Timelike Compton Scattering from a Longitudinally Polarised Target with CLAS12 at Jefferson Lab

Kayleigh Gates

Submitted in fulfilment of the requirements for the
Degree of Doctor of Philosophy

School of Physics and Astronomy
College of Science and Engineering
University of Glasgow

2024

Abstract

Explorations into the internal dynamics of hadrons are constantly evolving, and the requirement for experimental results to verify theoretical models of hadron structure is paramount. A key area in this field is the study of Generalised Parton Distributions (GPDs), which are functions used to model the momenta of quarks and gluons within hadrons, and the methods to access GPDs experimentally. One such scattering process that allows access to these is Timelike Compton Scattering. TCS complements existing Deeply Virtual Compton Scattering experiments and allows investigation into the universality of GPDs through access to the real and imaginary parts of the parton helicity independent GPD H^q via beam spin asymmetries (BSA), and it provides novel access to the real and imaginary parts of the parton helicity dependent GPD $\tilde{H}^q$ through target polarisation asymmetries (TSA). This thesis work presents a comparative study with the first published BSA for TCS at the Thomas Jefferson National Accelerator Facility (JLab), alongside a first time extraction of a Target Spin Asymmetry with the Summer 2022 data taking run.

JLab hosts the Continuous Electron Beam Accelerator Facility (CEBAF) which provides a 12 GeV electron beam to four experimental halls. Hall-B contains the CEBAF Large Acceptance Spectrometer, which took data across three run periods on a longitudinally polarised NH_3 and ND_3 fixed target from 2022-2023, to extract measurements of electron-proton scattering, from which a TCS signal could be extracted. The thesis discusses work done to understand and eliminate contributions from the non-/low-polarised nuclear background, testing pre-established cuts to eliminate pion background from a dilepton (e^+e^-) final state and modifying them as needed for the new experimental run, and attempts to hone in on a clean TCS signal from which to extract the two asymmetry observables.

A comparison with existing BSA results was performed; however, the statistical errors are too large to draw a significant conclusion as to whether there is agreement across each bin. More data is needed for a multidimensionally binned extraction. A proof of principle was achieved in the TSA measurements, with two out of four kinematic bins showing preliminary agreement in shape with theoretical values. Again, the errors are significant due to the contributions from the nuclear background. To support these conclusions, a further study was done, which takes into account an estimate of the asymmetries with the full available dataset (this thesis is based

only on data taken in the summer set; at the time of writing processing was still being conducted for the final two datasets), as well as an estimate including additional future experiment days that were awarded in July 2024. Additional work was done on a secondary project exploring the feasibility of measuring TCS at the upcoming Electron Ion Collider, supporting the design proposal for the detector for the first interaction region and giving a positive outlook for the future of these types of measurements beyond JLab.

Contents

List of Tables

List of Figures

Acknowledgements

Firstly, I would like to express my gratitude for my supervisory team. To Daria Sokhan, for providing me with the opportunity to undertake this project, aiding me in extending it from a mini project to an entire thesis, for all the late-night email chains, explanations of concepts that were novel to me, and your encouragement and support throughout. To Rachel Montgomery, for always being around for a chat whenever I needed. Your advice and expertise throughout the EIC ECCE project was invaluable and your constant encouragement truly made me believe I could do it. Finally, a special thanks goes to Bryan McKinnon, whose guidance, direction, and grounding during the crucial final half of this PhD truly brought everything together. With the support from all of you, I felt I was absolutely able to make the most of this project.

I am endlessly grateful to everyone over at JLab for the discussions and advice, special thanks go to everyone in the Dilepton and RGC working groups for their support throughout my analysis. I want to make specific mention of Rafayel Paremuzyan and Pierre Chatagnon and thank them for the many email chain discussions and expertise regarding the TCS channel, and their work on the event generators used in my simulation studies. I am exponentially grateful also to Silvia Niccolai, for her endless support while I conducted my role as RGC chef and her continued encouragement and guidance throughout my analysis.

A huge thank you goes out to the Nuclear and Hadron physics group at Glasgow, both staff and students alike, for the helpful discussions in meetings, over email and teams, and for generally keeping me sane. Special thanks go to Derek Glazier, whose software tools and guidance on how to use them made this analysis possible. I would also like to extend a particular thanks to Paul Naidoo, your chats over various caffeinated drinks kept me sane and kept me afloat and I will be forever grateful. Thanks to all the guys in 414 for being stellar office mates, and to all of the rest of the group for the laughs and the chats throughout, couldn't have asked for a better bunch!

To my family back home, your unwavering belief in me is what got me here. Every exam result, every single step in my journey you have been there as my own personal cheer squad. You have had my back even when I have doubted myself the most. You always knew I could do it, and it was my absolute honour to prove you right.

Lastly, I can't even express the amount of gratitude I have for my partner Robert. I am beyond lucky to be sharing my life with someone who pushes me beyond even my own expectations.

Your patience, kindness, love and motivation have been second to none, as too has the support from your family. Thank you for everything.

Declaration

All work performed in this thesis is the work of the author, unless otherwise explicitly stated/referenced, and has not been submitted for any other degree or qualification.

Chapter 1

Introduction

1.1 Hadronic Properties and Motivations

The constituent quark model explains that a proton consists of three valence quarks, *uud*, so the assumption naturally follows that the mass of a proton is the calculated sum of the masses of these three quarks. An up quark (u) has a mass of approximately 2.16 MeV and the down quark (d) approximately 4.67 MeV [1], so the total mass of the proton should be approximately 8.99 MeV. This is, in fact, less than 1% of the experimentally ascertained mass of the proton, 938 MeV. This is indicative of another source of mass other than the three valence quarks, the likely candidate arising from dynamical interactions between quarks and gluons within the nucleon [2, 3]. Another problem in the constituent quark model is the discrepancy between modelled and measured proton spin. Experimental data has shown that roughly 30% of the total spin of a proton comes from quarks, and up to 70% of spin comes from gluons and the orbital angular momentum of quarks [4–6].

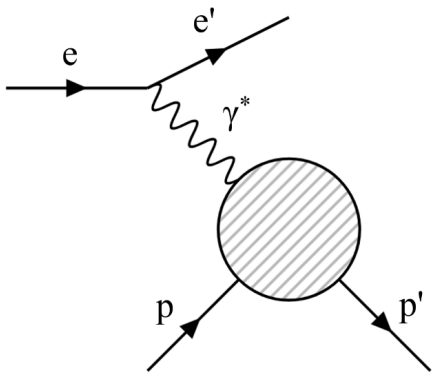
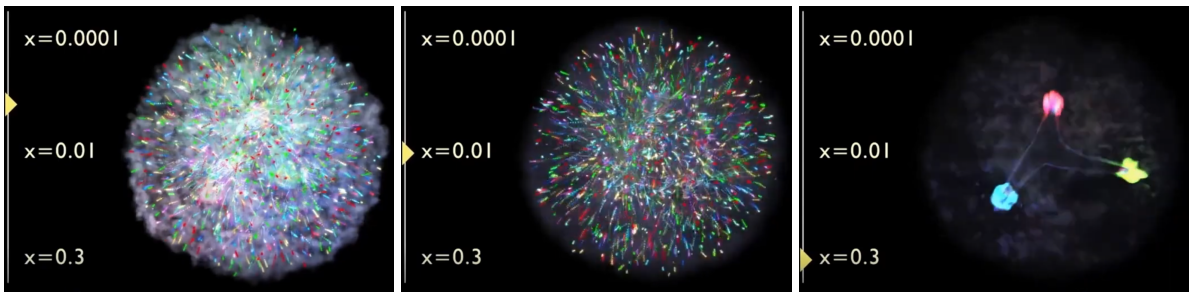


Figure 1.1: Diagram showing a scattering process where incoming electron e exchanges a virtual photon γ^* (four momentum = q) with a nucleon p.

At distances $\approx$ 1 fm, quarks and gluons are confined inside hadrons by the action of the strong



(a) (b) (c)

Figure 1.2: Visualisation of the proton structure at different x regions. (a) Low x region, where gluons dominate. (b) Mid x region, where the quark sea begins to become visible. (c) High x region, where valence quarks dominate. Taken from the Visualising Proton Structure video created by Rolf Ent and Richard Milner [8].

force. Quantum Chromo-Dynamics (QCD) describes this action; however, in the confinement regime, things become decidedly more complex because of the size of the strong coupling constant, α_S, at these energy scales. In scattering processes, such as in Figure 1.1 where a photon is incident on a nucleon, for large values of this incident photon virtuality, $Q^2 = -q^2$ (where q represents the four-momentum of this photon), α_S is small and so the associated physics remains in a more well understood and calculable perturbative QCD regime. This corresponds to small values of nucleon momentum carried by the partons, known as the longitudinal momentum fraction x. This is shown in a video put together by Richard Milner (MIT) and Rolf Ent (Jefferson Lab) which details the internal structure of the proton at various values of x, screenshots of which are shown in Figure 1.2. At the lowest values of x the gluons begin to dominate and it is hard to separate key features at all, but, as x increases, the sea quarks start to appear followed by the valence quark structure at $\mathcal{O}(10^{-1})$. Because of this, a non-perturbative QCD approach is required to describe the action [7].

The distribution of partons within a nucleon is theoretically modelled using distribution functions, however numerous different experimental scattering processes are required to not only access each parameter of such distributions but also to assess the extent to which they are process dependent, or, their universality. A wealth of information on the inner structure of hadrons can be extracted from scattering processes, one such example is the impact parameter, or transverse position (b) and longitudinal momentum fraction (x) of the up quark in both an unpolarised (top) and longitudinally polarised (bottom) proton, detailed using analysis of Deeply Virtual Compton Scattering experiments shown in Figure 1.3. From these distributions, information on the position of quarks as a function of their longitudinal momentum in an unpolarised proton can be extracted, as well as how the longitudinal polarisation of the quarks is distributed in the longitudinally polarised case.

Such experiments are performed at facilities worldwide, at a range of different energies, each

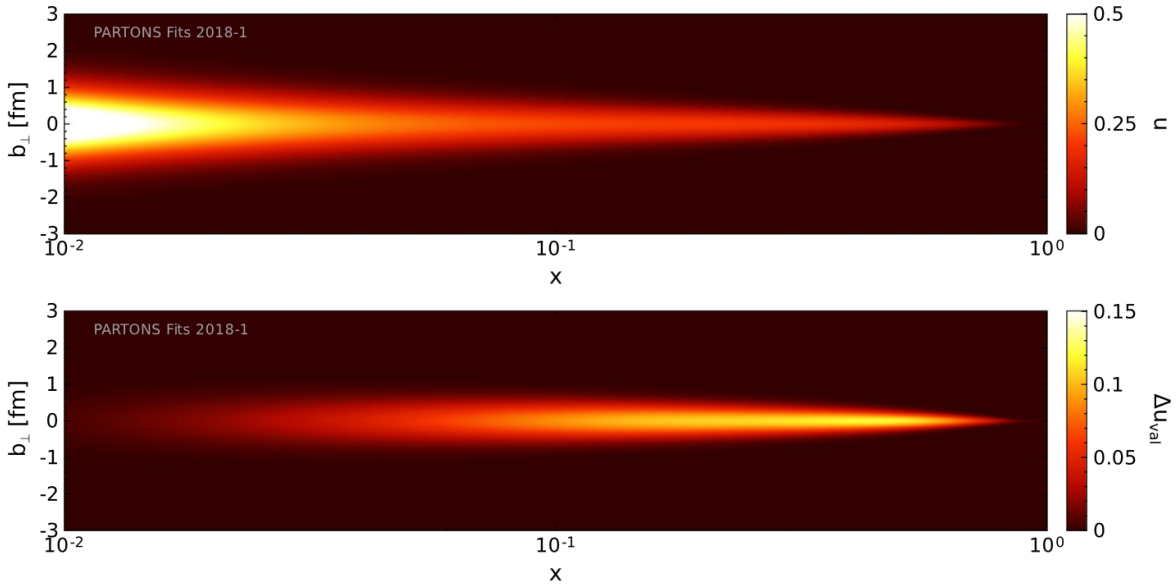


Figure 1.3: The position of the up quark in the proton at different momenta shown by impact parameter b vs x from Deeply Virtual Compton Scattering experiments. The top plot shows all up quarks and the bottom plot shows the longitudinal polarisation of only the valence quarks (dennoted u_{val}) as a function of b and x. Taken from [9]

attempting to improve the understanding of these internal distributions at different energy and momentum scales. Figure 1.4 shows a distribution of Q^2 vs x_B representing the kinematic phase space coverage of a range of different experimental facilities, including data from the two electron beam energy eras of the Thomas Jefferson National Accelerator facility (JLab), 6 GeV and 12 GeV.

The following thesis details an analysis performed on a longitudinally polarised ammonia (NH_3) target with the CLAS12 (CEBAF Large Acceptance Spectrometer at 12 GeV) detector at JLab - a fixed target electron accelerator facility in Virginia - with the aim of extracting Timelike Compton Scattering (TCS) asymmetries. These particular observables provide access to the real and imaginary parts of Generalised Patron Distributions (GPDs), which aid in providing a tomographic mapping of nucleon structure.

The document is split into the following sections. Chapter 2 details the theoretical background of the experiment, beginning with outlining the basic principles of elastic scattering and then extending this up to non-perturbative QCD, before detailing the key observables accessible experimentally and the phenomenology of TCS. Chapter 3 details the experimental setup, including CEBAF and the CLAS12 detector, along with details on the longitudinally polarised target. Chapter 4 discusses the analysis stages, including the Monte Carlo simulation, methods of nuclear background subtraction, corrections and the fitting of the data. Chapter 5 details the results of the analysis, including the asymmetries extracted from the data and information that can be gleaned from them, and finally Chapter 6 details the conclusions of the analysis and the future

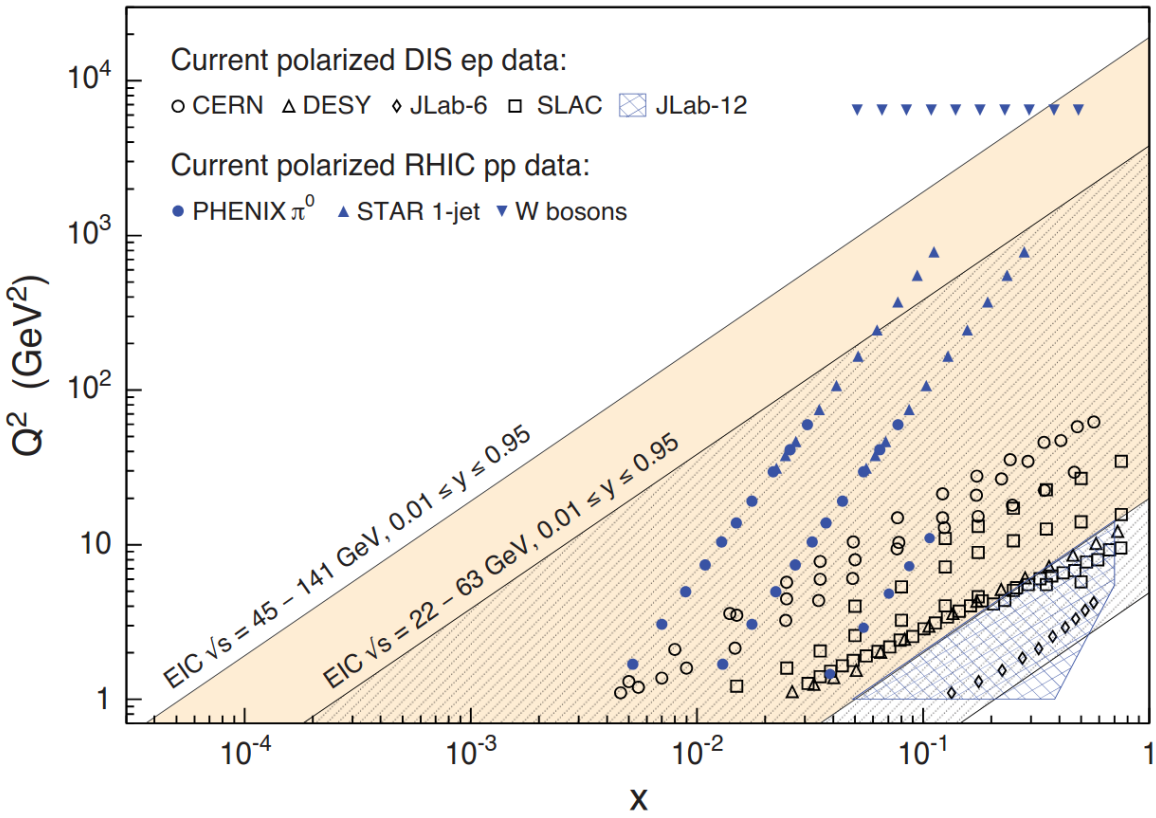


Figure 1.4: Phase space coverage of lepton-scattering and some proton-proton scattering experiments over recent years, with the upcoming Electron-Ion Collider (EIC) shown at two separate centre of mass energies [10].

outlook of these studies, discussing future data-taking runs at Jefferson Lab, alongside some preliminary studies of TCS feasibility at the future Electron-Ion Collider. Appendix A.1 shows a study done on monitoring the beam polarisation fluctuations throughout the data-taking run, and Appendix A.2 shows a binning study on the final extracted observables.

Chapter 2

Physics Motivation and Theoretical Background

2.1 Visualising structure

Fraunhofer diffraction patterns are observed when visible light diffracts through an aperture, producing an image. Taking the Fourier Transform of the aperture gives the diffraction pattern, and vice versa. This technique involves using light in the visible spectrum, that is, with a wavelength between 380-750 nm, to image objects of $\mathcal{O}(10^0)$ mm in size. To image smaller objects, the wavelength required also must be smaller; for example, in X-ray diffraction, a 0.1 - 10 nm wavelength probe is used to image objects at an atomic scale [11]. Taking this to the subatomic level, at scales of < 1 Fermi (10^{-15} m), a probe of comparable wavelength is required. Looking to quantum field theory, virtual particles such as photons, which come from a high-energy electron beam interacting with a nucleon target via photon exchange, can be used for this purpose. These photons are known as deeply virtual, which mathematically means their four-momentum transfer is significantly large ($Q^2 >> M_N^2$, where M_N^2 is the mass of the nucleon target), and that their mass is considered off shell [12].

2.1.1 Elastic Scattering

Exploring this concept deeper involves first looking at the principles of elastic scattering as a tool for exploring nucleon structure. The methodologies and equations in this chapter are taken largely from a mixture of [13], [14] and [15].

A formalism can be developed by first assuming that a non-relativistic point-like electron with charge Z_1, scatters off a point-like proton with charge Z_2, with the assumption that both particles are spinless, as shown in Figure 2.1. This incident electron is then scattered by the Coulomb field of the proton, with a differential cross-section of the form:

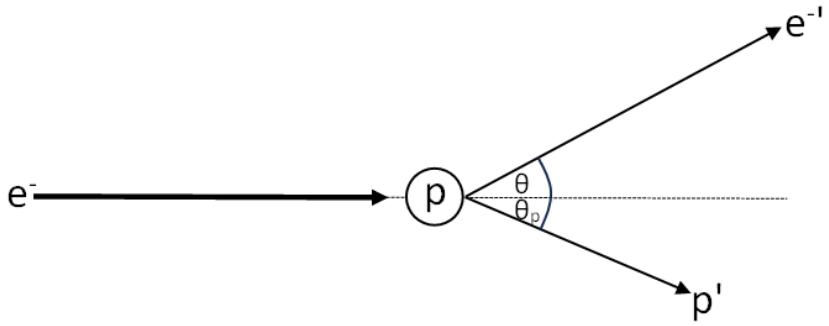


Figure 2.1: A representation of an incoming electron scattering from a proton, both particles assumed to be spinless and point-like.

.

$$\frac{d\sigma}{d\Omega}_{\text{Rutherford}} = \frac{\alpha^2}{16E_K^2 \sin^4(\frac{\theta}{2})} \tag{2.1}$$

where $\alpha = \frac{Z_1 Z_2 e^2}{4\pi\epsilon_0 \hbar c}$ = 1/137 (when Z_1=Z_2=1), θ represents the scattering angle of the electron, and E_K is its initial kinetic energy. This is referred to as the Rutherford cross-section, and comes from the reduced form of:

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2}\left(\frac{1}{m_p + E_e - E_e\cos\theta}\right)^2 \langle|\mathbf{M}|^2\rangle \tag{2.2}$$

where m_p is the mass of the proton, $\langle|\mathbf{M}|^2\rangle$ is the Lorentz invariant spin averaged matrix element, or more intuitively the interaction probability, and the reduced form is obtained by assuming the electron (energy E_e) is non-relativistic and the proton is at rest.
Extending this to account for the fact that the electron is a relativistic, spin $1/2$ particle, but maintaining the assumption of point-like structure of the proton gives the Mott cross-section;

$$\frac{d\sigma}{d\Omega}_{\text{Mott}} = \frac{\alpha^2}{4E_e^2 \sin^4(\frac{\theta}{2})}\cos^2\left(\frac{\theta}{2}\right) \tag{2.3}$$

or if the proton recoil is considered;

$$\frac{d\sigma}{d\Omega}_{\text{Mott}} = \frac{\alpha^2}{4E_e^2 \sin^4(\frac{\theta}{2})}\frac{E_{p'}}{E_e}\left(\cos^2\left(\frac{\theta}{2}\right) - \frac{q^2}{2m_p^2}\sin^2\left(\frac{\theta}{2}\right)\right) \tag{2.4}$$

where m_p is the proton mass, E_e and $E_{p'}$ are the incident electron and scattered proton energies respectively, and q^2 is the four-momentum transfer of the virtual photon exchanged in the interaction.

2.1.2 Form Factors

To further develop these concepts, an additional term must be introduced that accounts for the fact that the proton is not point-like but instead has structure:

$$\frac{d\sigma}{d\Omega} = \frac{d\sigma}{d\Omega}_{\text{Mott}} |F(Q^2)|^2 \tag{2.5}$$

where the additional factor $|F(Q^2)|^2$ (note that $Q^2 = -q^2$) is known as the proton form factor. This form factor can be shown as a representation of the key elements of proton structure, in terms of distribution of charge and magnetic moment in the Rosenbluth Formula:

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E_e^2 \sin^4(\frac{\theta}{2})} \frac{E_{p'}}{E_e} \left(\frac{G_E(Q^2)^2 + \tau G_M(Q^2)^2}{1+\tau} \cos^2\left(\frac{\theta}{2}\right) + 2\tau G_M(Q^2)^2 \sin^2\left(\frac{\theta}{2}\right) \right) \tag{2.6}$$

where $G_E(Q^2)$ represents the electric form factor, $G_M(Q^2)$ represents the magnetic form factor and $\tau = \frac{Q^2}{4m_p^2}$. These two terms are the Fourier transforms of the charge distribution and magnetic moment of the proton, respectively. Their relationship with the Dirac (F_1) and Pauli (F_2) form factors is as follows;

$$G_E = F_1 - \tau F_2$$
$$G_M = F_1 + F_2.$$

This interpretation of $G_E(Q^2)$ and $G_M(Q^2)$ as representations of charge and magnetic moment is only valid in the Breit frame, where the momentum of the scattered proton p' is exactly equal and opposite its initial momentum p, i.e. there has been no energy transfer, $p + p' = 0$.
Different values of Q^2 provide access to particular information by expressing the Rosenbluth formula in the limit where Q^2 is very small and very large. At small Q^2, $\tau << 1$, therefore the Rosenbluth formula becomes:

$$\left(\frac{G_E(Q^2)^2 + \tau G_M(Q^2)^2}{1+\tau} + 2\tau G_M(Q^2)^2 \right) \left(\frac{d\sigma}{d\Omega} \right)_{\text{Mott}} \approx G_E(Q^2)^2 \left(\frac{d\sigma}{d\Omega} \right)_{\text{Mott}} \tag{2.7}$$

and at large Q^2, $\tau >> 1$, so the Rosenbluth formula becomes:

$$\left(\frac{G_E(Q^2)^2 + \tau G_M(Q^2)^2}{1+\tau} + 2\tau G_M(Q^2)^2 \right) \left(\frac{d\sigma}{d\Omega} \right)_{\text{Mott}} \approx (1+2\tau) G_M(Q^2)^2 \left(\frac{d\sigma}{d\Omega} \right)_{\text{Mott}}. \tag{2.8}$$

This dependence intuitively relates to the dominance of the magnetic force ($1/r^3$) over the electric force ($1/r^2$) at short distances.

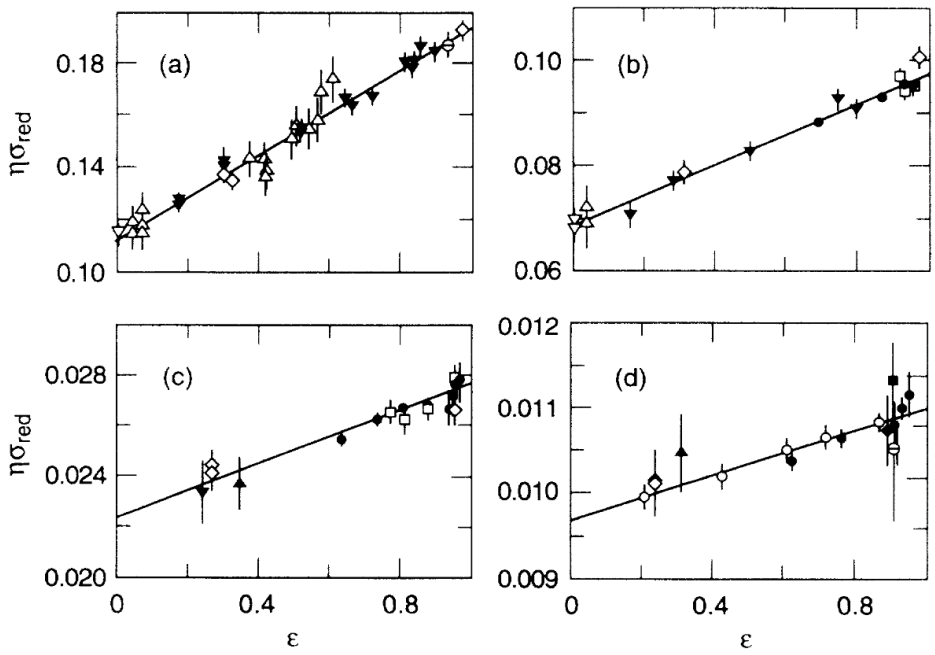


Figure 2.2: Rosenbluth seperation for (a) Q^2 =0.6, (b) Q^2 = 1.0, (c) Q^2 = 2.0 and (d) Q^2 = 3.0 [16].

