P_C^\gamma\,P_z^R$
$B_C\,T_L\,R_N$	$= 1 - E\,P_C^\gamma\,P_L^T$
$B_C\,T_L\,R_Y$	$= 1 + P\,P_y^R - E\,P_C^\gamma\,P_L^T + H\,P_C^\gamma\,P_y^R\,P_L^T$
	$-C_x\,P_C^\gamma\,P_x^R - C_z\,P_C^\gamma\,P_z^R + L_x\,P_x^R\,P_L^T + L_z\,P_z^R\,P_L^T$
$B_C\,T_T\,R_N$	$= 1 + T\,P_T^T \sin(\phi) + F\,P_C^\gamma\,P_T^T \cos(\phi)$
$B_C\,T_T\,R_Y$	$= 1 + P\,P_y^R - C_x\,P_C^\gamma\,P_x^R - C_z\,P_C^\gamma\,P_z^R$
	$+(\Sigma\,P_y^R + T - O_x\,P_C^\gamma\,P_z^R + O_z\,P_C^\gamma\,P_x^R)\,P_T^T \sin(\phi)$
	$+(F\,P_C^\gamma\,P_y^R - G\,P_C^\gamma\,P_y^R + T_x\,P_x^R + T_z\,P_z^R)\,P_T^T \cos(\phi)$
$B_L\,T_U\,R_N$	$= 1 - P_L^\gamma\,\Sigma\,\cos(2\phi)$
$B_L\,T_U\,R_Y$	$= 1 + P\,P_y^R - (\Sigma + T\,P_y^R)\,P_L^\gamma\,\cos(2\phi) - (O_x\,P_x^R + O_z\,P_z^R)\,P_L^\gamma\,\sin(2\phi)$
$B_L\,T_L\,R_N$	$= 1 - \Sigma\,P_L^\gamma\,\cos(2\phi) - G\,P_L^\gamma\,P_L^T \sin(2\phi)$
$B_L\,T_L\,R_Y$	$= 1 + P\,P_y^R + L_x\,P_x^R\,P_L^T + L_z\,P_z^R\,P_L^T$
	$-(\Sigma + T\,P_y^R + T_x\,P_z^R\,P_L^T - T_z\,P_x^R\,P_L^T)\,P_L^\gamma\,\cos(2\phi)$
	$+(F\,P_y^R\,P_L^T - G\,P_L^T - O_x\,P_x^R - O_z\,P_z^R)\,P_L^\gamma \sin(2\phi)$
$B_L\,T_T\,R_N$	$= 1 - \Sigma\,P_L^\gamma\,cos(2\phi) - P\,P_L^\gamma\,P_T^T \sin(\phi)\cos(2\phi)$
	$+T\,P_T^T \sin(\phi) - H\,P_L^\gamma\,P_T^T\,\cos(\phi)\sin(2\phi)$
$B_L\,T_T\,R_Y$	$= 1 - P_L^\gamma\,P_y^R\,P_T^T\,\sin(\phi)\cos(2\phi)\, + \Sigma(P_y^R\,P_T^T \sin(\phi) - P_L^\gamma\,\cos(2\phi))$
	$+P(P_y^R - P_L^\gamma\,P_T^T\,\sin(\phi)\cos(2\phi)\,) + T\,(P_T^T \sin(\phi) - P_L^\gamma\,P_y^R\,\cos(2\phi)\,)$
	$+(E\,P_y^R - H)\,P_L^\gamma\,P_T^T\,\cos(\phi)\sin(2\phi)$
	$+(C_x\,P_z^R - C_z\,P_x^R)\,P_L^\gamma\,P_T^T\,\sin(\phi)\sin(2\phi)$
	$-(O_x\,P_x^R\,\sin(2\phi) + O_z\,P_z^R)\,P_L^\gamma\,\sin(2\phi) + (T_x\,P_x^R + T_z\,P_z^R)\,P_T^T \cos(\phi)$
	$+(L_x P_z^R - L_z\,P_x^R)\,P_L^\gamma\,P_T^T\,\cos(\phi)\cos(2\phi)$

The likelihood functions derived for each experimental configuration were heavily dependent on the reduced cross sections shown in Table 3.4.

Asymmetries between experimental settings provide the basis for the derivation of likelihood functions. The procedure for developing these functions differs according to the number of elements that are polarised in the experiment. For the simpler cases, where either the beam or the target is polarised (but not both), the asymmetry is straightforward.

$$A = \frac{\sigma^a - \sigma^b}{\sigma^a + \sigma^b} \tag{3.5.2}$$

where the superscripts a and b indicate the two relevant polarisation settings.

For those cases where both the beam and the target are polarised, there are three asymmetries that are possible. For the purposes of clarity and simplicity, nomenclature associated with a linearly polarised photon beam will be assumed (parallel and perpendicular polarisation settings) but can be adapted for the case of circular beam polarisation. The target polarisation settings will be expressed as positive and negative. All cross-sections refer to the reduced cross-section equations.

$$A_1 = \frac{(\sigma_{\parallel}^{+} + \sigma_{\perp}^{-}) - (\sigma_{\parallel}^{-} + \sigma_{\perp}^{+})}{(\sigma_{\parallel}^{+} + \sigma_{\perp}^{-}) + (\sigma_{\parallel}^{-} + \sigma_{\perp}^{+})} \tag{3.5.3}$$

$$A_2 = \frac{(\sigma_{\parallel}^{+} + \sigma_{\parallel}^{-}) - (\sigma_{\perp}^{+} + \sigma_{\perp}^{-})}{(\sigma_{\parallel}^{+} + \sigma_{\parallel}^{-}) + (\sigma_{\perp}^{+} + \sigma_{\perp}^{-})} \tag{3.5.4}$$

$$A_3 = \frac{(\sigma_{\parallel}^{+} + \sigma_{\perp}^{+}) - (\sigma_{\parallel}^{-} + \sigma_{\perp}^{-})}{(\sigma_{\parallel}^{+} + \sigma_{\perp}^{+}) + (\sigma_{\parallel}^{-} + \sigma_{\perp}^{-})} \tag{3.5.5}$$

The asymmetry A_2 is an asymmetry over the beam polarisation and A_3 defines an asymmetry over the target polarisation. A_1 can be viewed as an asymmetry of the convolution of beam and target polarisations.

For each configuration, the associated reduced cross section is rewritten in the following form:

$$\frac{\sigma_{Red}}{\sigma_0} = fP^{\gamma} + gP^{\gamma}P^{T} + kP^{T} + h \tag{3.5.6}$$

where f contains all terms related only to the beam polarisation, g comprises the coefficients of the product of beam and target polarisations, k involves all terms related only to the target polarisation, and h consists of the terms completely unrelated to any polarisation.

In cases where only one of the beam or target is polarised, this form reduces to the simple asymmetry as the g term and either the f term (if only the target is polarised) or k term (if only the beam is polarised) goes to zero.

Through expanding these cross-sections and making some algebraic substitutions, these asymmetries can be simplified.

$$A_1 = \frac{\overline{P}_\gamma \overline{P}_T g}{h + \overline{P}_\gamma \Delta P_\gamma f + \overline{P}_\gamma \Delta P_\gamma \overline{P}_T \Delta P_T g + \overline{P}_T \Delta P_T k} \tag{3.5.7}$$

$$A_2 = \frac{-\overline{P}_\gamma f + \overline{P}_\gamma \overline{P}_T \Delta P_T g}{h + \overline{P}_\gamma \Delta P_\gamma f + \overline{P}_\gamma \Delta P_\gamma \overline{P}_T \Delta P_T g + \overline{P}_T \Delta P_T k} \tag{3.5.8}$$

$$A_3 = \frac{\overline{P}_T \overline{P}_\gamma \Delta P_\gamma g + \overline{P}_T k}{h + \overline{P}_\gamma \Delta P_\gamma f + \overline{P}_\gamma \Delta P_\gamma \overline{P}_T \Delta P_T g + \overline{P}_T \Delta P_T k} \tag{3.5.9}$$

where $\overline{P}_\gamma$ is the mean polarisation of the beam, $\overline{P}_T$ represents the mean polarisation of the target, ΔP_γ is the asymmetry of the beam polarisation, and ΔP_T is the asymmetry of the target polarisation.

As the likelihood function involved in this analysis programme is in an event-by-event form, several simplifications to these asymmetries can be made. The concept of a mean polarisation can be discarded, as can the concept of a polarisation asymmetry. The mean polarisation terms can thus be simplified to the raw degree of polarisation for each event, and the polarisation asymmetry terms go to zero (as the polarisation asymmetry of one event is meaningless). Then the asymmetry equations can be expressed as shown in Equations 3.5.10 - 3.5.12.

$$A_1 = \frac{-P_\gamma P_T g}{h} \tag{3.5.10}$$

$$A_2 = \frac{P_\gamma f}{h} \tag{3.5.11}$$

$$A_3 = \frac{P_T k}{h} \tag{3.5.12}$$

In cases where the target is unpolarised, the asymmetries containing a target polarisation term reduce to zero. Similarly, for those cases when the beam is unpolarised, the asymmetries containing a beam polarisation term are likewise reduced to zero.

In the simplest case of an asymmetry, it is assumed that the number of events from each polarisation state are equal. In the case of A_2, for example, the simplest form would occur when the difference between the number of events in one beam polarisation state and the number of events in the other beam polarisation state is negligible. When this simplification does not apply, however, it is necessary to introduce another term, known as the luminosity asymmetry (ΔL), before the final form of a likelihood function can be determined.

Following the formalism described in [74], the measured asymmetries can be expressed in the following form:

$$\widetilde{A}_i = \frac{A_i + \Delta L_i}{1 + A_i \Delta L_i} \tag{3.5.13}$$

where a unique ΔL term exists for each of the three asymmetries.

The number of events in each setting (for each asymmetry) in a dataset can be expressed as N_a^i and N_b^i. Then, there must exist a ratio r of the number of events in one setting to the number of events in the second setting [74]. Then the probability of an event having setting a is equal to this ratio r, and the probability of an event having the other setting (b) must be equal to $(1 - r)$. It is then possible to express the probability of observing exactly N_a (N_b) events in setting a (b) as follows [74]:

$$\Pr(N_a, N_b | r) \propto r^{N_a}(1 - r)^{N_b} \tag{3.5.14}$$

This describes a beta distribution with $\alpha = N_a + 1$ and $\beta = N_b + 1$.
We then express an asymmetry as a scaled and translated ratio: $\widetilde{A} = 2r - 1$ [74]

and the above expression becomes:

$$\Pr(N_a, N_b|\widetilde{A}) \propto (1+\widetilde{A})^{N_a}(1-\widetilde{A})^{N_b} \tag{3.5.15}$$

We can introduce a normalisation constant, Z, to remove the proportionality:

$$\Pr(N_a, N_b|\widetilde{A}) = \frac{1}{Z}(1+\widetilde{A})^{N_a}(1-\widetilde{A})^{N_b} \tag{3.5.16}$$

This distribution then has a mean of $\frac{N_a-N_b}{N_a+N_b+2}$ and a mode of $\frac{N_a-N_b}{N_a+N_b}$.

In the event-by-event approach, each bin will contain either one or no events. Each event in the dataset can exist in one and only one asymmetry setting. Then, for a bin that consists of one event in setting a, the probability can be expressed as [74]:

$$\Pr(N_a = 1, N_b = 0|\widetilde{A}) = \frac{1}{2}(1+\widetilde{A}) \tag{3.5.17}$$

and for a bin consisting of one event in setting b, similarly:

$$\Pr(N_a = 0, N_b = 1|\widetilde{A}) = \frac{1}{2}(1-\widetilde{A}) \tag{3.5.18}$$

The total likelihood can be considered as the product of probabilities of each event in the dataset.

$$\mathcal{L} = \prod_i \Pr_i(E_i|\mathcal{O}) \tag{3.5.19}$$

for all events E_i.

For computing applications, it is not practical to use a likelihood function in this form. Instead, it is more convenient and effective to use the natural logarithm of likelihoods.

$$\ln \mathcal{L} = \sum_i \Pr_i(E_i|\mathcal{O}) \tag{3.5.20}$$

In those cases where double asymmetries are performed, the final likelihood value is given by the following equation.

$$\ln(\mathcal{L}_{total}) = \ln(\mathcal{L}_{\widetilde{A}_1}) + \ln(\mathcal{L}_{\widetilde{A}_3}) + \ln(\mathcal{L}_{\widetilde{A}_3}) \tag{3.5.21}$$

3.5.1 Extraction of Recoil Observable, P

It was found during the initial stages of development that the recoil observable, P, could not be extracted using the asymmetry alone. The recoil polarisation observable, P, features only ever in the denominator of the likelihood function, and as such, strongly influences the observables found in the numerator. This was also found to be the case in the maximum likelihood analysis technique [74] as well as in the binned fitting approaches used in previous analyses [75]. The solution implemented in previous analysis programmes was adapted for the nested sampling method successfully. The extraction of P can be performed without an asymmetry, and is dependent on the direction cosine, $\cos(\theta_y)$. The likelihood equation associated with P alone is derived from the distribution of intensity, and as there is no asymmetry calculation, is dependent on the acceptance - a term used to denote the fraction of genuine events that are detected/recorded by the detector system. This is generally found through independent studies and simulations of the detector. The probability of an event whose direction cosine is $\cos(\theta_y)$ given P is shown in Equation 3.5.1 [74, 86].

$$\Pr(\cos(\theta_y) \mid P) = \varepsilon(\cos(\theta_y))\frac{1}{2}(1 + \alpha\cos(\theta_y)P) \tag{3.5.1}$$

where $\varepsilon(\cos(\theta_y))$ is the acceptance. An example of the calculation of acceptance is given in Section 4.3.

In the nested sampling programme's likelihood function, this is included as an additional likelihood term and combined with the standard likelihoods found from the equations derived earlier in Section 3.5.

3.6 Single Polarisation Analysis

The development of an analysis method for each potential polarisation configuration can be broken into two distinct sections: the development of single polarisation programmes and the development of double polarisation programmes (discussed in Section 3.7). Although the term "single-polarisation" generally refers to observables that are dependent only on one polarisation (i.e. either the beam polarisation, recoil polarisation or the target polarisation), such as Σ, in this case the expression is used to denote experimental configurations where only one of the beam or target is polarised. Similarly, "double-polarisation" generally refers to observables, such as C_x, that depend on multiple polarisations (such as beam and recoil polarisations), here it is used to in reference to experimental configurations involving two or more of these components being polarised.
The simulated data for each configuration were generated from one set of amplitudes. Each analysis was tested with 5000 data points and 50000 iterations. The degrees of beam and target polarisations were constant ($P_\gamma = 0.7$, $P_T = 0.7$), and the luminosity asymmetry (ΔL) was 0. Ten different sets of simulated data were used to test each analysis, although only one set of results are given here.

The simplest case occurs when only one element of the experimental configuration is polarised. Following the formalism discussed in Section 3.5, likelihood functions were derived for each of the possible single polarisation states. The likelihood equations and results from perfect simulated data are shown in this section. In all plots included in this section, there are three common features. A green histogram shows the normalised prior distribution, determined as discussed in Section 3.2.1. A blue histogram indicates the posterior distribution, the mean of which is taken as the extracted value of the observable. A vertical red line shows the 'true' value of the observable, in other words, the value used as an input in the event generation.

There are six single polarisation configurations that yield information about polarisation observables. The likelihood functions of all six programmes were derived as discussed in Section 3.5. Each single polarisation programme was only dependent

on one asymmetry (either A_2 or A_3), as only one of the beam or target was polarised in each case. The resulting posterior distributions of these programmes on simulated data are shown in Figures 3.13 - 3.18.

Table 3.5: Single polarisation configurations

Configuration	Beam	Recoil	Target
$B_L T_U R_N$	Linear	No	Unpolarised
$B_L T_U R_Y$	Linear	Yes	Unpolarised
$B_C T_U R_Y$	Circular	Yes	Unpolarised
$B_U T_L R_Y$	Unpolarised	Yes	Longitudinal
$B_U T_T R_N$	Unpolarised	No	Transverse
$B_U T_T R_Y$	Unpolarised	Yes	Transverse

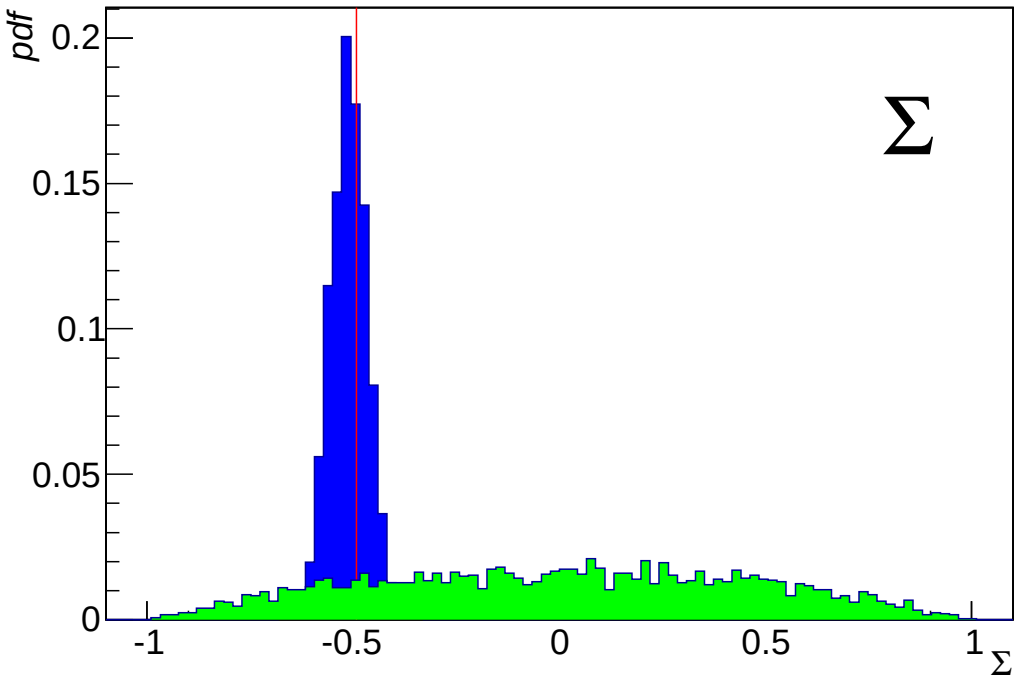


Figure 3.13: Using a linearly polarised photon beam, unpolarised target, and no recoil information, the only observable directly measurable is Σ.

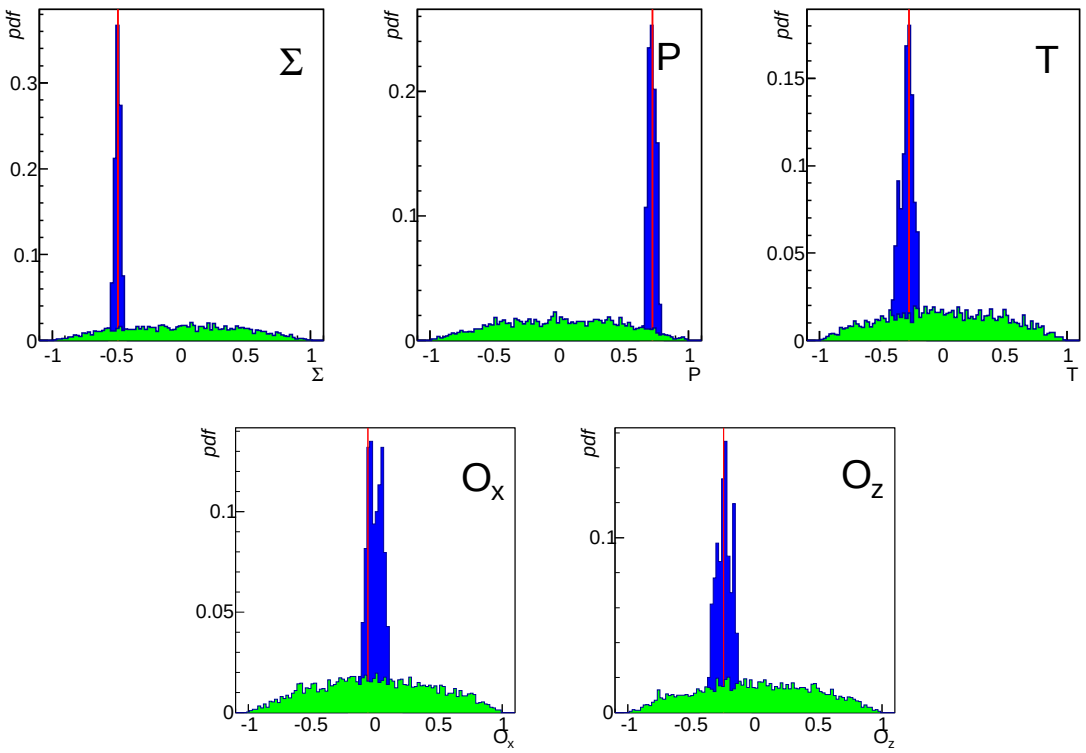


Figure 3.14: Using a linearly polarised photon beam, unpolarised target, with recoil information, five observables can be measured. Top: Σ, P, T. Bottom: O_x, O_z.

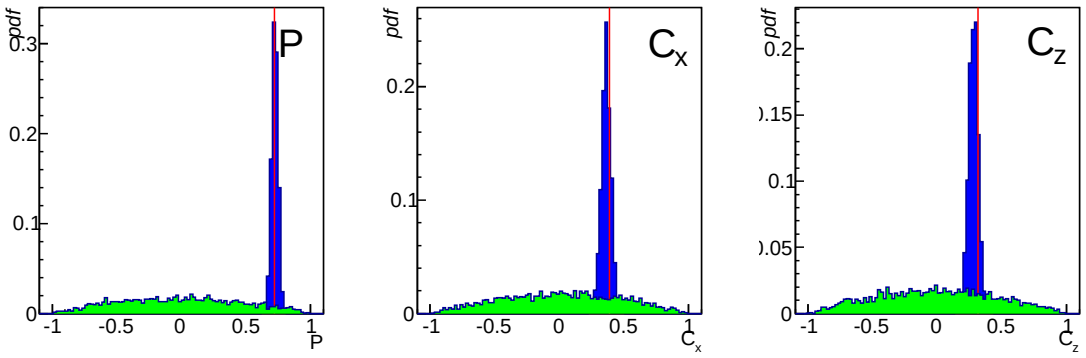


Figure 3.15: With a circularly polarised beam, unpolarised target and recoil information, P, C_x and C_z can be extracted.

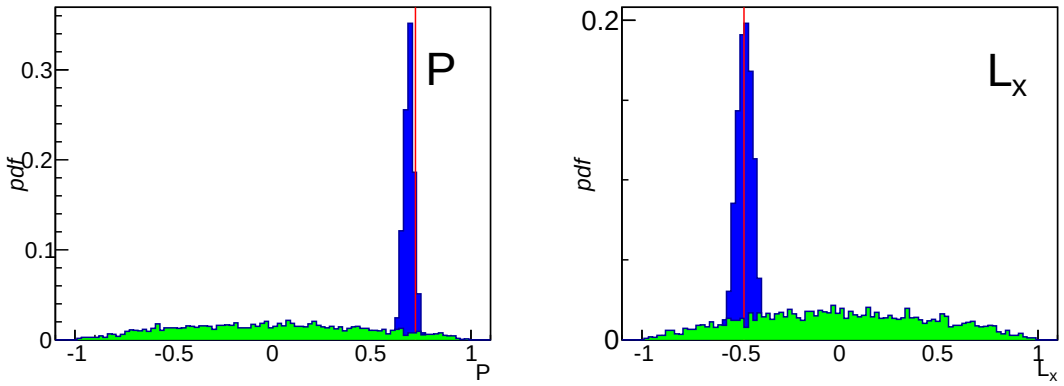


Figure 3.16: An unpolarised beam, longitudinally polarised target and recoil information yields P and L_x.

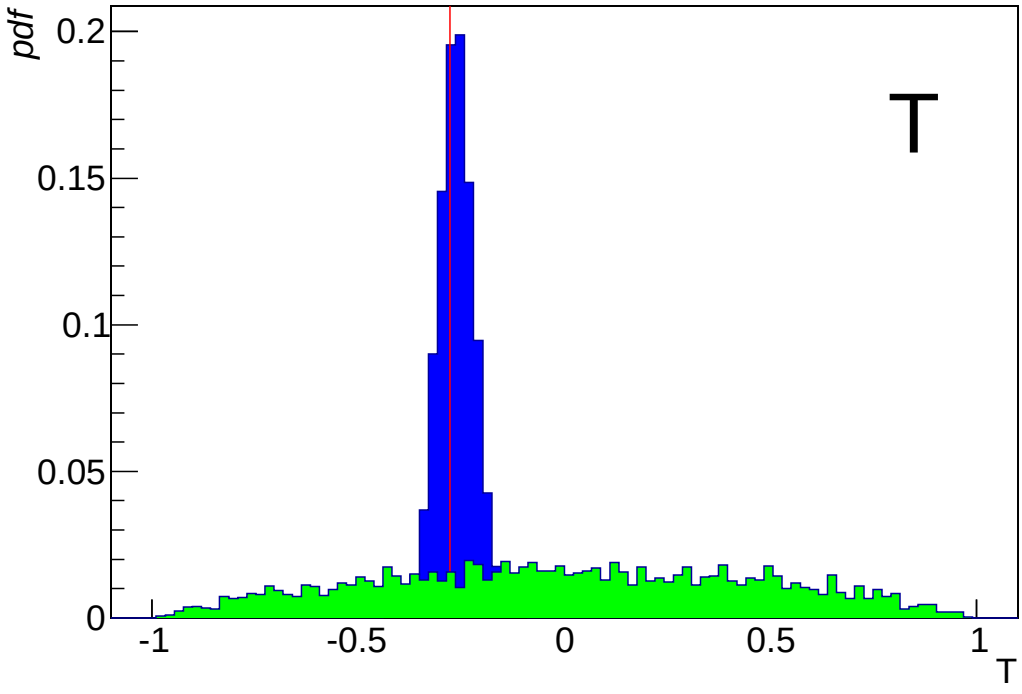


Figure 3.17: With an unpolarised beam, transversely polarised target and no recoil information, T can be measured.

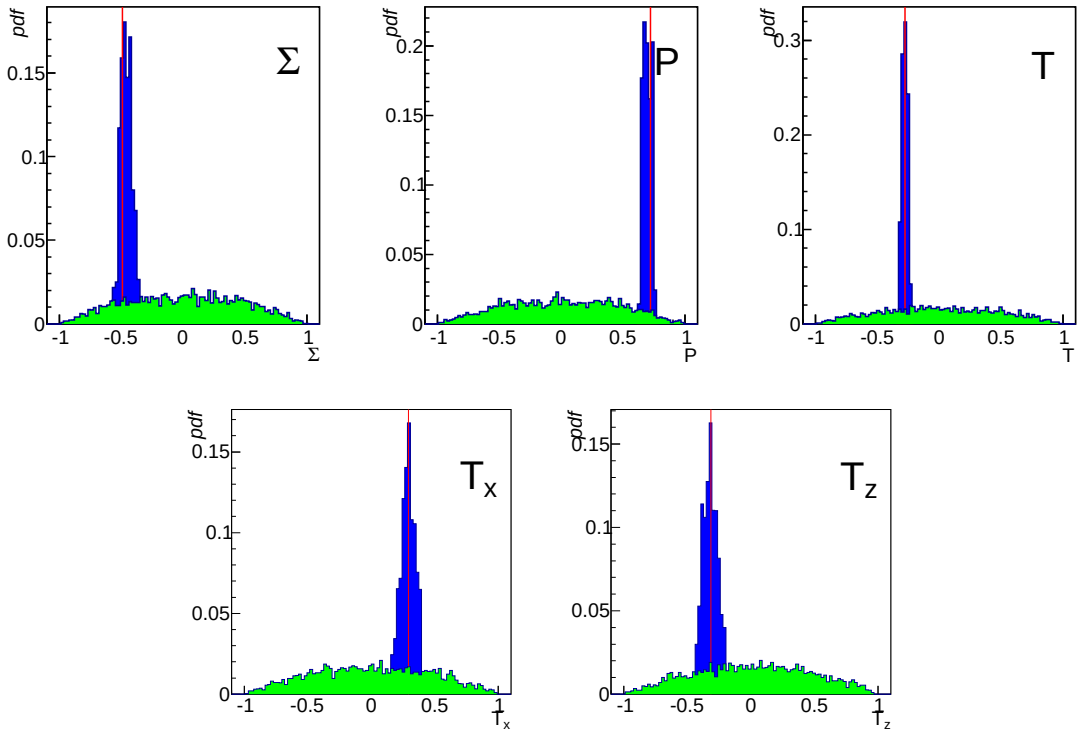


Figure 3.18: A configuration consisting of an unpolarised beam, transversely polarised target and recoil information yields Σ, P, T, T_x, T_z.

3.7 Double Polarisation Analysis

The cases where both the photon beam and target were polarised resulted in a more complicated form of likelihood function. In these cases, at least two asymmetry equations were used. Multiple asymmetries were combined by adding together each asymmetry's respective log-likelihood value. Similar to the single polarisation cases, the analyses for each possible configuration were written and tested. The resulting plots are shown in Figures 3.19 - 3.26.

Table 3.6: Double polarisation configurations

Configuration	Beam	Recoil	Target
$B_C T_L R_N$	Circular	No	Longitudinal
$B_C T_L R_Y$	Circular	Yes	Longitudinal
$B_C T_T R_N$	Circular	No	Transverse
$B_L T_L R_N$	Linear	No	Longitudinal
$B_C T_T R_Y$	Circular	Yes	Transverse
$B_L T_L R_Y$	Linear	Yes	Longitudinal
$B_L T_T R_N$	Linear	No	Transverse
$B_L T_T R_Y$	Linear	Yes	Transverse

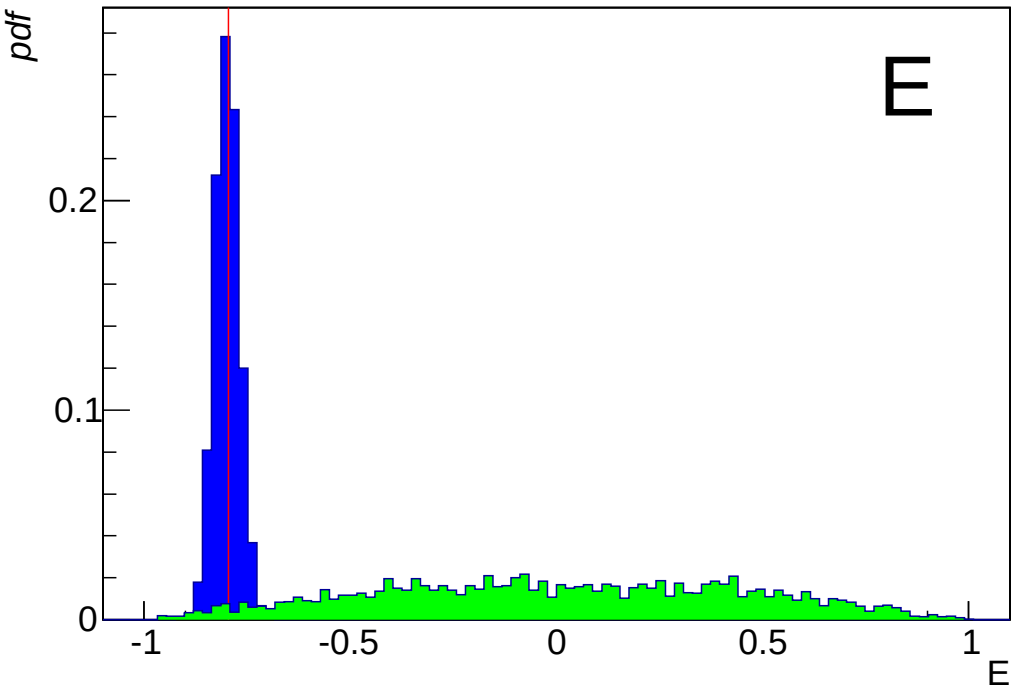


Figure 3.19: With a circularly polarised beam, longitudinally polarised target and no recoil information, E is the only measurable observable.