Simultaneous extraction of the electric and magnetic form factors can be achieved at low Q^2 by measuring the cross-section at a fixed value of Q^2 over a range of θ. This is expected to produce a linear fit, since the Rosenbluth formula can be split as:

$$\frac{d\sigma}{d\Omega} = \frac{d\sigma}{d\Omega_{Mott}}\left(\frac{G_E(Q^2)^2 + \tau G_M(Q^2)^2}{1+\tau} + 2\tau G_M(Q^2)^2 \tan^2(\frac{\theta}{2})\right)$$

$$y = mx + C$$

$$y = \frac{d\sigma}{d\Omega}, \qquad x = \tan^2\frac{\theta}{2}$$

$$m = \frac{G_E(Q^2)^2 + \tau G_M(Q^2)^2}{1+\tau}$$

$$C = 2\tau G_M(Q^2)^2$$

which is referred to as the Rosenbluth Separation. The resulting fit gives access to each form factor, an example of which is shown in Figure 2.2, which details the results of a series of elastic scattering experiments performed at the Stanford Linear Accelerator Centre (SLAC) between $0.1 < Q^2 < 3\,\mathrm{GeV}^2$. These distributions show the normalised cross-section $\eta\sigma$ versus the longitudinal polarisation of the virtual photon $\varepsilon = [1 + 2(1+\tau)\tan^2(\frac{\theta}{2})]$ which depends only on θ when Q^2 is fixed. These slight modifications to the original formula in (2.6) occur when the formula is expressed in its Lorentz-invariant form under the Born approximation [16].

2.1.3 Deep Inelastic Scattering

As higher values of Q^2 are approached, the dominant scattering interactions are inelastic rather than elastic, meaning that the target nucleon changes in the final state. This can mean the final state is excited, or in the case of Deep Inelastic Scattering, DIS, (even higher Q^2), it breaks up and some of the energy from the initial state goes into creating new particles. This is represented pictorially in Figure 2.3. It is in these processes one must introduce x_B or Bjorken x, which is defined as $\frac{Q^2}{2p\cdot q}$ (or $\frac{Q^2}{2p_2\cdot q}$ in Fig 2.3) and describes the longitudinal momentum fraction carried by the struck quark in the Bjorken regime, i.e. at large virtuality of the incident photon (large Q^2). It was experiments observing this scattering process that first led the path towards revealing the partonic structure of nucleons [17], [18]. Because of the introduction of this second variable x_B alongside Q^2, the scattering cross section must now be described as functions of both x_B and Q^2. These are known as structure functions ($F_1(x_B, Q^2)$, $F_2(x_B, Q^2)$) and describe internal nucleon structure in terms of momentum [19]. These functions are typically accessed using inclusive measurements, where only the final state electron is detected. More detail on this will be discussed in Section 2.3.

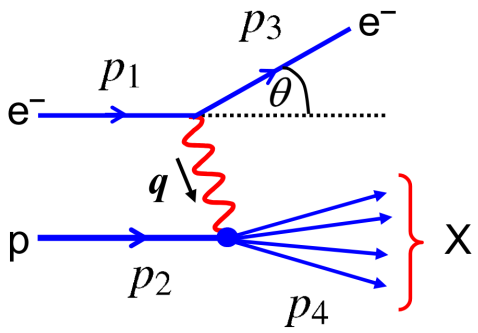


Figure 2.3: Inelastic scattering of electron off a proton showing the breakup of the proton in the final state [20].

2.2 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is a Quantum Field Theory (QFT), distinguishable from Quantum Electrodynamics (QED) in two key ways, one is the fact that in QCD, for processes that occur at relatively low values of q, the strong coupling constant, α_s, is large, and for high values, $q >$100 GeV, the strong coupling constant is small. This is known as asymptotic freedom. Typically, this means that calculations at low Q^2 in QCD are non-perturbative, meaning they cannot be solved analytically, and require frameworks such as Lattice QCD (LQCD) to solve them numerically. In the region where the strong coupling constant is small, quark-gluon and gluon-gluon interactions can be treated within a perturbation framework, for example, the interaction between a deeply virtual photon and a struck parton. The other way in which QCD and QED differ is that gluons, the mediators of the strong force, interact with themselves, which

is not the case for photons in QED. In QCD, quarks are established as carrying both colour and electric charge, and quark-gluon interactions can thus be described similarly to QED, with the substitution of the coupling constant by the strong coupling constant. However, gluons also carry colour charge, therefore there are added gluon-gluon interaction properties in the theory not present between photons in QED [11].
The QED Lagrangian is given by;

$$\mathscr{L}_{QED} = \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi + q\bar{\psi}\gamma^{\mu}\psi A_{\mu} - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.9}$$

where the first term represents the Dirac equation for an electron, the second term represents the electron current with charge $q = -e$, and the interaction with vector field A_{μ}. The third term represents the kinetic potential of the vector field A_{μ}, where $F_{\mu\nu}$ is the field strength tensor ($F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$). Finally, ψ is the quark field and m is the quark mass. From here, the QCD Lagrangian describing quark-gluon interactions can be derived, which is given by;

$$\mathscr{L}_{QCD} = \sum_{q}(\bar{\psi_i^q}(i\gamma^{\mu}\partial_{\mu} - m_q)\psi_j^q + g_s\bar{\psi}\gamma^{\mu}\psi t_{ij}^a G_{\mu}^a) - \frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu} \tag{2.10}$$

where g_s is the strong coupling constant (representing the color charge), t^a are the set of 3x3 matrices that make up the SU(3) group, also known as the Gell-Mann matrices. The sum over q represents the sum over the quarks, while the i and j terms represent color indices, 1-3, and the a term represents the color index, 1-8. G_{μ}^a and $G_{\mu\nu}^a$ are the gauge field and field strength tensors analogous to A_{μ} and $F_{\mu\nu}$ in QED; however, the QCD field strength tensor has additional terms ($G_{\mu\nu}^a = \partial_{\mu}G_{\nu}^a - \partial_{\nu}G_{\mu}^a - g_s f_{abc}G_{\mu}^b G_{\nu}^c$). It is these extra terms that reveal the additional interaction vertices present that correspond to gluon-gluon interactions, which are not found in QED [21].

2.2.1 Running Coupling Constant

In the perturbative regime, the strong coupling constant is defined as;

$$\alpha_S = \frac{g_s^2}{4\pi} \tag{2.11}$$

the form of which was first introduced in Section 2.1.2 as the fine structure constant $\alpha \approx 1/137$ for QED. The strong coupling constant is also dimensionless, as in QED, and can be understood as the strength of the strong force. The renormalisation of this QFT is done by allowing the strong coupling constant to acquire scale dependence and then performing a normalisation to a measurable value at the given scale. This procedure is how coupling constants that run at the energy scale of the virtual photon q^2 are defined. This can first be done for the QED coupling constant α by defining the renormalisation scale μ such that;

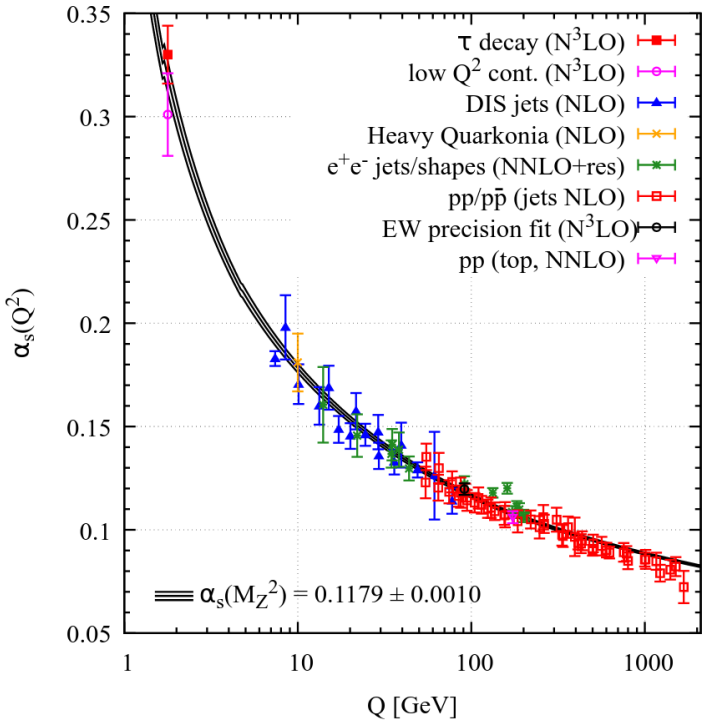


Figure 2.4: The current status of measurements of $\alpha_S(Q^2)$ with experiment type and the degree of perturbation theory used in the extraction [24]. NLO refers to Next-to-leading order (in the strong coupling constant), NNLO refers to next-to-next-to leading order, and so forth.

$$\alpha(q^2) = \frac{\alpha(\mu^2)}{1 - \alpha(\mu^2)\frac{1}{3\pi} ln(\frac{q^2}{\mu^2})} \tag{2.12}$$

This can then be modified for QCD and the strong coupling constant;

$$\alpha_S(q^2) \approx \frac{\alpha_s(\mu^2)}{1 + (\frac{11N_c - 2N_f}{12\pi})\alpha_S(\mu^2) ln(\frac{q^2}{\mu^2})} \tag{2.13}$$

where terms N_f and N_c have been introduced to represent the number of active quark flavours, and the number of colors (3) respectively.

Experimental measurements of $\alpha_S(q^2)$ at a series of values of Q are shown in Figure 2.4, quoted at a scale governed by the mass of the Z boson, M_Z = 91 GeV. At low energy scales, it is clear that the strong coupling constant has a value closer to order $\alpha_S \approx 1$, which defines the non-perturbative regime, but at higher energy scales it becomes smaller, so treatment of these QCD processes can be done within a perturbative framework. This scale dependence is termed the running of the coupling constant [22, 23].

2.2.2 QCD and Nucleon Properties

Utilising the QCD Energy Momentum Tensor (EMT) can give access to the mechanical properties of the nucleon. The EMT takes the form;

$$\hat{T}^{\mu\nu} = \sum_q \bar{\psi}_q \gamma^\mu i \overleftrightarrow{D}^\nu \psi_q + \frac{1}{4} g^{\mu\nu} F^2 - F^{c\mu\lambda} F^{c\nu} \tag{2.14}$$

where ψ_i is the quark field, $\overleftrightarrow{D}^\mu$ is the covariant derivitive, and as seen prior γ^ν are the gamma matricies, and $F^{\mu\alpha}$ represent the gluon field strength tensors. The first term represents the quark part of the EMT while the second and third terms correspond to the gluon parts. Taking the matrix elements of this for quarks and gluons ($\hat{T}_a^{\mu\nu}$ where a can be replaced with a q or G to represent quarks or gluons) allows a decomposition of the mass, spin and mechanical properties of the proton, as follows [25] [26];

$$\langle p', \vec{s}' | \hat{T}_a^{\mu\nu} | p, \vec{s} \rangle = \bar{u}' \left[A_a(t) \frac{\gamma_{\{\mu} P_{\nu\}}}{2} + B_a(t) \frac{i P_{\{\mu} \sigma_{\nu\}}{}^\rho \Delta_\rho}{4m} + D_a(t) \frac{\Delta^\mu \Delta^\nu - g^{\mu\nu} \Delta^2}{4m} + m \bar{c}_a(t) g^{\mu\nu} \right] u e^{i\Delta x} \tag{2.15}$$

where $\gamma_{\{\mu} P_{\nu\}}$ is a notation to represent $\gamma_\mu P_\nu + P_\nu \gamma_\mu$. The term Δ represents the four momentum transfer ($t = \Delta^2, \Delta = (p' - p)$) and $P = (p' + p)/2$, the average four momentum. Each term A, B, D and $\bar{c}$ are the Gravitational Form Factors (GFFs). A encodes information on the hadron's energy and momentum, B describes the gravitomagnetic angular momentum inside the hadron, different notations for the EMT decomposition show this term as J, which contains information on nucleon spin decomposition. D is referred to as the D-term, and relates to the pressure and shear forces inside of the nucleon. The final term, $\bar{c}$ relates to the trace anomaly. [27] [25] [26] .

2.3 Parton Distribution Functions and Generalised Parton Distributions

The distribution of partons within a nucleon in terms of position and momentum can be described through the use of structure functions. An example of these are Parton Distribution Functions (PDFs) which, at leading order, are tools that provide a one-dimensional, probabilistic interpretation of the longitudinal momentum fraction at which partons are found inside a hadronic structure. The data for such distributions is obtained through Deep Inelastic Scattering (DIS) and other hard-scale scattering experiments. Two distinct regions are revealed when plotting PDFs: the valence quark region in which constituent quarks (those that sum to give the net properties of the nucleon, as per the quark model) are found, and the quark-gluon sea where dynamical interactions between quark-antiquark pairs and gluons occur [28].

Looking at Figure 2.5, it is clear that when x is high (close to 1), the valence quark region

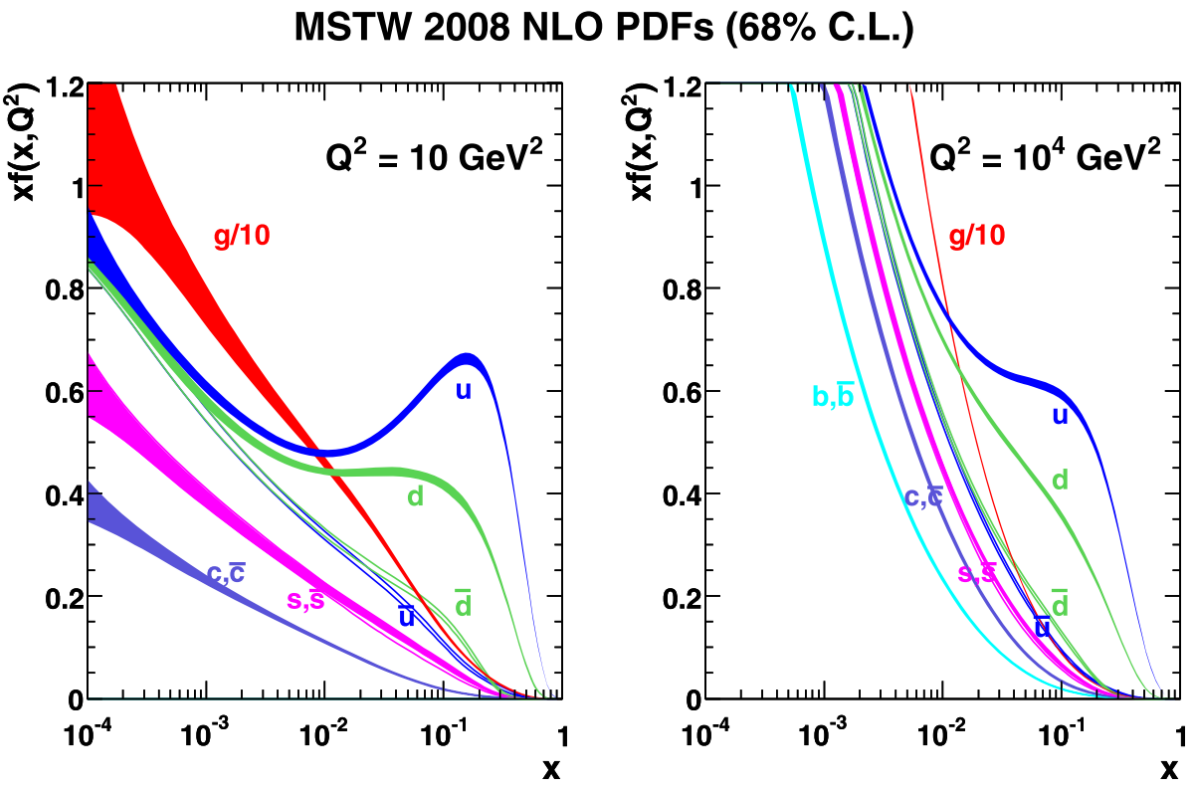


Figure 2.5: -to-leading order PDFs at a scale of $Q^2 = 10\,\mathrm{GeV}^2$ and $Q^2 = 10^4\,\mathrm{GeV}^2$ with a 68%, or 1σ confidence level shown by each coloured band [29].

is dominant, i.e that the up (u) and down (d) quarks associated with the quark model of the internal structure of the proton (uud) is revealed. When x is low (close to 0), there is a dominance of gluons, and strange (s) and charm (c) quarks and their associated antiquarks become more pronounced. This is indicative of the dominance of the quark-gluon sea region at this scale.

To further describe internal nucleon structure, Generalised Parton Distribution functions (GPDs) are utilised, which are accessible via exclusive (all final state particles are known/detected) processes that occur in a deeply virtual (high Q^2) limit [30]. At this limit, scattering amplitudes of processes can be factorised into what is referred to as a 'hard' part, representing processes describable in QED and perturbative QCD, and a 'soft' part where non-perturbative QCD describes the scattering amplitude. This 'soft' part is the description of the internal structure of the nucleon, and how it changes when the scattering process occurs. GPDs are functions of x, t and ξ, where ξ represents the skewness parameter, and represents half of the change in longitudinal momentum fraction of the struck/active quark in the process. Taking the Fourier Transform of GPDs with respect to momentum transfer t, i.e the difference squared in initial and final hadron four momenta (see Figure 2.8c), where $t = (q - q')^2 = (N' - N)^2$, gives access to the impact parameter b which describes the transverse position of the partons. Using this along with longitudinal momentum information will give access to a fuller tomographic image of the internal

structure of the nucleon [31,32]. GPDs can also be related to a wealth of structural information, as they provide indirect access, via sum rules, to GFFs. This allows access to information on pressure, mass and spin decompositions [27] and hence a broader understanding of them is invaluable to developing an understanding of the nucleon [33].

Generalised Parton Distributions are defined as matrix elements of non-local (describing an initial and final quark being created and annihilated at different points in spacetime) and non-diagonal (the initial and final nucleon momentum are different) quark and gluon operators in light-cone coordinates, and are given by the following expressions [34], [11], [35];

$$F_q = \frac{1}{2}\int \frac{dz^-}{2\pi} e^{ixP^+z^-} \langle p_2|\bar{\psi}^q(-\frac{z^-}{2})\gamma^+\psi^q(\frac{z^-}{2})|p_1\rangle = \frac{1}{2P^+}\left[\mathbf{H}^q(x,\xi,t)\bar{u}(p_2)\gamma^+u(p_1) + \mathbf{E}^q(x,\xi,t)\bar{u}(p_2)\frac{i\sigma^{+\alpha}\Delta_\alpha}{2m}u(p_1)\right] \quad (2.16)$$

$$\tilde{F}_q = \frac{1}{2}\int \frac{dz^-}{2\pi} e^{ixP^+z^-} \langle p_2|\bar{\psi}^q(-\frac{z^-}{2})\gamma^+\gamma_5\psi^q(\frac{z^-}{2})|p_1\rangle = \frac{1}{2P^+}\left[\tilde{\mathbf{H}}^q(x,\xi,t)\bar{u}(p_2)\gamma^+\gamma_5 u(p_1) + \tilde{\mathbf{E}}^q(x,\xi,t)\bar{u}(p_2)\frac{\gamma_5\Delta^+}{2m}u(p_1)\right] \quad (2.17)$$

where P^+ is the light cone momentum of the nucleon, γ^+ represents the light cone gamma matrix, $\gamma^\pm = \frac{1}{\sqrt{2}}\left[\left(\begin{smallmatrix}0 & I\\ I & 0\end{smallmatrix}\right) \pm \left(\begin{smallmatrix}0 & \sigma^3\\ -\sigma^3 & 0\end{smallmatrix}\right)\right]$, with I representing the identity matrix and σ^3 representing the third Pauli matrix, $z^- = \frac{1}{\sqrt{2}}(z_0 - z_3)$, Δ_α is the momentum transfer and m is the mass of the nucleon. More directly, F_q and $\tilde{F}_q$ are derived from first principles using Wigner distributions, which in a QFT inherently describe the particle system using spatial and momentum information.

Figure 2.6 shows the relationship between each type of distribution discussed in this section and outlines the general process of deriving GPDs from Wigner distributions via the Fourier transform and then integration over light cone momentum k. Also detailed is the accessibility of the impact parameter using GPDs or Wigner distributions. Figure 2.7 shows a representation of the changes between information acquired from form factors, parton densities and GPDs from the perspective of the internal structure of the nucleon. In the latter image the impact parameter b is referred to as r.

The four quark helicity conserving GPDs are $\mathbf{H}^q$, $\mathbf{E}^q$, $\tilde{\mathbf{H}}^q$ and $\tilde{\mathbf{E}}^q$. QCD also describes four further GPDs, $\mathbf{H_T}^q$ $\mathbf{E_T}^q$, $\tilde{\mathbf{H_T}}^q$ and $\tilde{\mathbf{E_T}}^q$ which represent a flipped helicity of the active quark, accessible through measurements on transversely polarised targets. Relating back to section 2.1.2, the integrals over x of GPDs $\mathbf{H}^q$ and $\mathbf{E}^q$ are representations of the Dirac and Pauli form

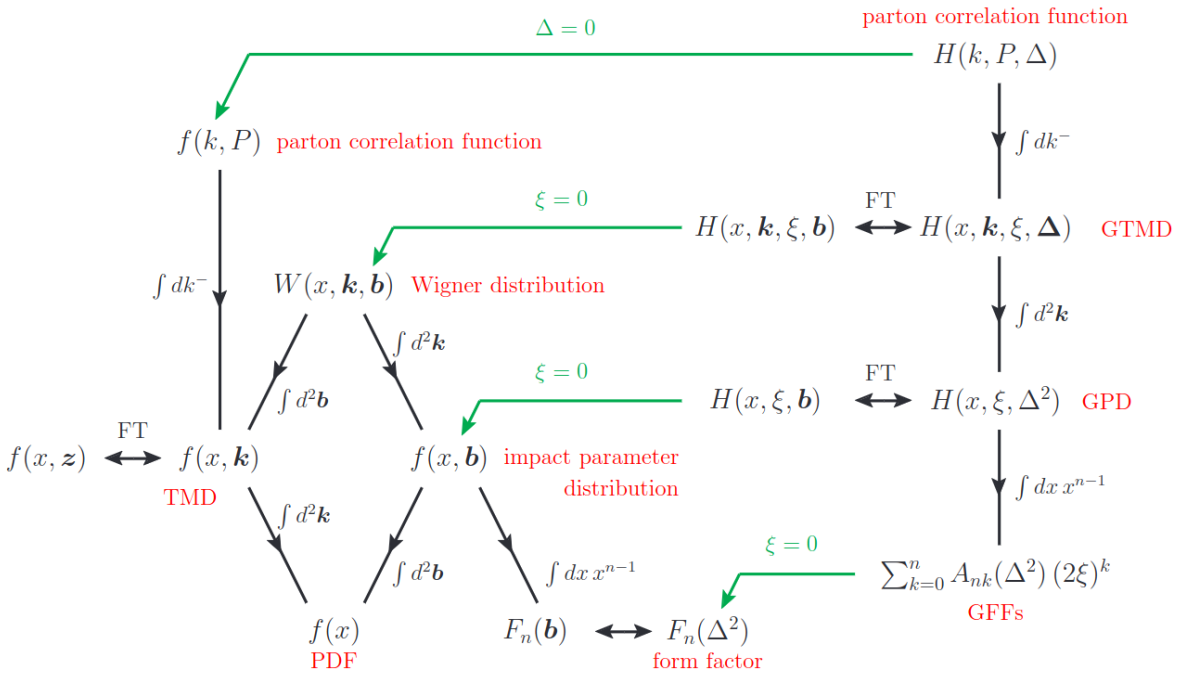


Figure 2.6: Relationships between each of the distribution functions, TMDs (Transverse Momentum dependent Distributions), PDFs (Parton Distribution Functions), GFFs (Generalised Form Factors), GPDs (Generalised Parton Distributions), GTMD (Generalised Transverse Momentum dependent Distributions). Note in this figure momentum transfer t is represented as Δ^2. Image taken from [33].

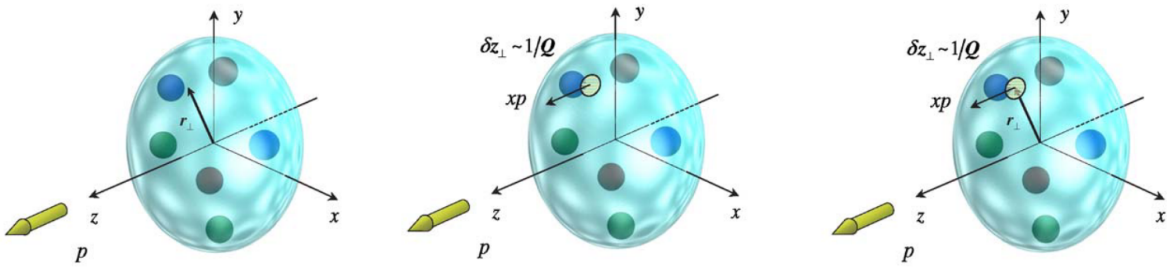


Figure 2.7: From left to right the information accessible via form factors, parton densities and GPDs from a nucleon structure perspective. Image taken from [34]

factors and the integrals over x of GPDs $\tilde{\mathbf{H}}^q$ and $\tilde{\mathbf{E}}^q$ are representations of the axial and pseudo-scalar form factors. The $\tilde{X}$ represents GPDs that are dependent on quark helicities, which are therefore polarised, and the absence of the ˜ represents GPDs that average over quark helicity and are unpolarised. GPDs of the form $\mathbf{H}^q$ and $\tilde{\mathbf{H}}^q$ do not encode a change in nucleon spin, but $\mathbf{E}^q$ and $\tilde{\mathbf{E}}^q$ are associated with a nucleon spin flip. When accounting for frame independence the helicity conserving GPD expressions become:

$$F_q = \frac{1}{2}\int \frac{d\lambda}{2\pi} e^{ixPz} \langle p_2|\bar{\psi}^q(-\frac{z}{2})\not{n}_- \psi^q(\frac{z}{2})|p_1\rangle$$
$$= \frac{1}{2Pn_-}\left[\mathbf{H}^q(x,\xi,t)\bar{u}(p_2)\not{n}_- u(p_1) + \mathbf{E}^q(x,\xi,t)\bar{u}(p_2)\frac{i\sigma^{\alpha\beta}n_{-\alpha}\Delta_\beta}{2m}u(p_1)\right] \quad (2.18)$$

where n_- may be any light-like vector [35, 36].
In the limit of $(\xi,t) \to 0$ the GPDs can be shown as directly related to quark PDFs:

$$\mathbf{H}^q(x,0,0) = q(x)\Theta(x) - \bar{q}(-x)\Theta(-x) \quad (2.19)$$

where $\Theta(x)$ is the Heaviside step function, and $q(x)$ and $\bar{q}(x)$ are the quark and anti-quark PDFs respectively [11, 36].

2.3.1 GPD Models

GPD models are used to constrain GPDs in a required limit or parameterise them in a way that is useful for the analysis of experimental data. A classification of one type of model is known as 'Double Distribution', formulated by Radyushkin and Muller [37] [38] and models of this type are used to parameterise the (x,ξ) dependence of GPDs. Parameterisations are achieved using an ansatz where terms α and β are introduced, such that $x = \beta + \alpha\xi$. This serves to separate out the absolute value of longitudinal momentum, as opposed to accessing the value as an average and also constrains GPDs to a polynomiality condition, in essence requiring that a given GPD moment must be a polynomial in ξ, or skewness.

The VGG model, named after its creators Vanderhaeghen, Guichon and Guidal [39], is a model that is used to parameterise each of the four GPDs ($\mathbf{H}^q$, $\tilde{\mathbf{H}}^q$, $\mathbf{E}^q$, $\tilde{\mathbf{E}}^q$) based on particular conditions, for example, $\mathbf{H}^q$ requires a particular formalism to model its t dependence to allow extrapolation to very small x, which is experimentally inaccessible. The constraint, known as Regge Theory, is based on relating the squared centre of mass energy of the system, s, t and the spin of constituents in that system.

The Goloskokov Kroll (GK) model, formulated by Sergey Goloskokov and Peter Kroll [40]

parameterises the GPD $\mathbf{E}^q$ in Deeply Virtual Meson Production (DVMP) from its relationship to $\mathbf{H}^q$, which is experimentally more well known. The t dependence in the GK model is also parameterised using Regge trajectories, however, it maintains its dependence on x, unlike the VGG model. This dependence is used to model both sea and valence quark GPDs at very low x and high Q^2 [36].

2.4 Accessing GPDs experimentally

2.4.1 Compton Form Factors

GPD's are functions of variables x, ξ and t, where as a reminder x is the longitudinal momentum fraction of the struck quark, ξ is the skewness parameter, and t is the four-momentum transfer squared. Experimentally, t and ξ are readily accessible, however, x is not directly measurable as it is integrated over. What is measured experimentally is instead a GPD at a given point ($x = \pm\xi, \xi, t$), or $\int_{-1}^{1} dx \frac{H(x,\xi,t)}{x \pm \xi}$. Each of these can be experimentally accessed through different processes [36]. Since there are 4 helicity-conserving GPDs, there must be 8 such quantities. These are known as Compton Form Factors (CFFs). CFFs are invariant quantities that play an analogous role to form factors for electromagnetism. The four CFFs H_1, E_1, $\tilde{H}_1$, $\tilde{E}_1$, can be written in terms of GPDs as follows [32]:

$$H_1(\xi,\eta,t) = \sum_q e_q^2 \int_{-1}^{1} dx \left(\frac{\mathbf{H}^q(x,\eta,t)}{\xi - x - i\varepsilon} - \frac{\mathbf{H}^q(x,\eta,t)}{\xi + x - i\varepsilon}\right) \tag{2.20}$$

$$E_1(\xi,\eta,t) = \sum_q e_q^2 \int_{-1}^{1} dx \left(\frac{\mathbf{E}^q(x,\eta,t)}{\xi - x - i\varepsilon} - \frac{\mathbf{E}^q(x,\eta,t)}{\xi + x - i\varepsilon}\right) \tag{2.21}$$

$$\tilde{H}_1(\xi,\eta,t) = \sum_q e_q^2 \int_{-1}^{1} dx \left(\frac{\tilde{\mathbf{H}}^q(x,\eta,t)}{\xi - x - i\varepsilon} - \frac{\tilde{\mathbf{H}}^q(x,\eta,t)}{\xi + x - i\varepsilon}\right) \tag{2.22}$$

$$\tilde{E}_1(\xi,\eta,t) = \sum_q e_q^2 \int_{-1}^{1} dx \left(\frac{\tilde{\mathbf{E}}^q(x,\eta,t)}{\xi - x - i\varepsilon} - \frac{\tilde{\mathbf{E}}^q(x,\eta,t)}{\xi + x - i\varepsilon}\right). \tag{2.23}$$

Compton Form Factors have a clear real and imaginary part as revealed by the $\xi + x - i\varepsilon$ in the denominator of each CFF, meaning there are 8 functions accessible experimentally. At x =+/-ξ (where a positive value relates to a quark and negative an anti-quark) the imaginary part of the CFF, expressed in terms of GPDs is extracted. It is clear that at all other values of x the real part of the CFFs are sensitive to GPDs.

2.4.2 Deeply Virtual Compton Scattering and Timelike Compton Scattering

Deeply Virtual Compton Scattering (DVCS) is a simple, key, process in the experimental determination of GPDs. It involves scattering of a deeply virtual photon off a single parton within the nucleon, re-emitting some of this energy in the form of a real photon in the final state [30]. The pure DVCS cross-section as written in [32] is:

$$\frac{d\sigma}{dQ^2 dt dy d\phi} \approx \frac{\alpha^3{}_{em}}{8\pi s_{ep}^2}\frac{1}{Q^2}\frac{1+(1-y)^2}{y^3}\sum_{\lambda,\lambda'}|M^{\lambda'+,\lambda+}|^2 \tag{2.24}$$

where: M is the matrix element of the process; λ is the incoming photon helicity; λ' is the outgoing photon helicity; y represents the inelasticity; and s_{ep} is the squared centre of mass energy of the electron-proton system.

Time-like Compton Scattering is a hard exclusive process, commonly referred to as the 'inverse' of the DVCS process. The TCS formalism that has the initial quasi-real photon (γ) produced by an electron scattering process is as follows;

$$ep \rightarrow l^+ l^- p^{'}(e^{'}) \tag{2.25}$$

where p and $p^{'}$ are the initial and final state four momenta of the nucleon, (in this case, a proton) and e and $e^{'}$ are the initial and final state four momenta of the electron. The scattered proton produces a virtual photon (γ^*) in the final state which decays into a lepton pair (l^+l^-). See Figure 2.8 for a pictorial representation of DVCS and TCS kinematics and associated variables.

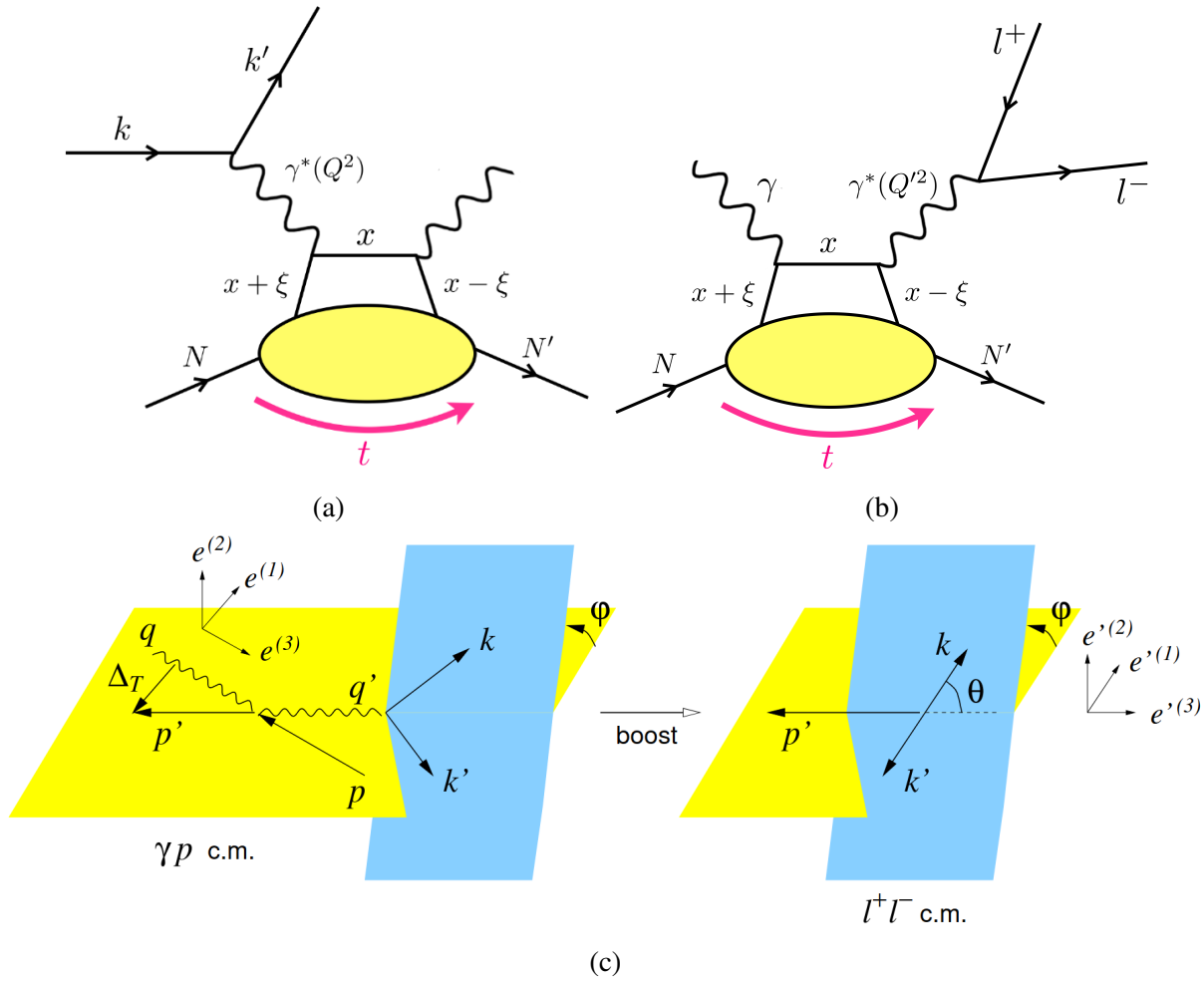


Figure 2.8:
(2.8a and 2.8b) Handbag diagrams of DVCS and TCS respectively. The incoming real photon is represented as γ, the virtual photon as γ^* and the target nucleon is represented as having initial and final momentum N and N', respectively. The produced lepton pair is represented as l^+ and l^-. Taken from [3].
(2.8c) Representation of TCS kinematics in the hadronic plane (yellow) and leptonic plane (blue) (separated by angle ϕ). The initial four momenta of the real photon and the nucleon are represented by convention as p, q and the final state four momenta are represented as p', q'. The angle between the k lepton and the scattering axis of the nucleon is represented as θ. Taken from [32].

The final state in TCS is defined by the scattered proton and the produced lepton pair, from which information about the scattered electron can be deduced. The scattered electron is often not directly detected experimentally, due to the quasi-real nature of the initial photon, meaning the scattered electron is deflected at very small angles, therefore reconstruction is a simpler method than direct detection.
The pure TCS cross-section as written in [32] is:

$$\frac{d\sigma}{dQ'^2 dt d(cos\theta) d\phi} \approx \frac{\alpha^3{}_{em}}{8\pi s^2} \frac{1}{Q'^2} \frac{1+cos^2\theta}{4} \sum_{\lambda,\lambda'} |M^{\lambda',-\lambda-}|^2 \tag{2.26}$$

where

$$\frac{1}{2}\sum_{\lambda,\lambda'}|M^{\lambda'-,\lambda-}|^2 = (1-\eta^2)(|H_1|^2+|\tilde{H}_1|^2) - 2\eta^2 Re(H_1^* E_1 + \tilde{H}_1^* \tilde{E}_1) \\ -(\eta^2 + t/4m_p^2)|E_1|^2 - \eta^2\frac{t}{4m_p^2}|\tilde{E}_1|^2. \tag{2.27}$$

TCS and DVCS can be related kinematically as follows;

$$\begin{aligned} \textbf{DVCS} &\rightarrow \textbf{TCS}, \\ Q^2 &\rightarrow Q'^2, \\ x_B = \frac{Q^2}{2p\cdot q} &\rightarrow \tau = \frac{Q'^2}{((p+q)^2 - m_p^2)}, \\ \xi &\rightarrow \xi, \\ t &\rightarrow t, \\ \phi &\rightarrow \phi \end{aligned}$$

where the variables Q'^2 and τ are introduced for TCS. The variable τ is the equivalent of the Bjorken variable x_B in DVCS, and is the ratio of the squared four-momentum transfer of the virtual photon to the squared centre of mass energy ($s = (p+q)^2$) of the system. The variable Q'^2 is the squared four-momentum transfer of the virtual photon in TCS, accessible using information from the detected final state dilepton pair.

The TCS cross-section is dominated by a similar process, known as the Bethe-Heitler (BH) process, in which an incoming photon (γ) produces a lepton pair (l^+l^-), and one of these leptons goes on to exchange a photon with the target, giving the same final state as in TCS [41, 42]. A representation of the BH processes that dominate TCS is shown in Figure 2.9, and its cross-section, in the limit of $Q'^2 \gg t$ and small target masses takes the form;

$$\frac{d\sigma}{dQ'^2 dt d(cos\theta) d\phi} \approx \frac{\alpha^3{}_{em}}{-2\pi s^2 t}\frac{1+cos^2\theta}{sin^2\theta} \times \left[\left(F_1^2 - \frac{t}{4M^2}F_2^2\right)\frac{2}{\tau^2}\frac{\Delta_T^2}{-t} + (F_1+F_2)^2\right] \tag{2.28}$$

The BH dominance can be understood by comparing the phenomenology of TCS and DVCS, as the DVCS cross-section also has a BH process which contributes to the total cross-section. The term y in DVCS is known as the inelasticity and is defined as $y = \frac{(k-k')\cdot p}{k\cdot p}$. Its analogous term in the TCS cross-section is θ, the polar angle of the produced electron in the centre of mass frame

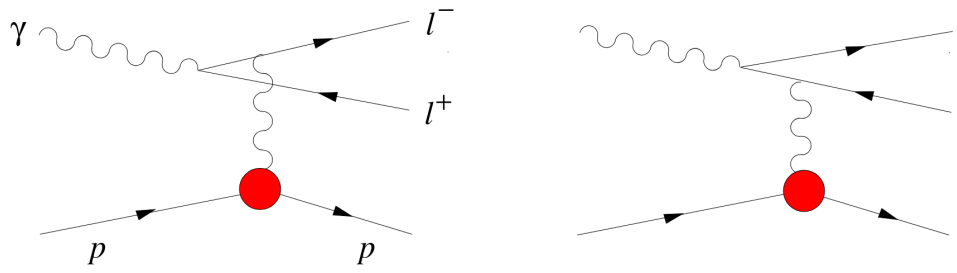


Figure 2.9: Feynman diagrams showing the Bethe Heitler process, the left diagram shows the photon exchange between the proton and lepton l^- and on the right the exchange between the proton and lepton l^+. Taken from [32].

of the lepton pair. The two terms relate as follows;

$$\frac{1-cos(\theta)}{2} \approx \frac{1}{y} \tag{2.29}$$

So at small y an enhancement of the DVCS cross-section is seen with respect to the BH background, as it includes a $1/y$ term (see Equation 2.24), however, the equivalent TCS term is fixed to lie between -1 and +1 due to the $cos(\theta)$ term (see Equation 2.26), and so the TCS cross-section is always dominated by the BH process [35], [32].

Measurements of DVCS observables allow access to the imaginary part of CFFs. The real part is important for accurate constraints on the GPD model and can be accessed via DVCS double spin observables and cross sections. Measurement of interference between TCS and Bethe Heitler allows access to four Compton Form Factors, $\mathrm{Re}\,H_1$, $\mathrm{Im}\,H_1$, $\mathrm{Re}\,\tilde{H}_1$, $\mathrm{Im}\,\tilde{H}_1$ the former two give access to GPD $\mathbf{H}^q$ and the latter two to GPD $\tilde{\mathbf{H}}^q$, see Equations (2.20-2.23). Notably, TCS can also access the real part of the CFFs, and thus can be used to place constraints on GPDs [32,43]. The interference term is accessible via asymmetries due to the fact that the lepton pair in TCS is produced in a C (charge parity) odd state and in BH they are in a C even state, the asymmetry observable has a sign change between the two lepton momenta and thus the interference between the two processes is projected out [35].

The first measurement of TCS was published in 2021 [42], one of the key observables extracted was a measurement of Beam Spin Asymmetry or BSA on an unpolarised hydrogen target, shown in Figure 2.10. The BSA is calculated as follows:
The BSA for an unpolarised target is calculated as follows:

$$A_{\odot U} = \frac{N^+ - N^-}{Pb \times (N^+ + N^-)} \tag{2.30}$$

where N^i represents the number of counts in each orientation of beam (i$=\pm$) and P_b represents the beam polarisation. In the case where the target is polarised, the BSA becomes;

$$A_{\odot U} = \frac{P_t^-(N^{++} - N^{-+}) + P_t^+(N^{+-} - N^{--})}{Pb \times (P_t^-(N^{++} + N^{-+}) + P_t^+(N^{+-} + N^{--}))} \tag{2.31}$$

where N^{ij} represents the number of counts in each orientation of beam (i) and target polarisation (j) and $P_t^{\pm}$ is the target polarisation. This difference is due to the fact that the polarisation of the target must be effectively 'cancelled out' in the calculation in order to accurately obtain an asymmetry based purely on beam helicity. The notation $A_{\odot U}$ refers to a circularly polarised beam $\odot$ and an (U)npolarised target. It is sometimes depicted as $A\phi$ denoting an asymmetry binned in the kinematic variable ϕ

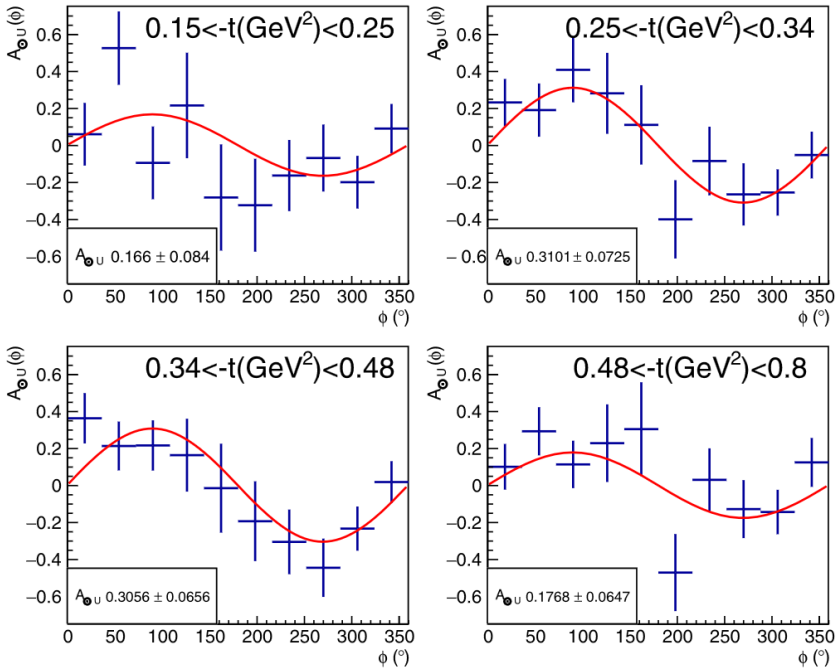


Figure 2.10: BSA extracted by Pierre Chatagnon using data on an unpolarised Hydrogen target, expressed in 4 bins of t. Taken from [42].

This measurement can give access to GPD $\mathbf{H^q}$, however, measurement on a longitudinally polarised target allows extraction of two further observables, a Target Spin Asymmetry (TSA) and a Double Spin Asymmetry (DSA). This thesis will focus on an extraction of the BSA, for complementarity with previously published results, and novelly a TSA, which is an observable relating the polarisation of the beam and target in each orientation, as shown;

$$A_{UL} = \frac{N^{++} + N^{-+} - N^{+-} - N^{--}}{Df \times (P_t^-(N^{++} + N^{-+}) + P_t^+(N^{+-} + N^{--}))} \tag{2.32}$$

where DF is the dilution factor, discussed further in Section 4.1.2. The notation A_{UL} represents an (U)npolarised beam (i.e. not taking into account the BSA), and a (L) ongitudinally polarised target. This is also sometimes referred to as A_{Uz} representing a target polarisation in the z-plane, which is the same plane as the beam and therefore a longitudinal polarisation, whereas a transversely polarised target would be polarised in the x-y plane.

These polarised observables are particularly interesting due to their sensitivities to particular GPDs. TSA extractions, for example, can give access to GPDs $\mathbf{H}^q$ and $\tilde{\mathbf{H}}^q$. The GPD $\tilde{\mathbf{H}}^q$ is less well known than the $\mathbf{H}^q$ GPD, though DVCS data exists for both (for example in [44], [45], [46] and [47]). Experimental access to these quantities are invaluable for studies of GPD universality, i.e, whether the theoretical models for GPDs are independent of the scattering process used to determine them.

Figures 2.11 and 2.12 show the theoretical curves for the BSA and TSA respectively, taken at a low t and central invariant mass in the TCS phase space. The BSA curve is taken from [48] and the TSA curve is taken from [49], where the GPD parameterisations used are the VGG model, modelling the (x, ξ) dependence using a Radyushkin double distribution ansatz, and the t dependence modelled on a Reggeized ansatz. The BSA curve details contributions as calculated at Leading Order (LO) and Next-to-Leading Order (NLO), where the order refers to that of the strong coupling constant. The TSA curve details contributions from different GPDs, showing how for the TSA, the experimentally less well-known GPD $\tilde{\mathbf{H}}^q$ is readily accessible.

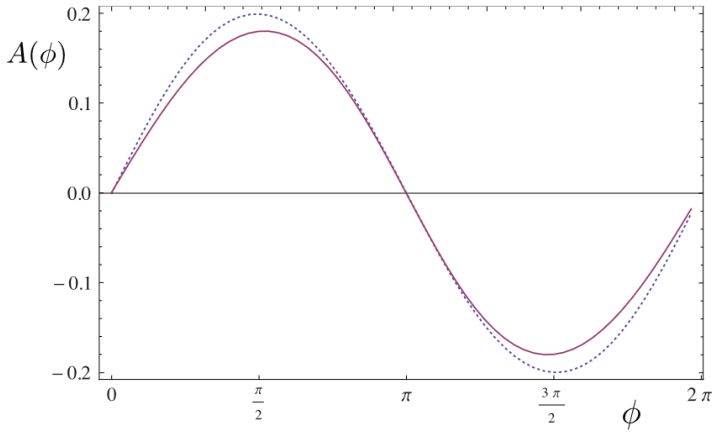


Figure 2.11: BSA theory curve showing LO in solid and NLO in dashed. Taken from [48].

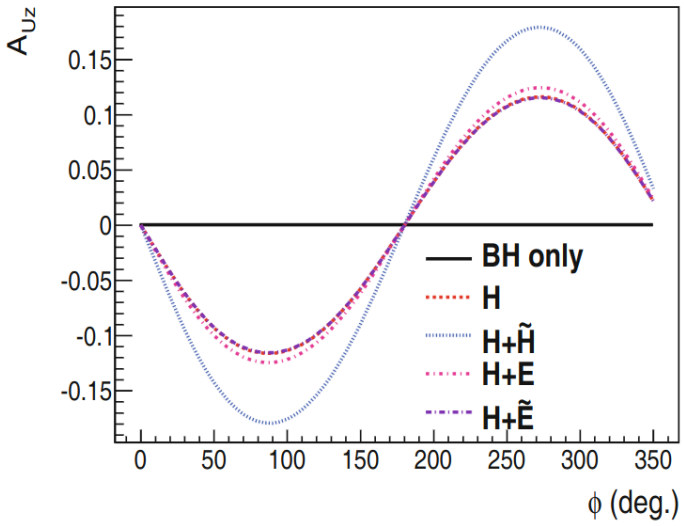


Figure 2.12: TSA theory curve calculated with different GPD contributions. Taken from [49].

Chapter 3

Experimental Setup

The process of acquiring suitable data for analysis involves a series of steps, starting with the production of the electron beam, which is directed to the target material, where it interacts, producing particles that are detected. The detector systems lead to a read-out procedure before the resulting data is processed for analysis. This chapter will detail the experimental setup, including the CEBAF accelerator, the CLAS12 detector, and the polarised target (which was a specific addition to this data taking run), before moving on to discuss the readout and data-processing procedures.

3.1 CEBAF and the Beamline

The Thomas Jefferson National Accelerator Facility (JLab) is a fixed-target facility in Newport News, Virginia. At the core of JLab is the Continuous Electron Beam Accelerator Facility (CEBAF), which provides an electron beam to four main experimental halls, A, B, C, and D, see Figure 3.1. Electrons begin at the injector, in a three-beam bunch structure at a frequency of 500 MHz. The beam is then passed through two **LIN**ear **AC**cellerators (LINACs), each providing $\approx$1.1 GeV to passing electrons and operating at a frequency of 1500 MHz, encapsulating the three beam bunches. Electrons make five passes of the full 'racetrack', giving a total beam energy of up to 12.1 GeV, see Figure 3.2. The full 12 GeV electron beam is delivered only to Hall D, as it is situated a further half-pass of the accelerator away from the other three Halls, to which a maximum of 11 GeV can be provided after the five accelerator passes. Simultaneous running of all four halls is achieved using Radio Frequency (RF) separators, positioned at the end of the South Linac. The first of these separators operates at 750 MHz and splits each of the three beams in the clump structure into two, 250 MHz halves. The first half of each beam is pumped toward Halls A-C and passed through a 500 MHz RF separator, allowing these three beam halves to be to be separated from each other, one for each hall. The second half of each beam is directed back to the accelerator, where it is recirculated through the North Linac, for half a pass, bringing the beam energy up to 12.1 GeV, and directed to Hall D [50].