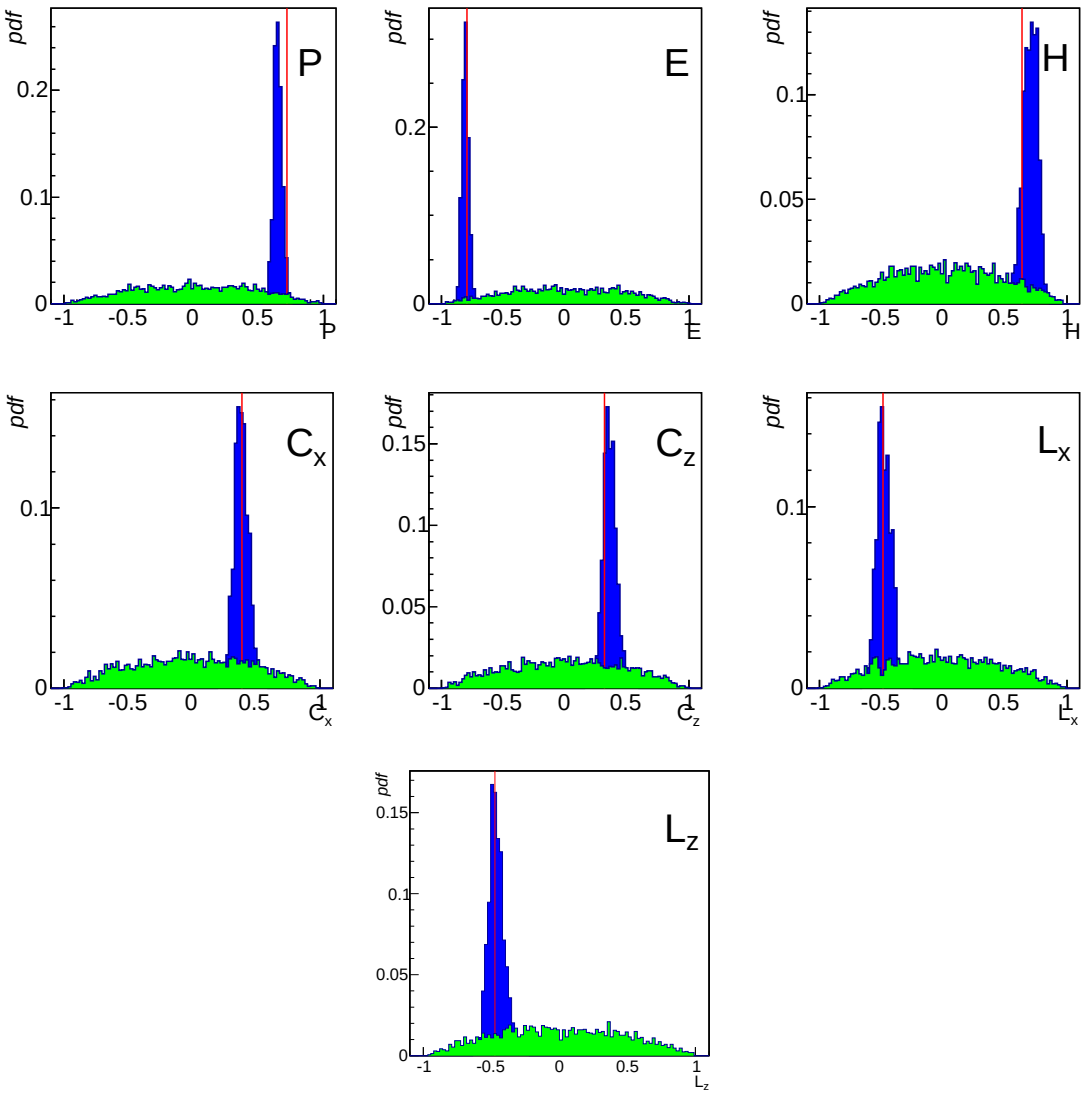


Figure 3.20: When recoil information is included with a circularly polarised beam and longitudinally polarised target, P, E, H, C_x, C_z, L_x and L_z can be extracted.

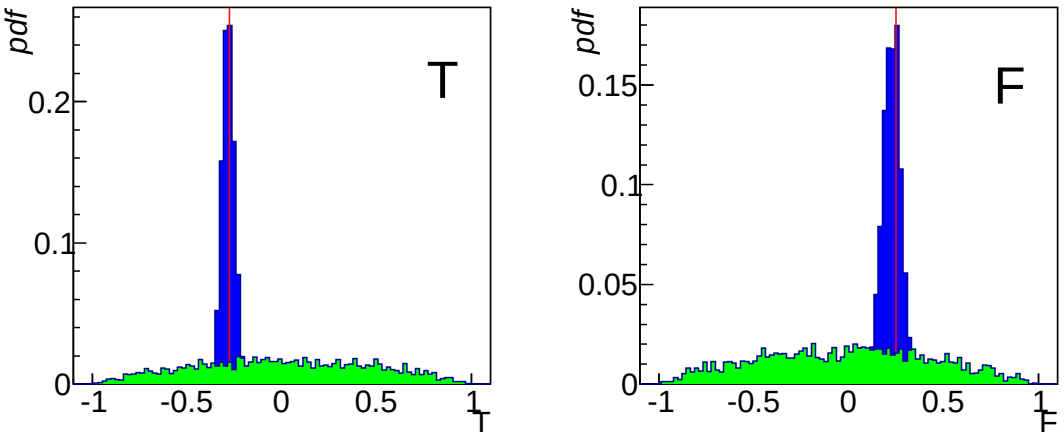


Figure 3.21: A circularly polarised beam, transversely polarised target with no recoil information yields T and F.

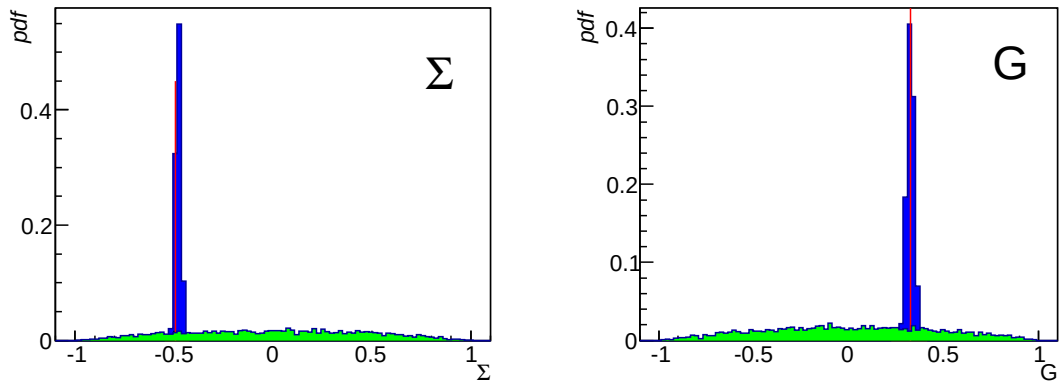


Figure 3.22: A linearly polarised beam and longitudinally polarised target configuration with no recoil information is sensitive to Σ and G.

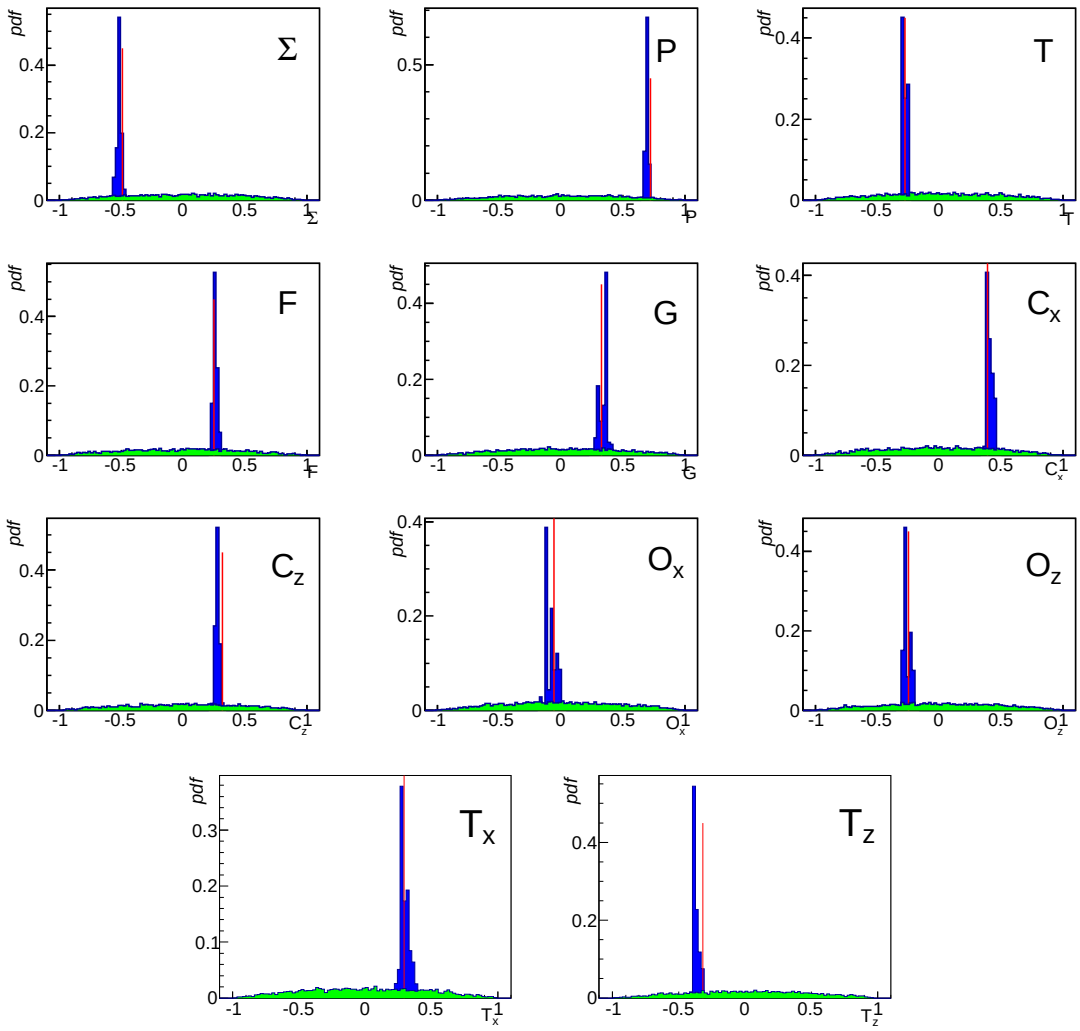


Figure 3.23: With a circularly polarised beam, transversely polarised target and recoil information, it is possible to measure Σ, P, T, F, G, C_x, C_z, O_x, O_z, T_x and T_z.

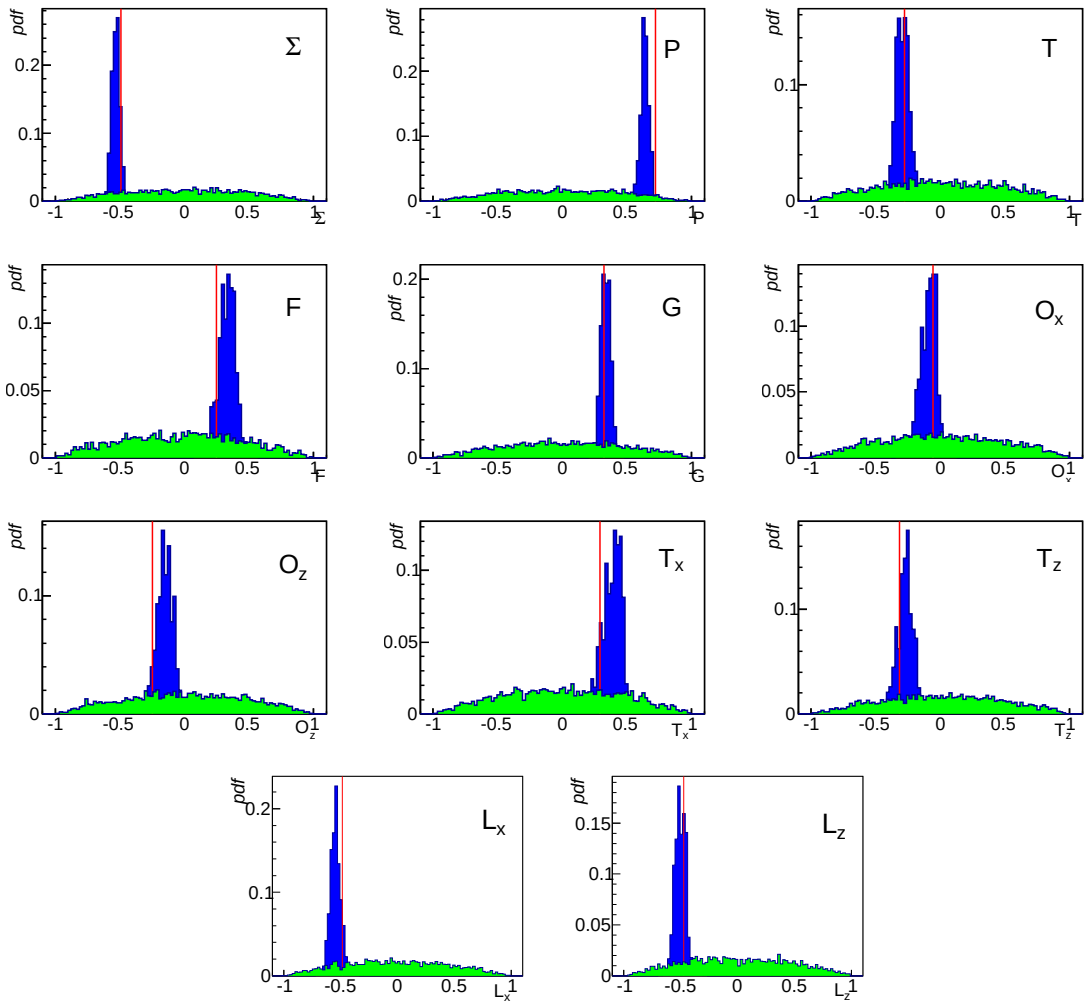


Figure 3.24: Using a linearly polarised beam, longitudinally polarised target and including recoil information, Σ, P, T, F, G, O_x, O_z, T_x, T_z, L_x and L_z can be extracted.

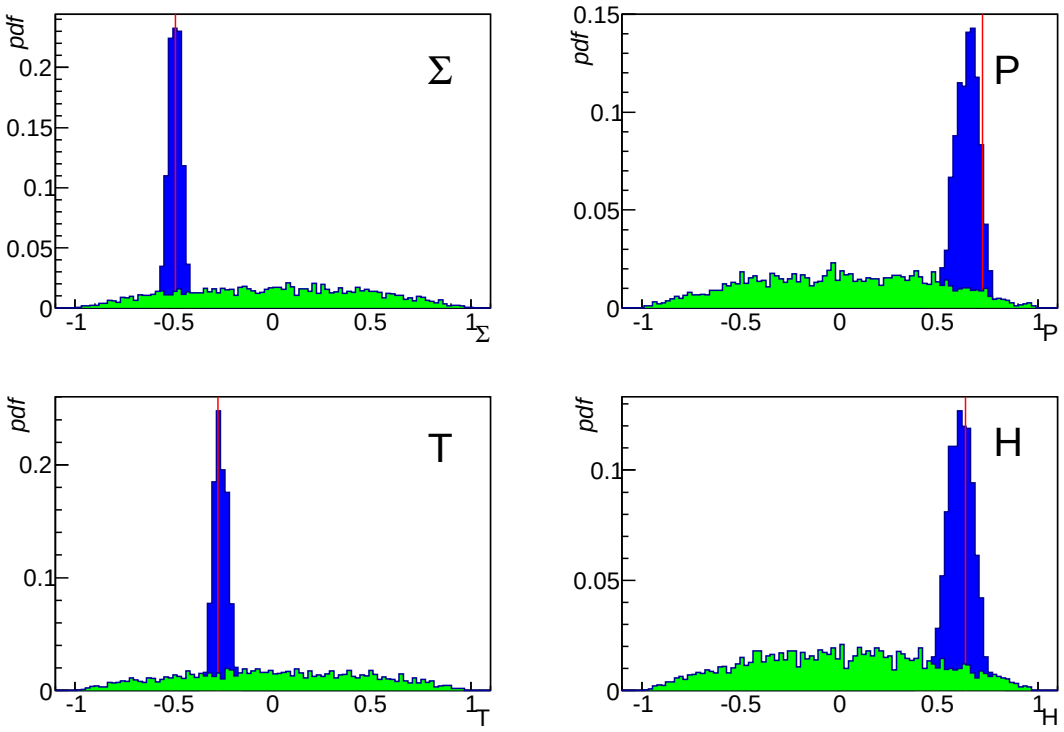


Figure 3.25: A linearly polarised beam and transversely polarised target configuration with no recoil information yields Σ, P, T and H.

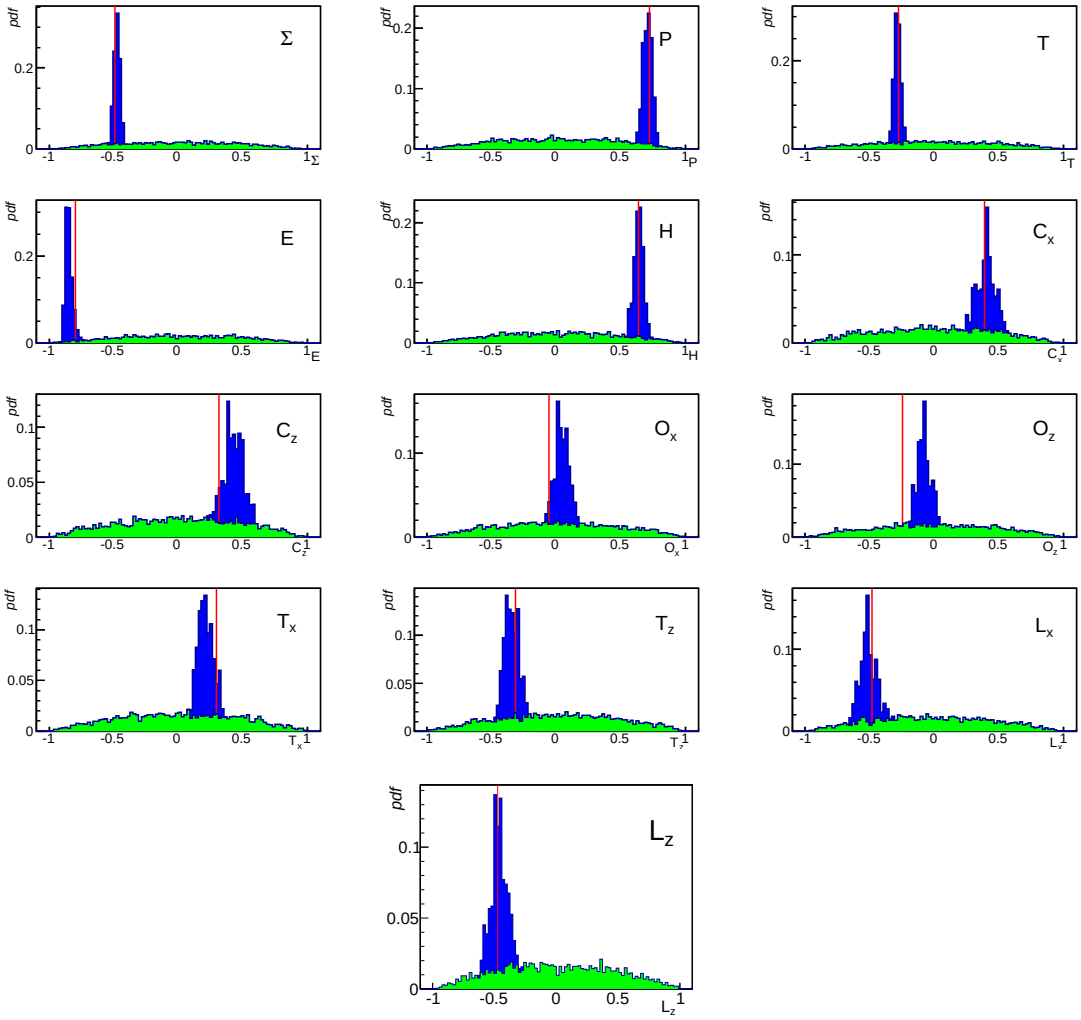


Figure 3.26: With a linearly polarised beam, transversely polarised target and recoil information, it is possible to extract Σ, P, T, E, H, C_x, C_z, O_x, O_z, T_x, T_z, L_x and L_z.

3.8 Combining Data from Multiple Experiments

One of the most unique features of the nested sampling analysis programme is its ability to analyse data from multiple experiments simultaneously and in a completely consistent manner. This significant advantage arises because of the exploration of amplitude space, rather than observable space, and because of the event-by-event element involved in the algorithm. The incorporation of data from multiple experiments is surprisingly simple, both conceptually and in terms of programming implementation. All that is required is an additional data term for each event in the data file(s) indicating which experiment produced it and likelihood equations for each experiment involved. When the likelihood method is called and the events are looped over, this additional datum will act as an indicator selecting the appropriate likelihood equation. Any background effects, systematic variables or experiment-dependent constant multipliers can be included in the particular likelihood equation.

By combining data from different experiments, extracting polarisation observables in such a consistent manner and exploring the amplitude space, the amount of information extracted can be significantly increased. It is possible to constrain possible values of other observables (those to which none of the included experiments are sensitive). This surprising result is illustrated in the examples that follow. Using simulated data, this principle was first tested with data from a linearly polarised photon beam and unpolarised target configuration with recoil information ($B_L T_U R_Y$), and a circularly polarised photon beam with unpolarised target configuration with recoil information ($B_C T_U R_Y$). The second example features simulated data from a linearly polarised photon beam - unpolarised target configuration with recoil information ($B_L T_U R_Y$) with simulated data of a linearly polarised photon beam and longitudinally polarised target configuration, with no recoil information ($B_L T_L R_N$).

3.8.1 Combining $B_L T_U R_Y$ and $B_C T_U R_Y$

It was known at the time of development that real data from $B_L T_U R_Y$ and $B_C T_U R_Y$ experiments would be available for analysis. Based on this, these two types of experiment were initially chosen to develop and test the analysis of combined datasets. The first experiment, involving a linearly polarised photon beam, unpolarised target, and recoil information, is sensitive to five polarisation observables: the beam asymmetry observable Σ, recoil observable P, target observable T, and two beam-recoil double polarisation observables, O_x and O_z. The second experiment type consisted of a circularly polarised photon beam, unpolarised target and recoil information. It can be used to directly measure three observables: the recoil observable P and another two beam-recoil double polarisation observables, C_x and C_z. Prior to the datasets for both experiments being combined, they were separately analysed and plotted to include the information known for all seven measurable observables. Figure 3.27 shows that the observables measurable by the $B_L T_U R_Y$ experiment are well resolved, but C_x and C_z are not determined. The amplitude posterior distributions after analysing only the $B_L T_U R_Y$ experiment are shown in Figure 3.28. Similarly, Figure 3.29 shows the results of analysing only the simulated data from the $B_C T_U R_Y$ experiment, and it can be seen that only those three observables to which it is sensitive can be extracted. The posterior distributions of the remaining four observables are not conclusive. The resulting plots of the complex amplitudes after analysing only the $B_C T_U R_Y$ experiment are given in Figure 3.30.

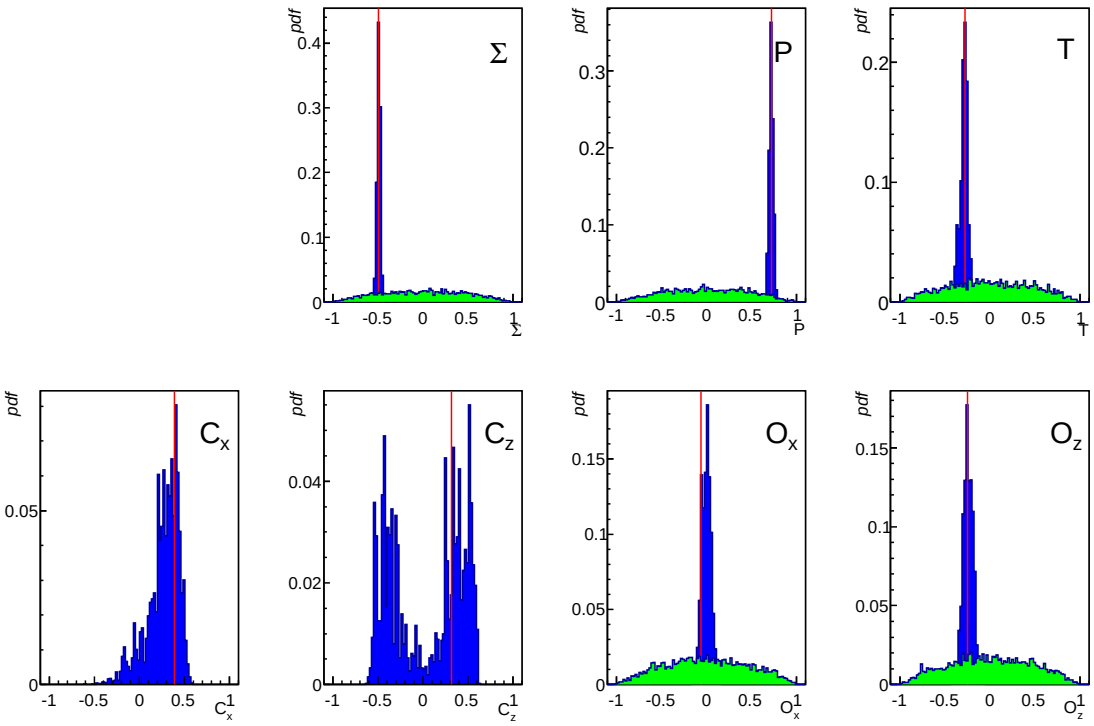


Figure 3.27: Using only simulated data from a $B_L T_U R_Y$ experiment, the above observables (Σ, P, T, C_x, C_z, O_x and O_z) were extracted. This set-up is not sensitive to C_x and C_z. In each of the above plots, the prior distribution is plotted (green), the posterior distribution is shown (blue), with a red line indicating the input value of the observable in the event generator.

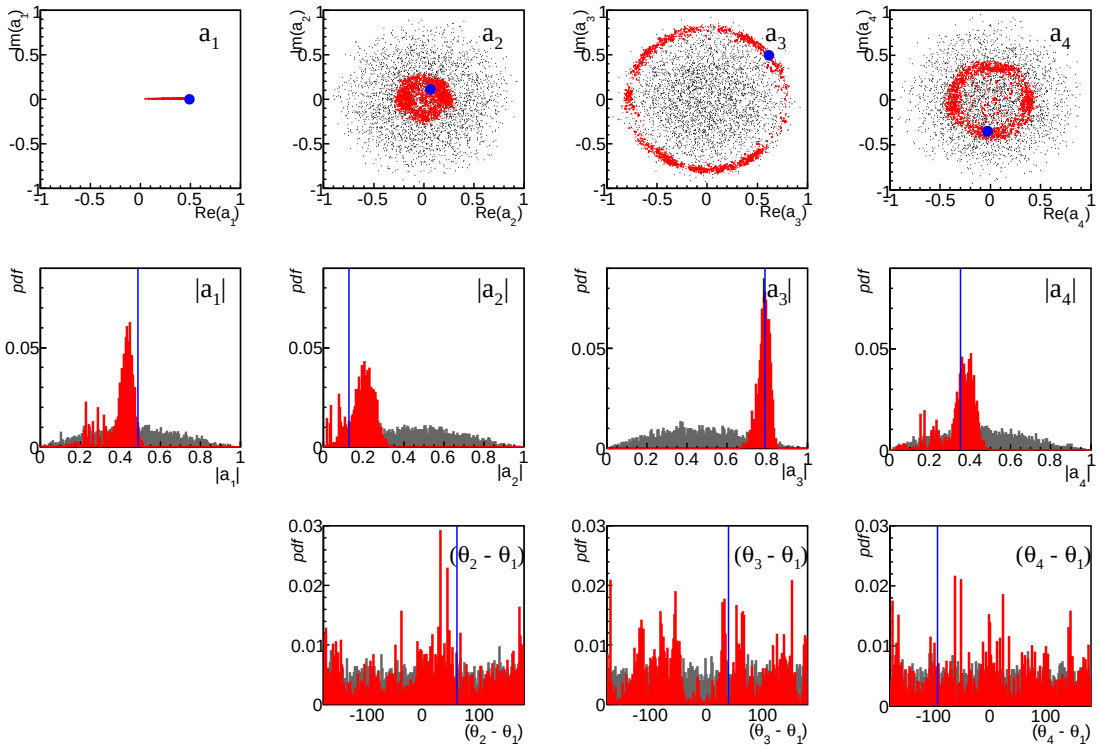


Figure 3.28: Posterior distributions of the complex amplitudes, after analysis of simulated data of a $B_L T_U R_Y$ experiment. The first row shows an Argand plot for each amplitude, a_i, relative to the a_i. The black scatter plots represent the prior distributions used, and the red scatter plots overlaid depict the posterior distributions. The second row of plots shows the magnitude of each amplitude, and the third row shows the relative phase between θ_2 to θ_1, θ_3 to θ_1, and θ_4 to θ_1, respectively. The blue lines and points in all plots indicate the values input into the event generator.