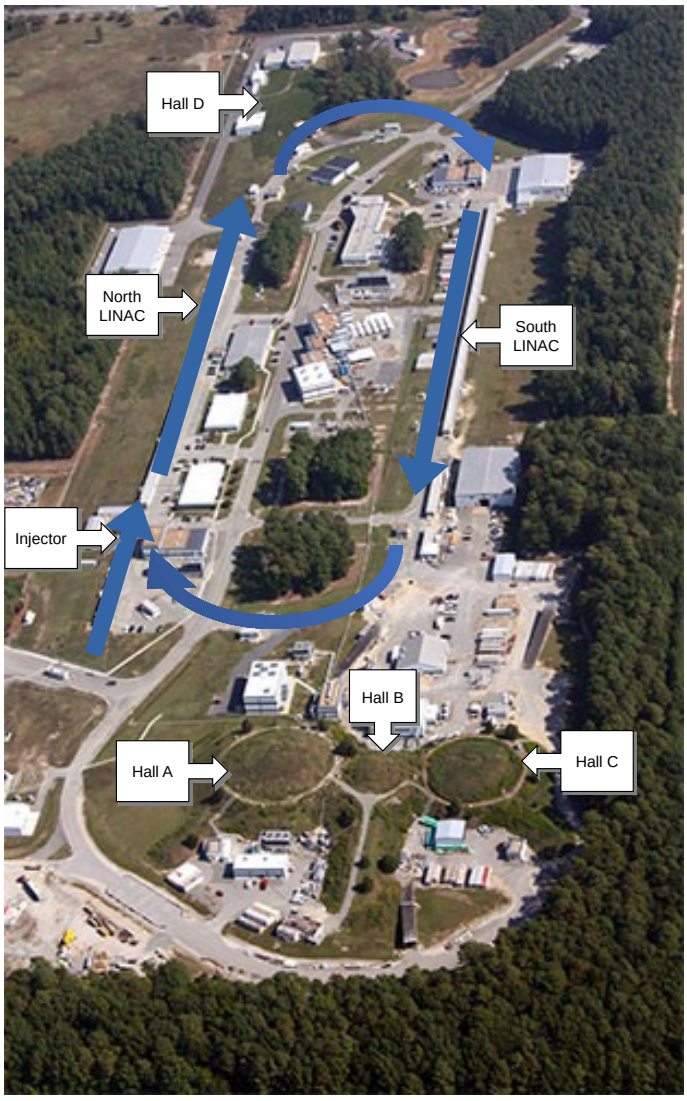


Figure 3.1: An aerial view of the Thomas Jefferson National Accelerator Facility (JLab), with blue arrows showing the path of accelerated electrons through CEBAF and each Experimental Hall (A-D) labelled. Modified from [51].

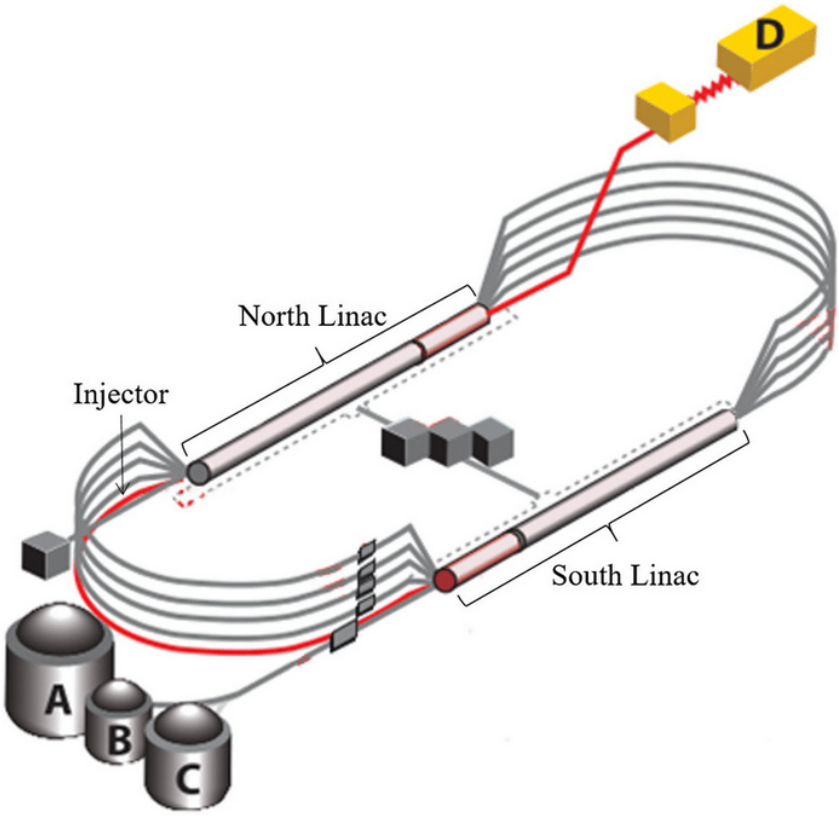


Figure 3.2: A schematic of CEBAF showing the North and South LINACs, the injector, and the five arcs through which the beam circulates before arrival at the experimental halls.

3.1.1 Beamline

Multiple systems work together to ensure that the beam delivered to each experimental hall is physics-ready. A schematic of these subsystems in Hall B is shown in Figure 3.3.

Beam Position and Tuning

Nanoamp Beam Position Monitors (nA-BPMs) 2C21, 2C24 and 2H01 measure the beam position in the x, y and z plane with an accuracy of 50 μm. A feedback system, known as the Orbit Lock, uses these values to ensure the beam position on the target is stable. Halo counters, situated just downstream of the target, monitor beam-induced background in the hall and form a part of the Fast Shut Down (FSD) procedure, to protect sensitive detectors in the hall from excessive backgrounds. During running, the halo counters are closely monitored and will show heightened rates if the beam is scraping the edges of the beam pipe. Tungsten wires called 'harps' are moved into the beam at a 45° angle, where information from the beam hitting the wire is combined with Halo Counter information to form a description of the beam profile. These 'harp scans' are conducted at the beam tuning stage to establish reference x and y positions for the beam, and then also periodically throughout the run to determine beam quality [53].

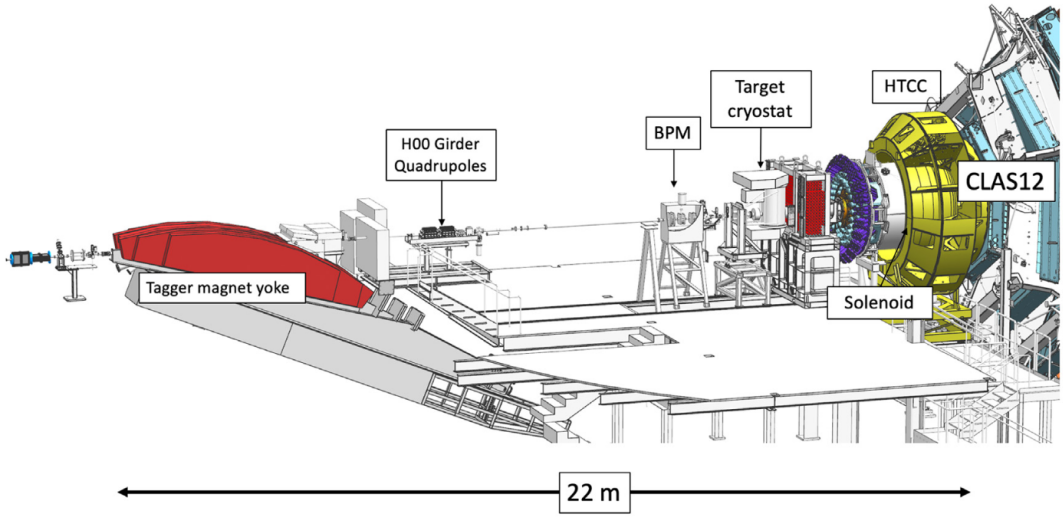


Figure 3.3: Schematic of the CLAS12 Beamline, showing the tagger magnet, the halo counters, beamline position monitor (BPM) and the CLAS12 detector [52].

Beam Current

The beam current is measured using a combination of information from nA-BPMs, the Synchrotron Light Monitor (SLM) and the Faraday Cup (FCup). The SLM is constructed out of a mirror and a prism and is used to split the synchrotron light generated by the incoming electron beam into two images, one is used to provide a beam image view, and the other is directed to a photo-multiplier tube (PMT) for measuring beam current, specifically the beam charge asymmetry [54] which is an asymmetry calculated on the difference between positive and negative accumulated charge Q for each helicity state ($\pm$), i.e. $A_{BCA} = \frac{Q^+ + Q^-}{Q^+ + Q^-}$. The FCup provides the most accurate beam current information instantaneously. To protect the FCup from high beam power, it is positioned behind the beam-dump, through which some current leaks out. The leakage is measured by the FCup and its value is a scaled value of the true beam current [53].

Beam Polarisation

The electron beam is circularly polarised by the CEBAF photocathode gun which uses lasers to create the three beam structure outlined previously [50]. The polarisation of the beam is monitored using the Møller Polarimeter, which works according to the principles of electron-electron scattering ($e + e \rightarrow e + e$). Møller runs [55] are performed at least once per configuration change (when magnet polarity or target type is changed, for example). These polarisation values are calculated by measuring the beam helicity gated coincidence rates of scattered Møller electrons on scintillation detectors [53]. For the dataset this thesis analyses, Møller runs that were taken had an average recorded polarisation value of 82.7$\pm$ 1.5 %, where the error given is the absolute error, see Figure 3.4. Note that this figure takes the absolute value of the beam polarisation, runs taken with the half-wave plate in have a negative beam polarisation, and runs with the half-wave

plate out have a positive beam polarisation. Two runs are marked in red, which represent runs taken during a spin dance procedure, which is a method that uses the inherent spin of electrons to discern how well the polarimeter is measuring their polarisation at various angles (in this case, angles of 0, 19.24, 28 and 64.5 degrees) where these angles are a measure of how much rotation is necessary to compensate for spin precession effects from the accelerator. Outliers marked in blue represent Møller runs that were taken to assess the detriment of beam bleed-through from a neighbouring experimental hall, Hall-C. A more detailed study of the beam polarisation can be found in Appendix A.1.

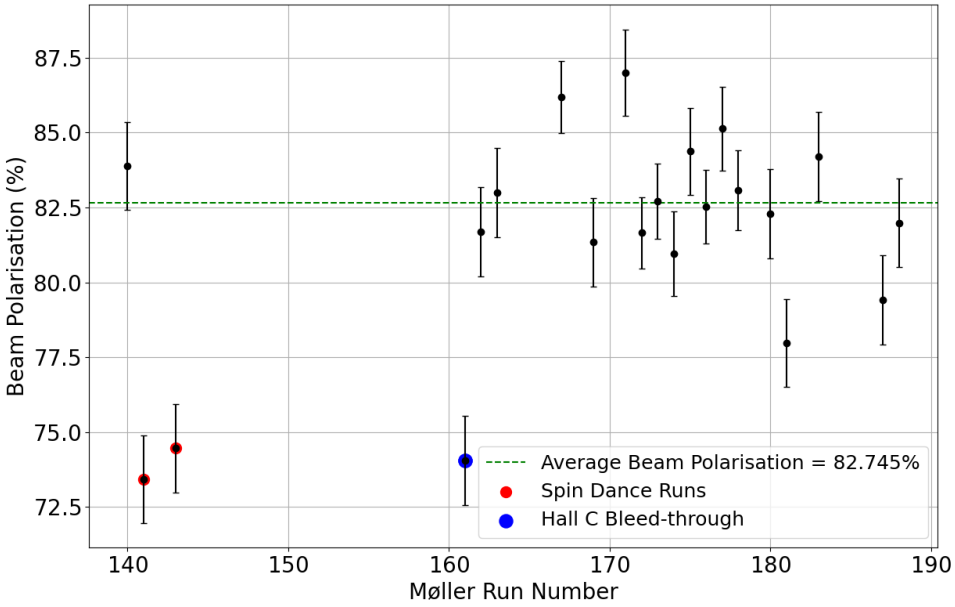


Figure 3.4: The absolute value of beam polarisation for RGC as taken by Møller runs across the full run period. The green line represents the mean polarisation value after the removal of bad runs, and the error bars at each point represent the absolute error.

3.2 The CLAS12 Detector

Hall B houses CLAS12 (the CEBAF Large Acceptance Spectrometer at 12 GeV) [52]. CLAS12 was an upgrade to the CLAS spectrometer, which was built for the JLab 6 GeV era. Some of the key features of CLAS12 are its large solid angle and azimuthal (ϕ) angular coverage, a polar angle (θ) coverage between $35^\circ - 125^\circ$ provided by its central solenoid magnet and detector, and the forward polar angle range up to 35° provided by its forward superconducting torus magnet and detector systems. For a full view of the CLAS12 spectrometer see Figure 3.5.

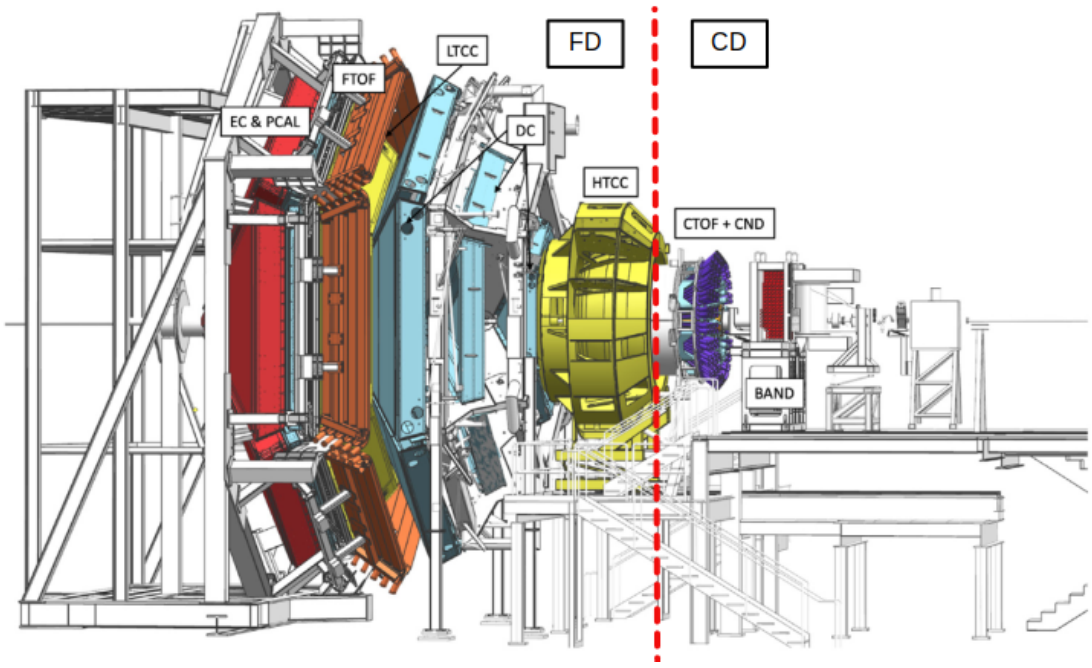


Figure 3.5: A schematic of the CLAS12 spectrometer showing the split between Forward and Central detector regions, with each subsystem coloured for distinction. The subsystems of the Central Detector are surrounded by the solenoid, coloured here in grey between the HTCC and the CTOF+CND [52]. The beam direction in this image runs from right to left.

3.2.1 Forward Detector

The CLAS12 Forward Detector (FD) has six sectors, split into multiple subsystems. The superconducting torus magnet [56] provides each sector of the FD with the same perpendicular (with respect to the beam direction) magnetic field, with the aim of deflecting charged particles to establish track momentum based on curvature. Each of the six torus magnet sectors hosts three independent, 36-layer Drift Chambers (DCs) [57] used to determine the trajectory of charged particles. Particle identification for particles at high energies is achieved in the FD via the use of Time Of Flight (TOF) [58], Cherenkov counters and Electromagnetic Calorimeters (ECALs) [59].

Drift Chambers

The combination of drift chambers and the forward torus are collectively known as the Forward Tracking system. This system is used for the identification of charged particles with forward polar angle, between $5^\circ - 40^\circ$, and with momentum greater than 200 MeV/c^2. This lower bound is placed due to the presence of the magnetic field, which will cause lower momentum particles to be swept out of the tracking system. The physical structure of the drift chambers is detailed in Figure 3.6 where we see three distinct regions, R1 positioned just before the torus field, R2 inside the torus field, and R3 just after the torus field. Each region has six sectors, and each sector has two 'superlayers', known as SLs, which are a grouping of 6 layers of wires each [57].

The drift chambers, paired with the forward micromegas vertex tracker (see Section 3.2.1) are used to discern particle momentum, calculated using track curvature in the torus field [58].

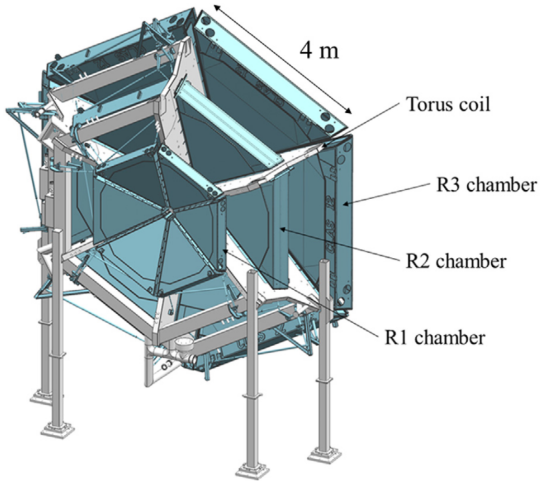


Figure 3.6: Schematic of the CLAS12 forward tracking system showing the relative positions of the three Drift Chamber regions (R1-3) [57].

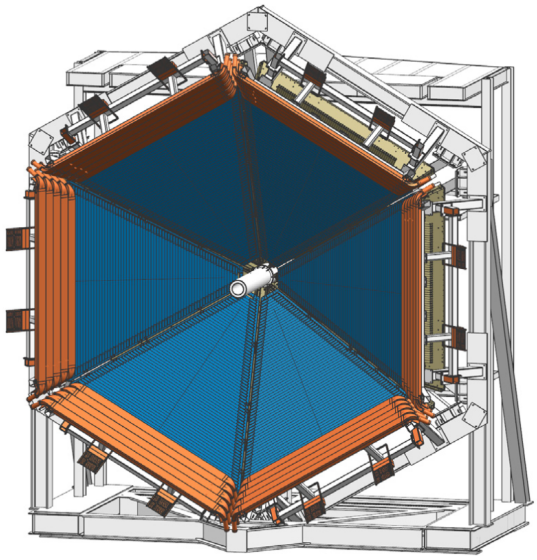


Figure 3.7: The CLAS12 Forward Time of Flight system. The panel 1b counters are in blue, and the panel 2 counters are shown in orange, the panel 1a counters are positioned just downstream of the perspective of this image.

Cherenkov Counters

The High Threshold Cherenkov Counter (HTCC) [60] is positioned in front of the drift chambers, upstream of the torus, and is used to detect electrons at forward polar angles, between $5^\circ - 35^\circ$. The HTCC is comprised of 48 mirrors, which direct Cherenkov photons produced via particle interactions with carbon dioxide in the detector volume, towards a series of photomultiplier tubes (PMTs). Out of the six torus sectors, four are equipped with Low Threshold gas Cherenkov Counters (LTCCs) [61], built similarly to the HTCC. These detectors are positioned between the DC and FTOF and are primarily used for π/K (pion/kaon) separation, achieved through the fact that pions are much lighter than kaons and thus leave a signal in the LTCC. The other two sectors are equipped with Ring Imaging Cherenkov (RICH) [62] detectors to further improve the π/K separation. These detectors comprise a series of focusing mirrors and PMTs much like the LTCC and HTCC, but they also include aerogel radiators and spherical mirrors for both direct and reflected (for larger angles) Cherenkov light detection, depending on the polar angle of the incident particle.

Forward Time of Flight

Each of the six torus sectors is also equipped with a three-layered series of scintillation counters, known as the Forward Time of Flight system (FTOF) [58]. The FTOF is used for the Particle Identification (PID) of charged particles, for example, π, p and K with momenta up to 5 GeV. The layers of counters are denoted panel 1a, panel 1b (both covering the θ region $5^\circ - 35^\circ$) and panel 2 (covering the θ region $35^\circ - 45^\circ$). Hit information from panels 1a and 1b are taken together to discern the velocity information for the charged particles, and, since $\beta = v/c$, it is from this information along with particle momentum that particle mass ($m = \frac{p}{\beta\gamma}$, where γ is the Lorentz factor), and subsequently its identity can be determined.

Electromagnetic Calorimeters

The CLAS12 Forward Electromagnetic Calorimeter [59] consists of six calorimeters, one for each sector of the FD. Each calorimeter houses a combination of the CLAS Electromagnetic Calorimeter (EC), and a pre-shower calorimeter (PCAL) added in the CLAS12 configuration to extend the radiation length of the detector. The CLAS12 ECAL uses scintillating light to detect electromagnetic showers produced by electrons up to 12 GeV, as well as photons and neutrons, to allow for analysis of the hit position and energy deposition of both charged and neutral particles at high momentum. Shower detection is done by pinpointing the hit position and timing information of clusters produced when a particle interacts with the calorimeters. Accurate reconstructions of this information allow separation between neutral particles such as neutrons and photons, based on their respective velocities. The detector is layered longitudinally, with incident particles seeing the PCAL first, then the inner EC (ECIN) and the outer EC (ECOUT). This aids in PID as the PCAL can be used to separate electrons from pions, and the ECIN and ECOUT can be used to separate electrons from photons due to the material radiation length and absorption of shower energy.

This is important for photoproduction processes, such as DVCS and TCS, in which one source of background is pion production, which must be distinguishable from photons/leptons in the final state. This segmented structure also allows some discrimination between pions and muons since muons will interact via absorption in the ECAL and emission of a photon, whereas pions will interact via ionization or they will produce hadronic showers. These showers produce a notable tail in the energy distribution of inner and outer EC plots.

Each sector is in a triangular configuration to maximise space with the hexagonal shape of the CLAS12 detector. Each layer has a set of scintillator bars denoted U,V and W, with specific orientations parallel to each side of the triangle sector, 120 degrees apart, shown in Figure 3.8. Lead sheets are placed between each consecutive layer to create the EM showers.

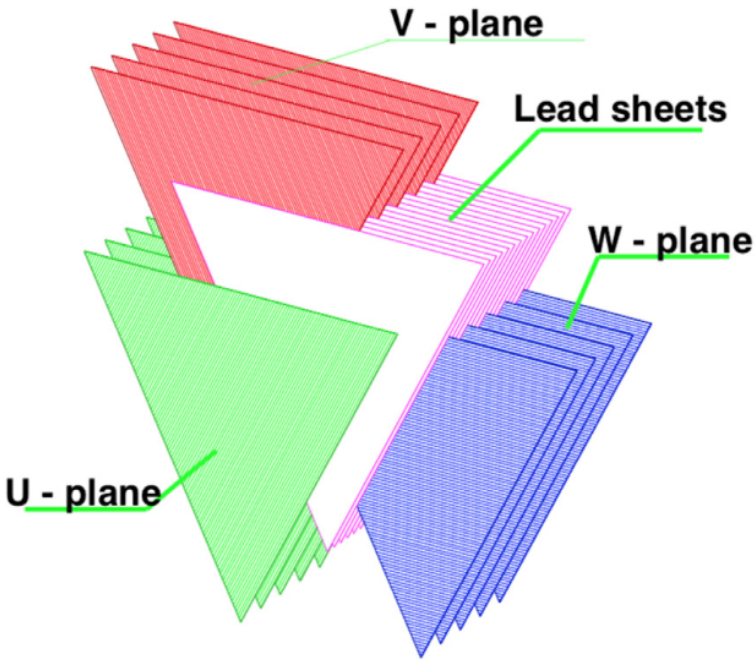


Figure 3.8: A representation of the CLAS12 Forward Electromagnetic Calorimeter showing the U,V and W oriented scintillator bars of the calorimeter. The U is shown in green, the V in red and the W in blue. The interwoven lead sheets are shown in pink/white [59].

Forward Tagger

The forward tagger is composed of an electromagnetic calorimeter, a hodoscope, and Micromegas trackers. Its main purpose is to detect electrons and photons at low polar angles (around 2° - 5°) with a low momentum transfer. This is a key element for fully exclusive processes, where the measurement of the scattered electron is required. There are three distinct variables of the scattered electron that the forward tagger can assist in detecting; the energy, in the calorimeters via the electron's electromagnetic shower and the azimuthal angle and polar angle, measured in the Micromegas trackers. The scintillation hodoscope is required for the separation of electrons and photons [63].

3.2.2 Central Detector

The solenoid magnet [56] surrounds the subsystems that make up the CLAS12 Central Detector (CD), generating a 5T magnetic field to measure charged particle momentum and charge. The magnetic field produced by the four inner coils acts as a shield to Møller electrons, directing these events to the beam dump. Møller electrons are produced when beam electrons scatter off atomic electrons from the target, which can provide significant background to exclusive processes, where the scattered electron is required for reconstruction [52]. The solenoid is self-shielding which means that it has a fifth superconducting coil, positioned outside the main set of four coils, which generates a field in the opposite direction to the others protecting sensitive

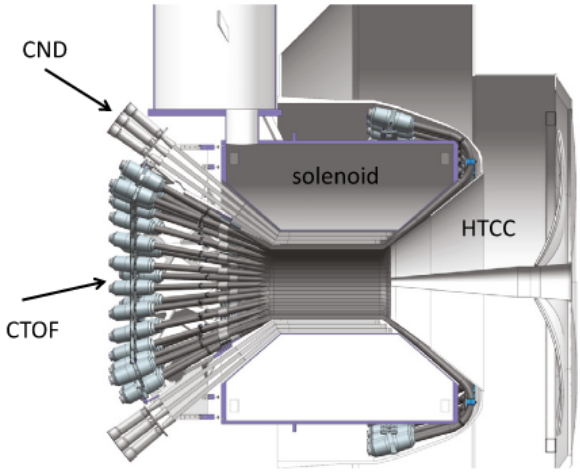


Figure 3.9: The CLAS12 Central Time of Flight system showing a side-on view of the barrel structure, comprising 48 scintillator bars, with respect to the position of the CND, the HTCC and the solenoid [64].

components near the solenoid from stray magnetic fields [56]. The central solenoid covers the polar angular range $35^\circ - 125^\circ$.

Central Time of Flight

The Central Time of Flight (CTOF) [64] system consists of a 25 cm radius, hermetic barrel of thin scintillation counters that surround the target. Each scintillator counter is read out by photomultiplier tubes (PMTs) positioned on the end of each scintillator bar. A schematic of the design and its location with respect to other central subsystems is shown in Figure 3.9. The CTOF system is used for charged particle timing measurements and separation of $\pi/K/p$ in the momentum range around 300 MeV $< P <$ 2 GeV, the lower bound is in place due to the solenoid field curving low momentum tracks away from the detector acceptance.

Central Vertex Tracker

The Central Vertex Tracker (CVT) is comprised of two subsystems, the Silicon Vertex Tracker (SVT) [65] and a six-layer Barrel Micromegas Tracker (BMT) [66] that surrounds it. The two systems are used for the tracking of low momentum particles, such as recoil baryons and large angle $\pi/K/p$, with high efficiency. The SVT provides a transverse momentum resolution on these particles of $\frac{\delta p_T}{p_T} < 5\,\%$ and the BMT provides a polar angle resolution of $\delta\theta <$ 10-20 mrad, within the angular acceptance of the CD. Tracks from the CVT are paired with hits in the CTOF for the purpose of particle identification. A schematic of the CVT design can be seen in Figure 3.10.

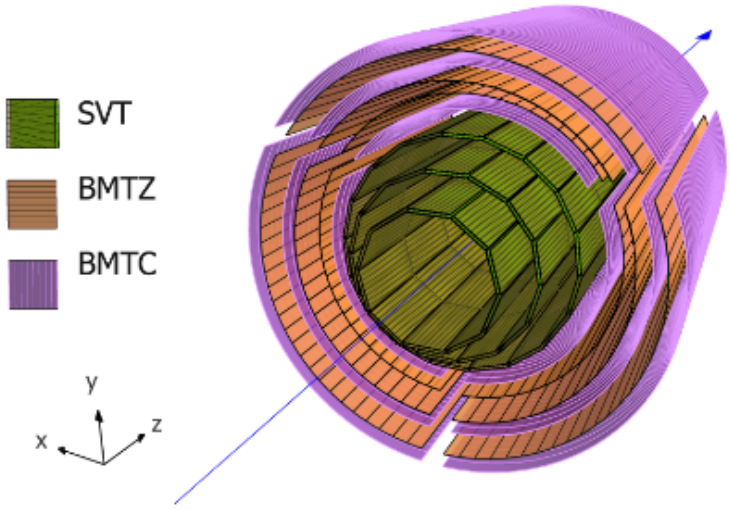


Figure 3.10: The CLAS12 Central Vertex Tracker system showing the SVT surrounded by the BMT. BMTZ refers to sectors of the BMT oriented parallel to the beamline for azimuthal particle trajectory, and BMTC refers to sectors that curve around the beamline for longitudinal measurements of particle trajectory [67].

Central Neutron Detector

The Central Neutron Detector (CND) [68] is a scintillator barrel that makes up the outermost layer within the solenoid of the central region of CLAS12. It was designed to detect 0.2 -1 GeV neutrons at laboratory polar angles from $40^\circ - 120^\circ$, and to improve neutron detection efficiency in the central region. This system is crucial for interaction channels involving electron scattering off a neutron, for example, those on a deuterium target.

3.2.3 Trigger and DAQ

Two systems, the Trigger, and the DAQ (Data Acquisition) work together to record data taken by CLAS12. The trigger fires a signal to the DAQ to record an event when it receives a signal from the CLAS12 detector subsystems, HTCC, DC, FTOF, ECAL, CTOF, CND and FT. At the first stage of event identification, there is a check for clusters produced in the ECAL that have energy above or below a defined threshold, depending on the particular experiment being conducted. The HTTC is then used to identify the nature of the clusters from the calorimeters, based on charge and number of photoelectrons (electrons produced from Cherenkov light hitting the PMTs), and hits in the 6 layers of the DC are then used to establish particle track geometry. PMTs in the FTOF, CTOF and CND trigger upon a hit over a defined energy threshold value and, if electrons are scattered at very small angles, the FT employs a cluster-finding algorithm to determine a hit. Once a hit is registered, the second stage of event ID involves checks on the time coincidence for each of the six Forward Detector segments, and then again separately

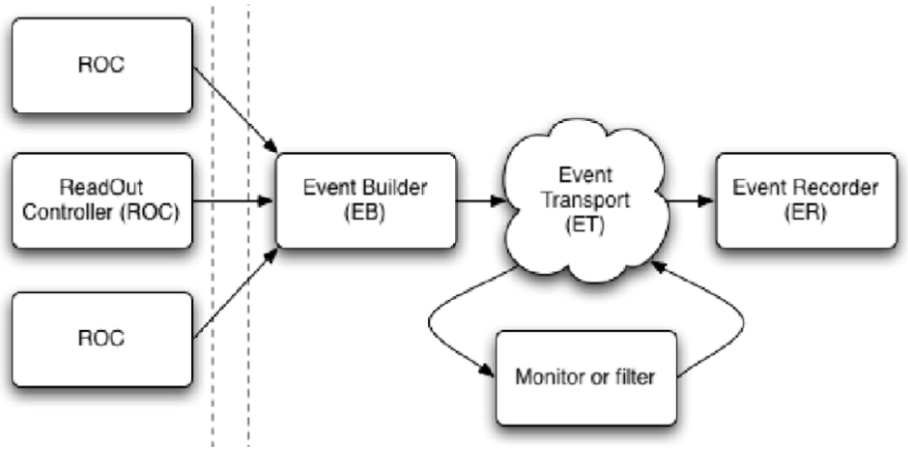


Figure 3.11: Flowchart describing the DAQ system from Read Out Controllers (ROCs) to Event Readout (ER) [70].

for each component of the Central Detector. The final stage combines all readouts into a set of trigger bits that can be read out by the DAQ system [69].

The process of Data Acquisition and readout begins at the Read Out Controllers (ROCs) which consist of various components such as JLab-designed processor boards, Linux servers, and VME (Versa Module Eurocard)/VXS (VME with serial extension) crates. These components allow for a scalable and modular design, making the DAQ system relatively easy to upgrade should the need arise. Each component has a shared 250 MHz clock, supplied by optical fibres, which is used for synchronisation. These fibres also distribute trigger, reset and busy signals. The ROCs are connected to the Event Builder (EB) which is responsible for the assembly of the data from each ROC into a single event. Events are then transferred to the ET (Event Transfer) ring, so-called as its memory storage is organised in a circular structure, allowing access by multiple data processing programs in real-time. From here, the data is then passed to the Event Recorder (ER) where data is written to disk, see Figure 3.11 for a flowchart describing this process [70].

3.3 Data Processing

3.3.1 Data Cooking Procedure

Data is stored on tapes in the form of .evio files. In this form the data is still in a raw state, so to analyse this data and perform calibrations/corrections the data needs to first be reconstructed. Data reconstruction, or ‘cooking’ requires a chef, which formed part of the work of this PhD. The author held the chef role for one year, and during that time was responsible for cooking data for subsystem experts all the way through the calibration stages of the dataset used in the thesis, up to the point where the full dataset was assessed for final complete cooking. The author was also involved in some of the early stage processing of the Fall 2022 and Winter 2022/2023 run

periods, which were, at the time of this thesis, still being calibrated.

The cooking procedure is done by first configuring a .yaml file, which includes all the geometry variations and calibration constants for each detector system, as well as magnet configurations, and the inclusion of services such as AI tracking or denoising algorithms to clean up the data signals. A configuration file is then produced, with the settings for the cooking itself, for example, how many runs, which software framework (called COATJAVA) version to use, and which CLAS12 reconstruction architecture service (called CLARA) to use. Once all the configurations have been established, a workflow script is initiated, which submits the cooking stages as a series of jobs per data-taking run on the JLab computing farm. Cooking for analysis typically has three stages per run, see Figure 3.12 for a visual representation of these steps. Firstly, there is the decoding stage, where the .evio files are converted to .hipo (High Performance Output) format [71], which is a format suitable for large datasets, with fast, high-performance compression. Next comes the reconstruction stage, where hits are reconstructed as particle tracks and sorted into data banks, suitable for analysis. Finally comes the analysis stage which takes the reconstructed files and applies a script called a 'wagon' script, which skims the data into subsets, called trains, appropriate for each particular analysis. The specific train used in this thesis is discussed in further detail in Section 4.

Depending on the cooking type there are multiple things that can occur after this stage. If the data is fully calibrated, the cooking is referred to as a 'pass1' which is where every run is fully cooked for analysis and trains are produced. Prior to this, the majority of processing tasks are calibration cooks, which are done to ensure each detector subsystem has the appropriate calibration constants uploaded to the Calibration Constants Database (CCDB). The way the latter is achieved is by producing 'timelines' which are a series of fitted histograms for each run. These histograms are published with a set of preset boundaries and checks are performed to see if the fits applied to each histogram fall between these boundaries. If this is not the case, a second iteration of calibrations will be performed, the data will be recooked based on the new constants, and then timelines will be produced on the new cook. These timelines are referred to as Quality Assurance (QA) timelines, and allow the distinction between 'good' and 'bad' runs across each period. More detail on the calibration procedure is given in the following sections.

3.3.2 Subsystem Calibrations

Calibrating each individual subsystem is a long and complex process, requiring rigorous multi-stage algorithms and careful checking of the data by software and detector experts. Calibrations can be very broadly classified into two categories, pre-beam and in-beam. Pre-beam calibrations are done before the experimental run, with the aim of preparing the detectors for data taking, often involving tests using cosmic muons to monitor detector response and establish the nec-

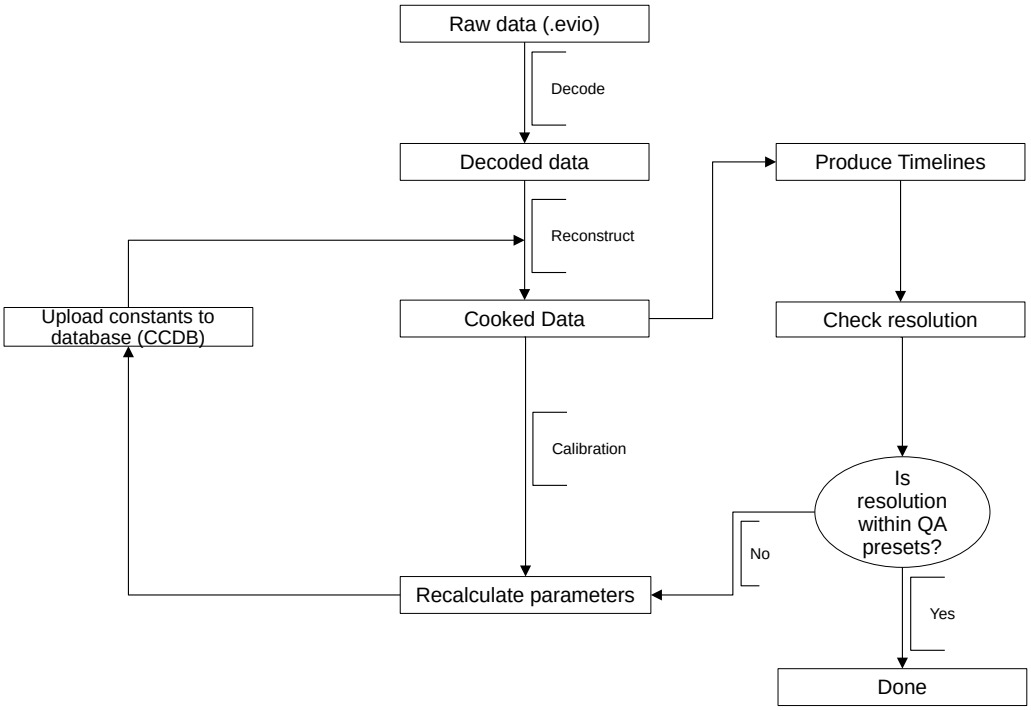


Figure 3.12: A flowchart showing the data processing procedure that takes data from raw state to calibrated analysable state. QA stands for Quality Assurance, which is a set of boundaries within which detector subsystem resolutions should fall to be considered 'good' runs, and CCDB stands for Calibration Constant Database, which is where calibration constants for each detector subsystem are uploaded.

essary voltage limits for operation. In-beam calibrations are required for preparing the dataset for analysis and aim to ensure things like track lengths and coordinates, timing and energy are accurately recorded, ready to pass to the event reconstruction software. The following sections introduce some of these calibration processes for each detector subsystem.

DC

Calibrating the drift chambers is done by taking the raw TDC (Time to Digital Conversion) values and correcting them to account for drift time of tracks and gas ions through the detector, and follows the form [57];

$$t_{drift} = t_{tdc} - t_{start} - t_0 - t_{flight} - t_{prop} - t_{walk} \tag{3.1}$$

where: t_{tdc} is the raw time; t_{start} is the time at which the trigger particle (usually the electron) left the target, and is given by the Forward Time of Flight system; t_0 is the wire time delay; t_{flight} is the time it took for the particle to travel from its interaction point to the wire; t_{prop} is the time it took for the signal to propagate through the wire; and t_{walk} is the extra timing correction required to account for the distance of the ion track to the wire. After the raw TCD time is corrected to account for the drift time, this time value is converted to a Distance of Closest Approach (DOCA). To achieve this conversion, the distance of the closest approach of a track to the wire (TRKDOCA) versus time is fitted with a 4^{th} order polynomial called the drift velocity function, whose parameters are: drift time, saturated drift velocity, distance to the point of minimum electric field strength, and the drift velocity at this point. These values are then calibrated by running the fit multiple times to obtain optimised values. Since the objective is to get a DOCA as a function of time, and the polynomial is an expression of time as a function of distance, the drift velocity function is inverted and it is in this way that the tracks are effectively reconstructed [57].

FTOF

Calibrations for the Forward Time of Flight (FTOF) detectors follow a multi-stage process. Firstly, timing offsets between left and right PMTs are determined and accounted for using hit points and times in the TDCs. Next, the ADC (Analogue to Digital Converter) recorded value is calibrated as a function of energy deposition. Prior to an experimental run, gain matching determines the required High Voltage (HV) setting for the Minimum Ionising Particle (MIP) peaks to match, such that during the calibration process the peaks are aligned and the energy loss of a particle track can be accurately determined for each counter. The third step is taking the measured ADC value as a function of hit coordinate to extract and calibrate the counter attenuation length. Next the left and right TDC times and track hit coordinate values are used to determine the effective velocity of light along each of the counters, before a time walk correction

is applied to the left and right side PMTs. This correction is determined from: the L/R offset in step one; the flight time from target to FTOF; the path length of the track; and the comparison of RF time from the accelerator to the reaction vertex of the target. After this, a correction is established to account for the timing difference between each counter which is required for hits that are recorded by multiple counters. This is done by first determining the offset between the beam bunch that produced the trigger particle and the measured start time of the event (i.e. which beam bucket was responsible for the recorded hit in the counter). Charged particle tracks that have a hit in multiple counters can then be traced back to the specific beam bucket, and any offsets in timing can be rectified. The final step is the timing calibration of the TDCs and correction of counter-hits for the difference between the actual track and the recorded hit point. These calibrations are initially performed for a single run, and closely following runs are checked for any notable shifts outside a defined allowance before a new set of runs is calibrated. The stages outlined above are completed in succession after the DC calibrations are complete, as there is a dependency on the accuracy of the track path length, and the stages are iterated and repeated due to inter-stage dependencies [58].

HTCC/LTCC

The Cherenkov counter calibrations also involve a gain-matching process, for the HTCC the position of the photoelectron peak is determined for each PMT, and the HV setting is adjusted to ensure the peaks are aligned. Timing calibrations for the HTCC are done by calculating the shift between the target interaction vertex and the channel response and applying this value as a correction to each channel. The HTCC response in each channel must also be corrected for physical factors such as dust, mirror distortion/imperfections, and PMT gain instability [60].
The LTCC gain matching process aligns the single photoelectron (SPE) peak position to an ADC value of 200, and the final reported number of photoelectrons out of the ADCs is corrected as $\frac{\text{ADC}}{\text{ADC}_{SPE}}$ [61]. This in essence ensures that the signal value for a single photoelectron depositing a signal on the anode is made as a baseline for which to gain match each ADC peak, and from this the number of photoelectrons can be extracted by calculating the ratio between the SPE ADC value of 200 and the measured, gain matched ADC value.

ECAL/PCAL

The detected energy in the inner, outer and pre-shower calorimeters (ECIN, ECOUT, PCAL) is determined from the energy measured by each PMT in a given module over each of the three orientations or views (U,V,W, see Figure 3.8), corrected by sampling fraction (energy deposition/particle momentum). The energy measured by each PMT is calibrated by determining the values of the attenuation function describing the effect of the distance between the cluster and the readout point. These values are determined by taking cosmic muon data runs before the experimental run. Timing calibrations of the calorimeters are performed using the clusters

recorded in each of the ECIN, ECOUT and PCAL, and taking the difference between measured and expected time. The expected time follows the form [59];

$$T_{\text{expected}} = t_{start} + \frac{\text{path length}}{\beta} + \frac{\text{readout distance}}{v_{eff}} \tag{3.2}$$

for each orientation, and the measured time is decomposed as follows [59];

$$T_{\text{measured}} = t_{TDC} - t_{ADC} \tag{3.3}$$

where each term in the latter has cable delay effects and component-specific corrections folded in. A χ^2 minimisation procedure is followed to minimise the difference between these two times, and, since the path length and readout distance from the cluster and start time are known, the effective velocity (v_{eff}) can be extracted [59].

FT

Pre-beam calibrations of the Forward Tagger involve testing the detector response over each layer of the detector using dedicated cosmic muon runs. Initially, gain matching is performed via an analysis of MIP signals across different high-voltage settings for the Hodoscope SiPMs (silicon photomultipliers). Energy calibration of the hodoscopes was also achieved through analysis of MIPs and comparison with simulation.

The energy calibration for the FT-Cal is achieved using analysis of elastic scattering at low beam energies and π^0 decay events ($\pi^0 \rightarrow \gamma\gamma$) for a wider spectrum of beam energies, up to 10.6 GeV. The electron scattering events with one cluster in the calorimeter were analysed to locate the single crystal in that cluster with the highest energy deposition from the electromagnetic shower, providing a reference point less prone to fluctuation for subsequent calibration. This crystal's energy, known as the 'seed' energy, was compared to the total energy of the cluster and, if found to be above a threshold value of 55% of the total energy, is plotted and compared with simulation to extract calibration constants. For the higher energy case with π^0 decay, the seed energy is compared to the known π^0 mass and reaction kinematics and plotted to extract calibration constants. These constants are checked by reconstructing the π^0 mass from the detection of the two photons and comparing them again to the known value. This process is repeated until the reconstructed mass is within a defined tolerance of the known value.

Timing calibrations are achieved using information from the FTOF on the particle start time, and comparing to a back projection of the recorded 'hit' time in the calorimeter and hodoscope to the interaction vertex [63].

CTOF

A gain matching procedure is also followed for the Central Time of flight system at the start of calibrations, however, an added consideration is put in place to handle the attenuation of light

reducing the signal along scintillator bars, as well as procedures to balance the signal between upstream and downstream PMTs. The attenuation function parameters are determined from the ratio of upstream and downstream PMTs versus counter hit position. The scintillator bars also have a timing correction factor applied associated with the fact that their edges are curved. The ADCs are calibrated using the energy deposition of MIPs much like in the FTOF. The PMTs must be calibrated for offsets in timing resulting from differing cable lengths in upstream and downstream PMTs. This allows the determination of a hit coordinate, providing a value for the effective velocity. The effective velocity is determined from TDC timing information and the distance between CTOF counters and central region trackers. Timing calibrations are performed to ensure that an accurate time of flight from the target to the CTOF counters is measured. One of the steps in this is to align each counter based on the average hit time in each counter from the TDCs, and correct for any offsets between each counter for the same event. As with the FTOF, the measured time by the CTOF must also be correlated with the RF of the accelerator, in order to ensure accurate timing measurements.

CND

The timing calibration of the CND is done by calculating and correcting for any measured timing offset between the event vertex time and the hit/path length information (L) recorded by the CND. The full equation to calculate this offset takes the form [68];

$$t_{\text{offset}} = \frac{t_L + t_R}{2} - t_0 - t_{tof} - \frac{\text{Path Length}}{2}\left(\frac{1}{v_{eff_R}} + \frac{1}{v_{eff_L}}\right) - \frac{1}{u_t} - \frac{t_{LR_{AD}}}{2} - TDC_j. \tag{3.4}$$

The terms t_{LR} and $t_{LR_{AD}}$ represent contributions from timing offsets between hits on coupled paddles in the CND, v_{eff} as usual is the effective velocity which is determined using the hit position and time differences across paddles. The 'u-turn' propagation u_t is added to account for the fact that structurally, the paddles are connected to each other by a semi-circle shaped light guide, akin to a 'u-bend', and therefore extra time is required for the signal to propagate through the paddles. Information from the FTOF is used to determine t_0 and t_{tof} is provided by the CVT. The energy calibration process for the CND involves accounting for the attenuation length of light across the scintillator bars, using ADC information, much like in the CTOF, and analysis of MIPs to extract energy conversion factors [68].

CVT

Cosmic muon data was initially taken to align and calibrate the BMT prior to beam time. During the experiment, tracking efficiency and cluster multiplicities were measured to determine the optimal high-voltage settings. Dedicated empty target, zero field runs combining data from the BMT and the SVT were used to monitor and improve vertex resolution.

During the calibration process of detectors after an experimental run, many iterations and tests need to be performed to check the quality of the data and to ensure the determined calibration constants are accurate. As part of the PhD the author performed some checks for proposed software changes to the reconstruction algorithm of the Central Vertex Tracker. The study took data from a 2018/2019 run, and compared it with the same data, reprocessed to include the updates to the Central Vertex Tracker. These updates aimed to improve the reconstruction of the scattered proton, thus increasing both the yield and resolution of the final state kinematic histograms.
The study (shown in Figures 3.13 and 3.14) looked at a dilepton and proton final state from ep scattering, and included the following cuts/definitions;

- 'FSI' - indicates Final State Inclusive, which represents only the requirement that the lepton pair is in the FD, but no other cuts applied. This is to show the effect of the CVT changes on the full dataset for a dilepton final state.

- RGA - indicates the requirement that the lepton pair is in the FD, the scattered proton is in the CD (as this is the region where the CVT is located) and some additional preliminary cuts from the 2018/2019 analysis note detailed below;

  - Sampling Fraction cut 3σ
  - Minimum energy deposition in the PCAL 60MeV
  - Cut on the intercept of Sampling fraction in the inner and outer calorimeters versus the sampling fraction in the pre-shower calorimeters of < 0.2
  - Electron z-vertex cut -13 cm $<$ z $<$ 12 cm
  - Loose fiducial cuts on the PCAL (one bar)

  Each of these cuts are discussed in more detail, pertaining specifically to the dataset in the thesis, in Section 4.1.1.

- 'Excl' represents the application of some preliminary exclusivity cuts on the 2018/2019 data, taking into account the fact that the study was performed on one single run so statistics were not sufficient for tight exclusivity.

  - Decay electron momentum > 1 GeV
  - Decay positron momentum > 1 GeV
  - Scattered electron Pt/P < 0.05

- 'Before' represents the run from the originally processed dataset and 'After' represents the run from the reprocessed dataset.

Looking to figures 3.13 and 3.14 which show the Missing Mass squared (MM^2) of the scattered electron (where $MM^2 = (e + p - e^+ - e^-)^2$), and the invariant mass of the dilepton pair, respectively, it is clear that the CVT changes improve both the resolution and yield of the reconstructed missing mass squared and invariant mass distributions. As a result, these tracking changes were implemented in the reconstruction of the dataset in this thesis.

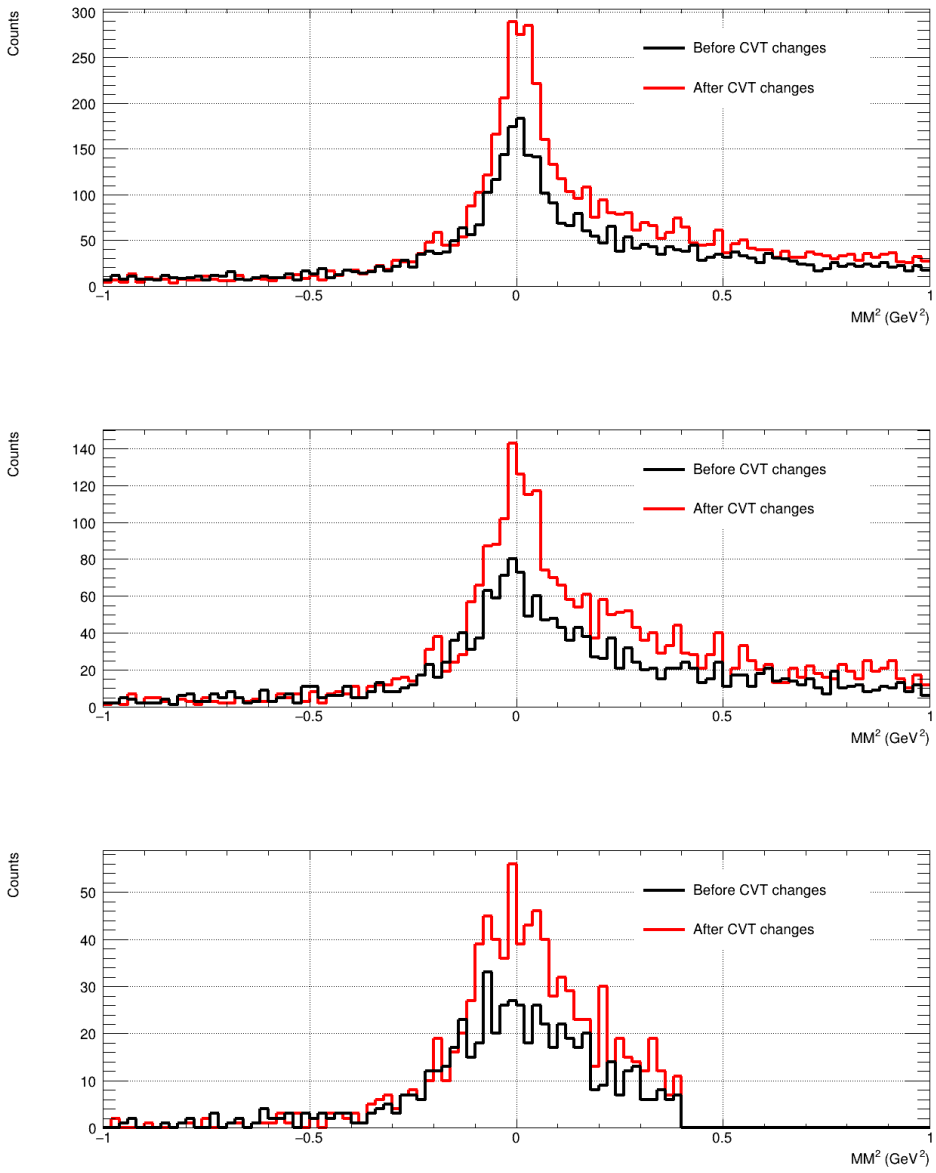


Figure 3.13: Reconstructed missing mass squared for before (blue, left) and after (red, centre) the CVT changes as well as both distributions overlaid (right)

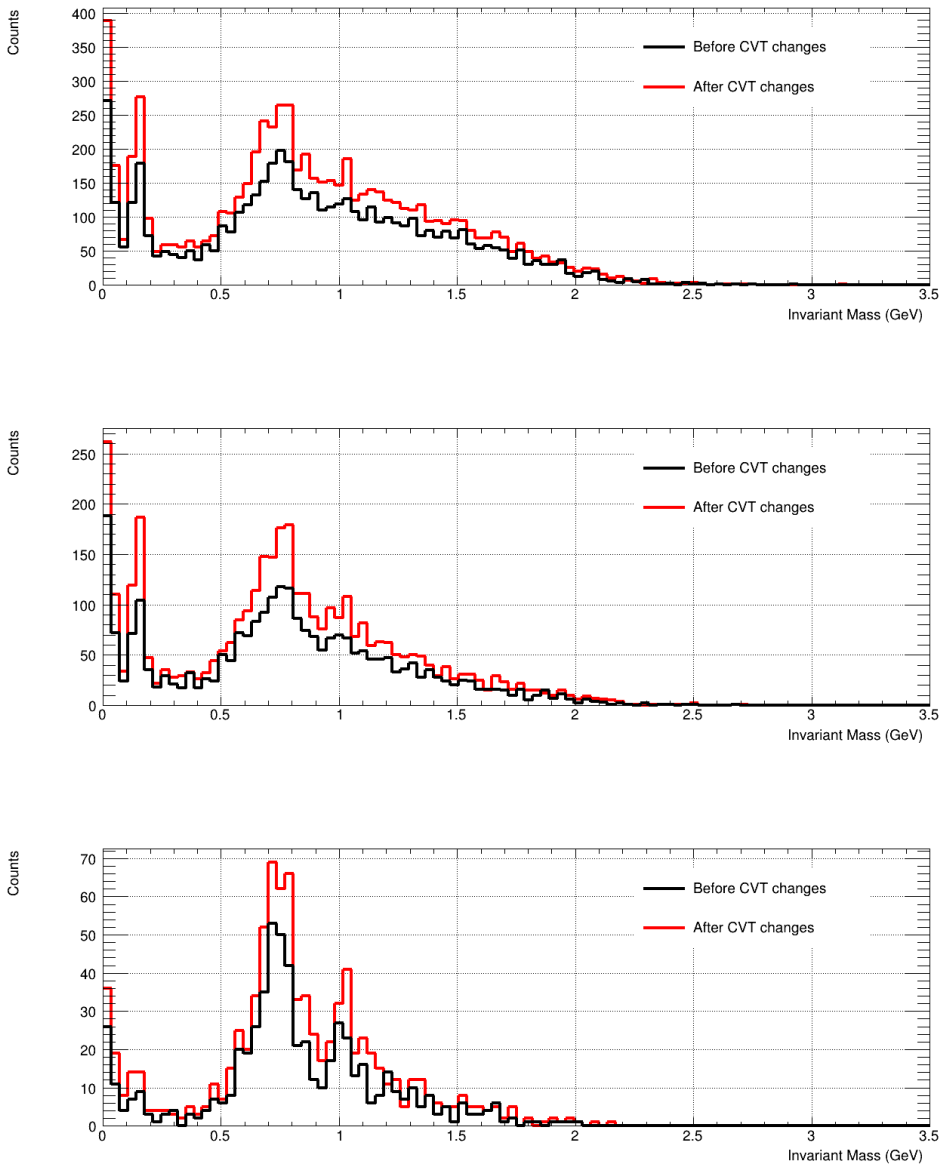


Figure 3.14: Reconstructed invariant mass of the dilepton pair before (blue, left) and after (red, centre) the CVT changes as well as both distributions overlaid (right)

3.3.3 Experimental Data

The first measurement of TCS on the proton used JLab experimental data from 2018 and 2019, with a 10.6 GeV polarised electron beam on an unpolarised hydrogen target [72]. The construction of a longitudinally polarised target, used in the experiment presented in this thesis, will allow the additional extraction of target polarisation asymmetries which can be used to further constrain GPD models.