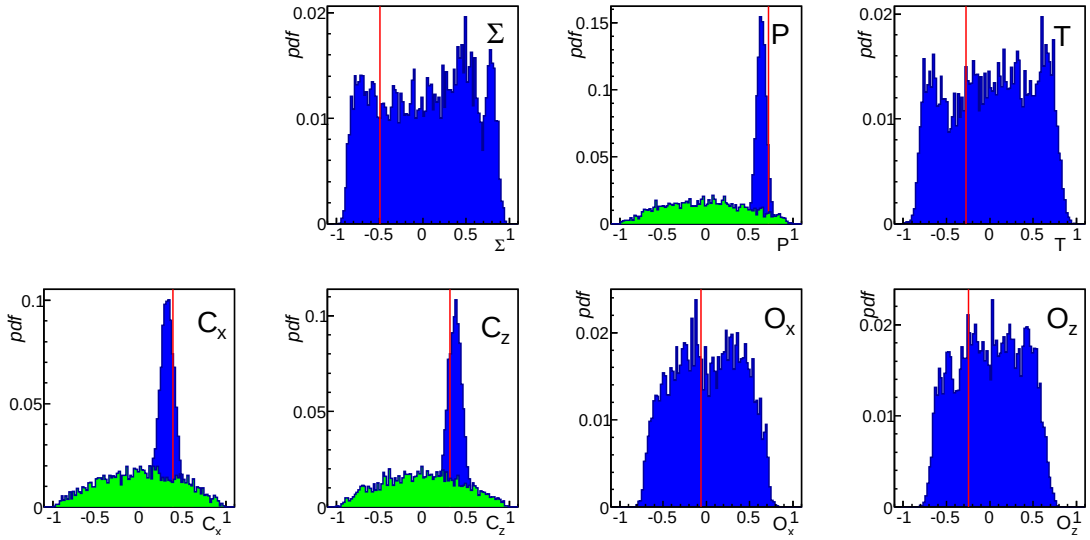


Figure 3.29: Using only simulated data from a $B_C T_U R_Y$ experiment, the above observables (Σ, P, T, C_x, C_z, O_x and O_z) were extracted. This set-up is only sensitive to C_x and C_z. In each of the above plots, the prior distribution is plotted (green), the posterior distribution is shown (blue), with a red line indicating the input value of the observable in the event generator.

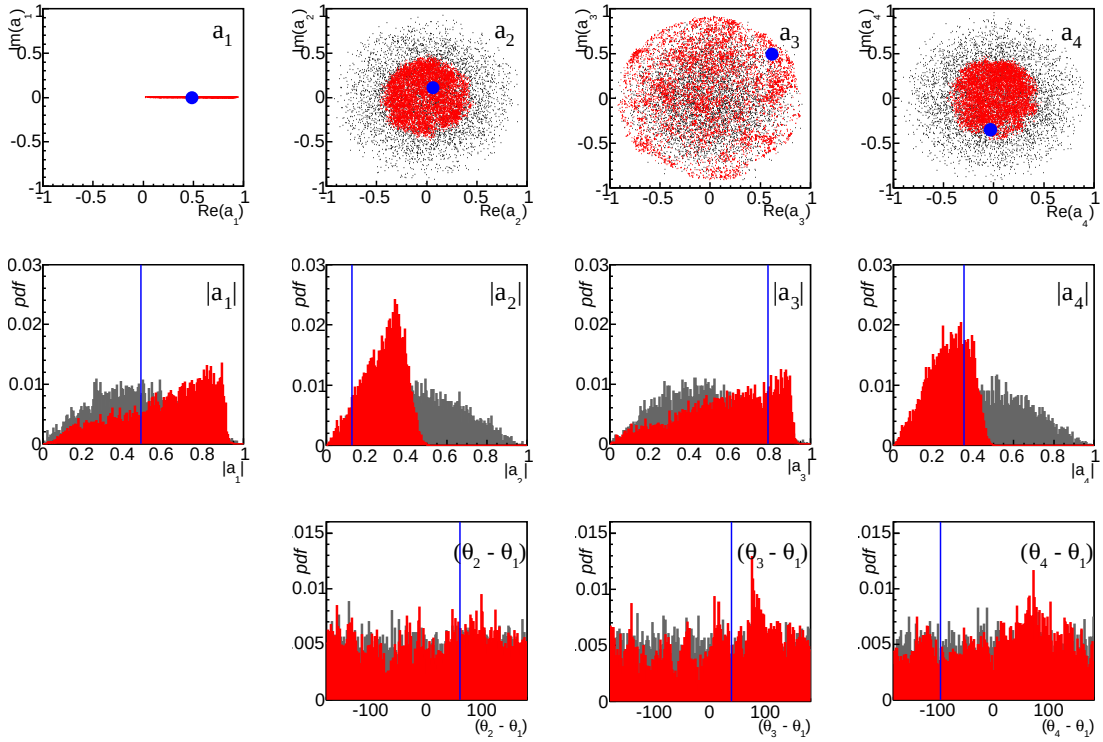


Figure 3.30: Posterior distributions of the complex amplitudes after analysis of simulated data of a $B_C T_U R_Y$ experiment. These plots follow the conventions described in Figure 3.28.

When the datasets are combined, it is possible to simultaneously extract all seven observables successfully. These results can be considered as a successful proof-of-concept.

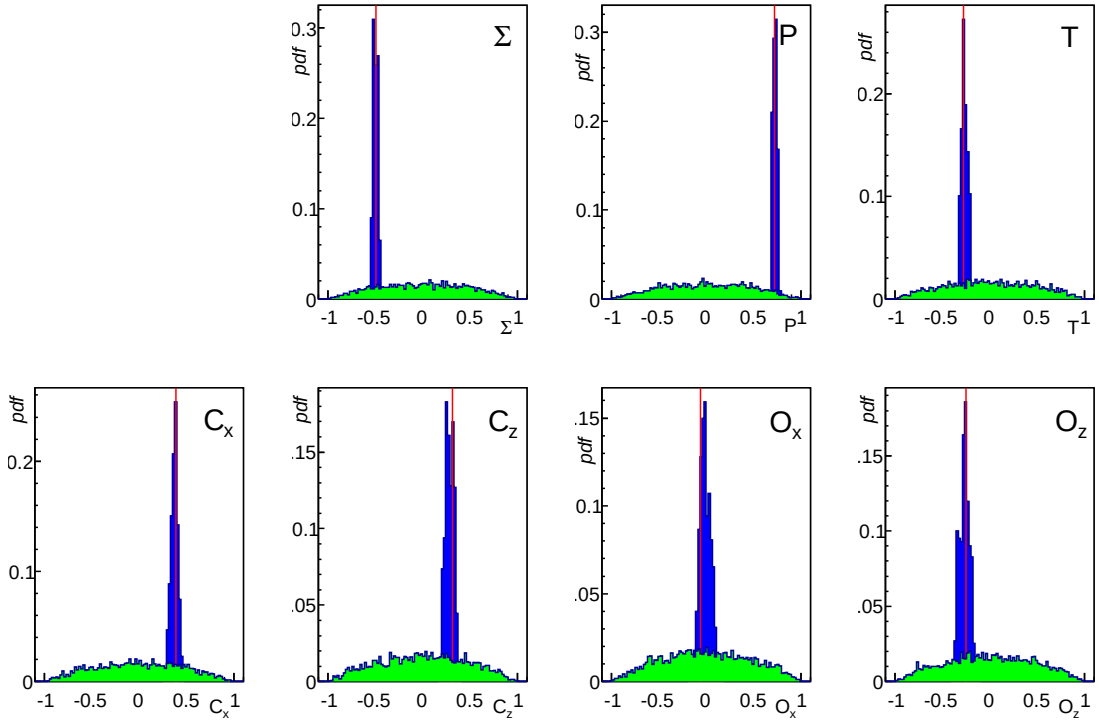


Figure 3.31: Using simulated data from both the $B_L T_U R_Y$ experiment and the $B_C T_U R_Y$ experiment, all seven observables (Σ, P, T, C_x, C_z, O_x and O_z) were extracted. In each of the above plots, the prior distribution is plotted (green), the posterior distribution is shown (blue), with a red line indicating the input value of the observable in the event generator.

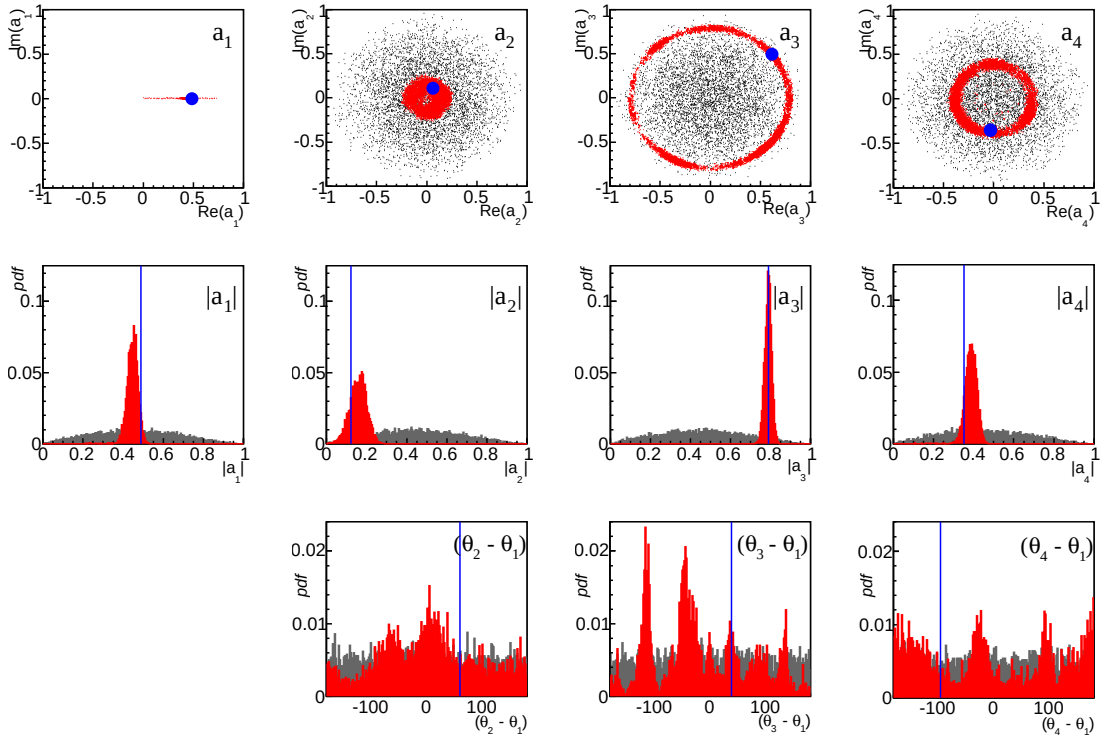


Figure 3.32: Posterior distributions of the complex amplitudes after combined analysis of simulated data of $B_C T_U R_Y$ and $B_L T_U R_Y$ experiments. These plots follow the conventions described in Figure 3.28.

The resulting amplitude plot, Figure 3.32, is further constrained with the addition of the $B_C T_U R_Y$ data.

3.8.2 Combining $B_L T_U R_Y$ and $B_L T_L R_N$

The further development of a combined analysis programme was focused on the inclusion of a polarised-target experiment. The $B_L T_L R_N$ type of experiment featuring a linearly polarised photon beam incident on a longitudinally polarised target and no recoil information, and provides a means to measure the beam asymmetry observable Σ and the beam-target double-polarisation observable G. Data simulated for a $B_L T_U R_Y$ type of experiment were analysed as described earlier in this chapter, and plotted to include the results for the G observable. This is given in Figure 3.33.

Figure 3.33 shows that even without the inclusion of any target polarisation, the use of amplitude space and observable correlations can provide access to a significant amount of information about observables that are not directly measurable. It was found that there was already a considerable amount of information concerning the value of the G observable. This was due to the nature of the analysis approach and the preservation of correlations between observables.

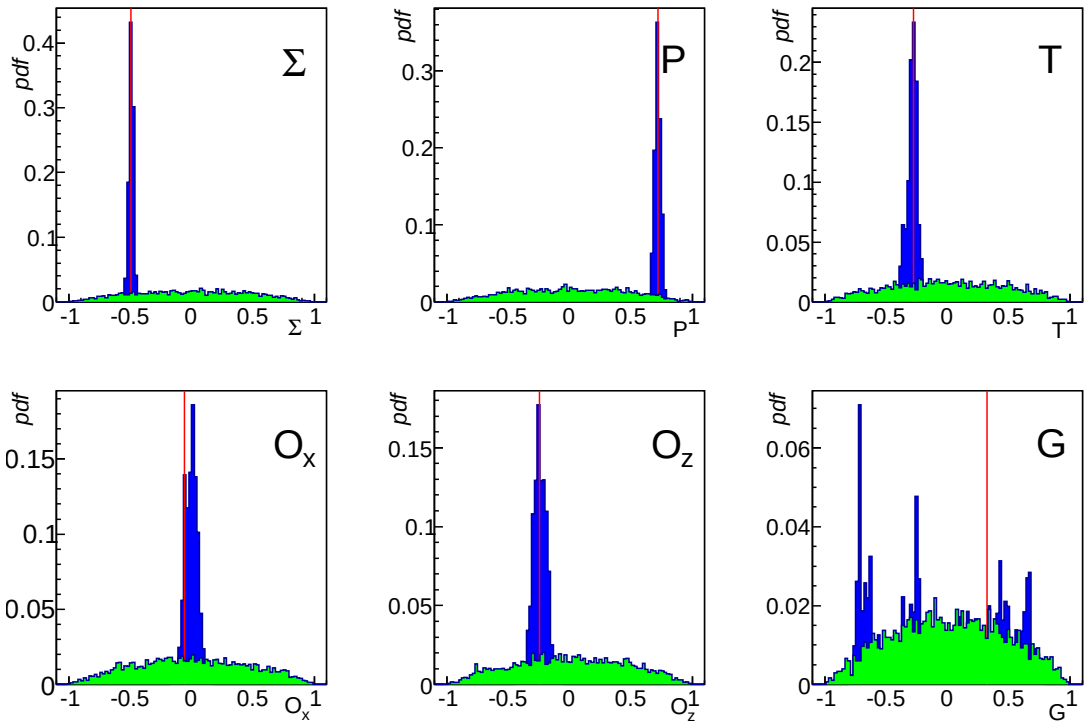


Figure 3.33: Using only simulated data from a $B_L T_U R_Y$ type of set-up, the above observables (Σ, P, T, O_x, O_z and G) were extracted. There is a clear ambiguity in the sign of G, an observable to which this experimental configuration is not directly sensitive. In each of the above plots, the prior distribution is plotted (green), the posterior distribution is shown (blue), with a red line indicating the input value of the observable into the event generator.

Similarly, simulated data from a $B_L T_L R_N$ style experiment was analysed and the posterior distributions of P, T, O_x and O_z were included in the results. It can be seen in Figure 3.34 that no conclusions can be drawn for the values of these observables from this experiment alone. The two observables to which the $B_L T_L R_N$

experiment is sensitive are well resolved in this analysis. The posterior distributions for the complex amplitudes were also plotted as they were in Section 3.8.1, resulting from the analysis of the simulated $B_L T_L R_N$ data. This is shown in Figure 3.35.

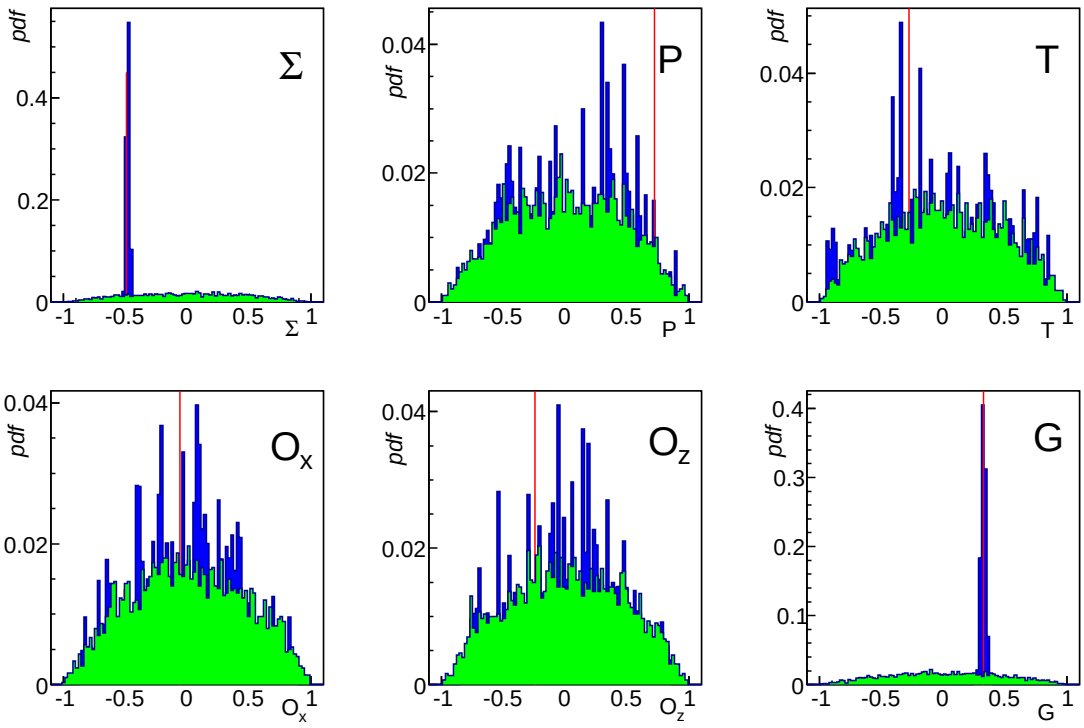


Figure 3.34: Using only simulated data from a $B_L T_L R_N$ style of set-up, the above observables (Σ, P, T, O_x, O_z and G) were extracted. This configuration is only sensitive to Σ and G, and very little information is found of the other observables shown here. In each of the above plots, the prior distribution is plotted (green), the posterior distribution is shown (blue), with a red line indicating the input value of the observable into the event generator.

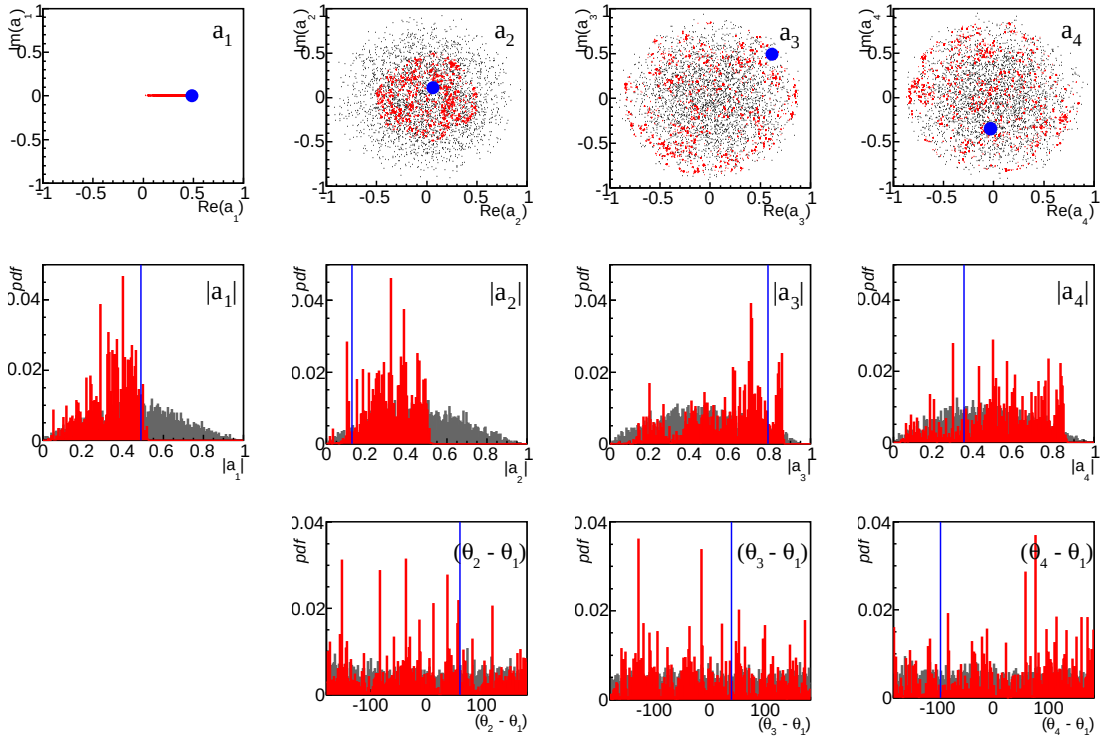


Figure 3.35: Posterior distributions of the complex amplitudes after the analysis of simulated data of a $B_L T_L R_N$ experiment. These plots follow the conventions described in Figure 3.28.

The simulated data from both experiments was amended to include an additional term, a flag term, that was used to select the appropriate likelihood equation for each event. The two data files were then concatenated and analysed together. The result of this simple concatenation approach is shown in Figure 3.36. The resulting amplitude posterior distributions are also shown, in Figure 3.37.

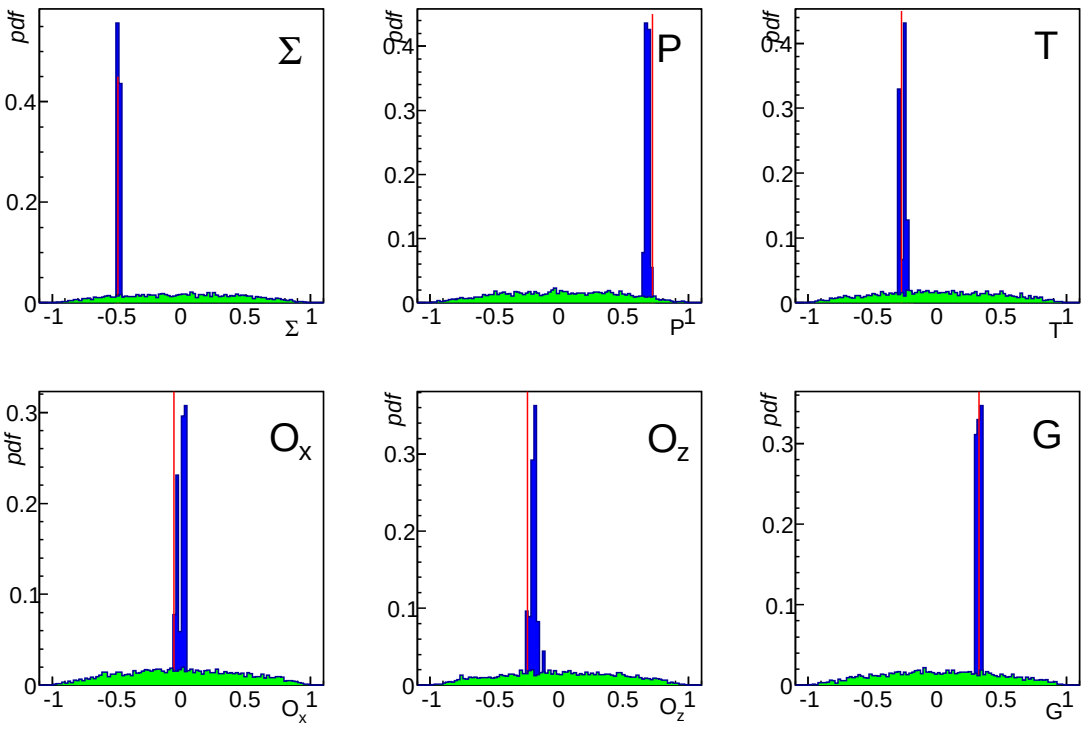


Figure 3.36: Using simulated data from both $B_L T_U R_Y$ and $B_L T_L R_N$ set-ups, the above observables (Σ, P, T, O_x, O_z and G) were extracted. Information on all observables to which at least one of the configurations is sensitive can be found. In each of the above plots, the prior distribution is plotted (green), the posterior distribution is shown (blue), with a red line indicating the input value of the observable into the event generator.

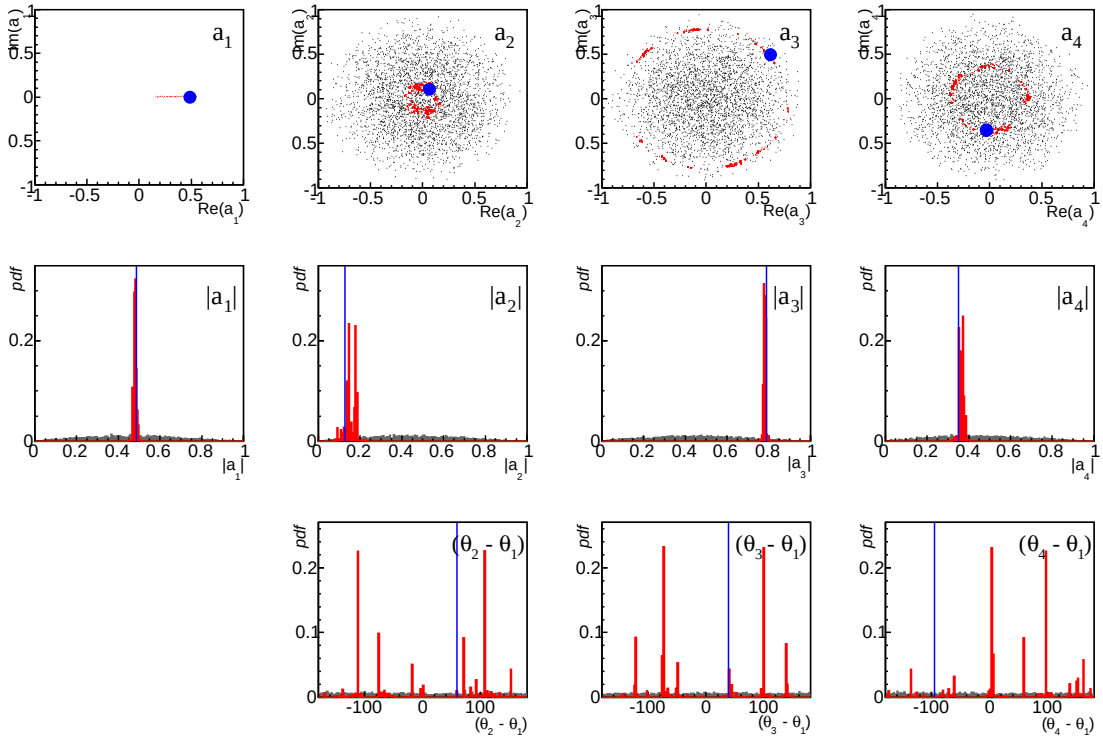


Figure 3.37: Posterior distributions of the complex amplitudes after the combined analysis of simulated data of $B_L T_L R_N$ and $B_L T_U R_Y$ experiments. These plots follow the conventions described in Figure 3.28.

This method proved to be effective for combining datasets from different experimental configurations and experimental runs and illustrated the wealth of information that can be shared from one experiment to another. It can be seen by the amplitude plots, Figures 3.28, 3.35 and 3.37, that the combined analysis of these two experiments further constrains the complex amplitudes, indicating an overall gain in information yield.

3.9 Summary

In this chapter, the development of the nested sampling was described. Currently used methods, including a binned χ^2 fitting method and a maximum likelihood approach, were discussed and compared to the Bayesian programme presented here. Neither method currently in use maximise the amount of information that can be

obtained from existing data, and this was the fundamental motivation for the development of a new method.

The prior distribution used in nested sampling analyses of pseudoscalar meson photoproduction experiments was chosen to be is amplitude space, which were then mapped to observable space for use in the likelihood function. Early benchmarking of this approach indicated that the nested sampling programme produced more accurate and precise results compared to those obtained by the more common binned fitting method, even for datasets with low statistics.

The derivation of the likelihood functions for all possible experimental set-ups was described and a table of all functions was given. These likelihood equations were based on asymmetry expressions between different polarisation states of the beam, target, or both. A luminosity asymmetry term was introduced to the equations to account for discrepancies between the number of events in one polarisation over another. These likelihood functions were tested with simulated data for each possible experiment type, and the results were successful.

The exploration of amplitude space rather than observable space ensured the preservation of constraints from observable correlations. It also allowed multiple experiments to be analysed together, by concatenating the data files. Two examples of this using simulated data were described, and it was shown that the concept of merging data files and analysing two independent experiments simultaneously produces results that are at least consistent with the results obtained by analysing each experiment individually. The amplitude plots produced by separate analyses and a combined analysis illustrated that a combined analysis resulted in more constrained complex amplitudes. It was shown that this method of merging data files and analysing multiple experiments simultaneously increases the amount of information extracted from a data set.

4

Analysis and Results of the g8b and g1c experiments

This chapter features the analysis of two experiments performed at CLAS. The g8b experiment consisted of a linearly polarised photon beam with an unpolarised proton target, and the g1c experiment was comprised of a circularly polarised photon beam, also incident on an unpolarised proton target. For each experiment, events from the $K^+\Lambda^0$ reaction channel were analysed. The results of each separate analysis are compared to results previously obtained and published, as well as to each other, with a compatible binning scheme employed. The aim of this chapter is to illustrate the effectiveness of the nested sampling method in obtaining results from real data and showing their consistency with accepted values, while ensuring that the observable values are permitted by the set of complex amplitudes.

Prior to the nested sampling programme being used to analyse data, the correct

events must be identified. This is done by identifying final-state particles and selecting events that match the desired reaction. For the $K\Lambda$ reaction channel, the final state particles required are a pion, a kaon and a proton. The missing mass distribution for accepted events is shown in Figure 4.1. The two prominent peaks show the two hyperons, Λ and Σ, with this final state. Further cuts are then made to select only Λ events. The full particle identification and event selection process is detailed in [75].

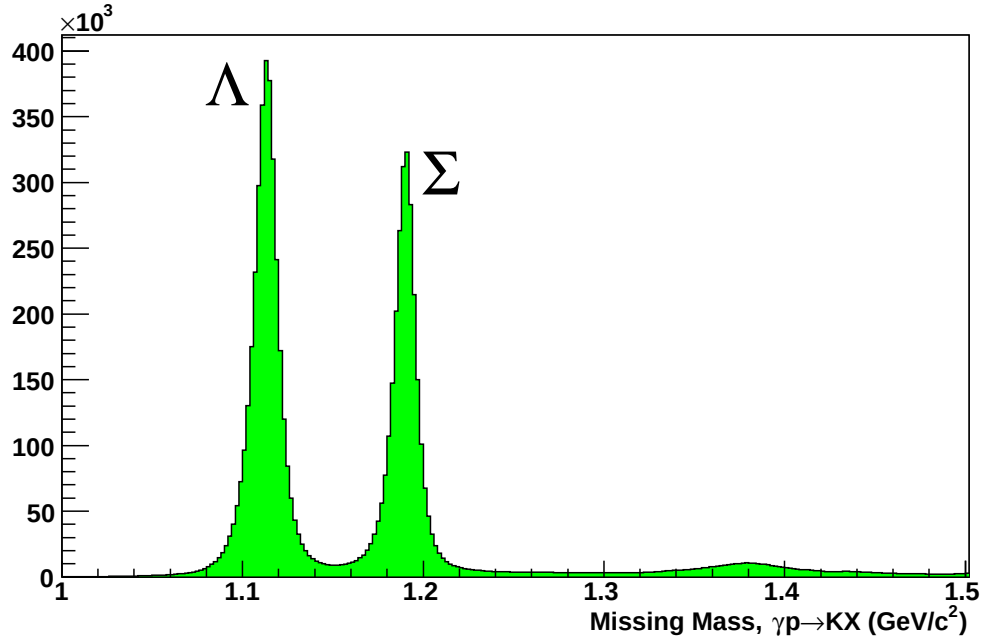


Figure 4.1: Plot of missing mass, used to identify hyperons [87].