Polarised Target

The NH_3 target itself is made by freezing ammonia into solid beads of variable size, between $1-4$ mm in diameter, and packing these into a liquid helium bath. The process of polarising the target was done by implanting free radicals (as a source of unpaired electrons) into the target material and polarising the mixture under the conditions of a high magnetic field of 5 T provided by the CLAS12 solenoid, and low temperature, maintained at approximately 1 K by a liquid helium refrigerator system.

Nuclear spin alignments in the target material at this temperature result in low polarisation, protons have a polarisation of around 0.5% and deuterons have a polarisation of around 0.1%. Electrons, however, are highly polarised at these conditions, around 99%.

When microwaves are applied to the sample, an electron spin and a nearby nucleon spin are flipped, and the electron polarisation is effectively 'transferred' to the nucleon in question (proton/deuteron in this case). The electron returns rapidly to its ground state, in around 100 milliseconds, but the nucleon remains in its flipped state. The electron can then be used for further nucleon flips, and in this way the whole sample can be polarised. A pictorial representation of this process is shown in Figure 3.15.

If the applied microwave frequency is below the electron resonance frequency of 140 GHz, it produces a positively polarised sample, and at a frequency higher than the electron resonance frequency a negative polarisation is produced. See Figure 3.16 for a visual representation of the target.

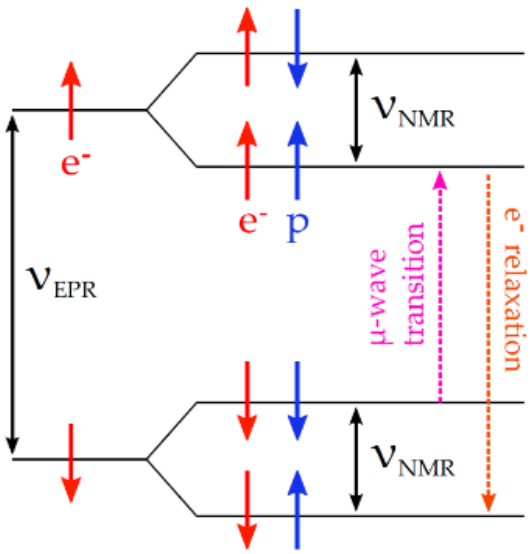


Figure 3.15: A pictorial representation of the Dynamic Nuclear Polarisation process, showing how applying microwaves to the sample can transfer electron polarisation to nucleons in the target material. Taken from [73]. The terms ν_{EPR} and ν_{NMR} represent the electron paramagnetic resonance frequency and the NMR frequency respectively.

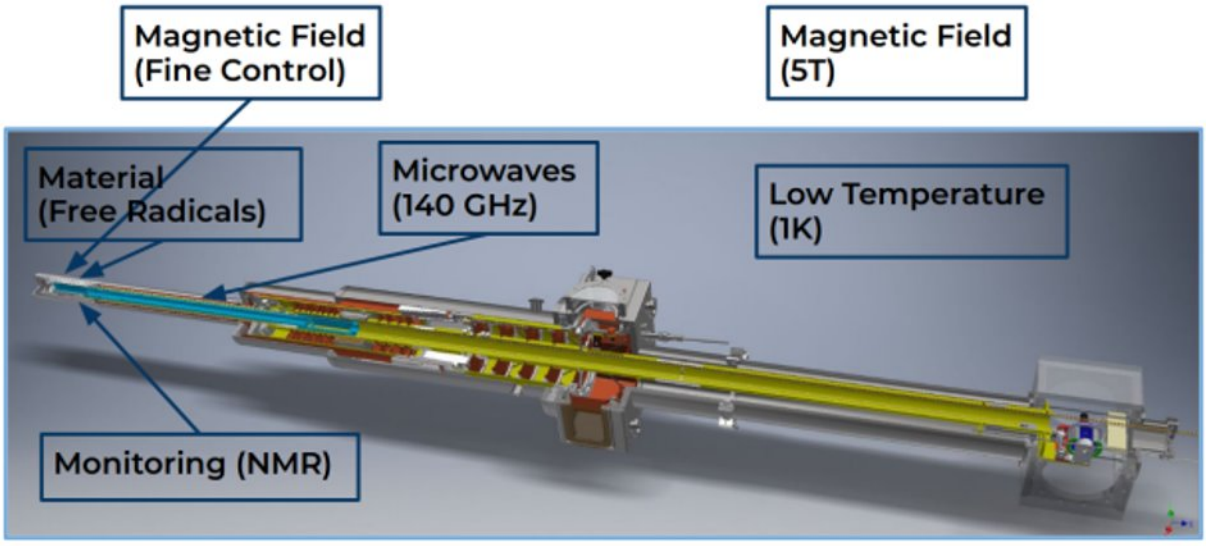


Figure 3.16: The RGC polarised target. Taken from an internal collaboration talk by Noémie Pilleux [74]

The polarisation of the target was measured throughout the experiment using Nuclear Magnetic Resonance (NMR), and the localised depolarisation effect of the high energy electron beam impinging on the target material was mitigated by uniformly moving the beam across the surface of the target material, in a process called rastering [73].

The raster system is comprised of two power supplies, two sets of dipoles, one to deflect the beam in the x-y plane, and a second to bend the beam back parallel to the z-axis (the beamline) and a raster driver, which controls the output current of the dipoles to produce the desired raster pattern. Figure 3.17 shows the x-y position of triggered events on the target, pre-calibration (so geometries should not be taken as completely accurate) but the distribution still shows occupancies being filled in a spiral pattern with a fixed gap distance and constant velocity. Note that this run was taken with a raster radius of 9 mm due to having larger target cell diameters, while the dataset used in this thesis had a raster radius of 6 mm.

The raster requires specific calibration, which is achieved by taking a tungsten wire target, which is comprised of the empty target shell with four, 0.051 mm thick tungsten wires arranged in a crosshair pattern across the end, as shown in Figure 3.18. This target is used to ascertain the optimal beam position and voltage settings that ensure the set raster pattern covers the maximal target surface area. The beam is directed towards each of the four crossing wires and its x and y position is recorded on the beam position monitor 2H01 (see Section 3.1.1), such that a link can be made between the physical point on the wire the beam touches, and its corresponding BPM value. This process also checks the centroid of the target, and ensures the maximal raster radius won't scrape the target cup edges. After data is taken, a further calibration step is necessary to account for any offsets in the measured position and is done by checking events with a two-

particle final state, and from these, reconstructing the x, y and z vertex position of the target.

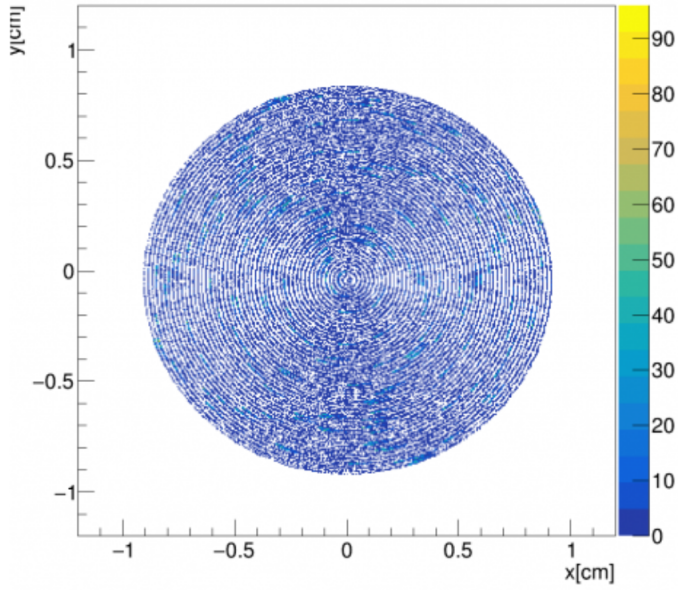


Figure 3.17: An image of the raster monitoring system for run number 17313 showing the spiral pattern in the x-y plane

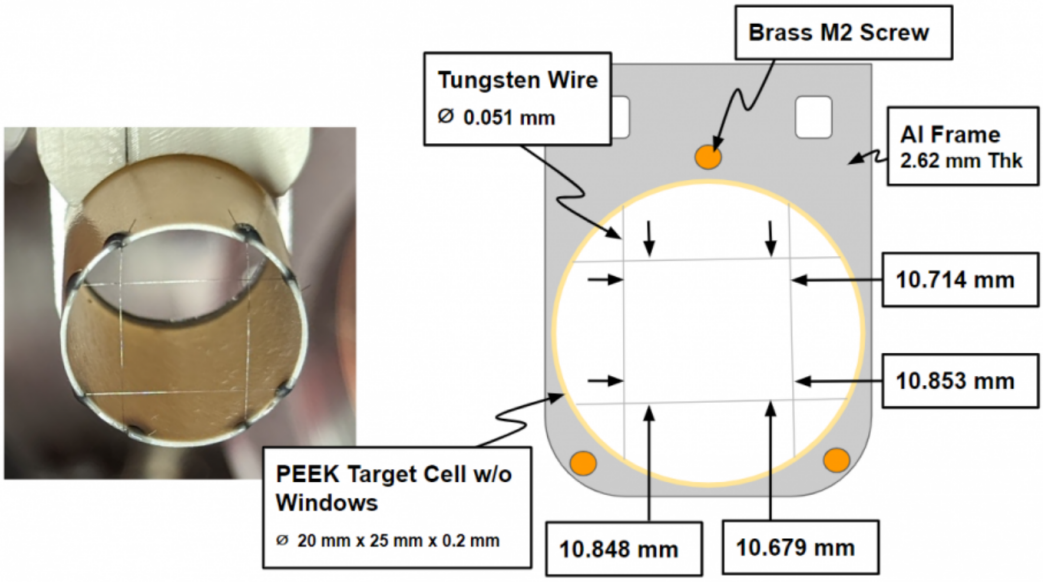


Figure 3.18: The geometry of the target used for calibration of the raster system. Taken from [75]

There were six target configurations throughout data taking, NH_3 and ND_3 were the key targets for analysis, and Carbon, CD_2, CH_2 and empty target runs were used for background estimates, accurate determination of the target density and calculation of the dilution factor. The dilution factor is the factor included in observable extraction which allows the exclusion of 'unpolarised' components, or at the very least, components which are not the target proton. This is discussed further in Section 4.

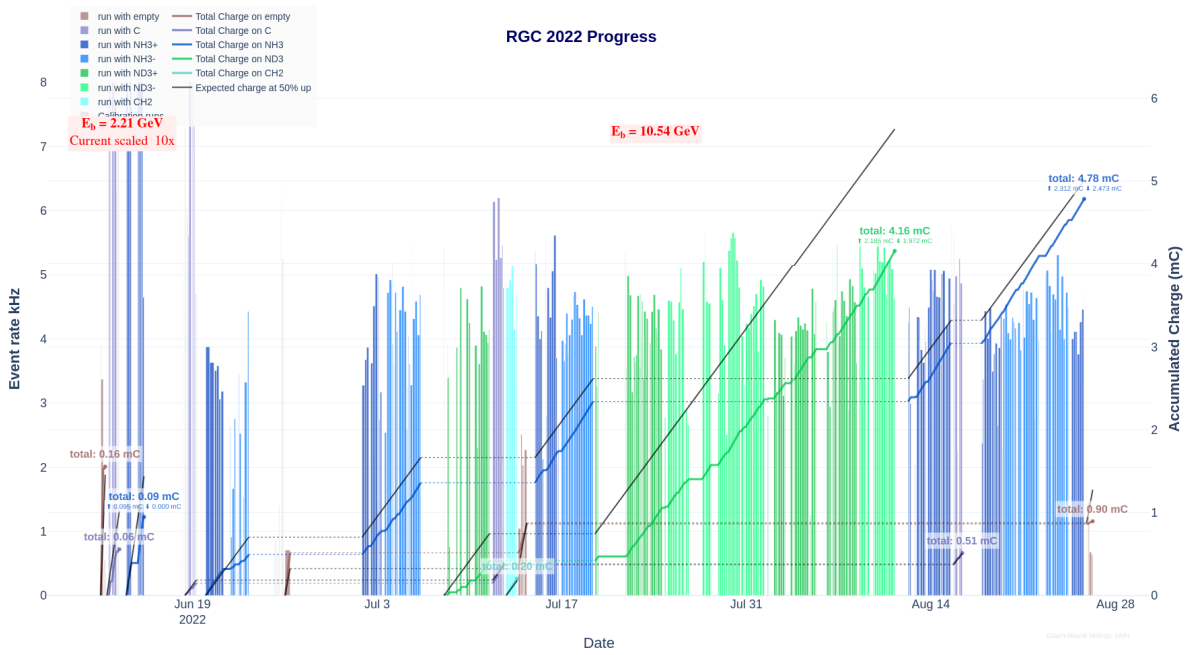


Figure 3.19: The Summer 2022 data taking run, showing the accumulated charge taken on all six target configurations, as well as runs for the purposes of calibration.

Data Taking Run

The experimental data-taking run at Jefferson Lab took data in Hall B using CLAS12 and ran between June 2022 and March 2023. The run plan included a period of time for data taken on longitudinally polarised NH_3, which allows the measurement of interactions on the polarised proton. The total accumulated charge taken on NH_3 amounted to 13.06 mC, a subset of which is shown in Figure 3.19. Extracting a measurement of TCS with this data is a first-time measurement at Jefferson Lab.

The analysis procedure selects the decay lepton pair (e^+e^-), and the scattered proton, with a missing mass squared study of the scattered electron;

$$\begin{aligned} MM_X^2 &= ((\vec{p}_e + \vec{p}_p) - (\vec{p_{p'}} + \vec{p_{e^+}} + \vec{p_{e^-}} + X))^2 \\ X &= \vec{p_{e'}} \end{aligned} \tag{3.5}$$

since the electron mass is very small, within detector resolution we expect to see a peak at zero, which is an indicator of a clean TCS signal. The proton can be selected to be detected in either the Forward Detector (FD) or the Central Detector (CD), presenting two topologies on which to perform the analysis. Asymmetries are extracted in the form of beam spin asymmetry (BSA) and target spin asymmetry (TSA), the latter a novel measurement available only when the target is polarised.

Chapter 4

Analysis

4.1 Data Analysis Procedure

During the data processing or 'cooking' stage, the raw data is processed and calibrated ready for analysis. The final stage of this involves taking the processed data, and performing what is called a 'skim' which in essence uses a particular script, one for each channel, to loosely select the final state particles of interest, (in this case, a proton, an electron, and a positron) and 'cut' the events that do not meet the conditions. There were several skims performed on the dataset and the one used for this analysis was the J/Ψ skim, which filters out events with a dilepton final state, so named due to its usage for J/Ψ analyses. The different possible filtering conditions in this skim are as follows:

1. The final state event must include an e^+e^-, e^-e^-, or e^+e^+ pair and at least one other positively charged particle.

2. The final state must include a $\mu^+\mu^-, \mu^-\mu^-,$ or $\mu^+\mu^+$ pair and at least one other positive particle in the Forward Detector (FD).

3. a high momentum (over a defined threshold of 2 GeV) $e^+e^-, e^-e^-,$ or e^+e^+ pair.

4. a high momentum $\mu^+\mu^-, \mu^-\mu^-,$ or $\mu^+\mu^+$ pair.

5. the number of electrons in the forward tagger (FT) > 0 and number of protons in FD > 0 and e^+ or e^- have a high momentum.

If any single one of the above conditions are met, the event is saved and is passed to the skim. If the event fails every condition, it is discarded. This ensures the maximal statistics are kept for analysis and the likelihood of discarding a 'good' event is minimised.
The next stage of analysis is to pass the skimmed data through chanser [76]. Chanser is a wrapper around a CLAS12 analysis software tool called clas12root [77] which takes information from particle banks in a .hipo file, and writes them to a ROOT [78] tree on which analysis can

be performed. A key feature of chanser is the fact that it can be used to perform a combinatorial analysis, meaning it can take a single event from the input file and perform all possible combinations of particles that fit the required final state. This is particularly useful when an event contains more particles than is required by the final state, or when tracks are not clearly identified [79]. It is within chanser where all four vectors are defined and both fiducial cut classes and masking/radiative corrections classes are introduced which can be applied in the next stage of analysis. The outputted files from chanser are passed through a ROOT macro which takes each element from the ROOT tree, applies fiducial cuts and, while iterating through all the events, applies exclusivity cuts and generates kinematic distributions (which are compared with Monte Carlo simulations) as well as the final state observables. Nuclear background contributions are also dealt with at this stage. More details on each of these stages will be discussed in the following sections.

4.1.1 Fiducial Cuts, PID and Corrections

A series of fiducial cuts and PID cuts to detector geometry and particle momenta had been previously established for the experiment which ran at JLab in 2018/2019 on the unpolarised liquid hydrogen target. These cuts and the procedures used to apply them were used as a starting point to establish whether any modifications were required for the 2022/2023 polarised nuclear target run. Further details on each of these fiducial cuts beyond what is described here can be found in [80]. The first series of cuts are those automatically included by the Event Builder when a particular particle identification (PID) is required. Using chanser, inherently in the topology of events, a PID is assigned to each final state particle before kinematics are calculated. For an electron, this automatically requires that there must be a minimum energy deposition in the PCAL (pre-shower calorimeter) of 60 MeV and that the total energy deposition divided by the momentum of the electron (which is referred to as the Sampling fraction, SF) must be within 5σ of the mean value for the detector. The final requirement is that the number of photo-electrons that register as having a hit in the HTCC must be greater than 2. Each of these will be considered in more detail, and in some cases modified, as the cuts are established.

Electron z-vertex

Fiducial cuts should be thought of sequentially. As the particle passes through CLAS12, it is accepted based on the criteria that it meets the fiducial geometry of each subsystem in turn. One of the first geometries to consider is that of the target itself.

The target was centred in physical lab coordinates at (0 cm, 0 cm, -3 cm), with translational and rotational deviations reported of the order $<$0.1 mm. Looking at Figures 4.1 and 4.2, which show the vertex position of the electron, there is a sharp peak in x, y and z at the reported target positions. The small bumps on the z-vertex distribution (Figure 4.2) to the right of the peak

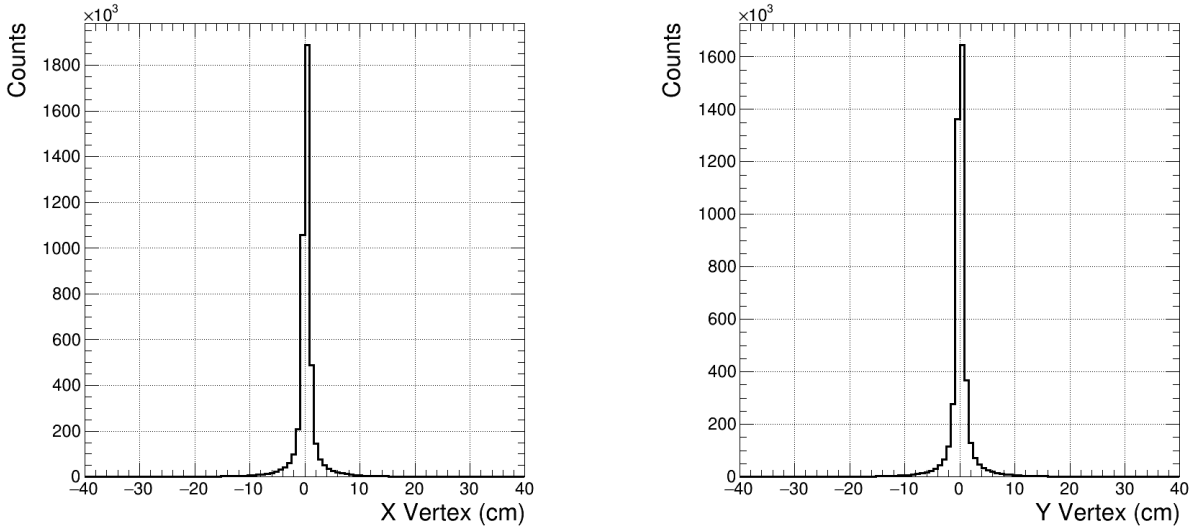


Figure 4.1: The electron x and y vertex, showing the target position centred at 0 cm in both coordinates.

marked by arrows dictate the position of a fiducial cut that should be applied to remove edge effects from the exit window of the target. Recalling from the previous section and Figure 4.18 (ensuring all positions on Figure 4.18 are shifted to reflect the physical centre of the target in the solenoid of -30 mm) this is where the close succession of exit windows appears, firstly at around 7 cm (10 cm in the schematic) then up to 11 cm (14 cm in the schematic). Thus, a cut to reject events outside of the window -10 cm to 5 cm is placed on this distribution, as shown by blue dashed lines.

HTCC

After the target and barrel subsystems, the first region seen by the interacting electrons is the HTCC, where there is significant pion contamination below a momentum threshold of approximately 4.5 GeV. To reduce the contamination up to this momentum value, a cut is applied which requires the minimum number of photo-electrons registered in the HTCC must be > 2, which is the minimum expected number of photo-electrons from an electron candidate track. The effect of this cut can be seen in Figure 4.3, where the large, negative pion dominated peak below 2 photoelectrons has been removed.

DC

The next region seen by the forward-going particles is the Drift Chambers (DC). There are two different DC fiducial cuts applied, one in the local x-y plane, which is applied to electrons, and one in the local $\theta - \phi$ plane, which is applied to protons. These cuts are based on a distribution of tracks in each sector and layer of the DC in each local coordinate plane (x-y or $\theta - \phi$), where slices of tracking χ^2/NDF are plotted, and fitted at a fixed constant. These cuts serve to remove

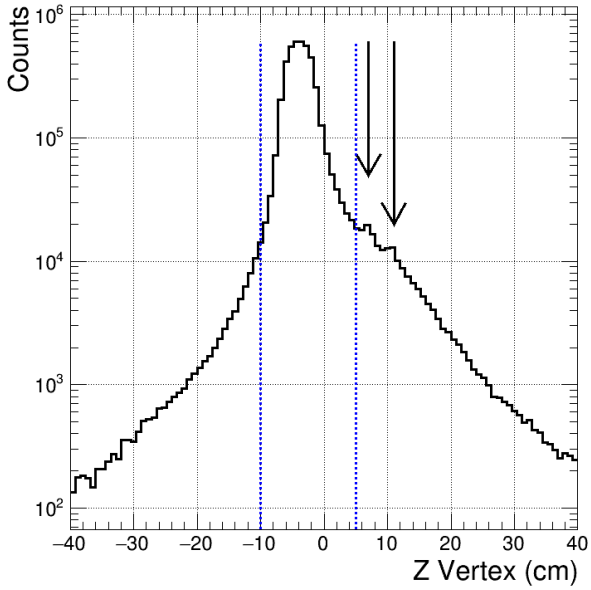


Figure 4.2: The electron z-vertex on a log scale, showing the target position centred at -3 cm and the bumps attributed to exit windows marked by arrows at 7 cm and 10 cm. The cut limits on the z-vertex are marked with blue dashed lines at -10 cm and 5 cm.

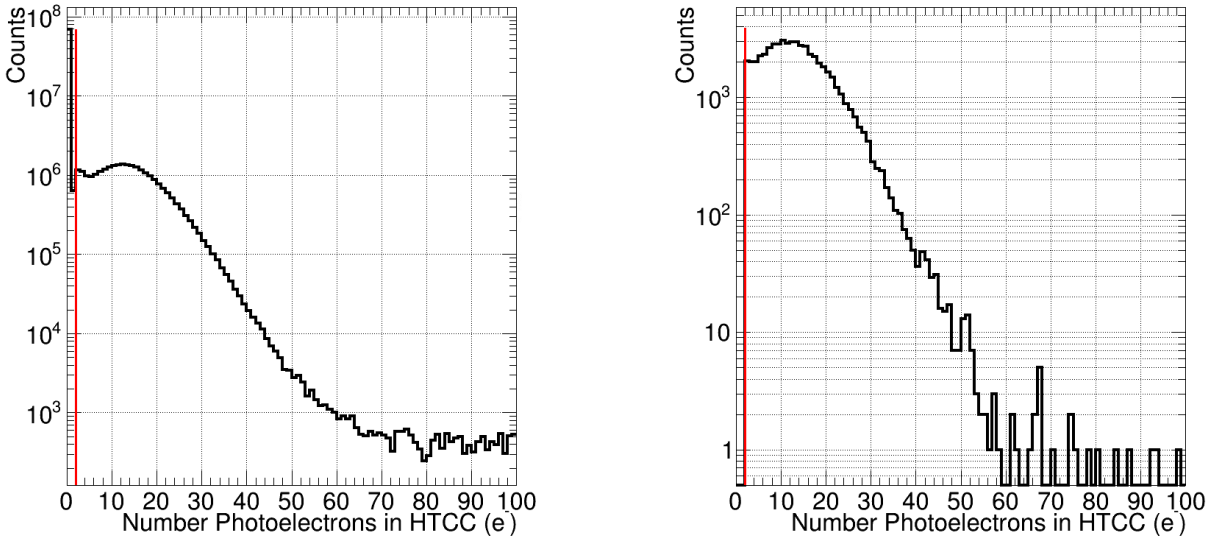
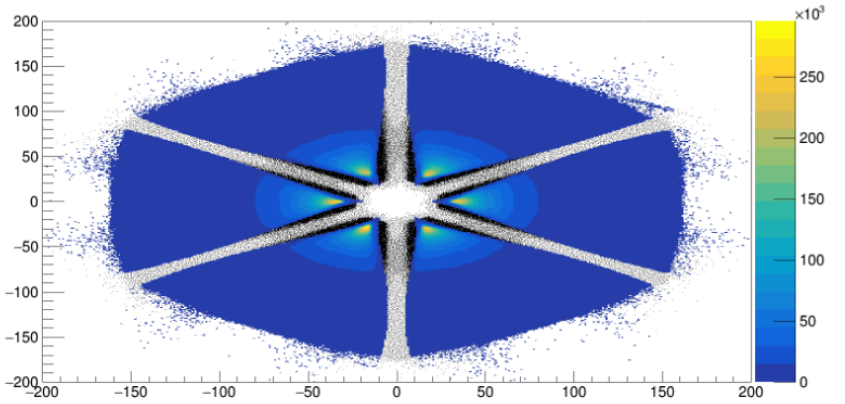
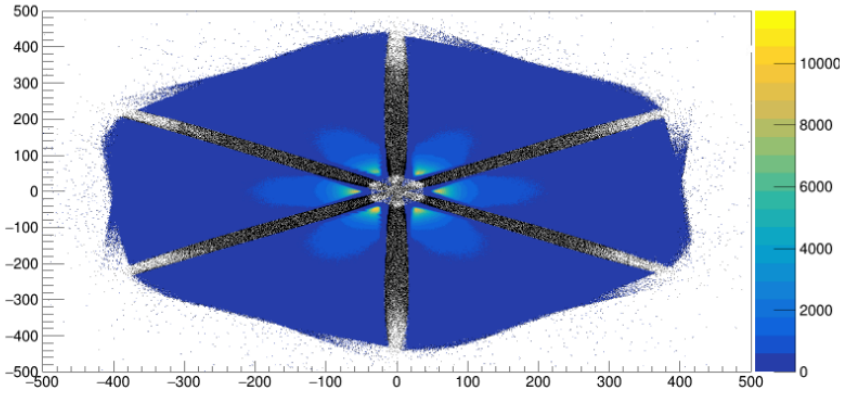


Figure 4.3: The number of photo-electrons registered in the HTCC for the electron candidate track, Left with no cut and Right with a minimum cut of 2.

poorly reconstructed tracks close to the edge of the detector. An example of the effect of these cuts for the 2018/2019 dataset is shown in Figure 4.4.



(a)



(b)

Figure 4.4: Figure 4.4a shows the effect of the DC fiducial cut in the x-y plane for electrons, and Figure 4.4b shows the effect of the DC fiducial cut in the $\theta - \phi$ plane for protons. The points cut out by each fiducial cut are shown in black. The x and y axis represent the DC geometry, in units of cm (each triangular sector has an outer width of 4 m in DC region 3, and 1.5 m in DC region 1). Taken from [80].

Calorimeters

Since the electron and positron are detected in the FD region, it is important to remove contamination from pions in the calorimeters. To do this, multiple cuts are used. Firstly there is a cut applied to the energy deposition in the pre-shower calorimeter (PCAL), this is set to be a minimum of 0.06 GeV in the event builder, which also proved sufficient in removing the pion 'hot-spot' in this analysis.

After this initial cut, the distribution looked like what is shown in Figure 4.5, where it is clear that a cut on the energy deposition in the ECIN and ECOUT is required, as there is a clear banded structure at a low EC energy deposition on the final distribution. After investigation, this was found to be attributed to particles that did not deposit energy in both the ECIN and ECOUT,

which can happen if the track is close to the calorimeter edge and therefore its energy not being fully deposited.

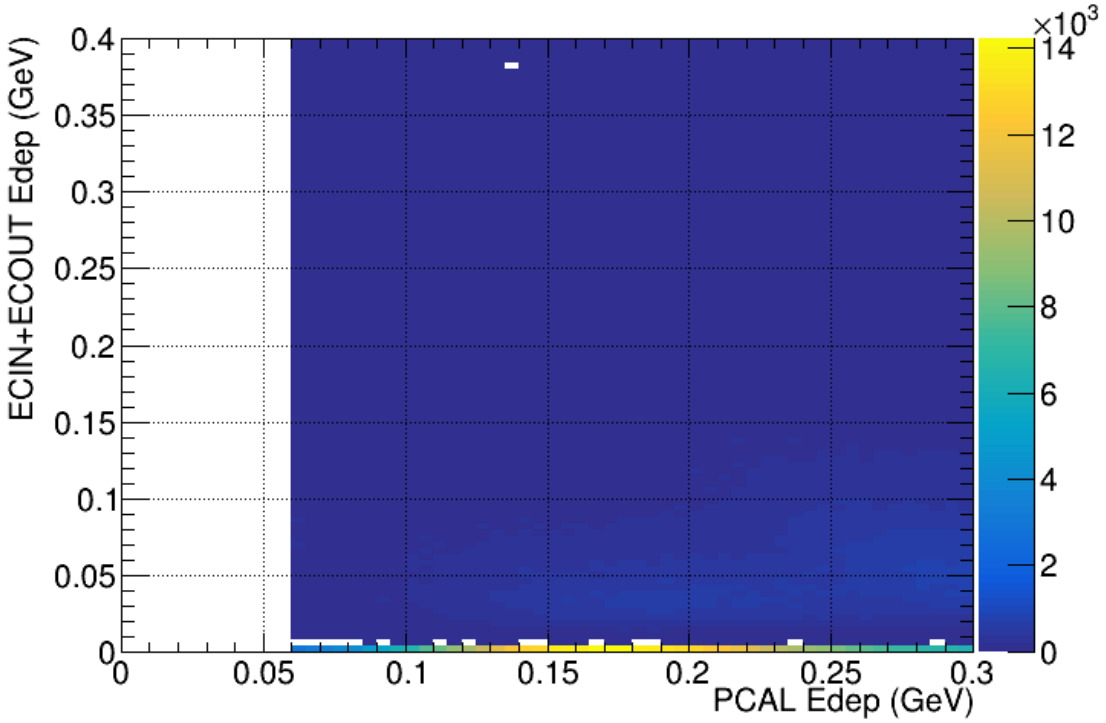


Figure 4.5: The result of applying a cut on the PCAL energy deposition of > 60 MeV. There is a clear band at low energy deposition in the EC which is swamping the signal events.

To ascertain that this is indeed the cause or if the particles are in fact pions, and in turn whether or not these events should be cut from the final stage observable, the invariant mass of the particles identified as the dilepton pair within the hotspot region (with an energy deposition in the inner and outer calorimeters of < 0.005) were plotted, with a series of exclusivity cuts that honed in on a dilepton final state (namely a low missing mass squared, quasi real initial photon and the dilepton pair in the forward detector). The distribution was then checked against a normalised simulation (more detail on this procedure is explored in Section 4.1.3) for the purposes of shape alone, and to observe the counts in the ϕ meson peak, where not only are the statistics are more plentiful, but the presence of this peak is a clear indication of e^+e^- events. This comparison is shown in Figure 4.6.

It is clear from Figure 4.6 that there is an agreement in shape for data and simulation in the hotspot across the full distribution, however when honing in on the ϕ peak region at around 1 GeV, two clear peaks are resolved, one at the ϕ mass and one at the $\rho(770)$ mass. The for-

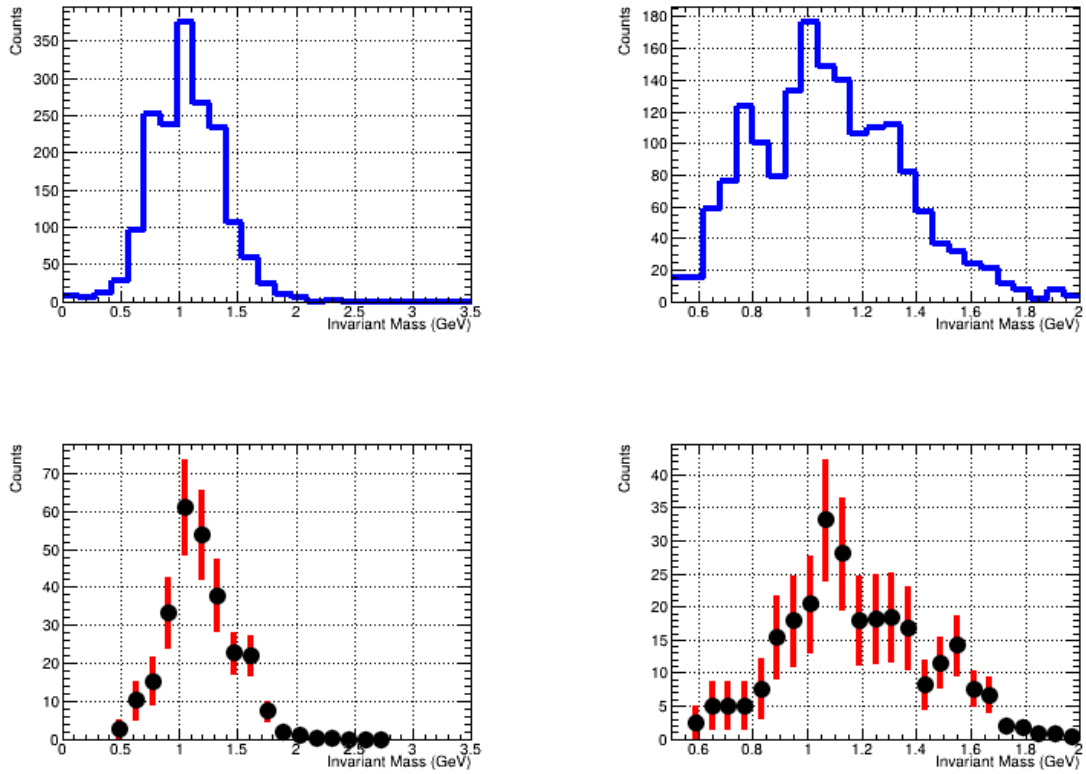
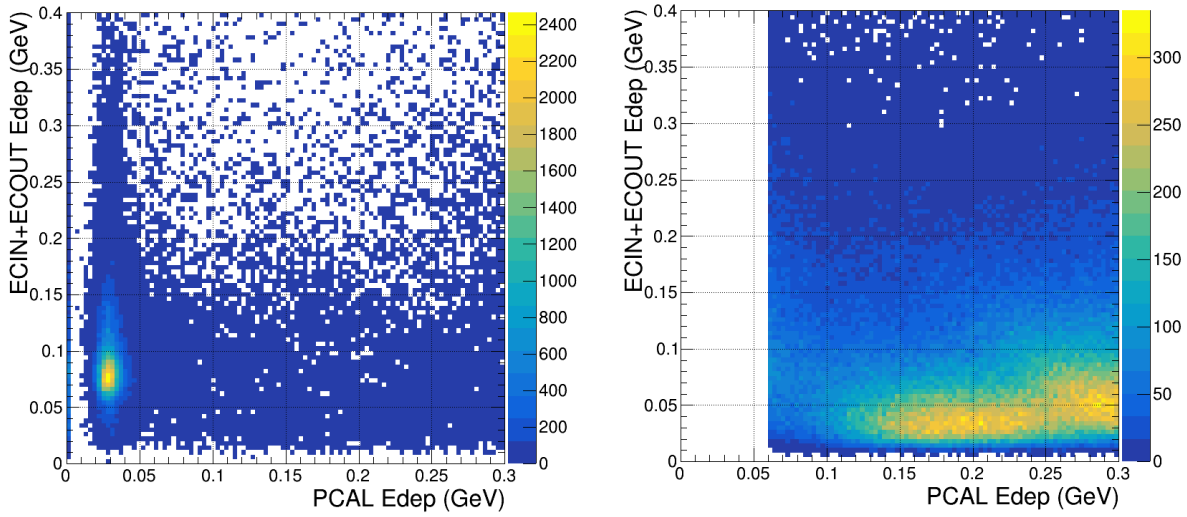


Figure 4.6: The invariant mass distribution of the dilepton pair in the hotspot region, with a series of exclusivity cuts applied to ensure a dilepton final state. The left plots show the full distribution, and the right show the the area around the ϕ meson peak, with the data in blue (top) and the simulation in red (bottom).

mer is indicative of a dilepton final state, and the latter likely belonging to pions as this peak is not present in the simulated e^+e^- events. Next, the possibility of the events belonging to e^+e^- tracks that have hit the edge of the calorimeter was investigated, to do this the same events were plotted, but with the added requirement that the energy deposition in the ECIN and the ECOUT were both greater than zero. This produced an empty plot, indicating these were all events from tracks that had missed one or both of the calorimeters and therefore their energy deposition was not fully reconstructed. This is a clear indication that these events should be removed from the final state observable, and a cut of $\text{ECIN}_{E_{dep}} + \text{ECOUT}_{E_{dep}} > 0.005$ was applied to the final analysis.

The effect of the PCAL cut intrinsic to the Event builder, alongside the additional requirement that $\text{ECIN}_{E_{dep}} + \text{ECOUT}_{E_{dep}} > 0.005$ can be seen in Figure 4.7, where Figure 4.7a shows the energy deposition in the PCAL versus the energy deposition in the sum of the inner and outer calorimeters, with only a cut requiring the final state lepton pair are in the forward detector and $\text{ECIN}_{E_{dep}} + \text{ECOUT}_{E_{dep}} > 0.005$. Figure 4.7b shows the same distribution after a 0.06 GeV



(a) (b)

Figure 4.7: Figure 4.7a shows the energy deposition in the PCAL vs the ECAL, with only cuts requiring an electron or positron in the FD, we see a clear region at < 0.05 GeV attributed to pion contamination. Figure 4.7b shows the same but with a minimum cut of 0.06 GeV in the PCAL, as well as all fiducial cuts established so far.

minimum energy deposition requirement, along with the previously established z-vertex, HTCC and DC cuts.

The next series of cuts involve the sampling fraction, which is set to be within 5σ of the mean value in the event builder. The sampling fraction is defined as the total energy deposition in each of the calorimeters, divided by the momentum of the particle. This analysis tightens the cut to 3σ of the mean value, for sampling fraction versus momentum, and sampling fraction versus energy deposition in the PCAL. The same cut is imposed for all six calorimeter sectors, and the effect of these cuts for sector 1 can be seen in Figures 4.8 and 4.9. Figures 4.8a and 4.9a show the sampling fraction vs momentum, and energy respectively for electrons with only the requirement of the final state lepton pair in the FD applied. The banded region in the sampling fraction vs energy deposition plot is attributed to pions with a low energy deposition in the PCAL, which are treated with the fiducial cut established in the previous section. Pions that shower upstream of the ECAL, however, can be seen as a deviation from the mean value in the sampling fraction vs momentum plot, so applying cut on 5σ (or, more tightly, 3σ) of the mean value is necessary to remove these pions, and we begin to see the candidate lepton tracks as a clearer defined region in figures 4.9c, 4.9b, 4.8c and 4.8b. In the final analysis, the 3σ cut was applied to both the sampling fraction vs momentum and sampling fraction vs energy deposition plots, as this was found to be the most effective at removing the pion contamination, without significantly impacting the signal region.

A further calorimeter cut involves the sampling fraction in the inner and outer calorimeters, with respect to the sampling fraction in the PCAL. This distribution effectively serves to show pions

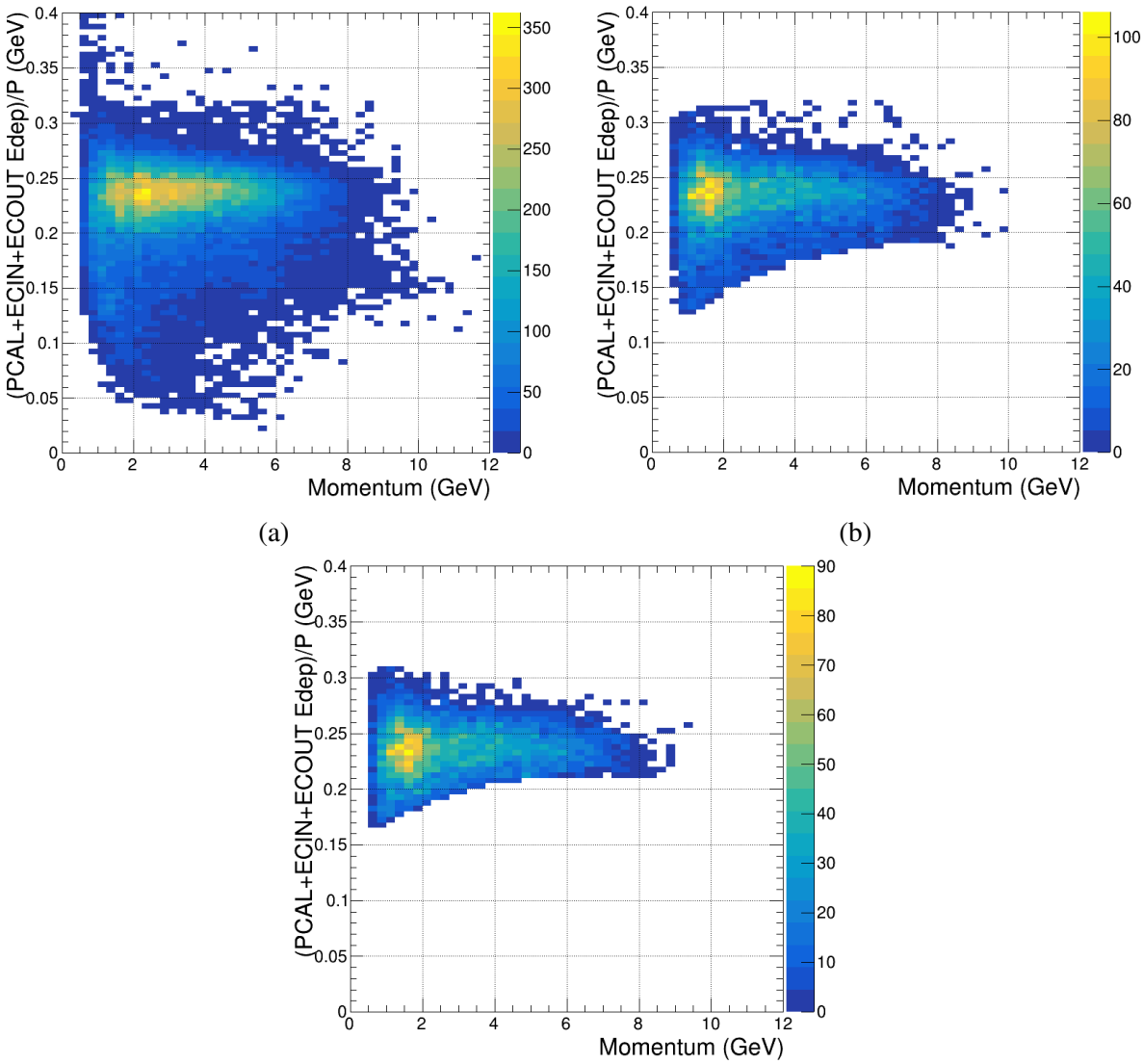


Figure 4.8: Figure 4.8a shows the sampling fraction vs momentum for electrons with no cuts applied, Figure 4.8b shows the same but with a 5σ cut applied, and Figure 4.8c shows the same but with a 3σ cut applied.

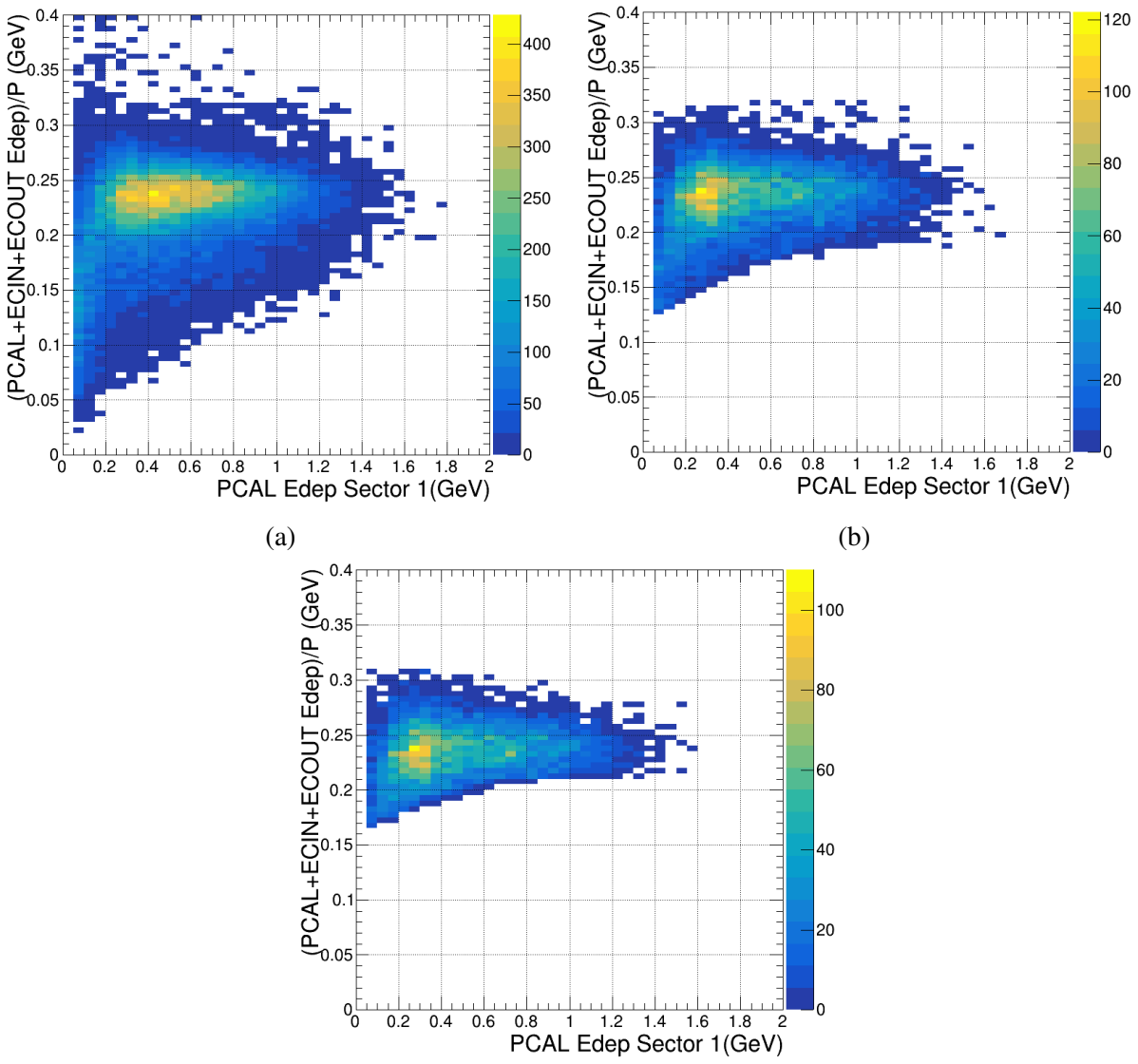


Figure 4.9: Figure 4.9a shows the sampling fraction vs energy deposition in the PCAL for electrons with no cuts applied, Figure 4.9b shows the same but with a 5σ cut applied, and Figure 4.9c shows the same but with a 3σ cut applied.

with momentum greater than the threshold of the HTCC (> 4.5 GeV), meaning they trigger the HTCC and are counted as electrons, and are therefore not identified/removed by the previous cuts. To visualise this, see the left side of Figure 4.10, which shows the described distribution of ECIN + ECOUT sampling fraction versus PCAL sampling fraction, with all cuts up to this point applied (albeit without the recently established sampling fraction cuts, so as to show the full distribution). A clear sloping pion tail is shown at lower sampling fraction values. The right hand side plot in Figure 4.10 shows the same distribution after the sampling fraction cut of 3σ is applied, as well as cut on the intercept between the correlation of the sampling fraction in the PCAL and the inner and outer calorimeters for any particle with momentum above 4.5 GeV. This latter cut more clearly requires that if $p > 4.5$ GeV, then a check is made that $\mathrm{SF}_{ECIN} < 0.2$ - SF_{PCAL} where the intercept, in this case, is set to 0.2. If the particle has momentum < 4.5 GeV, it is automatically allowed to pass the cut as it is assumed the HTCC will separate the pions an electrons at these momenta. It is clear that the tighter sampling fraction cuts along with the intercept cut help eliminate both pions registering in the PCAL, and those that affect upstream subsystems.