4.1 The g8b experiment

The g8b experiment was performed in Hall B at Jefferson Lab between June 20th and September 1st, 2005 [75]. The experiment involved a beam of linearly polarised photons incident on a stationary proton target of liquid hydrogen. An electron beam with an energy of 4.551 GeV was scattered using a diamond radiator, which produced the polarised photon beam with an energy ranging between 1.71 GeV and 2.19 GeV through the process of coherent bremsstrahlung, covering a $-0.75 < \cos(\theta_K^{CM}) <$

0.85 angular region. The reaction $\overrightarrow{\gamma}p \rightarrow K^+\Lambda^0$ was selected for analysis. Prior to the nested sampling programme being implemented, it should be noted that all particle identification had already been performed [75]. All analysis was performed by first sorting the data into smaller bins defined by the baryon mass (W) and angular region ($\cos(\theta_K^{CM})$) of the kaon polar angle in the centre of mass frame. This division of data was chosen based on previous analyses of the same data. This experiment is sensitive to five polarisation observables: the beam asymmetry observable (Σ), the recoil observable (P), the target observable (T), and two beam-recoil double polarisation observables (O_x and O_z).

4.1.1 Binning Scheme

The data were binned into smaller files corresponding to a region of baryon mass and kaon polar angle. It was decided to employ an identical binning scheme as was used in a previous analysis [75], as this would allow for the most direct comparison of results. The bin boundaries are given in Table 4.1. Although this binning scheme had been previously used, it was important to verify that the data populated each bin sufficiently. A two-dimensional plot of baryon mass W and kaon polar angle $\cos(\theta_K^{CM})$ of all data was created with the bin boundaries superimposed to illustrate this. This plot is shown in Figure 4.2. It can be seen in this figure that all bins are indeed sufficiently populated.

4.1.2 Luminosity Asymmetries

The likelihood functions developed for the nested sampling programme are derived from asymmetries that assume an equal number of events in each beam polarisation setting (parallel and perpendicular). The experimental data, however, often featured a disproportionate number of events in one setting. This can be taken into account by introducing a luminosity asymmetry term, ΔL_L, into the likelihood function as described in Section 3.5. Each luminosity asymmetry was calculated by performing a nested sampling fit across all events in a W bin (across the whole range

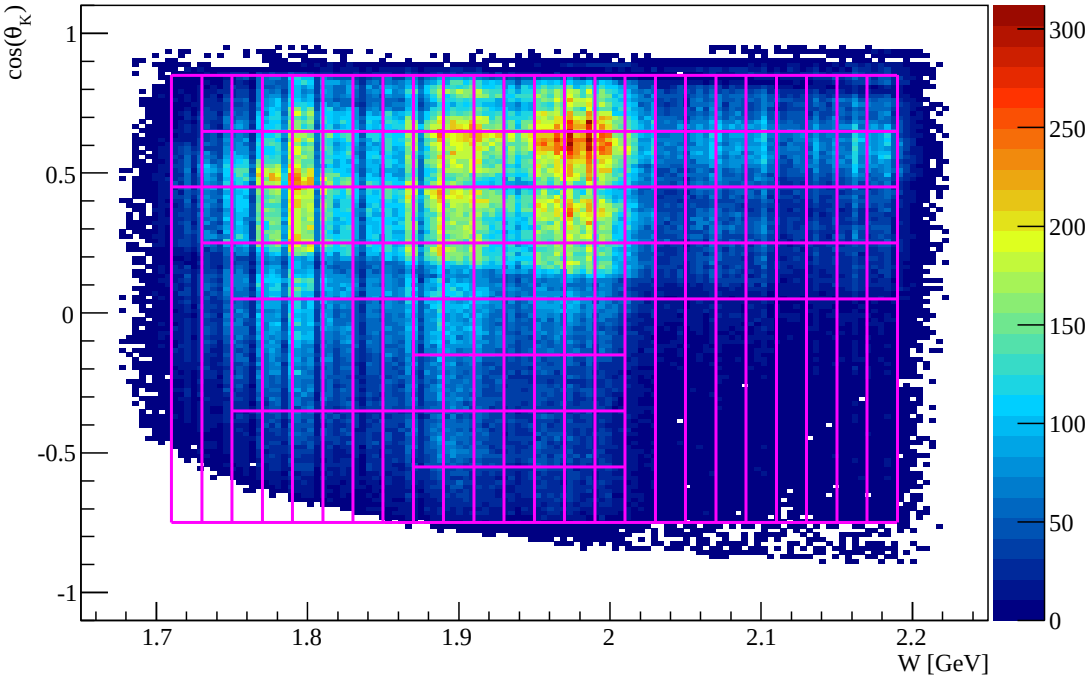


Figure 4.2: Distribution of all data used for g8b analysis. The magenta lines indicate the bin boundaries.

Table 4.1: Bin definitions

W bin (GeV)	$\cos(\theta_K)$ bin boundaries
(1.71, 1.73)	-0.75, 0.05, 0.45, 0.85
(1.73, 1.75)	-0.75, 0.05, 0.25, 0.45, 0.65, 0.85
(1.75, 1.77)	-0.75, -0.35, 0.05, 0.25, 0.45, 0.65, 0.85
(1.77, 1.79)	-0.75, -0.35, 0.05, 0.25, 0.45, 0.65, 0.85
(1.79, 1.81)	-0.75, -0.35, 0.05, 0.25, 0.45, 0.65, 0.85
(1.81, 1.83)	-0.75, -0.35, 0.05, 0.25, 0.45, 0.65, 0.85
(1.83, 1.85)	-0.75, -0.35, 0.05, 0.25, 0.45, 0.65, 0.85
(1.85, 1.87)	-0.75, -0.35, 0.05, 0.25, 0.45, 0.65, 0.85
(1.87, 1.89)	-0.75, -0.55, -0.35, -0.15, 0.05, 0.25, 0.45, 0.65, 0.85
(1.89, 1.91)	-0.75, -0.55, -0.35, -0.15, 0.05, 0.25, 0.45, 0.65, 0.85
(1.91, 1.93)	-0.75, -0.55, -0.35, -0.15, 0.05, 0.25, 0.45, 0.65, 0.85
(1.93, 1.95)	-0.75, -0.55, -0.35, -0.15, 0.05, 0.25, 0.45, 0.65, 0.85
(1.95, 1.97)	-0.75, -0.55, -0.35, -0.15, 0.05, 0.25, 0.45, 0.65, 0.85
(1.97, 1.99)	-0.75, -0.55, -0.35, -0.15, 0.05, 0.25, 0.45, 0.65, 0.85
(1.99, 2.01)	-0.75, -0.55, -0.35, -0.15, 0.05, 0.25, 0.45, 0.65, 0.85
(2.01, 2.03)	-0.75, 0.05, 0.25, 0.45, 0.65, 0.85
(2.03, 2.05)	-0.75, 0.05, 0.25, 0.45, 0.65, 0.85
(2.05, 2.07)	-0.75, 0.05, 0.25, 0.45, 0.65, 0.85
(2.07, 2.09)	-0.75, 0.05, 0.25, 0.45, 0.65, 0.85
(2.09, 2.11)	-0.75, 0.05, 0.25, 0.45, 0.65, 0.85
(2.11, 2.13)	-0.75, 0.05, 0.25, 0.45, 0.65, 0.85
(2.13, 2.15)	-0.75, 0.05, 0.25, 0.45, 0.65, 0.85
(2.15, 2.17)	-0.75, 0.05, 0.25, 0.45, 0.65, 0.85
(2.17, 2.19)	-0.75, 0.05, 0.25, 0.45, 0.65, 0.85

of kaon polar angles), and allowing the ΔL_L variable to be free and unconstrained. The luminosity asymmetry is a measure of the asymmetry of photon flux and is therefore only dependent on photon energy. There is no dependence on scattering angle or beam polarisation, so is constant across all $\cos(\theta_K)$ bins in a W bin. This value was then used for the analysis of each $\cos(\theta_K)$ bin within that W bin.

Initially, the value of ΔL_L obtained was used to define a Gaussian distribution from which a random value was generated. In the nested sampling programme, the value of ΔL_L was originally allowed to vary, but this had a negligible effect on the results when compared to using the value as a constant. A constant value was therefore used, simplifying the algorithm and reducing the computation time.

The luminosity asymmetry values for each W bin were obtained through the nested sampling extraction method and plotted. The results are shown graphically in Figure 4.3.

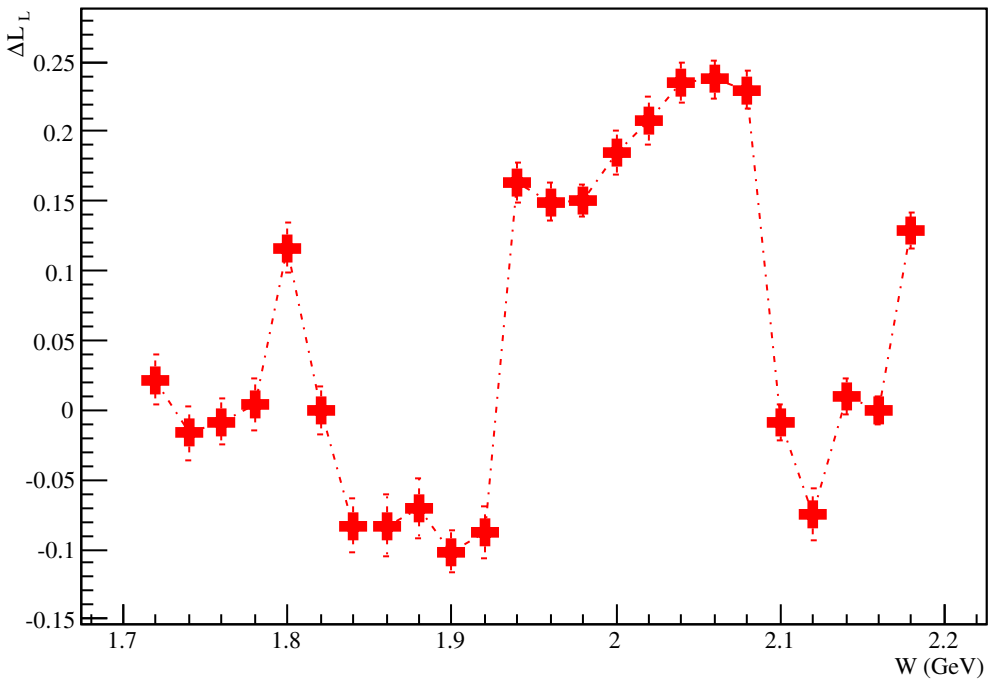


Figure 4.3: Luminosity asymmetry values obtained through nested sampling

4.1.3 Acceptance Correction

The extraction of the recoil observable, P, was not performed using the asymmetry-based likelihood functions derived in Section 3.5. This implies that acceptance terms would not be cancelled during an asymmetry calculation, and would have to be considered. The likelihood term associated with P was instead found by combining the probabilities of each event given a value of P, as shown in Equation 4.1.1.

$$Likelihood = \prod_i \Pr(\cos(\theta_{y_i}) \mid P) = \prod_i \varepsilon(\cos(\theta_{y_i}))\frac{1}{2}(1 + \alpha\cos(\theta_{y_i})P) \qquad (4.1.1)$$

where ε denotes the acceptance.

The value of acceptance used in this equation was dependent on the beam energy, W, and the kaon polar angle, $\cos(\theta_K^{CM})$ and was calculated independently of the analysis programme [75]. An event generator was used to create events, which were then passed to a simulation of the detector. Based on the physics of the detector, events were either accepted or rejected (i.e. not detected). The ratio of generated events to accepted events gives an acceptance ratio, η, for each data bin. The value of acceptance for bin i was then determined by Equation 4.1.2.

$$\varepsilon_i = \frac{1}{\eta_i} \qquad (4.1.2)$$

4.1.4 Consideration of Background

The measured events in the dataset contained a small proportion of background events in addition to the genuine signal used to extract polarisation observables. This background can be seen in Figure 4.1, under the two hyperon peaks. It has been shown that this background accounted for approximately 3% of the data [75,88], and is straightforward to take into account during analysis. It was assumed that the number of background events in the data taken when the photon beam was parallel-polarised was roughly equal to the number of background events present from a perpendicularly polarised beam. When an asymmetry of the cross-sections is taken, it can be seen that this background contribution can be treated as a small

dilution, β [75].

$$A = \frac{\sigma_\perp - \sigma_\parallel}{\sigma_\perp + \sigma_\parallel} = \frac{\sigma^e_\perp - \sigma^e_\parallel}{(1+\beta)(\sigma^e_\perp - \sigma^e_\parallel)} \tag{4.1.3}$$

where $\sigma^e_{\perp(\parallel)}$ denotes the cross-section due to true events.

The inclusion of this background fraction carries forth to the form of asymmetry equation used in the nested sampling programme's likelihood function as shown in Equation 4.1.4 below [75, 88].

$$A = \frac{P_\gamma g}{(1+\beta)f} \tag{4.1.4}$$

The value of β for each bin of data was established previously (independently) and was provided in [75].

4.1.5 Results of g8b

The results from analysing the g8b data using the nested sampling programme were compared to previous results obtained with a maximum likelihood fitting method [88]. The plots below show the comparisons for all observables that are directly measurable by g8b: Σ, P, T, O_x and O_z.

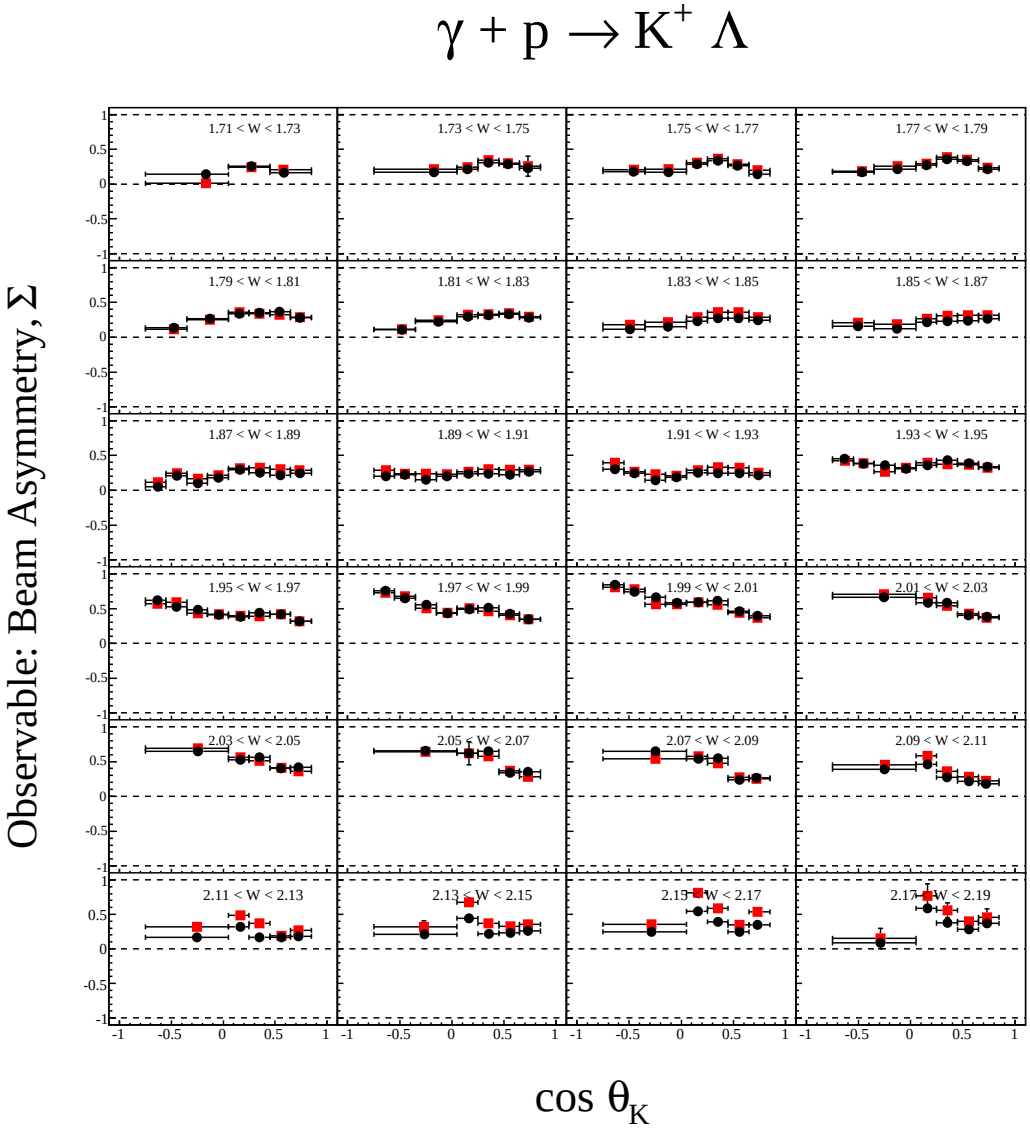


Figure 4.4: Results of beam asymmetry observable, Σ where red squares indicate previous results obtained through maximum likelihood analysis [88], and black circles show results from the nested sampling analysis.

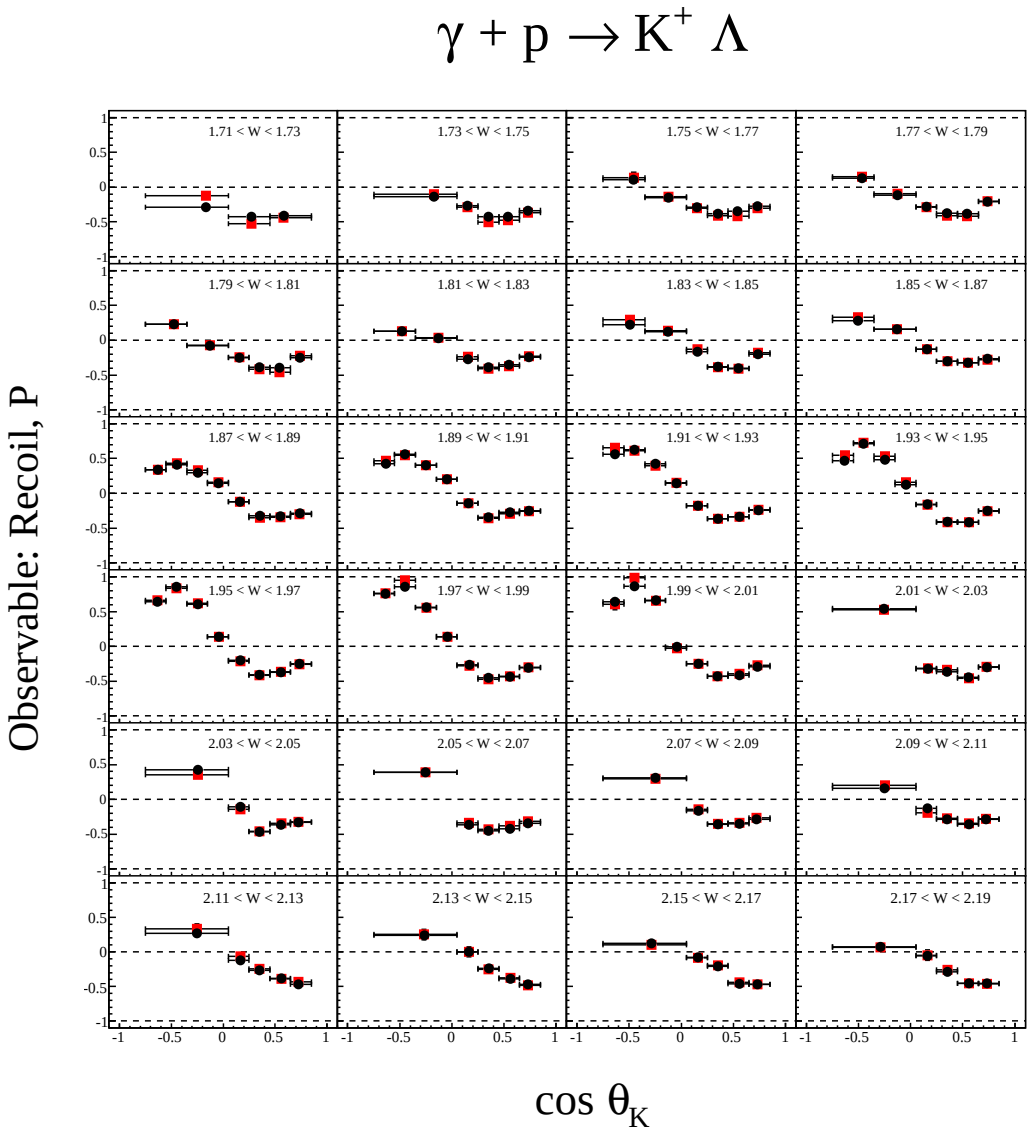


Figure 4.5: Results of recoil observable, P where red squares indicate previous results obtained through maximum likelihood analysis [88], and black circles show results from the nested sampling analysis.

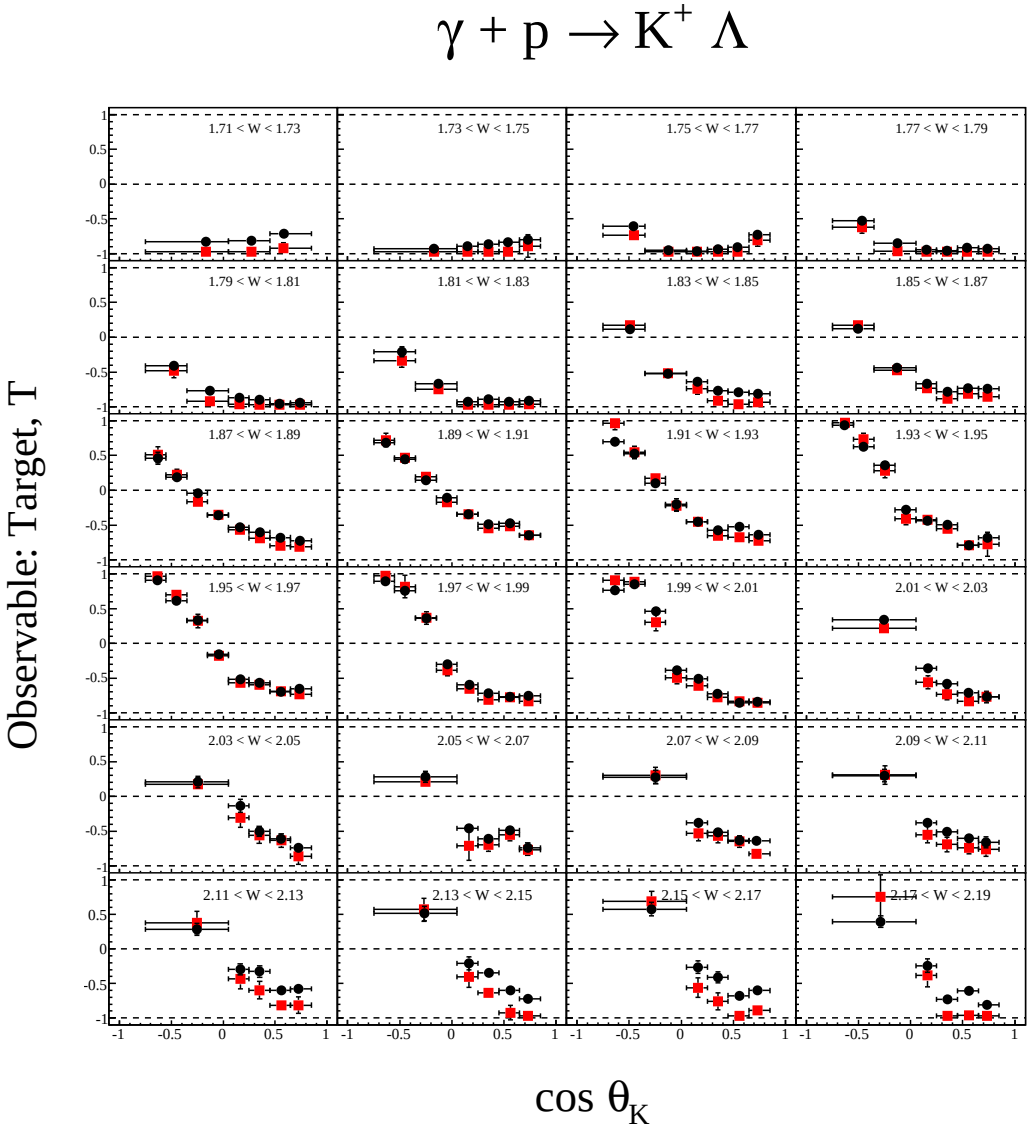


Figure 4.6: Results of target observable, T where red squares indicate previous results obtained through maximum likelihood analysis [88], and black circles show results from the nested sampling analysis.

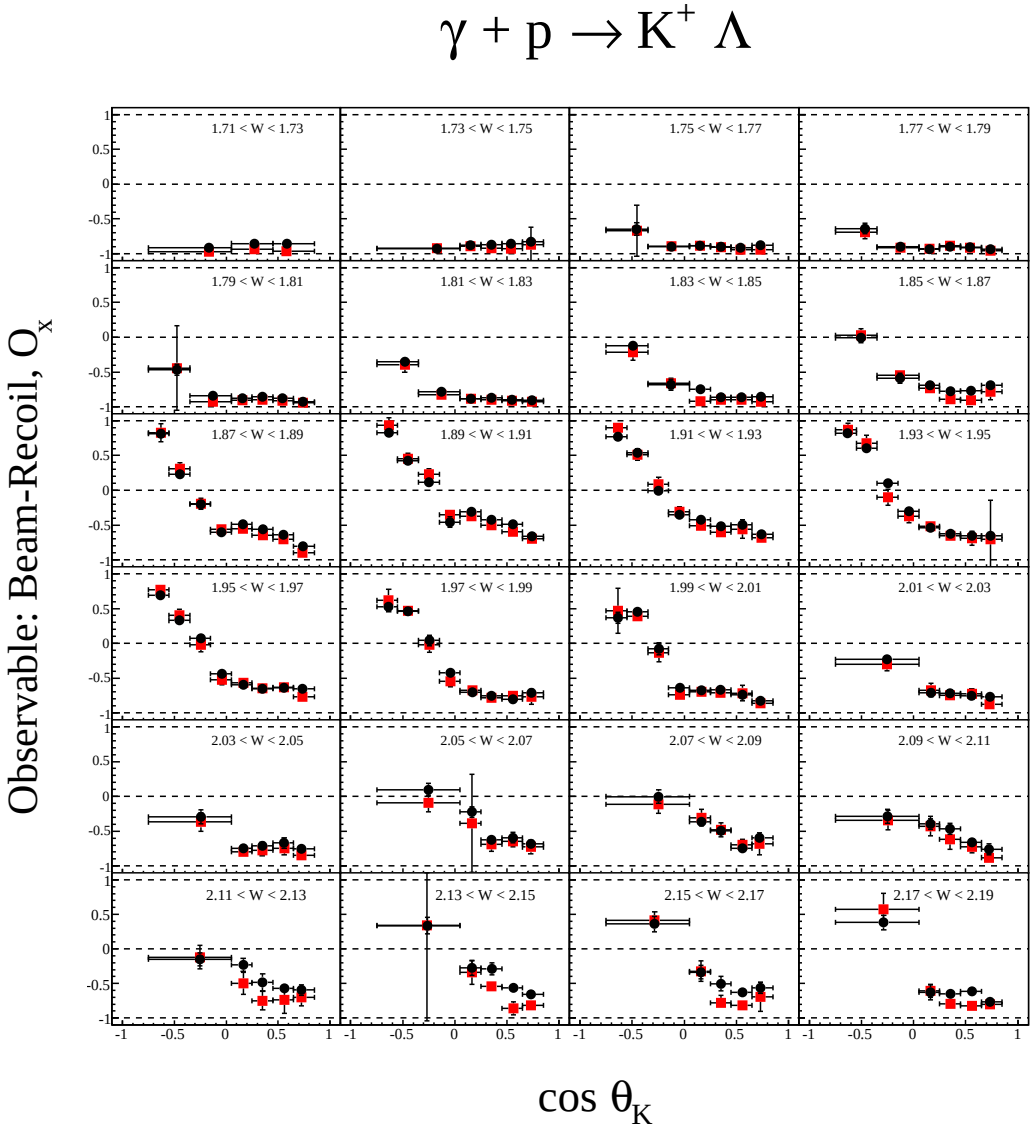


Figure 4.7: Results of the beam-recoil observable, Ox where red squares indicate previous results obtained through maximum likelihood analysis [88], and black circles show results from the nested sampling analysis.

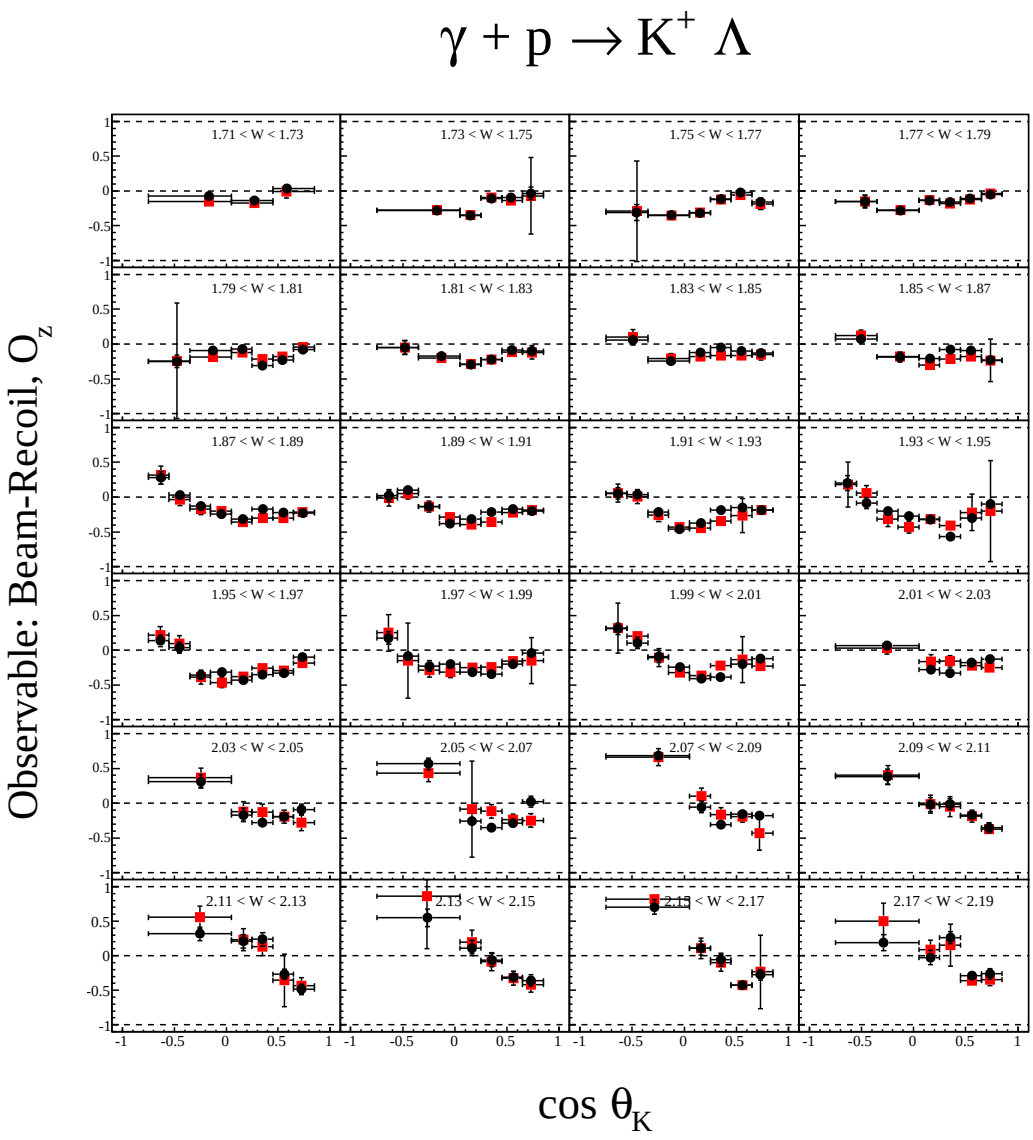


Figure 4.8: Results of the beam-recoil observable, Oz where red squares indicate previous results obtained through maximum likelihood analysis [88], and black circles show results from the nested sampling analysis.

It can be seen by these plots that the results from the nested sampling method agree strongly with previous values.

4.1.6 2-Dimensional Results for All Observables

One of the main features of the nested sampling programme is its use of amplitude space as the prior distribution. By constraining the amplitudes (through the use of the likelihood function(s)), information about observables that are not directly measurable by the experiment can be found. In order to visualise this information, two-dimensional likelihood plots were created for each observable, and for each bin. Regions of warmer colour indicate the more likely values of each observable in a particular bin. In many cases, it can be seen that observables that are not directly measurable by the experiment are constrained to a smaller region of observable space, or have a bi-modal probability density function indicating a reasonably resolved magnitude with a sign ambiguity. One such plot is shown in Figure 4.9. The plot shows the full results for one energy bin, and each cell of the plot corresponds to a polarisation observable. Each vertical band is the two-dimensional histogram of the observable's probability density function for a particular angular bin. The darker colours indicate the most probable values. Plots for all energy bins are included in Appendix B.

Amplitude plots similar to those introduced in Chapter 3 were produced for g8b data. Figure 4.10 shows the posterior distributions of the complex amplitudes for one bin in W and $\cos(\theta_K^{CM})$. The rings seen in the top row indicate that the magnitude of each amplitude is determined.

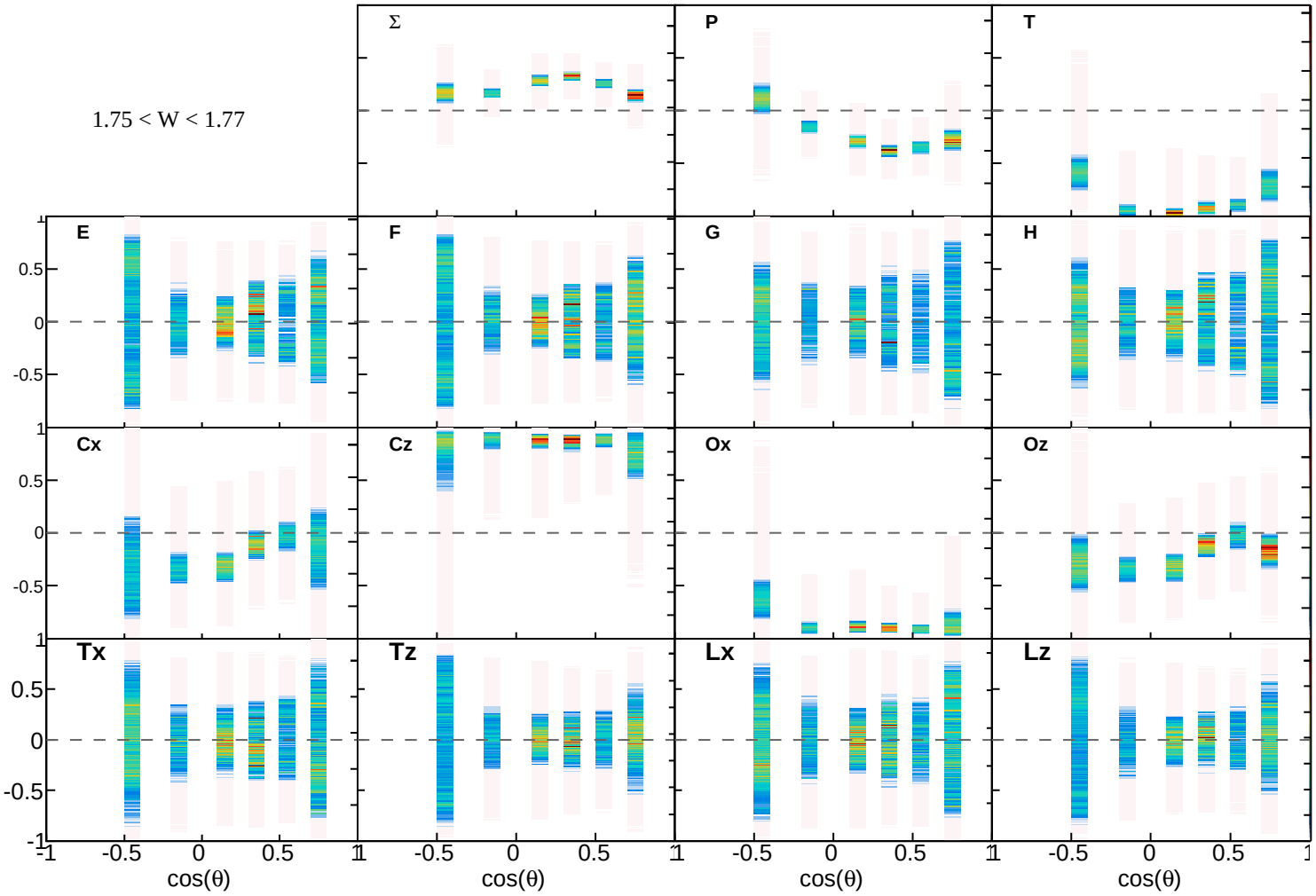


Figure 4.9: Plot showing likelihood distributions for all 15 polarisation observables in one energy bin, $1.75 < W < 1.77$ GeV.

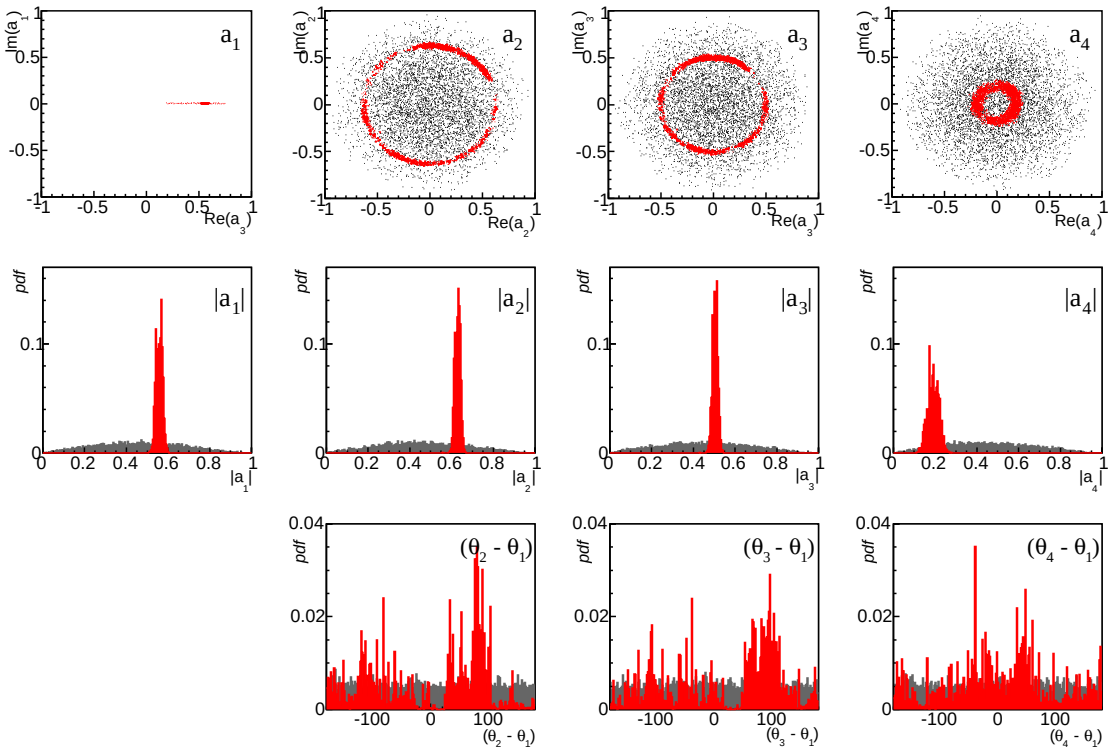


Figure 4.10: Posterior distributions of the complex amplitudes after the analysis of one bin of g8b data. These plots follow the conventions described in Figure 3.28.

4.2 The g1c experiment

The g1c experiment took place in Hall B of Jefferson Lab during late 1999 [89]. A beam of circularly polarised photons were incident on a stationary, unpolarised proton target, and the events from the $\gamma p \rightarrow K^+ \Lambda^0$ reaction channel were analysed. This experiment was sensitive to three polarisation observables: the recoil observable, P, and two beam-recoil double polarisation observables, C_x and C_z. An electron beam incident on a gold foil radiator was used to produce a beam of circuarly polarised photons through coherent bremsstrahlung [59]. The proton target used was an 18cm target cell containing liquid hydrogen.

During the g1c experimental run, three datasets were produced, corresponding to three electron beam energies: 2.445, 2.897 and 3.115 GeV [89]. The 2.445 and 2.897 GeV datasets involved the photon beam being circularly polarised, whereas the 3.115 GeV dataset did not. In this analysis, only the 2.445 GeV dataset was considered.

4.2.1 g1c data

A copy of the 2.445 GeV dataset was obtained from collaborators at Carnegie Mellon University (CMU). The data was stored in a format developed at CMU, and additional software (COBRA) was required to identify and extract events from the $K^+\Lambda^0$ reaction channel. This extraction involved the use of kinematic fitting to identify particles in the final state. The COBRA software [90] enabled relevant data to be easily identified from the full dataset. The desired final state particles (K^+ and p) were specified, and a kinematic fit on the missing mass of K^+p was performed to identify final-state pions. A confidence interval was then chosen, and verified by examining the effects of the kinematic fit.

The binning scheme used for analysing the g8b data was implemented here to allow for like-for-like future comparisons between the experiments. The angular and energy distribution of g1c data is shown in Figure 4.11, where the bin boundaries are featured.

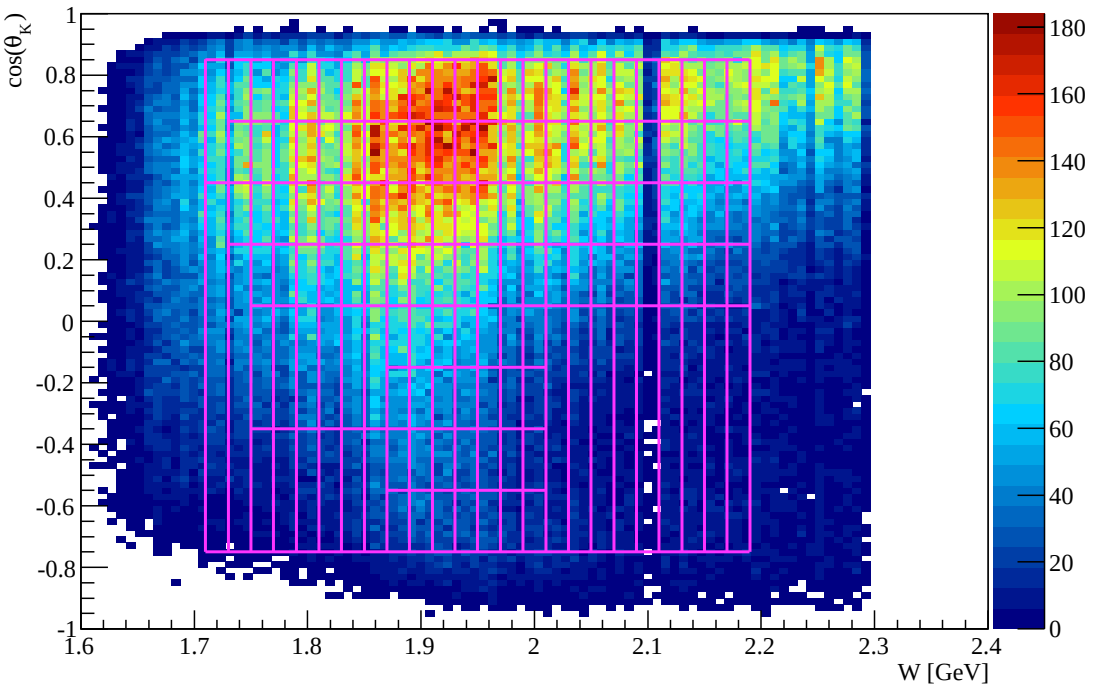


Figure 4.11: Distribution of all g1c data analysed. The magenta lines indicate the bin boundaries.

4.2.2 Luminosity Study of g1c

Before the analysis of g1c was carried out, luminosity asymmetry values were calculated in a similar manner to that described in Section 4.1.2. The value of ΔL_C was included in the nested sampling programme's likelihood function as a free variable, and determined using all events in an energy bin, across all values of $\cos(\theta_K^{CM})$. The results of this study are shown in Figure 4.12. These results deviate only slightly from zero, suggesting a relatively equal number of events in both polarisation settings (clockwise and anticlockwise circular beam polarisations).

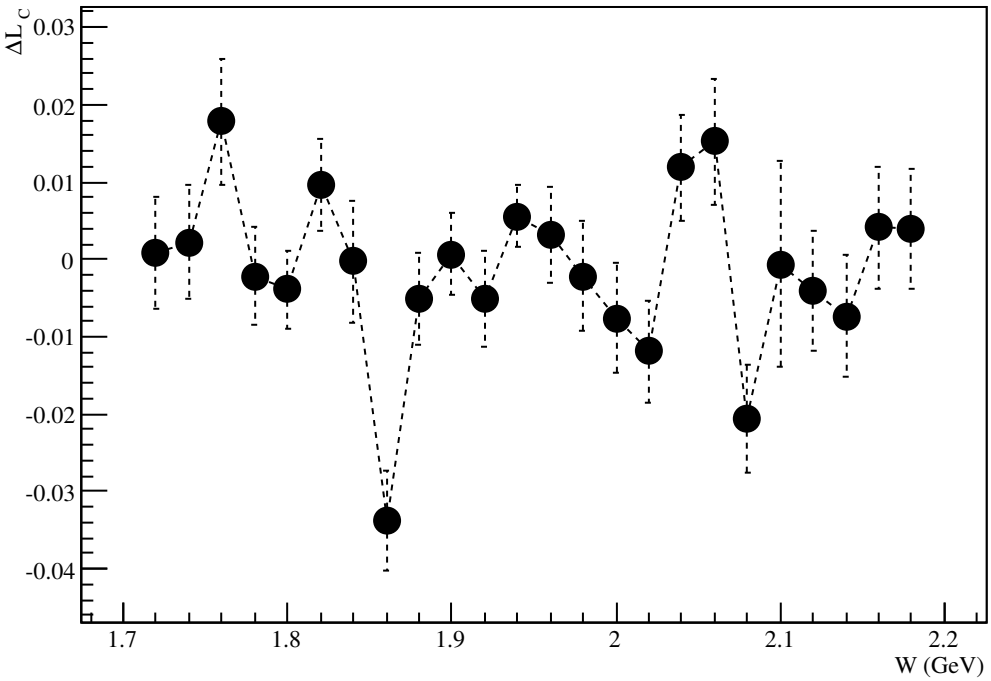


Figure 4.12: Plot of luminosity asymmetry values against photon beam energy, W for g1c data binned in the same bins as g8b.

4.2.3 Results of g1c Compared to Previous Analysis

A nested sampling analysis was used to extract the observables P, C_x and C_z. The values of the double polarisation observables (C_x and C_z) were compared to published results [59], and this comparison is shown in Figures 4.13 and 4.14. The extracted results of the recoil observable P were compared to the results from the analysis of g8b, using an identical binning scheme. The results from both experiments are displayed in the graph in Figure 4.15 below.

In general, the results for C_x and C_z appear strongly consistent with previous results, although there do exist some differences. There are several reasons for this slight discrepancy. Both sets of results correspond to the same experiment, but different datasets were analysed. The results from [59] were obtained by analysing a dataset where the electron beam used to produce photons had an energy of 2.897 GeV, and the dataset analysed with the nested sampling programme corresponded

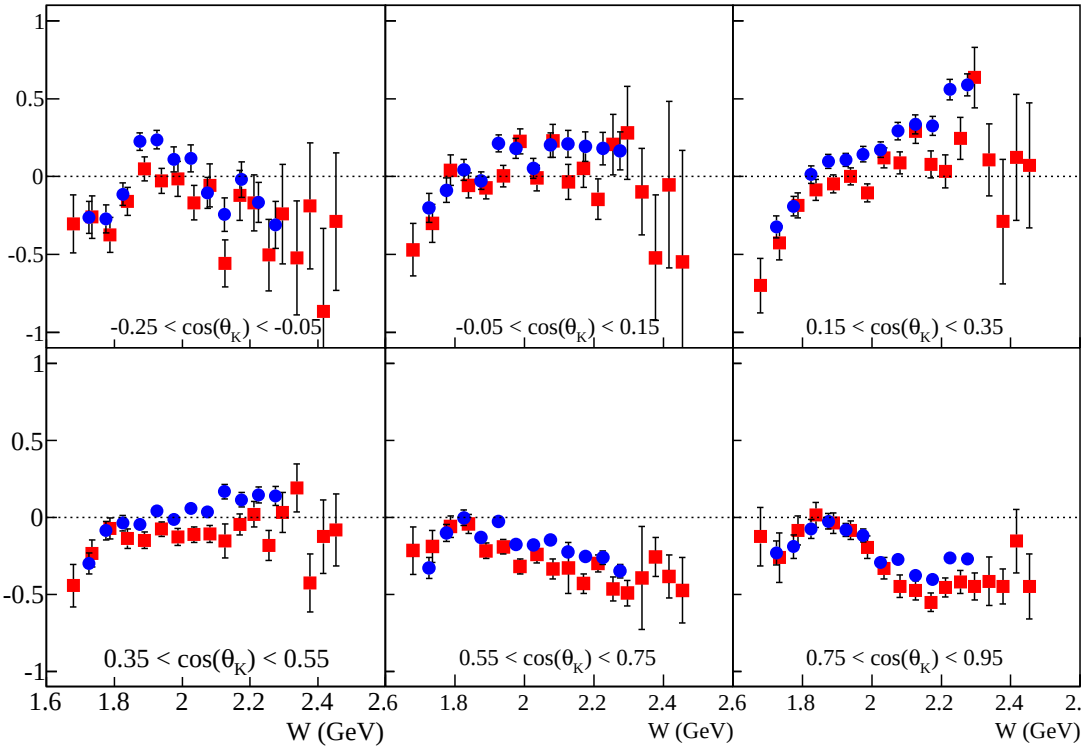


Figure 4.13: Comparison of previous results (red squares) with results from nested sampling (blue circles) for the C_x observable.

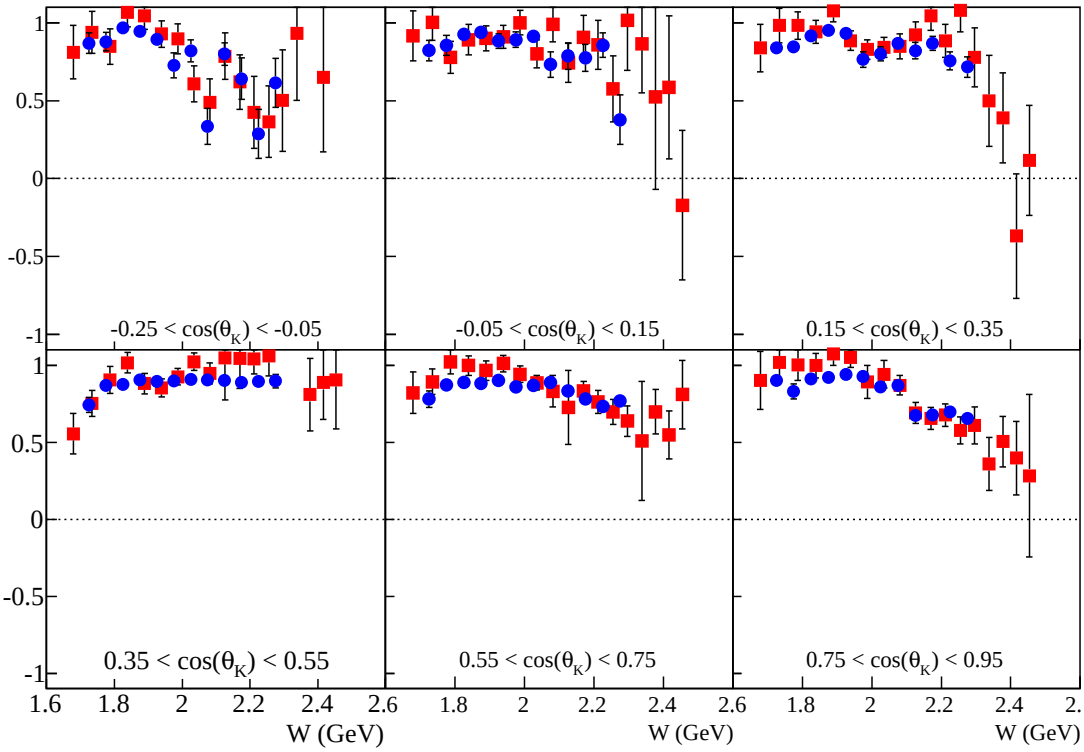


Figure 4.14: Comparison of previous results (red squares) with results from nested sampling (blue circles) for the C_z observable.

to the most stable dataset, stemming from an electron beam of energy 2.445 GeV. All run conditions were constant for both datasets, although the statistics of the datasets may slightly differ. The photon beam energies produced for both datasets covered an equivalent range. It should be stated that this analysis was performed with no acceptance corrections implemented. Initial comparisons of results agreed strongly with published values for g1c despite this lack of correction, and as the required acceptance information was no longer available, it was assumed that the acceptance ratio was sufficiently constant across the decay proton's direction cosine, $\cos(\theta_y)$, region in the Λ rest frame.

When compared to the results of P from g8b, it was also found that the values agreed. This suggested that the datasets were consistent with each other.

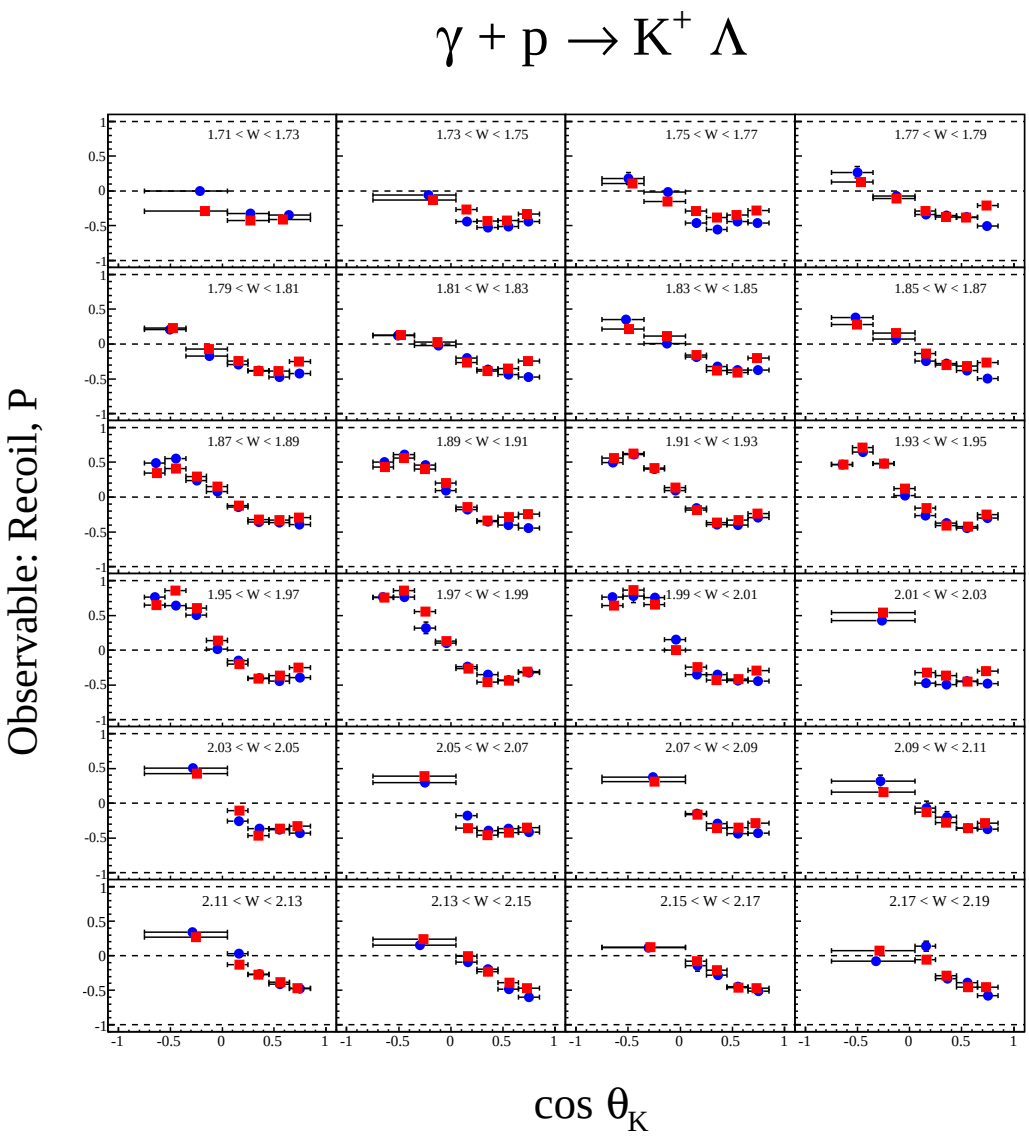


Figure 4.15: Comparison of results from g8b (red) with results from g1c (blue) for the recoil observable, P, where nested sampling was used in both analyses.

The results from g8b and g1c were also compared using one of the observable relations described in Section 1.3. Equation 1.7 was rearranged to the form shown in Equation 4.2.1 below.

$$C_x^2 + C_z^2 = 1 - \Sigma^2 - P^2 + T^2 - O_x^2 - O_z^2 \tag{4.2.1}$$

The values obtained for C_x and C_z from analysis of g1c data were used to cal-

culate a value for the left hand side of Equation 4.2.1, and the right hand side was determined from the values of Σ, P, T, O_x and O_z obtained through g8b analysis. This was done for each bin and plotted in order to verify that the results obey this relation. The comparison is shown in Figure 4.16.

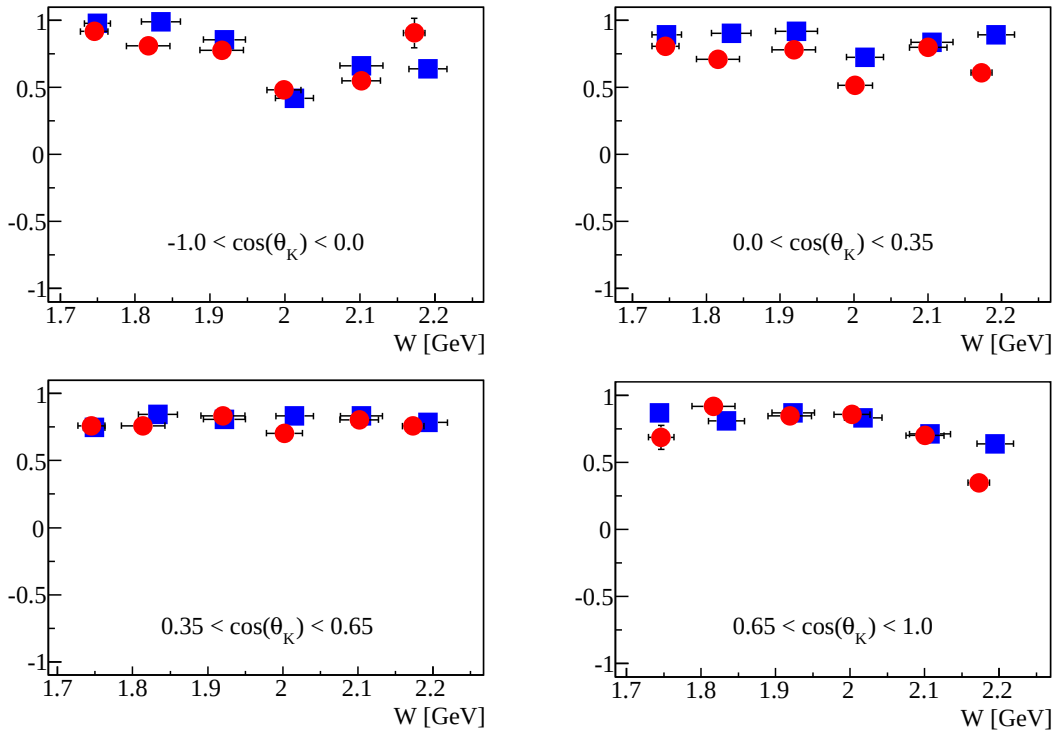


Figure 4.16: Comparison of g8b results (red) with results from g1c (blue) for the relation described in Equation 4.2.1 .

Full two-dimensional likelihood plots of all observables were produced for each energy bin, as they were for the analysis of g8b. Figure 4.17 shows the plot corresponding to the $1.75 - 1.77$ GeV energy bin. Plots for all energy bins can be found in Appendix B.

Again, amplitude plots showing the resulting posterior distributions of the complex amplitudes after analysis were produced. Figure 4.18 illustrates the amount of information learned about the amplitude space from the g1c data in a particular bin.

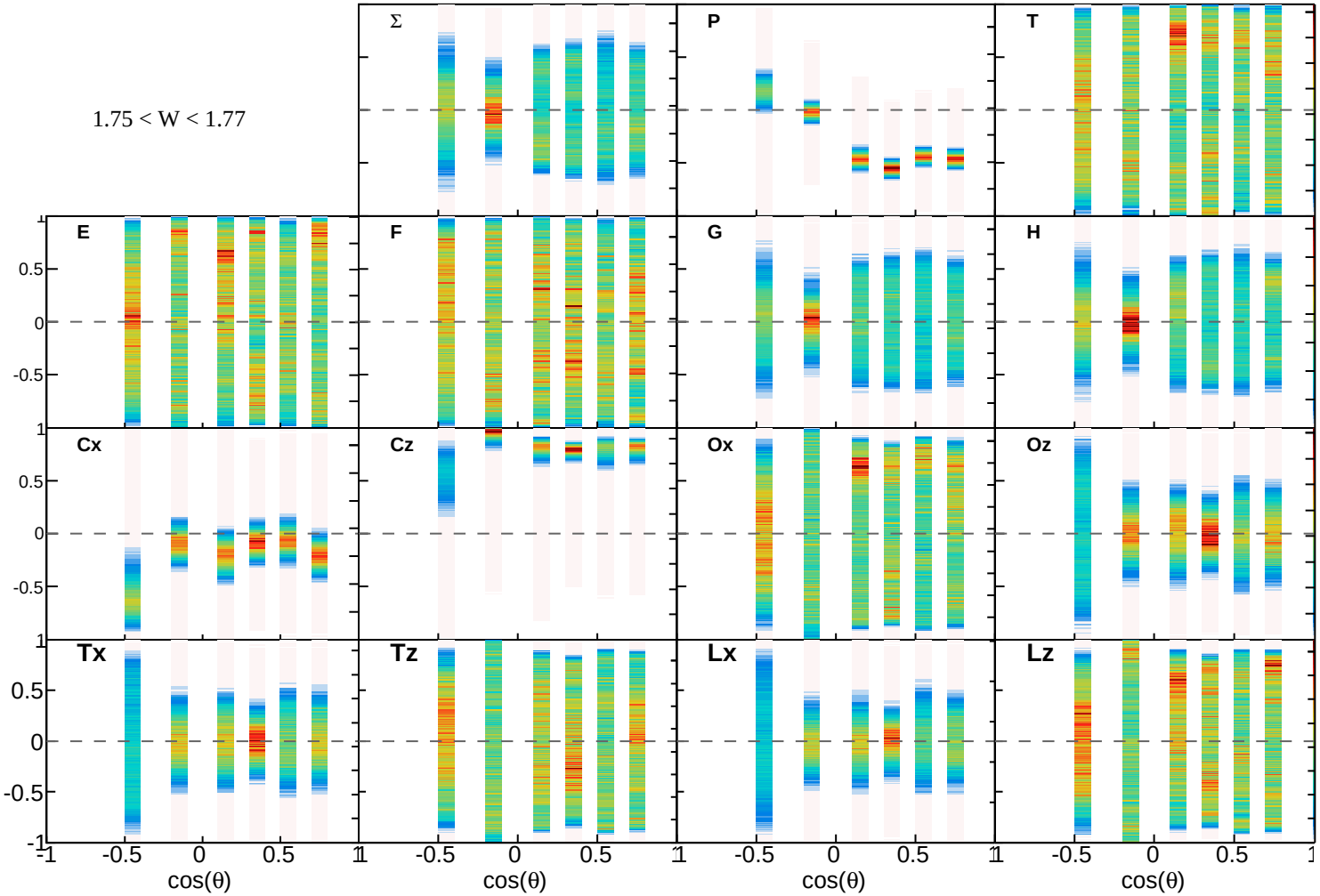


Figure 4.17: Plot showing likelihood distributions for all 15 polarisation observables in one energy bin, $1.75 < W < 1.77$ GeV.

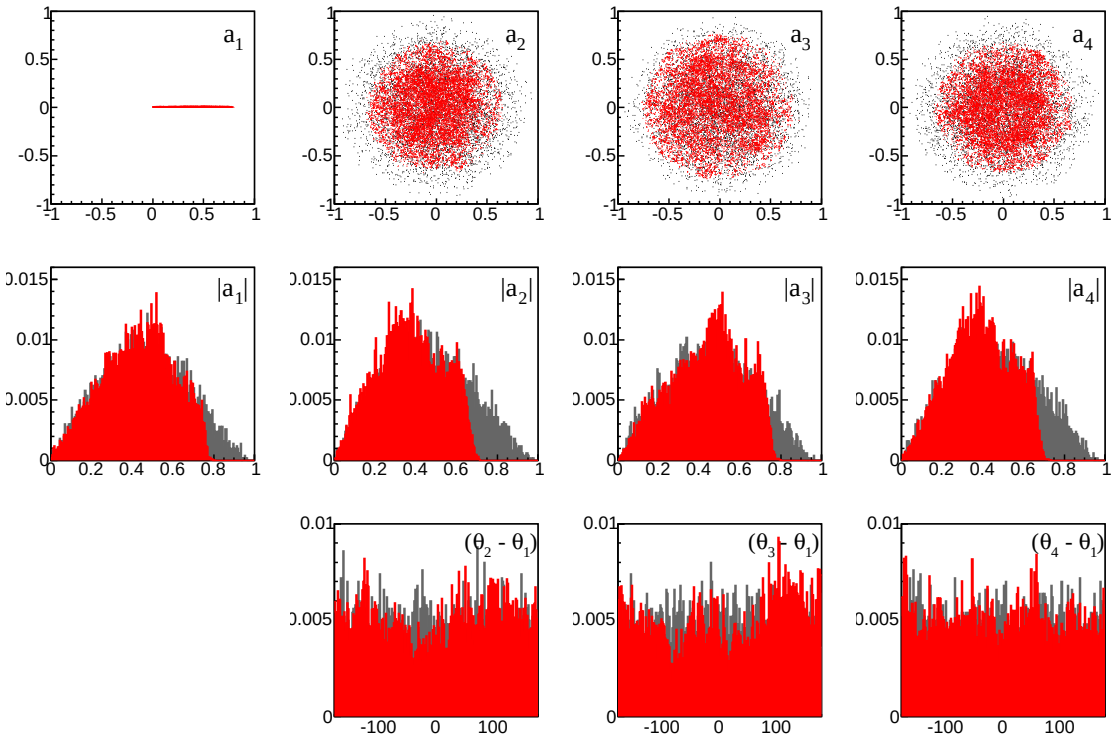


Figure 4.18: Posterior distributions of the complex amplitudes after the analysis of one bin of g1c data. These plots follow the conventions described in Figure 3.28.

4.3 Summary

The g8b experimental run at CLAS consisted of a linearly polarised photon beam incident on an unpolarised proton (liquid hydrogen) target. The data were binned in terms of photon energy (W) and kaon polar angle ($\cos(\theta_K^{CM})$). They were then analysed and compared to results from previous analyses of the same data sets. The asymmetry expression from which the likelihood function was derived was amended to include the effects of having unequal numbers of events in each of the beam polarisation settings. This was done by considering the luminosity asymmetry, ΔL, for each data bin. The values of ΔL were found by performing a nested sampling analysis over all data in an energy bin and extracting the ΔL term as an independent variable. The detector acceptance and background were considered using the techniques implemented in previous analysis methods.

A second experiment, g1c, was also analysed and compared to published results. Luminosity asymmetries for g1c were calculated using the same methodology that was employed for g8b luminosity asymmetry extraction. The values obtained through nested sampling were largely consistent with those previously published, despite a subset of the total data being used and no acceptance corrections. This was unsurprising as the run conditions were very much the same for all datasets. The agreement further suggested that the acceptance of the detector was fairly constant across the y-direction cosine of the decay proton in the Λ rest frame.

Including these considerations, it was found that the nested sampling programme produced results consistent with those of past analyses of the same data. These results further suggest that the values obtained for the five observables directly measurable by each experiment can be considered as "consistent", implying that they are allowed by the constraints imposed by the observable correlations. By using the amplitudes as a prior distribution, it was possible to extract information about all 15 observables, and this was shown using 2-dimensional likelihood plots. It can be seen that, although the true values of all observables cannot be determined from each of these experiments individually, they are constrained to some degree through the observable correlations.

5

Analysis and Results of Combined Data

In this chapter, one of the fundamental benefits of exploring amplitude space is investigated. The two datasets previously analysed separately, g8b and g1c, will be combined and analysed together. All seven observables that were measured in the two separate analyses will be extracted simultaneously from this combined dataset, ensuring that the obtained values are permitted by the constraints imposed by the observable correlations. The impact of this combined analysis on other observables will be shown graphically using two-dimensional likelihood plots.

Section 5.1 will describe how the datasets are combined and how the likelihood function was adapted. The next section will then show the results and compare them to the results obtained from the separate analyses.

5.1 Combining g1c with g8b data

The procedure for merging datasets from two experiments was straightforward. For each bin defined by W and $\cos(\theta_K)$, the dataset from each experiment was edited to include an additional term, an integer flag used to indicate the event's associated experiment. The two datasets were then simply concatenated.
The likelihood function was also revised to include both relevant likelihood equations. The newly included flag determined which likelihood equation (from those listed in Table 3.4) would be used for a given event, and after all events had been looped over, the likelihood components from each equation were combined. The luminosity asymmetry values given in Sections 4.1.2 (g8b) and 4.2.2 (g1c) were used in this analysis in their relevant likelihood equations. Although simple, this method uses information from both datasets to further constrain the values of all observables and ensure that all results are allowed by the restrictions introduced by the set of complex amplitudes.

5.2 Results of combining g8b and g1c

Aside from the introduction of a slightly more complex likelihood equation, the procedure for analysing the combined data was identical to the analyses described in Chapter 4. The results of the observables Σ, P, T, O_x, O_z, C_x and C_z proved to be in excellent accordance with the values obtained in the separate analyses of g8b and g1c, as illustrated in Figures 5.1 to 5.7.

The above plots indicate that it is possible and, indeed, straightforward to combine datasets from different experiments in this manner, and at the very least, obtain sufficiently similar results as separate analyses would yield. The distinguishing feature that is used in the combined analysis is that the observables are all extracted in a way such that the correlations imposed by the complex amplitudes on the observables are preserved.

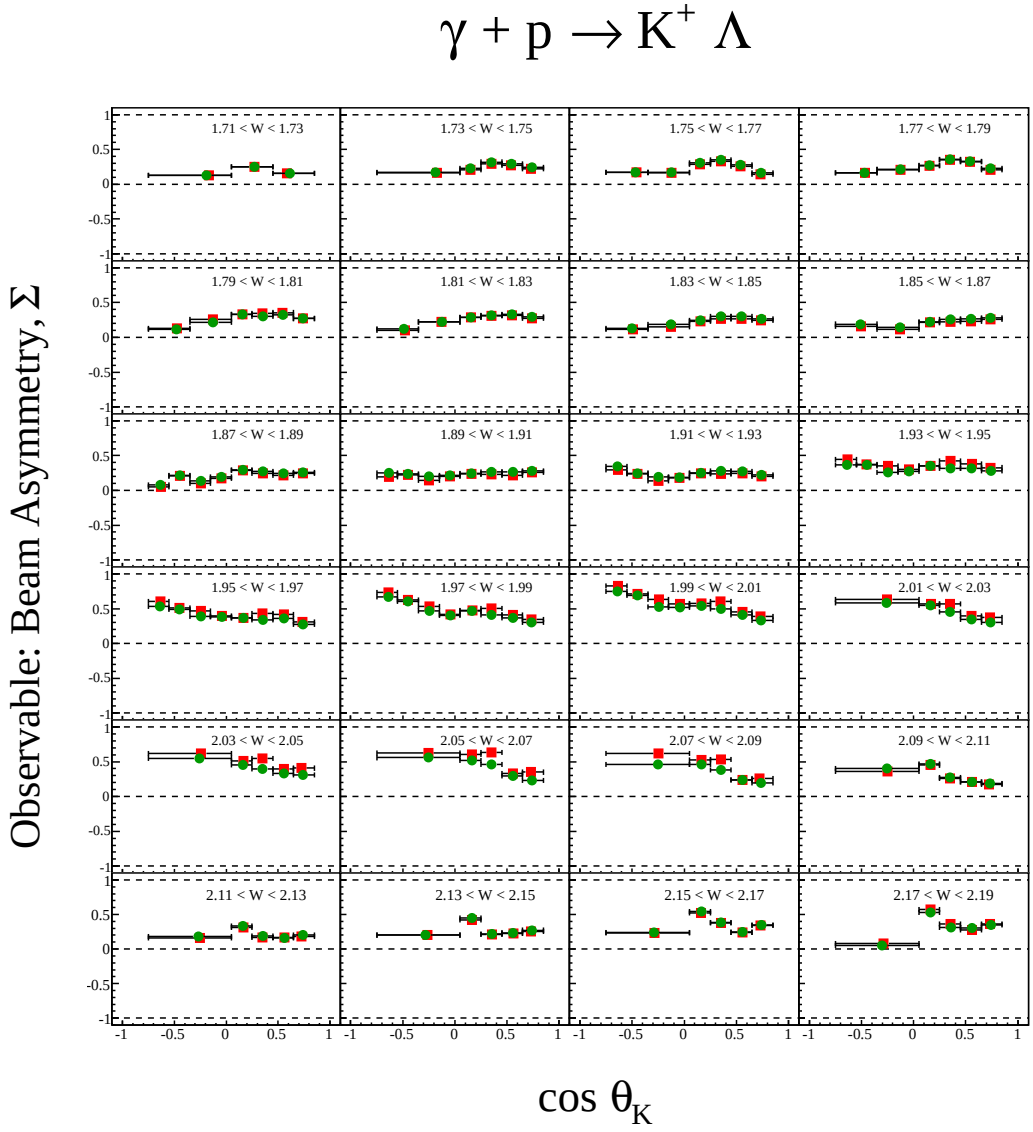


Figure 5.1: Results of combined analysis of g8b and g1c (green) compared to results from g8b analysis (red) of the beam asymmetry observable, Σ.

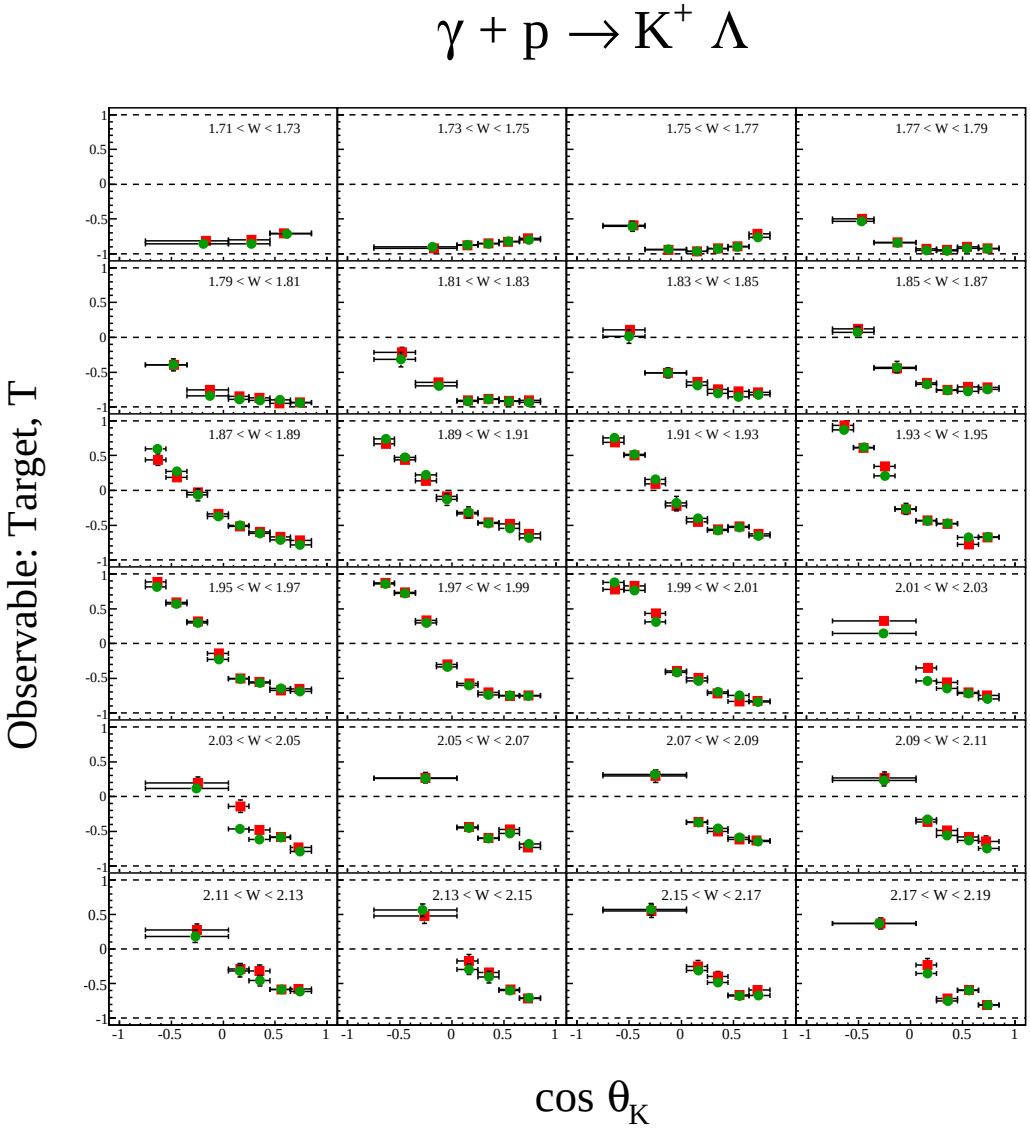


Figure 5.2: Results of combined analysis of g8b and g1c (green) compared to results from g8b analysis (red) of the target observable, T.

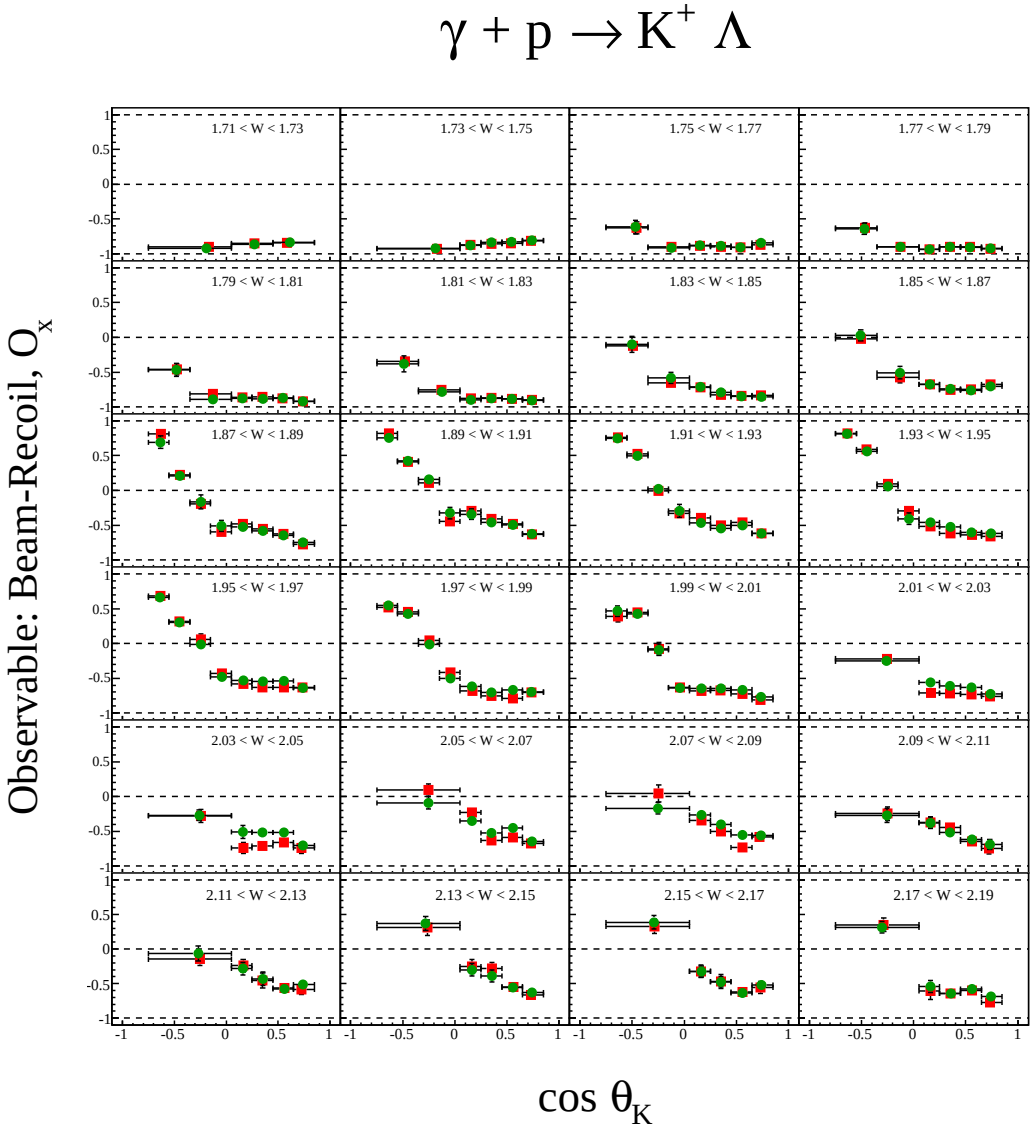


Figure 5.3: Results of combined analysis of g8b and g1c (green) compared to results from g8b analysis (red) of the beam-recoil double polarisation observable, O_x.

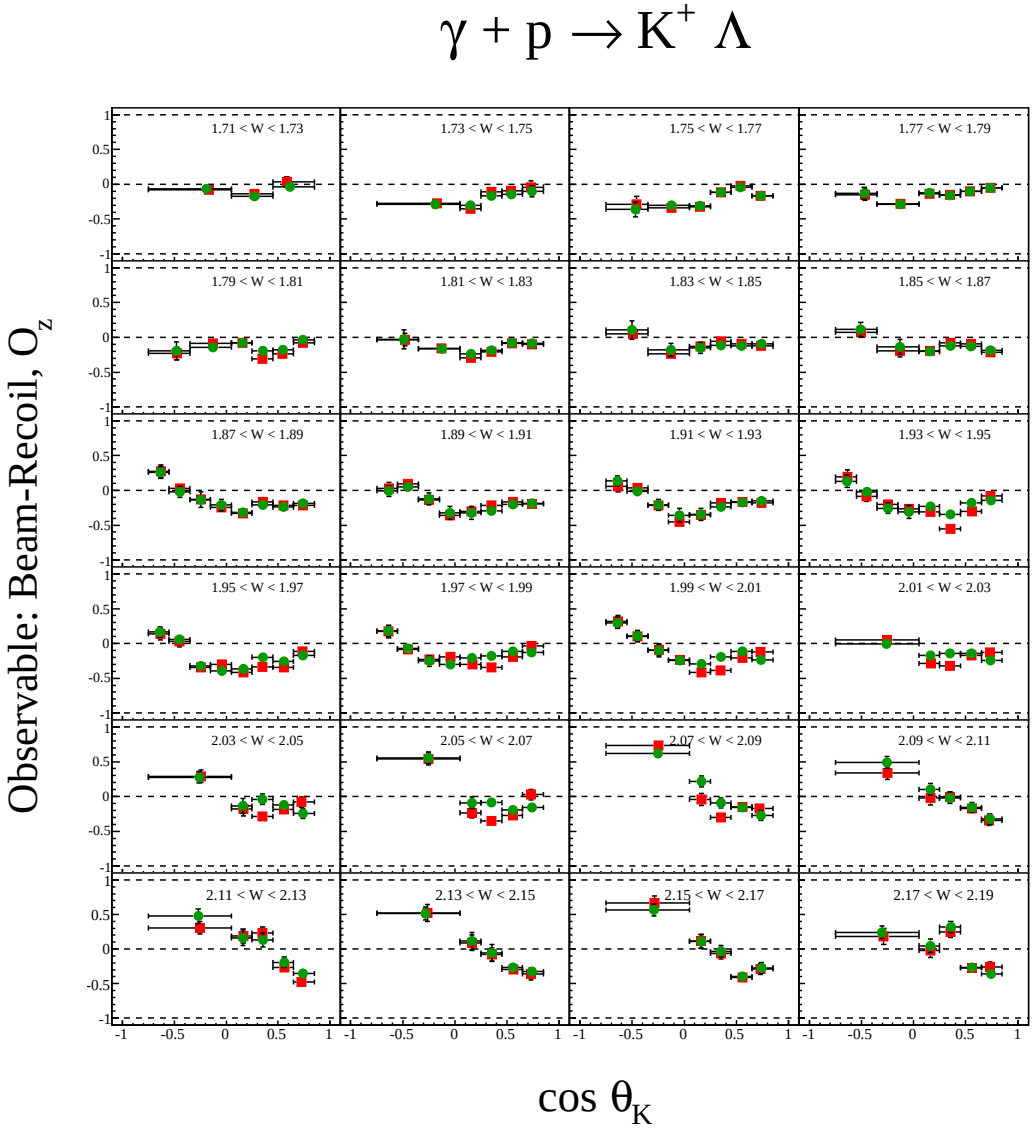


Figure 5.4: Results of combined analysis of g8b and g1c (green) compared to results from g8b analysis (red) of the beam-recoil double polarisation observable, O_z.

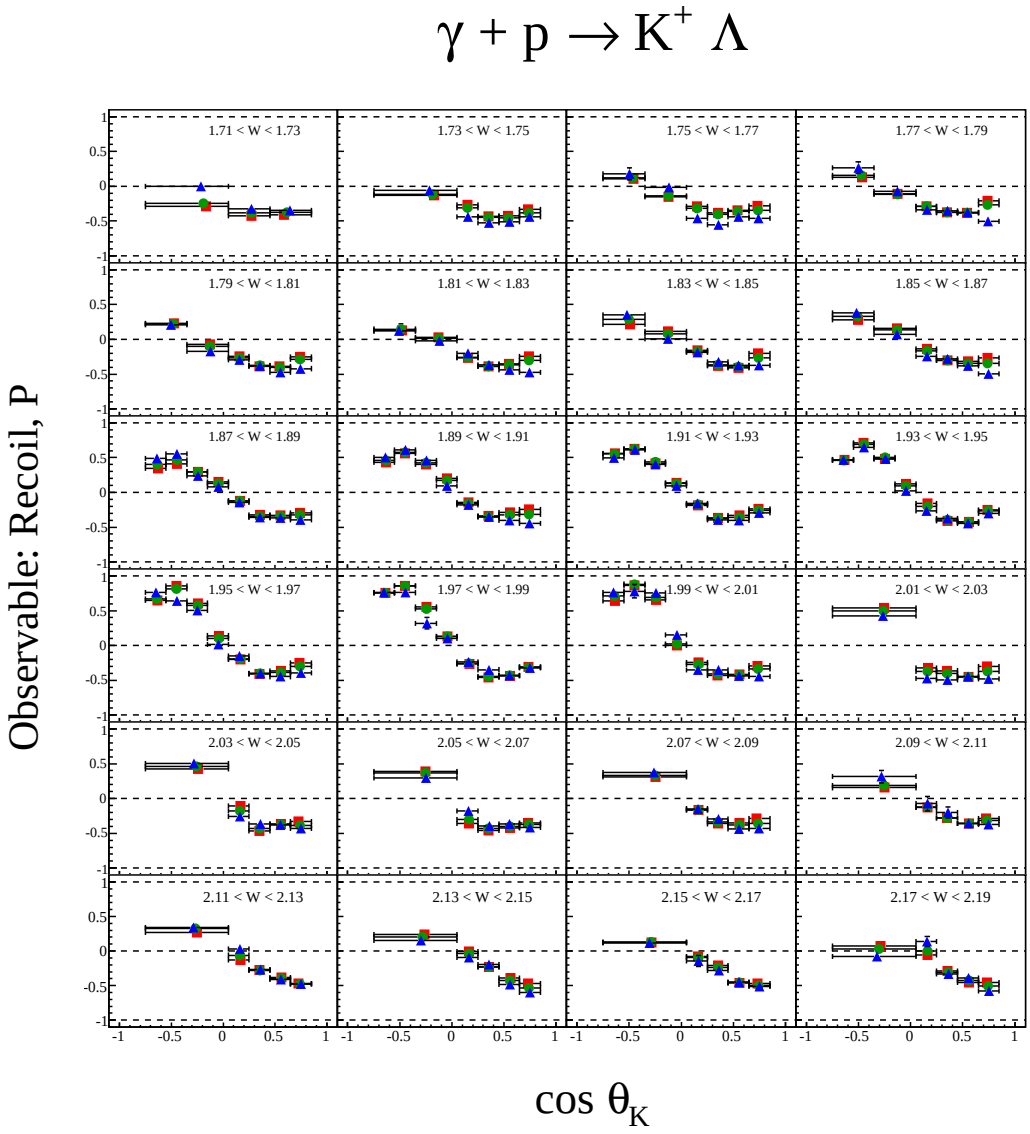


Figure 5.5: Results of combined analysis of g8b and g1c (green) compared to results from g8b analysis (red) and g1c analysis (blue) of the recoil observable, P.

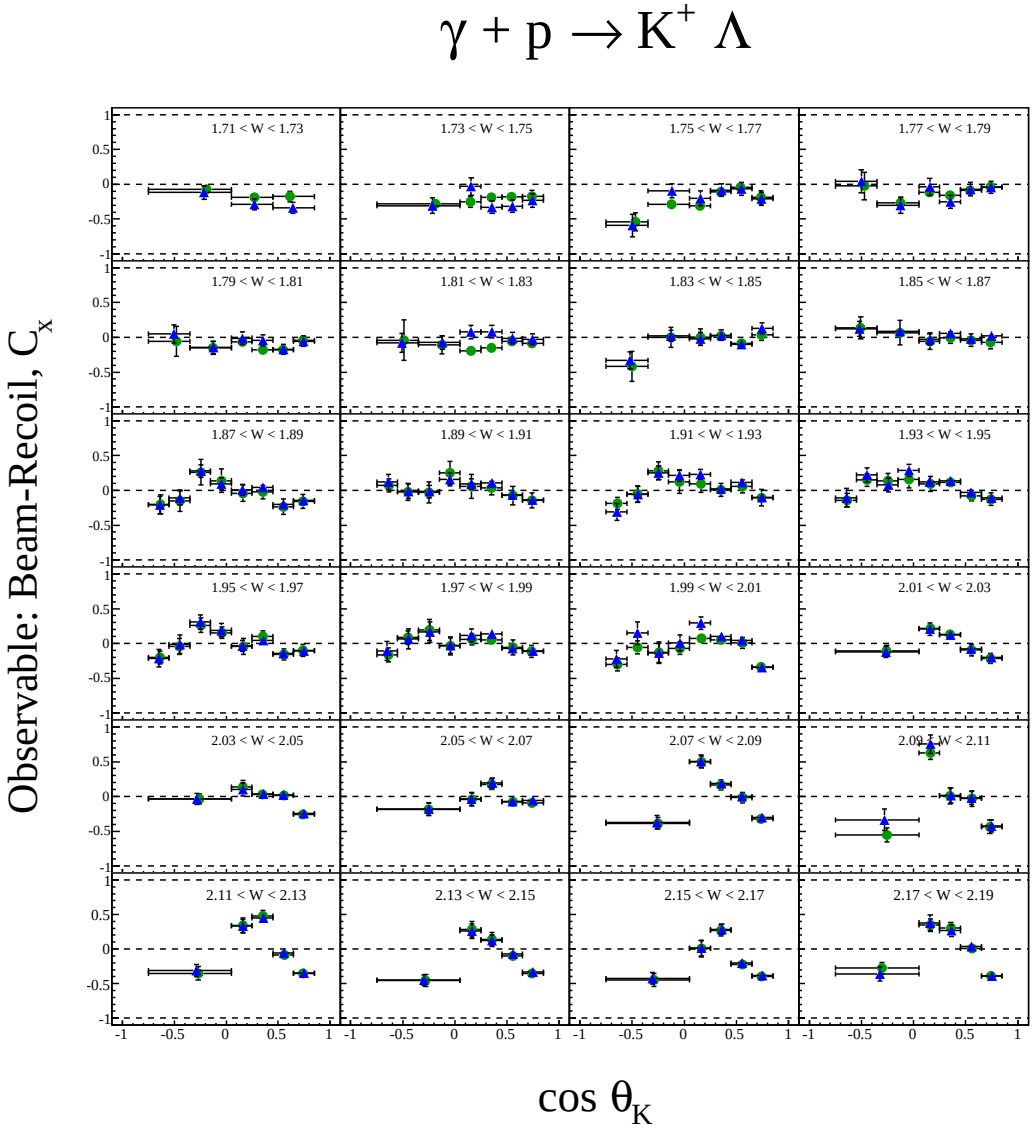


Figure 5.6: Results of combined analysis of g8b and g1c (green) compared to results from g1c analysis (blue) of the beam-recoil double polarisation observable, C_x.

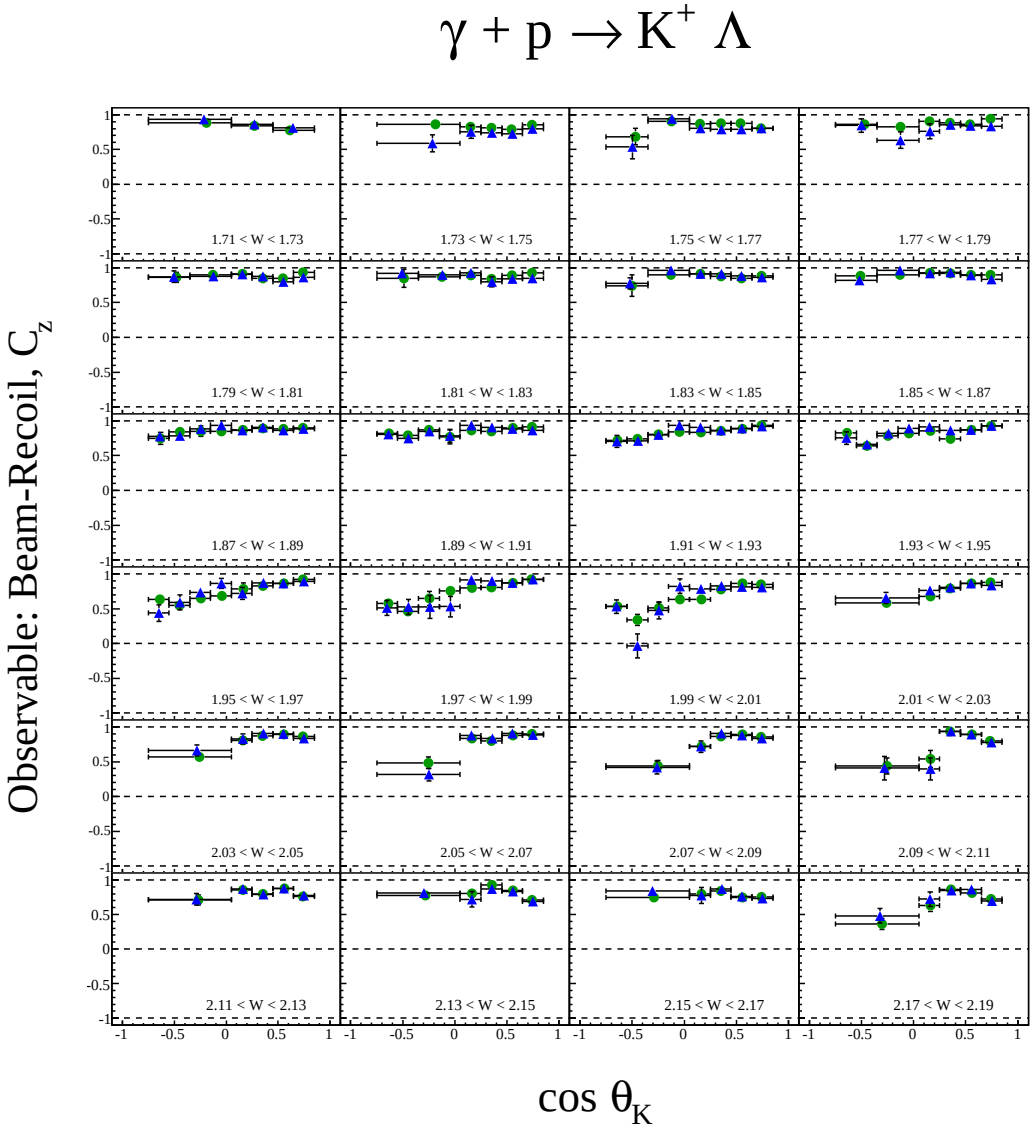


Figure 5.7: Results of combined analysis of g8b and g1c (green) compared to results from g1c analysis (blue) of the beam-recoil double polarisation observable, C_z.

Just as they were for the analyses presented in Chapter 4, two-dimensional likelihood plots of all fifteen observables for each energy bin were produced. The plot of one energy bin is given in Figure 5.8, and plots for all other energy bins can be found in Appendix B.

As was done in Chapters 3 and 4, amplitude plots depicting the posterior distributions of the four complex amplitudes were produced for the combined analysis of g8b and g1c. Figure 5.9 shows these results, which can be compared to Figures 4.10 and 4.18 to visualise the amount of information gained by analysing these two experiments together.

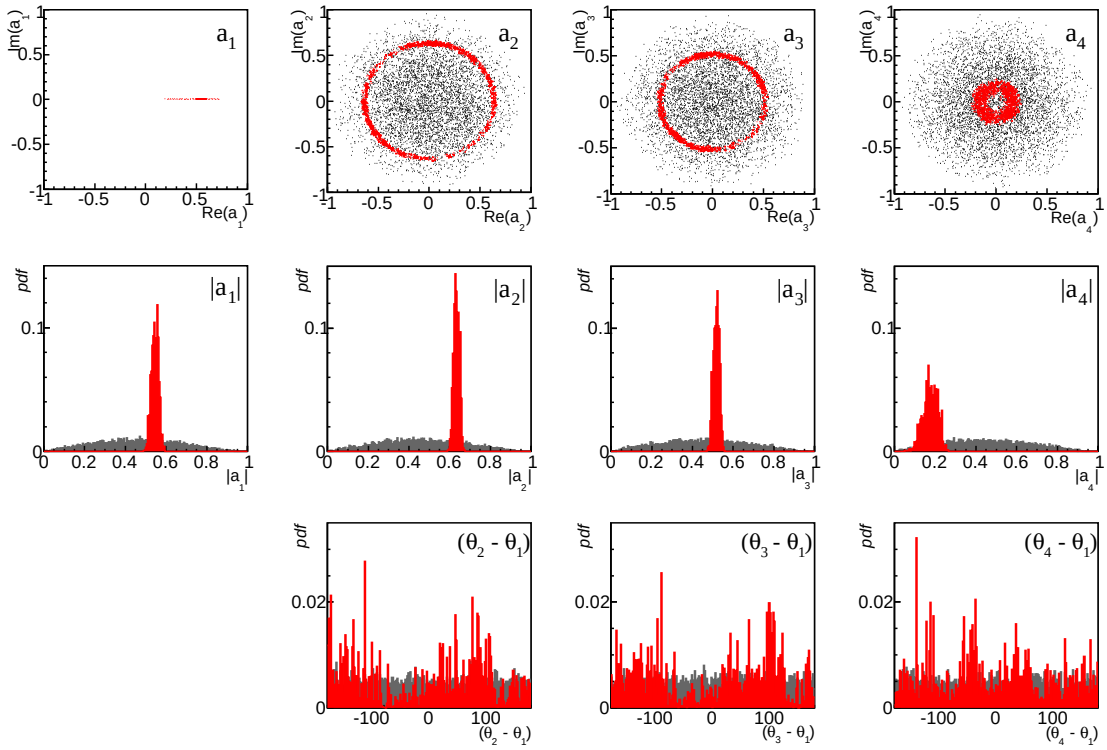


Figure 5.9: Posterior distributions of the complex amplitudes after the combined analysis of one bin of g8b and g1c data. These plots follow the conventions described in Figure 3.28.

In general, there is little difference between these results and the g8b two-dimensional plots. This is understandable as both experiments are beam-recoil experiments, and no additional information on target polarisation is introduced. It is therefore unsurprising that a significant amount of new information is not dis-

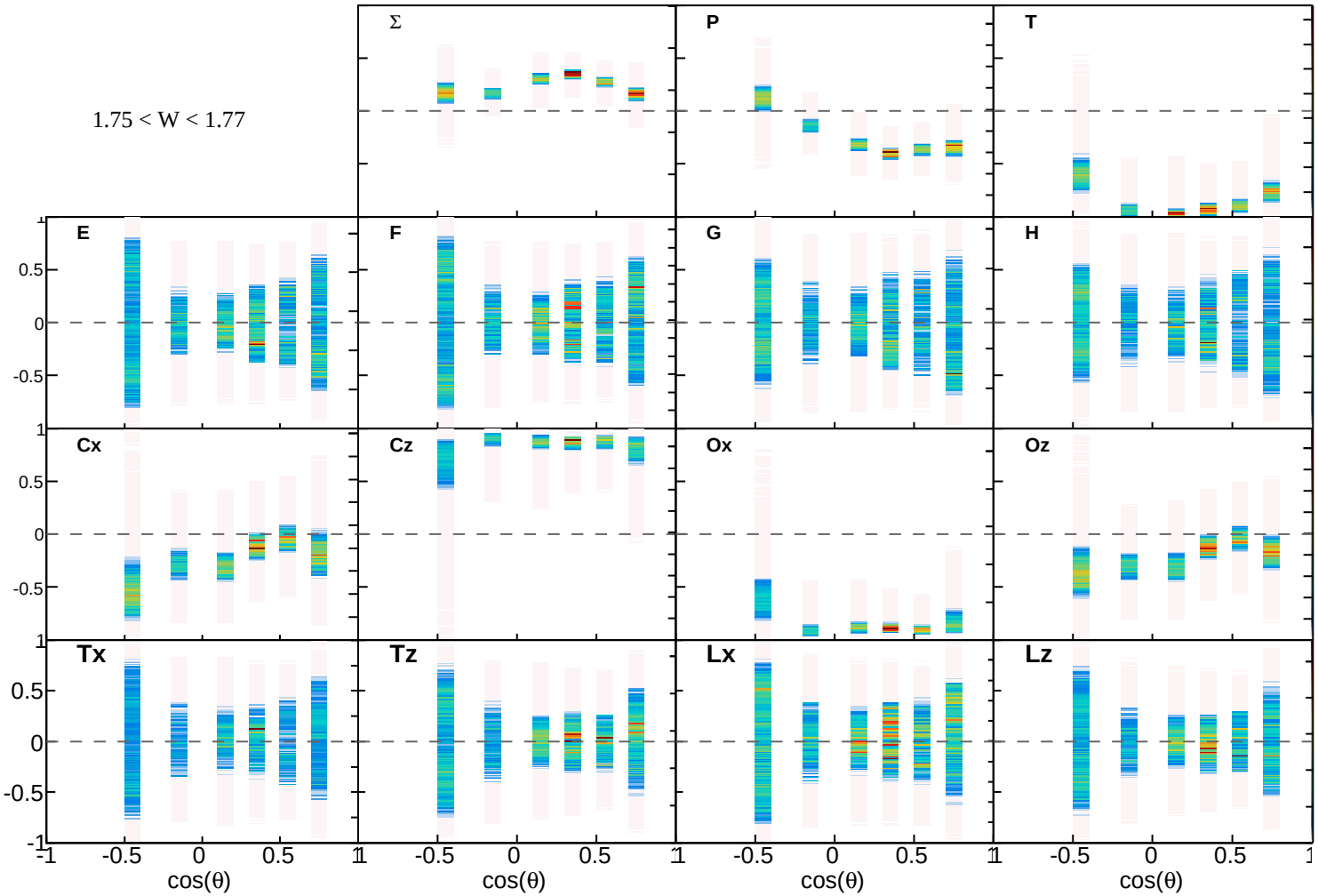


Figure 5.8: Plot showing likelihood distributions for all 15 polarisation observables in one energy bin, $1.75 < W < 1.77$ GeV, after combined analysis of g8b and g1c.

covered. These plots do, however, show that both analyses are compatible and well understood.

5.3 Summary

Datasets from two kinematically consistent experiments were combined and analysed simultaneously. This allowed the complex amplitudes in the prior distribution to be constrained by the likelihood functions associated with both experiments, increasing the amount of information extracted from the data. By extracting observables in this way, it is ensured that the values found are always compatible.

The results of the combined analysis were compared to those found in separate analyses. It was shown that these were sufficiently consistent with those found by analysing each experiment separately (as shown in Chapter 4). Two-dimensional plots of the combined analysis were created, and show how all fifteen observables are affected by analysing these merged datasets. It was found that there was not a substantial gain in information on those observables not directly measured in this case, but this was not unexpected. The two experiments used in this analyses were both beam-recoil experiments, implying that no information on target polarisation was included. This method of combining datasets can be easily generalised for any number and any type of experiments.

6

Summary and Outlook

The field of hadronic physics is focused on the understanding of the strong interaction in nucleons. Although QCD provides a suitable explanation in the perturbative region, the mass of a nucleon falls within the non-perturbative regime and there has not been sufficient developments in lattice QCD as of yet. A variety of quark models exist and vary based on the number of degrees of freedom and quark interactions. In order to learn which models are the most reliable, it is crucial to provide experimental evidence that can be used to support or oppose them. Each quark model predicts a unique nucleon mass resonance spectrum. We can develop our understanding of quark interactions by experimentally measuring resonances that are predicted by some models and not others. By merely examining a differential cross-section for meson photoproduction, it is impossible to determine all possible resonances. Instead, different reaction channels can be isolated and examined individually. It has been found that some channels have a greater sensitivity to certain

resonances, including the $K\Lambda$ strangeness channel.

Pseudoscalar meson photoproduction is one method of experimentally determining these resonances, and is completely described by four complex amplitudes. These amplitudes can be accessed experimentally through 15 polarisation observables and the cross-section. Each polarisation observable (and the cross-section) is a bilinear combination of the complex amplitudes, and by polarising different components of the experiment (photon beam and target) in different ways, it is possible to select a subset of these observables which can be extracted from data. There is a worldwide effort to analyse data from pseudoscalar meson photoproduction experiments, and several facilities around the globe dedicate many resources to this field. The CLAS Collaboration at Jefferson Lab in Newport News, Virginia is home to one such facility (the CLAS detector, housed in Hall B), and their excited nucleon (N*) program has resulted in many new measurements of polarisation observables over the last few decades.

The analysis of data from pseudoscalar meson photoproduction experiments is performed most commonly by a binned χ^2 fitting method. Data is binned into histograms based on the polarisation state, and an asymmetry between these histograms is found. A predefined fit is then applied and the parameters of such a fit are used to determine the values of various polarisation observables. This method has several disadvantages that result in a loss of information, including histogram binning and the absence of constraints inherently imposed by observable correlations. For this reason, a new method based on Bayesian statistics was developed to maximise the information yield from existing datasets. The method presented here, called nested sampling, consists of three fundamental components: a prior distribution, a likelihood function, and a posterior distribution. The prior distribution is made up of a set of points that cover the entire physical region. In its application to the analysis of pseudoscalar meson photoproduction, the prior distribution implemented was in amplitude space and consisted of the four complex amplitudes. These amplitude values were then used to map into observable space, inherently

maintaining the constraints imposed by the correlations between observables. The likelihood function used in this work was event-by-event, reducing the information loss further by removing the need for histogram binning. In the nested sampling algorithm, each point in the prior distribution is assigned a likelihood value based on this likelihood function. The point in the distribution with the lowest likelihood is determined, and overwritten with a copy of a surviving point. The value of this new point is slightly altered in a manner that ensures its likelihood is higher than that of the overwritten point. This is an iterative process, and after a sufficient number of iterations (determined by a termination condition), peaks form around the values of highest likelihood. This has several benefits over other numerical maximisation techniques, as it has no requirements of differentiability. It can also be used for multimodal problems, and the presence of multiple local extrema do not cause any complications. The posterior distribution is the resulting probabilility density function and is made up of the points that are deleted, over time developing peaks at regions of high likelihood.

The development of this programme led to the derivations of likelihood functions of any potential experimental configuration. Each likelihood function involved at least one asymmetry between cross-sections corresponding to different polarisation settings. A generic method of derivation was determined and a table of likelihood equations for each experimental configuration was compiled. Then, simulated data was used to develop and test the nested sampling programme for each experimental configuration. It was then hypothesised that the use of a prior distribution corresponding to amplitude space rather than observable space could conveniently lead to a method of analysing datasets from multiple (different) experiments. This was tested with simulated data by adding a flag to each datafile (used to select the appropriate likelihood equation for that event) and concatenating the two datasets. The simulation of this procedure proved successful, and it was shown that additional information pertaining to the amplitudes themselves could be discerned by this approach.

After its initial success on simulated data, the nested sampling programme was applied to the extraction of polarisation observables from experimental data. Two experiments were analysed separately, and then as a combined dataset. The first experiment, g8b, consisted of a linearly polarised photon beam incident on an unpolarised proton target. The polarisation of the hyperon was determined through its self-analysing properties. Five observables (Σ, P, T, O_x and O_z) were measured directly and the obtained values were compared to results from a previous analysis and found to be consistent. Two-dimensional likelihood distributions for all 15 observables were plotted for each energy bin, and it was shown that additional information was gained about observables to which the g8b experiment was not sensitive. Plots of the complex amplitudes, including magnitudes and relative phases were shown and were used to illustrate the amount of information acquired about the amplitudes.

Similarly, data from the g1c experiment consisting of a circularly polarised photon beam incident on an unpolarised proton target were analysed using the nested sampling programme. This experiment was sensitive to three polarisation observables, P, C_x and C_z. Once again, the results of this analysis were compared to published values and found to be consistent. The recoil observable, P, was compared to the values obtained from the g8b experiment, and again found to agree strongly. Two-dimensional likelihood plots of all 15 observables for each energy bin were again produced and illustrated the information gained about each observable. Amplitude plots were also produced, again indicating the amount of information known about the amplitudes after analysis of the g1c data.

The final analysis performed was that of a combined dataset, including events from g8b and g1c. An additional term was added to the data files, indicating the experiment from which an event was produced. The likelihood function was adapted to include likelihood equations from the two relevant experiments. This combined analysis proved to give values consistent with those that had been obtained in the separate analyses of each experiment, for all seven measurable observables. Once again, two-dimensional likelihood plots of the 15 observables for each energy bin were shown, and it could be seen that more information was learned through this

combined analysis. Plots of the amplitudes from the combined analysis compared to those from separate analyses also suggested an increase in information yield.

Development of a combined analysis programme for the analysis of the g9a experiment at CLAS is seen as the next milestone in the nested sampling project. The g9a experiment, known as FROST (FROzen Spin Target) [43, 91] consists of a linearly polarised photon beam and a longitudinally polarised proton target. This experiment would provide direct measurements of a multitude of polarisation observables: Σ, P, T, F, G, O_x, O_z, T_x, T_z, L_x and L_z, and would constitute a "complete measurement". The target is made up of butanol (CH_4H_9OH), with additional targets of carbon and polythene to aid in the consideration of background. The target polarisation is expected to be around 70%, however only approximately 20% of protons within the target are polarisable. This leads to considerable complications when attempting to account for the background. Currently, the background contributions are not sufficiently well understood to perform a rigorous analysis of the data. It is anticipated that once the target background is sufficiently well understood, the nested sampling programme will be able to extract these polarisation observables and reveal important information about the underlying complex amplitudes. It is also expected that the combination of data from the FROST experiment with data from g8b and g1c would lead to a highly reliable measurement of most of the 15 polarisation observables.

This analysis programme can also be applied to experimental data from other facilities. A similar experiment were performed recently at Mainz, where data from the $p\pi^0$ reaction channel will be analysed to extract the polarisation observable, G [92]. A polarised proton (butanol) target was used, and the experiment produced a high statistics data set. It is anticipated that nested sampling could play a part in the analysis of this experiment.

The nested sampling method of analysis described in this thesis can be applied to pseudoscalar meson photoproduction experiments from any experimental facility. This analysis technique has already been used to analyse data from CLAS and can contribute significantly to addressing the missing resonance problem in nuclear physics.

Glossary

API Application Programming Interface, 50, 159

CEBAF Continuous Electron Beam Accelerator Facility, 9

CLAS CEBAF Large Acceptance Spectrometer, 10

CPU Central Processing Unit, 149

FROST Frozen Spin Target, 135

GPGPU General-Purpose Graphical Processing Unit, 150

GPU Graphical Processing Unit, 150

Jefferson Lab Thomas Jefferson National Accelerator Facility, 9

MCMC Markov chain Monte Carlo, 26

MIMD Multiple Instruction stream Multiple Data, 153

MISD Multiple Instruction stream Single Data, 153

ML Maximum Likelihood, 35

OpenCL Open Computing Language, 50, 156

OpenMP Open Multiprocessing, 50, 159

PCI-e Peripheral Component Interconnect Express, or PCI-e, 152

pdf Probability distribution function, 23

PDG Particle Data Group, 6

Posterior The posterior distribution resulting from Bayesian analysis of data, 21

Prior A prior distribution, 21

QCD Quantum Chromodynamics, 2

SIMD Single Instruction stream Multiple Data, 153

SISD Single Instruction stream Single Data, 153

SQM Symmetric quark model, 4

Bibliography

[1] G.D. Rochester, C.C. Butler (1947), *Evidence for the existence of new unstable elementary particles*, Nature **160**, 855.

[2] B.R. Holstein (2000), *Weak and electromagnetic interactions of hyperons: A chiral approach*, Int. J. Mod. Phys. E **9**(05), 359-405.

[3] V.D. Hopper, S. Biswas (1950), *Evidence concerning the existence of the new unstable elementary neutral particle*, Phys. Rev. **80**, pp 1099.

[4] H.L. Anderson, E. Fermi, E.A. Long, D.E. Nagle (1952), *Total cross sections of positive pions in hydrogen*, Phys. Rev. **85**, pp 926.

[5] K.A. Brueckner (1952), *Meson-nucleon scattering and nucleon isobars*, Phys. Rev. **86**, pp 106.

[6] M.K. Sundaresan (2010), *Handbook of Particle Physics*, CRC Press, pp 129.

[7] M. Gell-Mann, Y. Ne'eman (1964), *The eightfold way*, WA Benjamin New York.

[8] M. Gell-Mann (1956), *The interpretation of the new particles as displaced charge multiplets*, Il Nuovo Cimento **4** (2), pp 848–886.

[9] S. Sakata (1956), *On a composite model for the new particles*, Prog. Theoret. Phys. (Kyoto) **16** (6), pp 686–688.

[10] J.L. Rosner (2001), *The eightfold way*, arχiv preprint hep-ph/0109241

[11] G. Zweig (1964), *An SU_3 model for strong interaction symmetry and its breaking*, No. CERN-TH-412, CM-P00042884.

[12] L. Hoddeson, L. Brown, M. Riordan, M. Dresden (1997), *The rise of the standard model: Particle physics in the 1960s and 1970s*, Cambridge Univ. Press, New York.

[13] E.D. Bloom, D.H. Coward, H. DeStaebler, et al. (1969), *High-energy inelastic e-p scattering at* 6° *and* 10°, Phys. Rev. Lett. **23** (16), pp 930–934.

[14] M. Breidenbach, J.I. Friedman, W. Kendall, R.E. Taylor, et al. (1969), *Observed behavior of highly inelastic electron-proton scattering*, Phys. Rev. Lett. **23** (16), pp 935.

[15] F. Halzen, A.D. Martin (1984), *Quarks and leptons: An introductory course in modern particle physics*, John Wiley and Sons.

[16] S. Narison (2004), *QCD as a theory of hadrons: from partons to confinement, Vol. 17*, Cambridge Univ. Press.

[17] D.J. Gross, F. Wilczek (1973), *Ultraviolet behavior of non-Abelian gauge theories*, Phys. Rev. Lett. **30**, pp 1343–1346.

[18] D.J. Gross, F. Wilczek (1973), *Asymptotically free gauge theories, I*, Phys. Rev. D. **8**, pp 3633–3652.

[19] H.D. Politzer (1973), *Reliable perturbative results for strong interactions*, Phys. Rev. Lett. **30**, pp 1346–1349.

[20] H.D. Politzer (1973), *Asymptotic freedom: An approach to strong interactions*, Phys. Rep. **14**, pp 129–180.

[21] K.G. Wilson (1974), *Confinement of Quarks*, Phys. Rev. D. **10**, pp 2445–2459.

[22] M. Bhagwat, P. Tandy (2006), *Analysis of full-QCD and quenched-QCD lattice propagators*, AIP Conf. Proc. **842**, pp 225–227.

[23] C.D. Roberts (2011), *Opportunities and challenges for theory in the N* program*, arχiv:nucl-th/1108.1030

[24] M. Arima, K. Masutani, T. Sato (2000), *Baryon resonances in a constituent quark model*, Prog. Theoret. Phys. Suppl. **137**, pp 169–188.

[25] S. Capstick, W. Roberts (2000), *Quark models of baryon masses and decays*, Prog. Part. Nucl. Phys. **45**, pp S241–S331.

[26] D.B. Lichtenberg, W. Namgung, E. Predazzi, J.G. Wills (1982), *Baryon masses in a relativistic quark-diquark model*, Phys. Rev. Lett. **48**, pp 1653.

[27] M. Anselmino, E. Predazzi, S. Ekelin, S. Fredriksson, D.B. Lichtenberg (1993), *Diquarks*, Rev. Mod. Phys. **65**, pp 1199–1234.

[28] J. Ferretti, A. Vassallo, E. Santopinto (2011), *Relativistic quark-diquark model of baryons*, Phys. Rev. C **83**, 065204.

[29] E. Santopinto (2005), *An interacting quark-diquark model of baryons*, Phys. Rev. C **72**, 022201.

[30] J. Beringer *et al.* (Particle Data Group) (2012), *Review of particle physics*, Phys. Rev. D **86**, 010001.

[31] G.Y. Chen, S. Kamalov, S.N. Yang, D. Drechsel, L. Tiator (2003), *Excitation of S_{11} resonances in pion scattering and pion photoproduction on the proton*, Nucl. Phys. A., **723** 3-4, pp 447–463.

[32] S. Schadmand (2005), *Photoabsorption and photoproduction on nuclei in the resonance region*, Proceedings of the first workshop on quark-hadron duality and the transition to pQCD.

[33] L.S. Cardman (2006), *Physics at the Thomas Jefferson National Accelerator Facility*, Eur. Phys. J. A **28**, s01, pp 7–17.

[34] B.A. Mecking *et al* (2003), *The CEBAF large acceptance spectrometer (CLAS)*, Nucl. Instr. Meth. A **503**, pp 513–553.

[35] The AIACE Collaboration (2000), *The AIACE collaboration (Genova-Frascati)* [online], Available at: <http://www.ge.infn.it/∼aiace/Aiace/aiace.html> [Accessed January 2014]

[36] R. Schumacher *et al* (1989), *Electromagnetic production of hyperons*, Jefferson lab experimental proposal E-89-004.

[37] F.J. Klein *et at* (2003), *Search for missing nucleon resonances in the photoproduction of hyperons using a polarised photon beam and a polarised target*, Jefferson lab experimental proposal E-02-112.

[38] N. Benmouna *et al* (2003), *Pion photoproduction from a polarised target*, Jefferson lab experimental proposal E-03-105.

[39] M. Dugger, E. Pasyuk *et al* (2004), *Measurement of polarisation observables in eta-photoproduction with CLAS*, Jefferson lab experimental proposal E-05-012.

[40] V.D. Burkert (2009), *The N* physics program at Jefferson Lab*, Chin. Phys. C **33** (12), 1043.

[41] K. Livingston (2009), *The Stonehenge technique. A method for aligning coherent bremsstrahlung radiators*, Nucl. Instr. Meth. A **603** (3), pp 205–213.

[42] A.M. Sandorfi, F.J. Klein *et al* (2006), *N* resonances in pseudoscalar meson photoproduction from polarised neutrons in* $\overrightarrow{H}$ $\overrightarrow{D}$ *and a complete determination of the* $\gamma n \rightarrow K^{circ}\Lambda$ *amplitude*, Jefferson lab experimental proposal E-06-101.

[43] C.D. Keith (2008), *The JLab frozen spin target*, Proc. of 18^{th} international spin physics symposium.

[44] V.D. Burkert (2008), CLAS12 and its science program at the Jefferson Lab upgrade, arχiv preprint hep-ph/0810.4718v2

[45] R.W. Gothe *et al* (2009), *Nucleon resonance studies with CLAS12*, Jefferson lab experimental proposal E12-09-003.

[46] A. Jankowiak (2006), *The Mainz microtron MAMI - past and future*, Eur. Phys. J. A **28** s01, pp 149–160.

[47] B.A. Mecking (2006), *Twenty years of physics at MAMI - What did it mean?*, Eur. Phys. J. A **28** s01, pp 209–219.

[48] W. Hillert (2006), *The Bonn electron stretcher accelerator ELSA: past and future*, Eur. Phys. J. A **28** s01, pp 139–148.

[49] G. Källé n (1964), *Elementary Particle Physics*, Addison-Wesley.

[50] A. Bettini (2008), *Introduction to Elementary Particle Physics*, Cambridge University Press.

[51] G.F. Chew, M.L. Goldberger, F.E. Low, Y. Nambu (1957), *Relativistic dispersion relation approach to photomeson production*. Phys. Rev. **106** (6), pp 1345.

[52] A.M. Sandorfi, S. Hoblit, H. Kamano, T.-S. H. Lee (2011), *Determining pseudoscalar meson photoproduction amplitudes from complete experiments*, J. Phys. G: Nucl. Part. Phys. **38**, 053001

[53] I. Barker, A. Donnachie, J. Storrow (1974), *Analysis of π^0 photoproduction at intermediate and high energies*, Nucl. Phys. B **79** 3, pp 431–360.

[54] D. Ireland (2010), *Information content of polarization measurements*, Phys. Rev. C 82, 025204

[55] T. Vrancx, J. Ryckebusch, T. Van Cuyck, P. Vancraeyveld (2013), *The incompleteness of complete pseudoscalar-meson photoproduction* arχiv preprint: arχiv:1303.2936v2 [nucl-th]

[56] I. Barker, A. Donnachie, J. Storrow (1975), *Complete experiments in pseudoscalar photoproduction*, Nucl. Phys. B 95, pp 347–356.

[57] R.A. Adelseck, B. Saghai (1990), *Kaon photoproduction: Data consistency, coupling constants, and polarisation observables*, Phys. Rev. C **42**, 108.

[58] W-T. Chiang, F. Tabakin (1997), *Completeness rules for spin observables in pseudoscalar meson photoproduction*, Phys. Rev. C 55 4, pp 2054–2066.

[59] R.K. Bradford, R.A. Schumacher *et al* (2007), *First measurement of beam-recoil observables C_x and C_z in hyperon photoproduction*, Phys. Rev. C **75**, 035205.

[60] J. Skilling (2006). *Nested sampling for general Bayesian computation*, Bayesian Analysis 1 4, pp 833–859.

[61] T. Bayes, R. Price (1763). *An Essay Towards Solving a Problem in the Doctrine of Chances*, Philosophical Transactions (1683-1775) 53, pp 370–418

[62] D. S. Sivia, J. Skilling (2006). *Data Analysis: A Bayesian Tutorial, 2nd ed.*, Oxford Science Publications, pp 3–10.

[63] J. Neyman (1935). *On the Problem of Confidence Intervals*, Ann. Math. Statist. 6 3, pp 111–116.

[64] J. Aldrich (2008). *R. A. Fisher on Bayes and Bayes' Theorem*, Bayesian Analysis 3, 1, pp 161–170.

[65] B. Efron (1978). *Controversies in the Foundations of Statistics*, American Mathematical Monthly 85 4, pp 231–246.

[66] F.J. Samaniego (2010), *A Comparison of the Bayesian and Frequentist Approaches to Estimation*, Springer Series in Statistics.

[67] D. J. Bartholomew (1965). *A comparison of some Bayesian and frequentist inferences*, Biometrika 52 (1-2), pp 19–35.

[68] W. R. Gilks, S. Richardson, D. J. Spiegelhalter (1996). *Markov chain Monte Carlo in practice*, London Chapman and Hall, pp 1–19.

[69] B. P. Carlin, T. A. Louis (2009). *Bayesian Methods for Data Analysis, 3rd ed.*, Texts in Statistical Science, CRC Press, pp 120–152.

[70] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, E. Teller (1953). *Equations of State Calculations by Fast Computing Machines*, Journal of Chemical Physics 21 6, pp 1087–1092.

[71] S.M. Lynch (2007), *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*, Springer.

[72] S. Jackman (2000). *Estimation and Inference vis Bayesian Simulation: An Introduction to Markov Chain Monte Carlo*, American Journal of Political Science, 44 2, pp 375–404.

[73] J. Skilling (2004). *Nested Sampling*, AIP Conf. Proc. 735, pp 395–405.

[74] D. Ireland (2011), *Extracting Polarization Observables in Pseudoscalar Photoproduction Reactions*, CLAS-NOTE 2011-010

[75] C.A. Paterson (2008), *Polarization Observables in Strangeness Photoproduction with CLAS at Jefferson Lab*, PhD thesis, University of Glasgow.

[76] K. Livingston, Private Communication.

[77] J. H. Pollard (1977). *Numerical and Statistical Techniques*, Cambridge University Press, pp 211.

[78] R. J. Barlow (1989). *Statistics: A Guide to the use of Statistical Methods in the Physical Sciences*, John Wiley & Sons, pp 81–90.

[79] ROOT (1995) *Root Manual: TMinuit* [online] Available at: <http://root.cern.ch/root/html/TMinuit.html > [Accessed July 2013]

[80] W. Vanderbauwhede, S. Lewis, D. Ireland (2013), *Implementing data parallelisation in a Nested-Sampling Monte Carlo algorithm*, High Performance Computing and Simulation (HPCS) International Conference Proceedings, pp 512–518.

[81] OpenMP (2011) *OpenMP Application Program Interface* [online] Available at: <http://openmp.org/wp/openmp-specifications/> [Accessed February 2012]

[82] A. Grama, A. Gupta, G. Karypis, V. Kumar (2003), *Introduction to Parallel Computing, 2nd ed.*, Pearson Addison Wesley, pp 310–330.

[83] Khronos Group, 2011. *The OpenCL Specification, Version 1.2* [online] Available at: <http://www.khronos.org > [Accessed November 2011]

[84] W. Vanderbauwhede (2013), *OclWrapper Class* [online] Available at: <https://github.com/wimvanderbauwhede/OpenCLIntegration> [Accessed September 2013]

[85] G. van Venrooij, F. Orderud, I. Cullinan (2012). *CLCC - The OpenCL Kernel Compiler* [online] Available at: <clcc.sourceforge.net> [Accessed March 2013]

[86] B. Dey, M. E. McCracken, D. G. Ireland, C. A. Meyer (2011), *Polarization observables in the longitudinal basis for pseudo-scalar meson photoproduction using a density matrix approach*, Phys. Rev. C 83, 055208

[87] B. McKinnon, Private Communication.

[88] D. Ireland, Private Communication.

[89] R. Bradform (2005), *Measurement of Differential Cross Sections and C_x and C_z for $\gamma p \to K^+\Lambda$ and $\gamma p \to K^+\Sigma^0$ Using CLAS at Jefferson Lab*, PhD thesis, Carnegie Mellon University.

[90] M. Williams (2005), CLAS Objective ROOT Analysis [online] Available at:<http://www-meg.phys.cmu.edu/cobra/> [Accessed November 2013]

[91] S. Fegan (2011), *Polarisation Observables from Strangeness Photoproduction on a Frozen Spin Target with CLAS at Jefferson Lab*, PhD thesis, University of Glasgow.

[92] A. Thomas (2012), *Polarisation Observables from the Photo-disintegration of the Deuteron*, Mainz Microtron proposal Exp.-Nr. A2-01/2012.

[93] L. Nyhoff, S. Leestma (1996), *FORTRAN77 for Engineers and Scientists, 4th ed.*, Prentice Hall.

[94] S.M. Mueller (2004), *Upgrading and Repairing PCs, 15th ed.*, Que Publishing, pp 308–320.

[95] J.Y. Hsu (2001), *Computer Architecture - Software Aspects, Coding and Hardware*, CRC Press, pp 79–82.

[96] A.R. Brodtkorb, T.R. Hagen, M.L. Sætra (2013), *Graphics processing unit (GPU) programming strategies and trends in GPU computing*, J. Parallel Distrib. Comput., **73**, pp 4–13

[97] D. Luebke, G. Humphreys (2007), *How GPUs Work*, Computer, **40** 2, pp 96–100

[98] X. Guan, H. Wu (2010). *Leveraging the power of multi-core platforms for large-scale geospatial data processing: Exemplified by generating DEM from massive LiDAR point clouds*, Computers and Geosciences, **36** (10), pp 1276-1282.

[99] D. Q. Ren (2011). *Algorithm level power efficiency optimization for CPUGPU processing element in data intensive SIMD/SPMD computing*, Journal of Parallel and Distributed Computing, **71** (2), pp 245–253.

[100] J.P. Hayes (1998). *Computer Architecture and Organization, 3rd ed.*, McGraw-Hill, pp 542–560

[101] W. D. Hillis and G. L. Steele, Jr. (1986), *Data parallel algorithms*, Commun. ACM, **29**, pp 1170–1183.

[102] D. Andrade, B.B. Fraguela, J. Brodman, D. Padua (2009). *Task-parallel versus data-parallel library-based programming on multicore systems*, Euromicro Conference on Parallel, Distributed, and Network-Based Processing, pp. 101-110

A

Parallel Computing

The extremely quick developments in the field of computing science have resulted in a reduced focus on efficient programming techniques. In many cases seen in physics, programmes require a negligible amount of time to provide results and any speed-up gained from employing an efficient programming style would go unnoticed. There are some algorithms, however, that can benefit greatly from shrewd programming and, in some cases, data parallelisation.

In order to take full advantage of the available computing power, it is important to recognise how some of the basic computer hardware components interact. Section 1 will detail the key aspects of computer hardware interaction, Section 2 will provide some theory on data parallelisation, and Section 3 discusses the software tools and implementations relevant to this work. Much of this appendix has been published in [80].

A.1 Fundamentals of Computer Architecture and Hardware

It is crucial to understand the components and various interactions that occur within the computer to exploit the performance of the hardware available. Some of the key components have undergone significant developments in the last decade, and these evolutions have resulted in massively increased computing power. The development of the multi-core processor has led the way in advancing the performance capabilities - but apart from speeding up computers, the fundamental interactions occurring within the computer's hardware have changed. Realising how these components interact can have a great impact on programme development and can save a staggering amount of time.

A computer can be described by a collection of interconnected hardware components. The standard computer architecture is the von Neumann machine (Figure A.1), shown below.

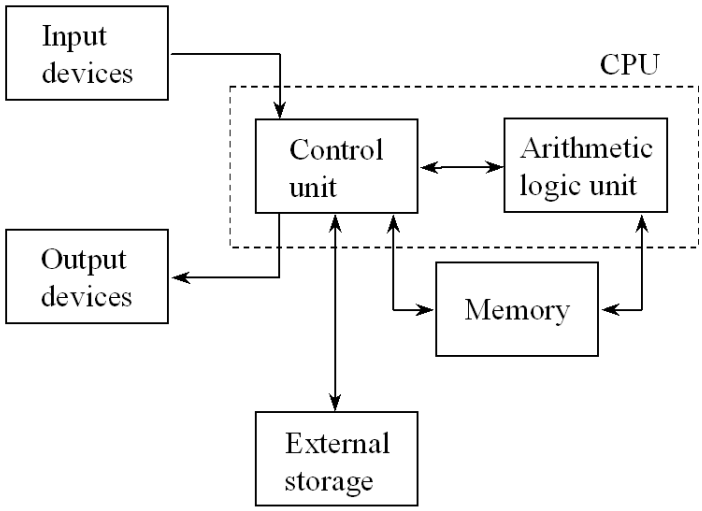


Figure A.1: Basic von Neumann computer architecture [93]

The central processing unit, or CPU , can be considered the "brain" of a computer. It is essentially a highly optimised serial von Neumann processor, designed

to execute sequential operations quickly [96]. This is generally where data are processed, and interactions with memory are frequent during the execution of a programme. This memory interaction takes place via an interface known as the front-side bus [94]. A bus is a collection of copper wires that are used to transmit data, memory addresses and control signals [95]. A variety of buses are present in a standard computer, and they can have significant differences. Properties such as clock frequency can vary from one type of bus to another, and this can have an impact in programme performance.

Up until the early 2000s, computers were improved by increasing the processor frequency. As frequency was increased, the processor speed also improved. This trend hit a wall however, due to the amount of power required by the CPU [96]. The power consumption per unit area was on track to match that of a nuclear reactor. It was at this point that no further improvement to the CPU could realistically be obtained and other methods of computing improvements had to be developed. Parallelism in computers then became the focus - from multicore CPUs to GPU (graphical processing unit) computing. Graphics cards lend themselves naturally to parallel programming.

The fairly recent introduction of multi-core CPUs has resulted in the bulk of all computing devices (from desktop computers to mobile phones) reaping the benefits of multi-core processing. A multi-core CPU can be defined as an individual component containing at least two independent processing units, which are generally referred to as cores (shown in Figure A.2). These cores can process data, execute instructions and interact with memory concurrently. The simultaneous use of multiple processing units is particularly well suited for parallel programming.

The developments made in graphical processing are generally due to the growing demands made by the gaming industry, but the benefits are increasingly being exploited within the scientific community. The GPU was specifically designed to render increasingly detailed images. Images are broken down into simple geometric units, usually triangles. In gaming, the graphics must be able to adjust quickly to new camera angles and positions. This requires the GPU to quickly transform these triangles from one coordinate system to a new one [97]. Each pixel is treated as

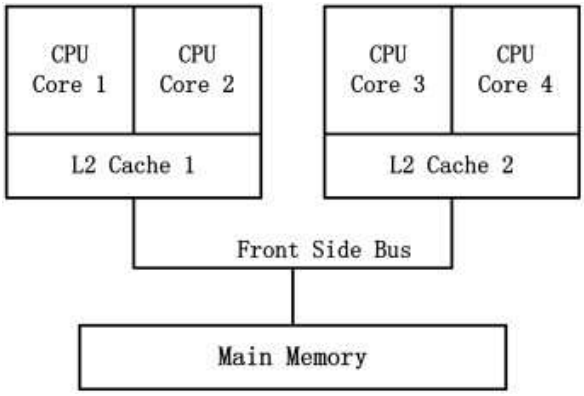


Figure A.2: Multi-core CPU diagram (quad-core Intel processor) [98]

an independent object, and the calculations required to transform each pixel into a new coordinate system are generally performed in parallel. Once the "power wall" was reached with raw CPU power, efforts were made to generalise the approach taken by most GPUs in order to improve computer performance. The transition from treating pixels as independent objects to parallelising operations or data for programming structures or objects is fairly intuitive. As GPUs became more mainstream in scientific disciplines, hardware vendors released higher level, non-graphics languages to allow for more straightforward parallel programming. This led to a substantial amount of research and development into general-purpose GPU (GPGPU) computing. There are hundreds of smaller cores within a GPU, with each simple core able to handle thousands of hardware threads, each of which identifies a task or programme that can be run concurrently [96] - and the number of smaller cores is increasing steadily. Performance improvements through GPGPU computing are therefore sustainable and increasingly inexpensive.

The GPU's memory is broken down into three types - the registers, the shared memory, and the global memory. These are illustrated in Figure A.3. Of these memory types, the registers are the fastest units. One register file exists for each multiprocessor, and is split amongst the threads [96]. They provide private memory for each thread, so one thread cannot access another thread's memory. The GPU's shared memory is accessible to all threads within the core. Data can be shared between threads, but access to shared memory is not as fast as access to a register

file. The global memory is considerably slower than both the register files and the shared memory [96]. It forms the main memory of the GPU and as such is not private to any threads or blocks. It has a high latency, which can decrease the effectiveness of parallel programming if not taken into consideration. The latency can be hidden by making appropriate choices of parallelisation parameters, such as number of threads and blocks over which to parallelise data or instructions. This is generally dependent on the algorithm(s) as well as the hardware architecture of the system.

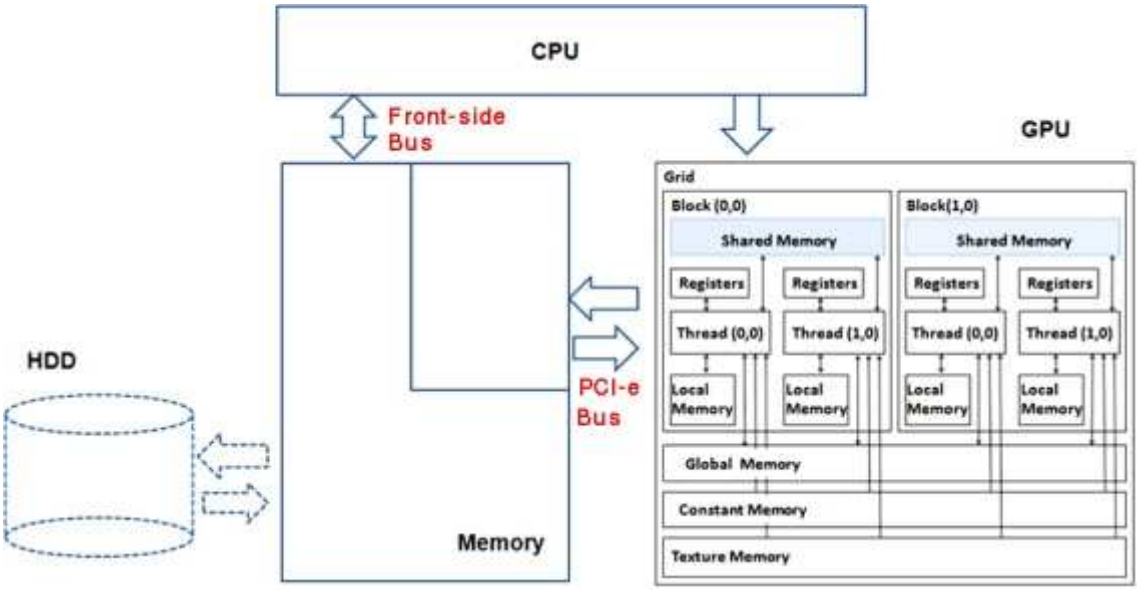


Figure A.3: Abstract computer architecture showing bus connections between components [99]

Data from the main memory is transferred to the cache of a compute unit in quantised amounts, known as *cache lines*. The transfer of instruction and data from memory to the CPU occurs via the front-side bus. The GPU is connected to the global memory by the PCI-express (PCI-e, or peripheral component interconnect express) bus [96], which is considerably slower than the front-side bus. The main difference between the two buses is the data transfer time, which is significantly longer with the PCI-e bus. This means that transferring data to the GPU has much more overhead than transferring it to the CPU would have. Data transfer across the PCI-e bus for a single element of data is equivalent to transferring an entire cache

line, so there are cases where running parallelised code on a GPU is not the most effective method.

A.2 Parallelising Data

Since the performance of a serial CPU has reached its pinnacle, parallel programming has become the focus of computing performance development. The use of massively multi-core GPUs has become mainstream in everything from scientific computing and state-of-the-art gaming technology to standard desktop computers and laptops. There is now a sustainable path to improving computing technologies for the foreseeable future. Although the spotlight is currently on GPGPU computing, it must be remembered that all programmes and algorithms must include some amount of sequential code, even if they exist solely to execute kernel functions or perform or some standard initialisations. In most cases, these serial sections of a programme create bottlenecks that no amount of parallelisation can avoid. For this reason, *heterogeneous computers* are favoured - that is, the combination of highly optimised serial CPU cores with the massively parallelisable GPU cores. These two main components of a computer must complement each other - if the CPU is outdated and obsolete, any speed-up obtained from a high-end GPU will be hidden by the slow processing at one of these bottlenecks. In order to make the most of the available hardware, both components must be taken into consideration.

A.2.1 Flynn's Taxonomy

A processor fetches instructions and data from memory. Conceptually, this can be considered as two separate streams of information being transferred between the processor and the memory (as shown in Figure A.4).

Traditionally, hardware systems have been classified using Flynn's taxonomy, which categorises systems into four types: Single Instruction stream Single Data (SISD) , Single Instruction stream Multiple Data (SIMD) , Multiple Instruction stream Single Data (MISD) and Multiple Instruction stream Multiple Data (MIMD) [100].

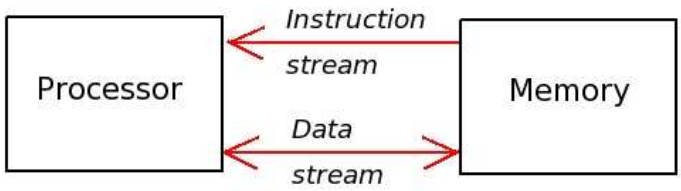


Figure A.4: Interaction between processor and memory [100]

In the first instance, SISD, a computer contains a single CPU solely capable of simple, sequential algorithms. This type of machine can perform one operation on one piece of data at a time and there is no scope of parallelisation. The second category, SIMD, is much more common. In this case, there exists one processor, but with multiple means of independent execution. MISD systems are extremely rare. They are parallel computers with multiple processors operating on one datum. These are generally only useful if fault-tolerance is the main objective of a system. Lastly, MIMD systems are currently the most common and consist of multiple processors capable of running multiple programmes simultaneously [100].

Although initially used to categorise the various types of computer hardware systems, Flynn's taxonomy has recently been applied to software and algorithms in a similar manner. SISD refers to a sequential algorithm (one piece of data being used in a single operation), SIMD includes programmes that perform the same operation on many data concurrently, MISD (although highly uncommon) would describe an algorithm that performs many operations on a single datum, and MIMD indicates a programme that can perform a multitude of operations on multiple data simultaneously.

SIMD algorithms are at the heart of data parallelisation. There are many cases in data analysis where an algorithm performs operations on many data points independently. For these situations, a considerable speed-up can be achieved by dividing the data set into smaller sections and performing the same operations in parallel, with one thread handling one section of data. The results from each thread can then be combined, as illustrated in Figure A.5.

It is important to note that not all algorithms can be parallelised; recursive and

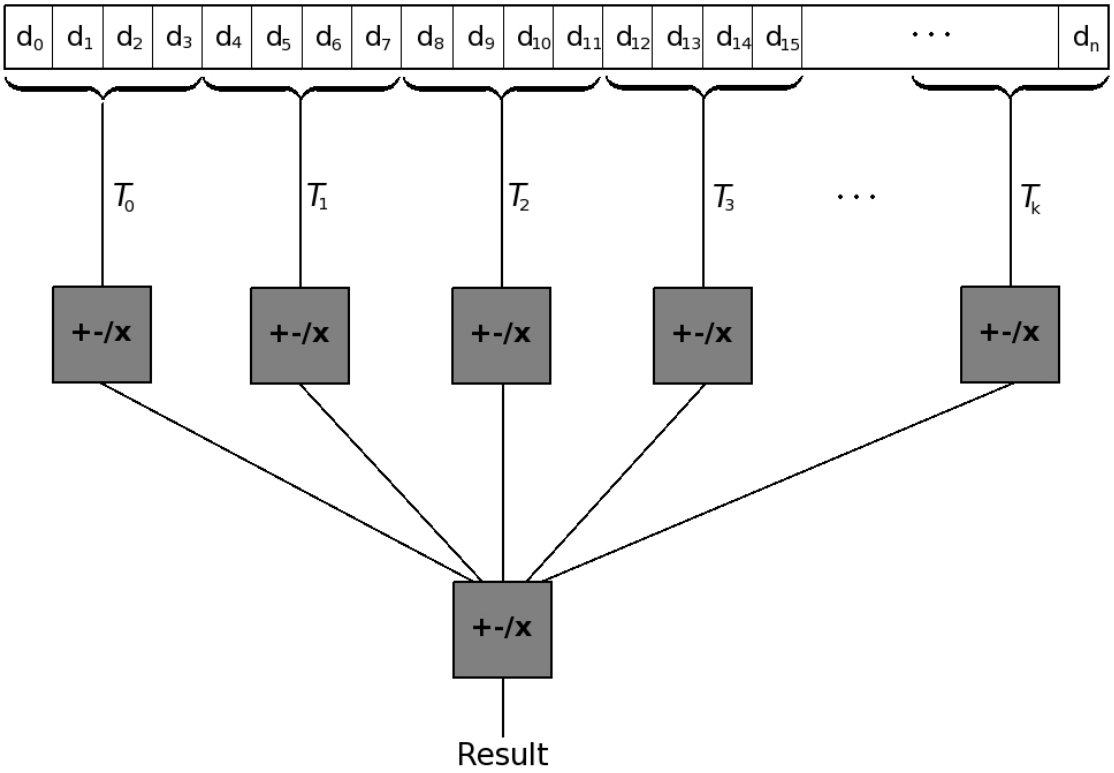


Figure A.5: Visual representation of a SIMD parallelised algorithm

sequential programmes, or even serial sections of code, can form bottlenecks that impede the run-time of a programme. There are some cases where parallelising data over multiple threads or cores can result in a slower run-time as no speed-up is gained and time is lost during data transfer or thread initialisations. Even in cases where an algorithm lends itself naturally to parallelism, it is crucial to understand exactly where and how it should be done in order to obtain the greatest benefit.

A.3 Software and Implementations

In this work, efforts were made to make analysis algorithms as efficient as possible. Two types of implementations were explored, in addition to the standard (unthreaded) programming methods. The software used for this will be introduced and explained here.

A.3.1 OpenCL

Until recently, writing programmes to run on a GPU required a substantial amount of initialisations and "boiler-plate" programming. The software tools and packages required were dependent on the manufacturer, and creating portable programmes was non-trivial. In 2008, the Khronos Group developed OpenCL (Open Computing Language) , a framework that allowed programmes to be executed on heterogeneous systems. With OpenCL, a programme can be parallelised and run on a GPU or indeed on a CPU (single- or multi-core).

A programme written using the OpenCL framework is comprised of two parts: the *host* and the *kernel*. The host contains the bulk of the programme, essentially everything that is not parallelised. There are convenient bindings for several standard programming languages, including C++ [83]. The kernel contains the function(s) that will be parallelised and is called by the host. The host programme passes all necessary data and variables to the kernel in each call. Using OpenCL, it is possible to run a programme over a specified number of cores (*compute units*) and a specified number of threads (*processing elements*), which can be passed as arguments to the kernel as well. The kernel is written in a language based on C-99,

and can occasionally be limited by the available hardware (in many cases, a particular GPU does not support the 'double' typedef, for example). Each instance of the kernel is known as a *work-item*, and can be identified by a global ID. Each work-item is allocated some predetermined amount of data, over which it performs the same calculation(s). These work-items are organised into *work groups*, in which each work-item is assigned a unique local ID. The work groups themselves are also assigned a work-group ID. These IDs are required for combining results of calculations in a clear and robust manner. Within a work group, the work-items run concurrently on the processing elements within a compute unit [83].

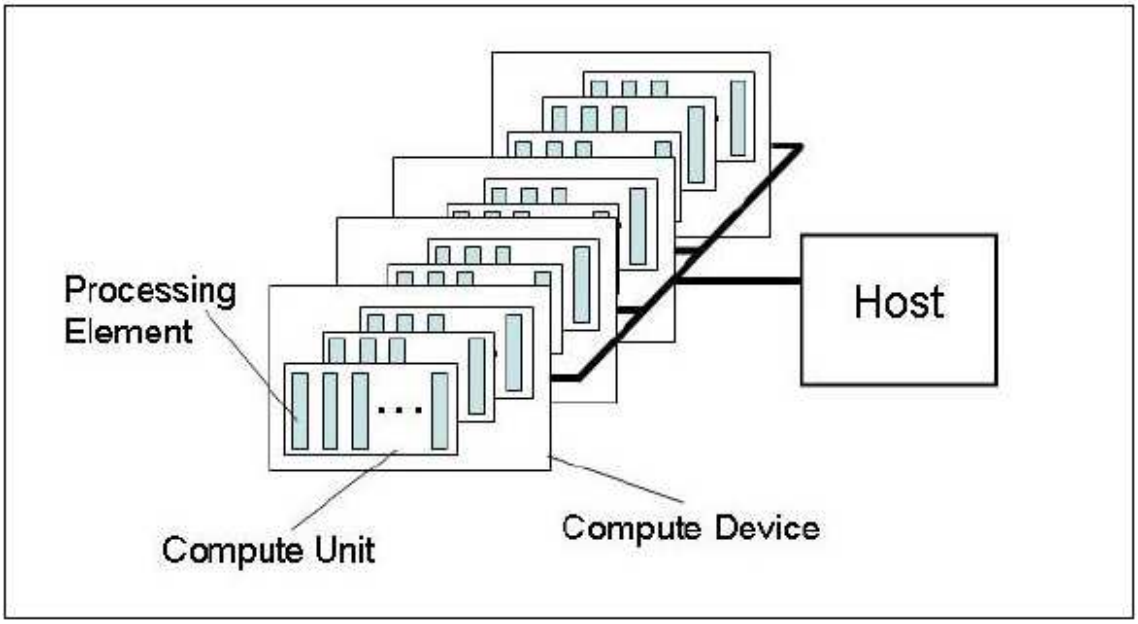


Figure A.6: Platform model showing of OpenCL architecture [83]

OpenCL provides support for both data parallel programming as well as task parallel programming. Data parallel programming involves performing the same calculation or programme over a wide set of data. The data is distributed across a number of threads, and all threads execute the same code over the subset of data [101]. The task parallel programming model is focused on the principle of running multiple programmes on different threads simultaneously [102]. In the field of nuclear physics data analysis, it is more common to encounter situations where the same instructions are carried out over large datasets, so the task parallel programming

model will not be discussed in any more detail here.

It is only recently that OpenCL has included C++ bindings in the software development kit. The introduction of these bindings has done much to simplify the intialisations and instantiations required when building an OpenCL programme. Despite this, there is still a considerable amount of tedious programming involved.

The structure of an OpenCL programme involves many explicit definitions and initialisations that are not usually considered when programming in C or C++. First, a platform must be declared. The platform includes the host programme and a list of devices and allows the sharing of resources and execution of kernel functions on included devices [83]. Next, a context is set up, which refers to an environment where kernels are executed and memory management is prescribed. Devices then must be found - typically, a device would refer to a GPU or multicore CPU. More formally, a device can be considered as a list of compute units. Once the devices are found, the kernel can be loaded and built. OpenCL handles arrays of data slightly differently to C or C++, in the sense that specific buffer objects must be created and initialised. A buffer object stores a collection of data and can be used as a pointer in the kernel. In the declaration of a buffer object, the memory type must be explicitly set - either a buffer object is a *read* buffer (where the contents cannot be changed in the kernel), a *write* buffer (where the contents are written in the kernel and cannot be read inside a kernel), or a *read-write* buffer (where the contents can be both read and changed in the kernel) [83]. Any other arguments to be passed to the kernel function must then be set. A command-queue is then created which contains a list of commands that are to be executed on a given device. The arrays from the CPU are then pushed to the appropriate compute device and the kernel can then be executed. Once the kernel is executed successfully, the results can be read from the device and, once the command-queue is finished, the programme can exit [83].

Most of these explicit initialisations described above are very similar from one programme to the other, and for this reason a very helpful class (*OclWrapper*) was developed to take care of most of the more standard operations that are required in OpenCL. The *OclWrapper* class was developed by W. Vanderbauwhede [84], a

collaborating colleague from the University of Glasgow's School of Computing Science. The main aim of the class was to simplify the arduous initialisations required by OpenCL. It also provides a platform-portable build script, allowing any OpenCL programme to be run on a variety of system architectures (assuming the appropriate software is installed).

There are many benefits of using OpenCL over other available software development kits and frameworks (such as CUDA). OpenCL is supported on most of the widely distributed platforms, including NVidia, Intel and AMD, on Microsoft, Apple and Linux operating systems. This results in a portability of parallelised programmes that has not been available prior to OpenCL's development. It is also possible to specify the compute unit used (e.g. GPU or CPU). Running a parallelised programme on a multicore CPU can provide significantly faster runtimes, even compared to running the same programme on a GPU. This is seen to be the case in situations where the dataset is fairly small and the kernel is called many times. The data is passed to the kernel in each iteration, and significantly more time is spent transferring data from the host to the GPU in comparison to the CPU.

Despite the numerous advantages of OpenCL over other implementations, there are some important disadvantages that need to be discussed. The fact that each manufacturer (NVidia, Intel, etc.) provides their own software development kit means that the level of support available varies from one manufacturer to another. Also, the programming required is generally nontrivial. In many situations, easier and more efficient solutions may exist. In cases as described previously where a relatively small dataset is being passed to the kernel many times, other software tools, such as Open Multiprocessing (OpenMP) , can prove to be much more successful.

A.3.2 OpenMP®

It is often more straightforward to parallelise a programme (or function) by multithreading on a CPU instead of using the more powerful software like OpenCL. OpenMP is an application progamming interface (API) that can be used to manipulate threads and parallelise functions [82]. It was designed to allow programmers to use threads without the hassles associated with creating, destroying and synchro-

nising individual threads [81]. It is supported by multiple platforms and is easily portable. In contrast to OpenCL, it is relatively simple to incorporate OpenMP into existing C, C++, or FORTRAN code. There is no accommodation for GPU programming (or running code on different devices) with OpenMP as it is restricted to the CPU. In C and C++, OpenMP is essentially a collection of #pragma compiler directives. Using these compiler directives, it is possible to specify the variables used, their scopes, and the number of threads over which the region of code is to be parallelised [82]. The handling of race conditions, where multiple threads access shared memory at the same time (either to read it or write to it), can be managed in a straightforward manner as well. OpenMP provides locks and semaphores for this reason. Despite not being able to use on other compute devices, OpenMP can provide significant speed-ups with relatively few (and simple) adjustments to pre-existing code. It is particularly suited to problems where the parallelised region of code is called iteratively or where the dataset size is relatively small (in most cases less than one million events can be considered small). Unlike OpenCL, OpenMP is already supported by most platforms - it is included in more recent versions of the gcc/g++ compilers, for example. This allows greater portability from system to system, and it is not necessary to install software development kits to enable its use [81].

B

Additional Plots

All additional plots can be found online at:

http://nuclear.gla.ac.uk/∼stefl/Thesis/slewis_thesis_appendixB.pdf