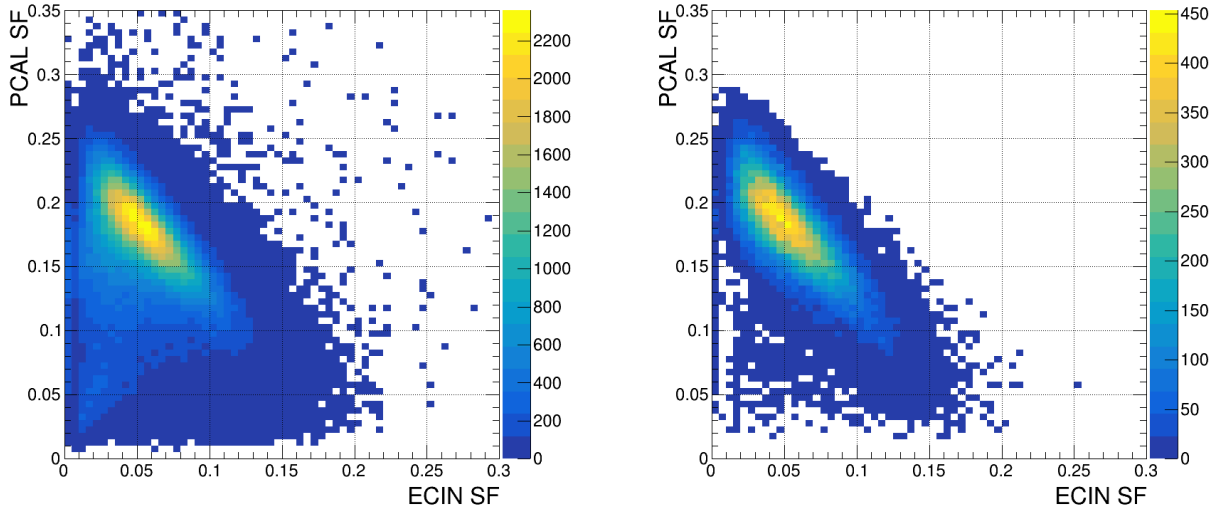


Figure 4.10: The correlation between PCAL and ECAL, on the left with all previously established cuts up to and not including sampling fraction applied, and to the right with SF cut of 3σ and a cut on the intercept < 0.2 when $p > 4.5$ GeV applied.

The last fiducial cuts placed on the calorimeter detector volume were on the PCAL, these cuts serve to ensure that the electromagnetic shower produced was from a hit in the detector volume, if the hit occurred at the edge of the PCAL, there will be a mis-reconstruction of the energy deposition, and therefore the sampling fraction itself will be incorrect. Removing these particles that hit right on the edge of the calorimeter means electrons will be more accurately reconstructed. The calorimeters are made up of bars, and at least one full bar is required to reconstruct a particle hit. The PCAL however has one extra bar compared to the ECIN and

ECOUT, so in order to ensure good reconstruction, an energy deposition at least two bars from the edge of the detector volume are required. Looking at the local PCAL coordinates versus sampling fraction in Figure 4.11, it is clear that the first two bars make up approximately 10 cm of the detector width. After about 14 cm, the correlation between sampling fraction and width becomes flattened at each sector, so it is at this point the cut is applied, as shown in Figure 4.11b.

The effect of this cut can be seen in Figure 4.12, where the PCAL detector volume is shown on the left after all cuts up to this point, and on the right after the fiducial cut is applied to remove these edge effects, overlaid with the original distribution, to clearly show where the electron candidates have been removed.

A test was done to ascertain the pion rejection of these cuts, achieved by performing a chanser analysis of the same dataset, with a $\pi^+\pi^-$ final state in the forward detector and in the target area, and splitting the events into those with both pions having momentum < 4.5 GeV (i.e. identified by the HTCC as almost assuredly pions) and events identified with final state pion momentum > 4.5 GeV (above the HTCC threshold).

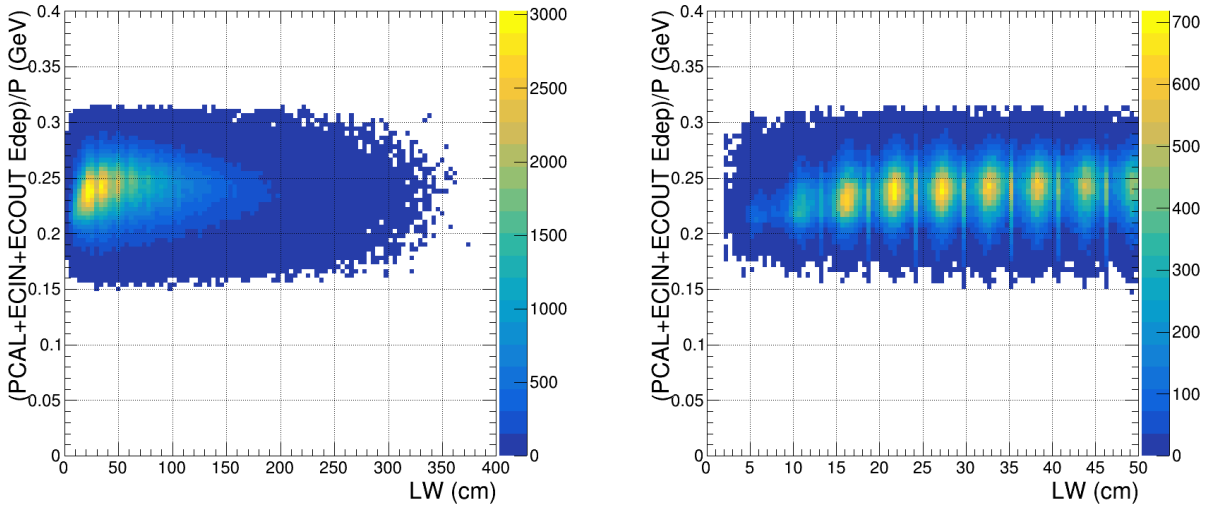
The resulting rejection percentage for pion events (calculated by taking $\frac{(N_{\text{total}}-N_{\text{passed}})}{N_{\text{total}}}$) for cut independently is shown in table 4.1 for $p < 4.5$ GeV and table 4.2 for $p > 4.5$ GeV. When all cuts are taken together, the overall pion rejection for pions with momentum < 4.5 GeV is 99.88%. Note all events will survive the triangle cut due to the fact that it allows all particles with momenta < 4.5 GeV to pass. For momentum > 4.5 GeV, when all cuts are taken together, the rejection was found to be 98.4%.

Table 4.1: Pion rejection for each cut for $p < 4.5$ GeV

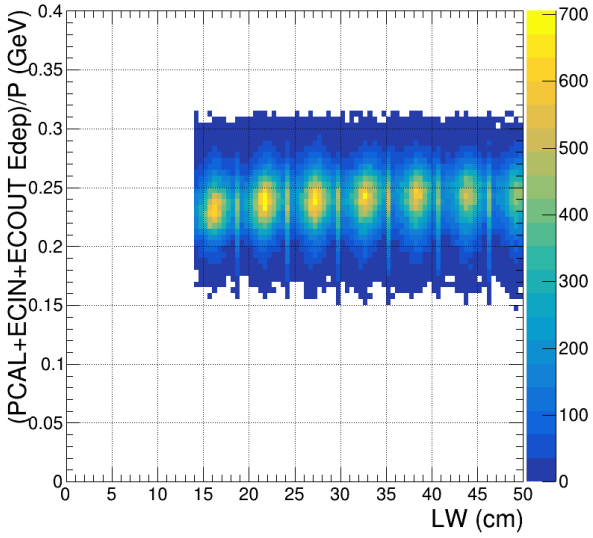
Cut	rejection percentage
Number of photoelectrons > 2	59%
Minimum PCAL energy deposition > 60 MeV	53%
SF = 3σ	92%
Triangle cut intercept = 0.2	0%
DC fiducial cuts	6%
PCAL fiducial cuts	15%

Table 4.2: Pion rejection for each cut for for $p > 4.5$ GeV

Cut	rejection percentage
Number of photoelectrons > 2	4%
Minimum PCAL energy deposition > 60 MeV	2%
SF = 3σ	91%
Triangle cut intercept = 0.2	97%
DC fiducial cuts	3%
PCAL fiducial cuts	30%



(a)



(b)

Figure 4.11: Figure 4.11a shows the correlation between the local LW coordinate plane of the PCAL and the sampling fraction across the full detector volume (left) and zoomed in to show the bar structures (right). Figure 4.11b shows the application of a 14 cm cut to ensure that reconstruction is performed when a particle deposits energy across at least the first two bars.

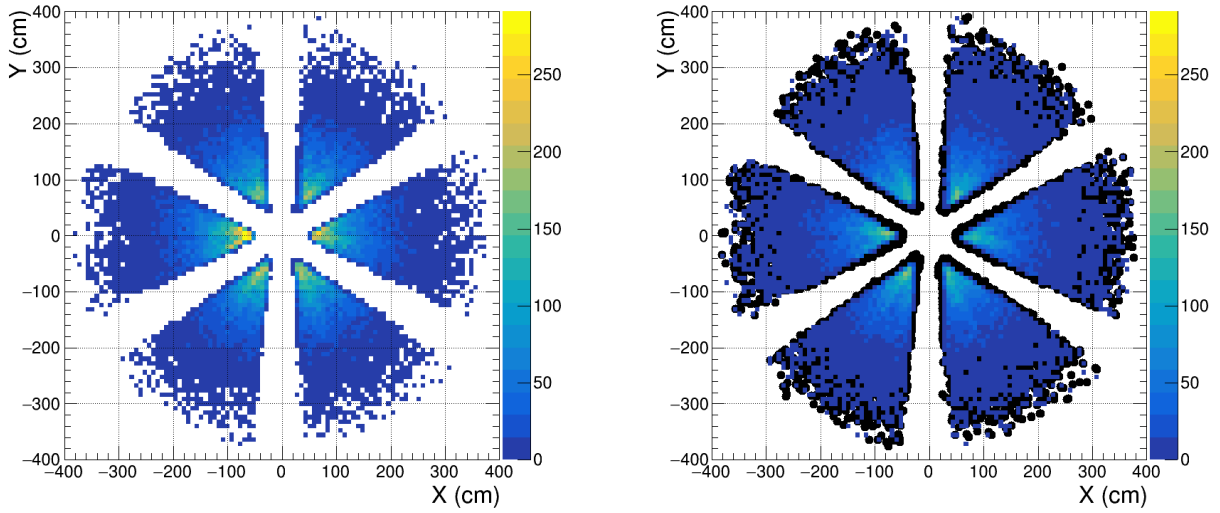


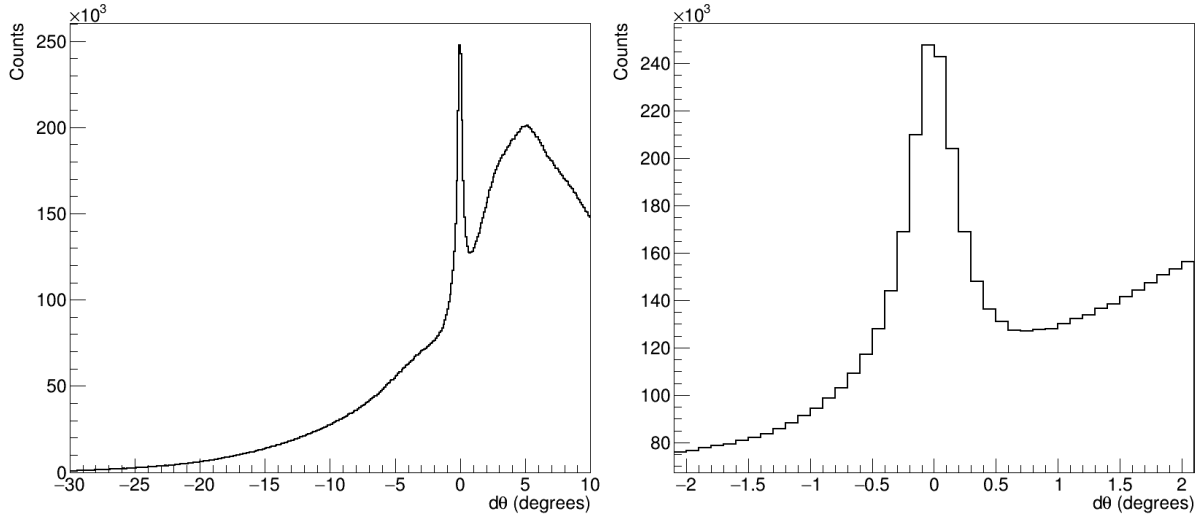
Figure 4.12: The detector volume of the PCAL showing each of the six sectors, on the left after all cuts up to this point, and on the right after the 14 cm fiducial cut is applied, overlaid with the original distribution (shown as black points at the edges).

Radiative Corrections

Since TCS involves the exchange of real and quasi-real photons with the target nucleon, there is the potential for radiated photons to be produced from the final state electrons deflecting in the forward torus field. The detection and reconstruction of these can lead to an incorrect reconstruction of the electron momentum and energy deposition. In order to correct for this, the angle between the electron and the detected photon is calculated. When this is small, below a defined threshold, one can make the confident assumption that this photon was radiated from the electron, and therefore carries some of its momentum. Rather than cutting these events out and affecting the statistics of the dataset unnecessarily, the momentum of these electrons is corrected by the momentum carried by those photons produced at a small angle. This region is shown in Figure 4.13, where it is clear the corrections should be applied to the $\delta\theta < \pm\, 0.7^{\circ}$ region.
From this point forward, unless otherwise declared, it can be assumed that all the fiducial cuts established above including radiative corrections are applied to all datasets.

4.1.2 Nuclear Background Subtraction

Since the data was taken on a longitudinally polarised NH_3 target, and the interaction of interest occurs on one free proton within the target sample, there is a significant background contribution from the Nitrogen in the target and surrounding materials such as the H_e in the target bath that must be understood and taken care of in the calculation of the observable. Throughout the analysis, three distinct methods were established to account for these contributions. The first method was assessing whether it would be possible to use a fitting algorithm modelled on the background data to extract a signal and background weight to supply to all kinematic



(a) (b)

Figure 4.13: Figure 4.13a shows the distribution of $\delta\theta$ between the electron and the photon, and Figure 4.13b shows the same but zoomed in to show the peak region attributed to radiative photons, between $\delta\theta < \pm\, 0.7^{\circ}$, where radiative corrections are applied.

distributions, as well as the final state observables. The second method was used to gain a visual understanding of how much of an acquired 'signal' in the dataset comes from a single proton, for the purposes of direct comparison with simulated data, which was only available on a free proton target. This method also allows a visualisation of additional effects such as Fermi motion. The third method follows the formalism established in [81] and [82], which outline the method of providing a correction factor, known as a dilution factor, to both the beam spin and target spin asymmetries that will remove the contribution from unpolarised materials.

Method 1

Understanding the contributions of the nuclear background to the NH_3 dataset was estimated by using data taken on the Carbon target, which, within a small ratio factor, can be likened to the contribution one can expect from the Nitrogen in NH_3. The first requirement for this study was to ascertain the total accumulated charge of the available Carbon data. There were a total of 27 Carbon runs available for analysis, and determining the accumulated charge was achieved by extracting information from the helicity scalar banks using clas12root. These are split between positive and negative accumulated charges in the Faraday cup ($FCup_+$ and $FCup_-$) and taking their sum gives $FCup_{total}$, which is used to normalise the carbon counts. Once the normalisation was appropriately calculated, the data, along with the NH_3 data was supplied to brufit [83], which is a maximum likelihood fitting package based on RooFit (a fitting tool in ROOT). With brufit, an sPlot (a Likelihood fitting tool that employs a weighting technique to separate out signal and background [84]) fit of the MM^2 of the scattered electron can be generated, as shown in Figure 4.14, where the data PDF represents all 157 runs worth of NH_3

data, with some preliminary exclusivity cuts, driven by the TCS analysis in [72].

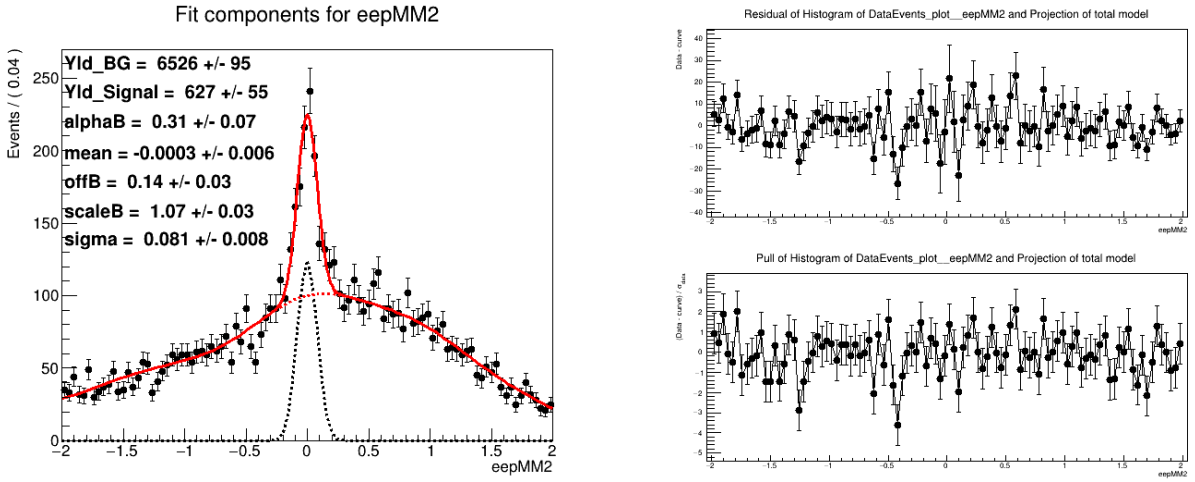


Figure 4.14: An sPlot fit of MM^2, with the requirements that the initial photon is quasi-real, MM^2 (called eepMM2 on the x-axis) is between +/-2 GeV^2 in order to get a good model for the background, and the dilepton pair in the forward detector. The histogram on the top right shows the residual (Data - curve) and on the bottom right the pull ((Data-curve)/mean), both shown across MM^2 of +/-2 GeV^2. The pull should oscillate around zero for a good fit, the y axis scale on this figure is between +3 and -5, and the oscillation is indeed around zero.

The background PDF is directly modelled by 27 carbon runs taken during the summer period, fed with three extra parameters: an alpha, representing a convolution of the background PDF with the data; an offset, representing any shift of the background PDF with respect to the data; and a scale representing a broadening or narrowing of the peak. An ideal fit should have an alpha and offset of zero, and a scale of 1, and this fit shows each value very close to this. It is also expected that the pull of the histogram (difference between the data points and the curve, shown on the bottom right of Figure 4.14) to be an oscillation around zero, which is again true. From this the signal data can be extracted, as well as an estimate of the nuclear background contribution. The issue with this method is the fact that due to not detecting the scattered electron, the missing mass squared becomes a parameter that is correlated with all other variables. This can be visualised in equation form as follows;

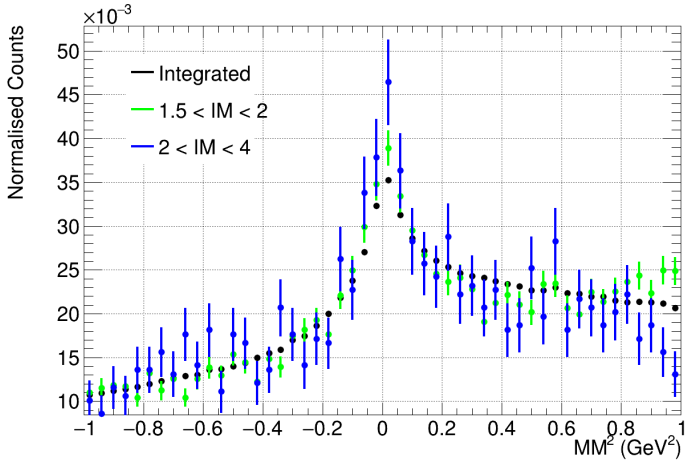
$$MM^2 = ((\vec{p}_e + \vec{p}_p) - (\vec{p_{p'}} + \vec{p_{e^+}} + \vec{p_{e^-}}))^2 \tag{4.1}$$

where the variables are the momenta of the beam electron, target proton, proton in the final state, and the two leptons. In order to calculate the invariant mass of the dilepton pair, we need to use the following equation;

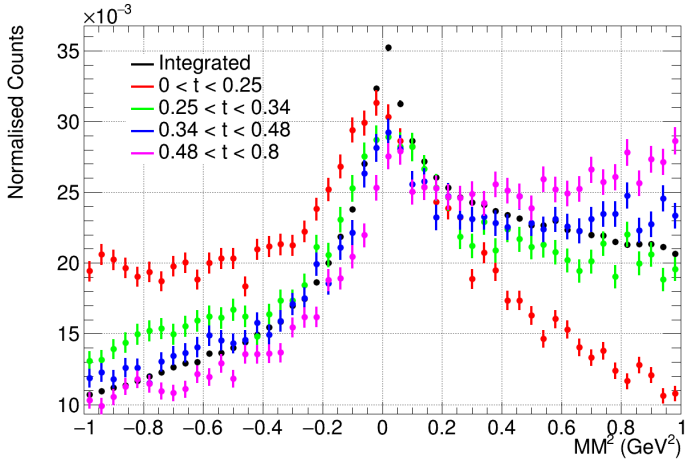
$$M_{inv} = \sqrt{(\vec{p_{e^+}} + \vec{p_{e^-}})^2} \tag{4.2}$$

so it is clear that the two variables are correlated, as information is required on the dilepton pair for both. In this same way, the correlation between MM^2 and t can be illustrated. A visualisation

of these correlations is shown in Figure 4.15, where in Figure 4.15a the MM^2 is plotted in bins of the invariant mass, one at the lower invariant mass region and one at the higher. Correlations of these two variables aren't completely clear, especially in the phase space of interest for TCS at a low MM^2, however when looking at Figure 4.15b there is a very clear dependence of MM^2 on each t bin, most notably between the first and last bins.



(a)



(b)

Figure 4.15: Figures 4.15a and 4.15b show the correlations between MM^2 and IM ($M_{e^+e^-}$), and MM^2 and t respectively.

Two mathematically uncorrelated variables, as an example, would perhaps be assumed to be the invariant mass of the dilepton pair, and the momentum transfer to the struck quark (t). This is because the momentum transfer to the struck quark is calculated using the initial and final state hadrons $t = (p' - p)^2$, and the invariant mass of the dilepton pair is calculated using the final state

leptons. Unfortunately, these two variables cannot be used for an sPlot fit for two reasons, one is that neither have a defined peak or 'bump' for TCS, there is no meson in the final state which the dilepton pair characteristically sum to, and there is no signal bump in the t distribution. The second reason comes from observing the correlation between these two variables in Figure 4.16 where, despite there being no apparent shared variables in the calculation, there is still a slight correlation that the lower the t bin, the lower the invariant mass, this comes directly from the physics of the interaction, since t represents momentum transfer to the struck quark, the photon released by the nucleon in the final state directly produces the lepton pair, so there is an inherent physics dependence despite the surface level maths. Therefore the most accurate way to perform the nuclear background subtraction in the final state observables is with the use of the dilution factor (Method 3). Method 1 however is a good cross-check for the scale of the dilution factor value, it is simply the case that the weights themselves cannot be used for a full background subtraction as in other channels.

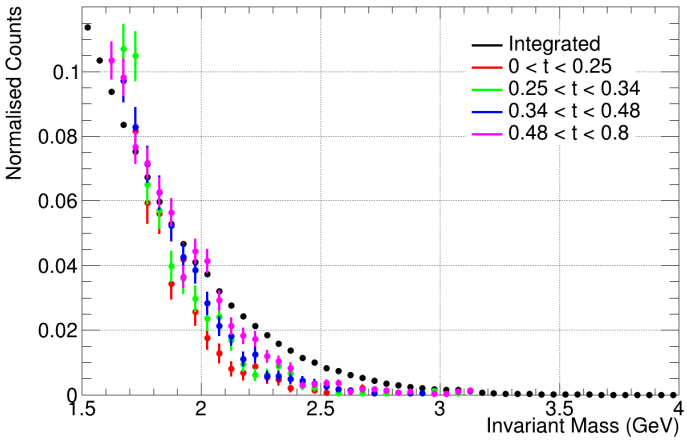


Figure 4.16: The correlation between t and IM ($M_{e^+e^-}$).

Method 2

The sPlot technique can also reveal a method by which data on a nuclear target, and Monte Carlo on a free proton target can be compared. The sPlot fit extracts a number of counts in the 'Signal' peak, which can be directly compared with counts from a Monte Carlo simulation of the same kinematics. More detail on this will be discussed in Section 4.1.3.

Method 3

This method of understanding the background is much more of an analytical exercise and provides a correction factor with which all asymmetries must be corrected, in order to extract the fraction of events that are attributed to the pure proton signal. This method is not affected by

correlations as it uses the physical target parameters alongside kinematic counts normalised by Faraday cup accumulated charge to extract a simple factor, which represents the ratio of signal data to the nuclear background. The equation to calculate the dilution factor is as follows [81]:

$$DF = \frac{9(n_{NH_3}-n_{Empty})\rho_{NH_3}(l_C L(-n_{CH}+n_{Empty})\rho_C+l_C l_{CH}(n_{Foils}-n_{Empty})(\rho_C-\rho_{CH})+l_{CH}L(n_C-n_{Empty})\rho_{CH})}{n_{NH_3}(9l_C L(-n_{CH}+n_{Empty})\rho_{NH_3}\rho_C+2l_{CH}L(n_C-n_{Empty})\rho_{NH_3}\rho_{CH}+l_C l_{CH}(n_{Foils}-n_{Empty})(9\rho_{NH_3}\rho_C-2(\rho_{NH_3}+3\rho_C)\rho_{CH}))} \tag{4.3}$$

This equation, in essence, is the full form of the ratio;

$$DF = 1 - \frac{n_C + n_{Empty} + n_{Foils}}{n_{NH_3}} \tag{4.4}$$

where each n_i is a representation of the number of counts taken on each target type, within the kinematic space of the final state, with dependence on each target density (ρ_i) and length (l_i). It is clear that a good understanding of all targets in use is required to perform this study, and to that end Table 4.3 details each target and its parameters, along with a pictorial representation in Figures 4.17 and 4.18.

Target Material	Density g/cm^3	Length (cm)
NH_3	0.867	5.0
Carbon	1.790	1.5
CH_2-CH_2	0.940	3.0
Empty (no LH_e)	0	5.0
Optical (with LH_e)	0.145	5.0

Table 4.3: Table of each target type and parameters used in Dilution Factor calculation

The carbon and polyethylene (CH_2-CH_2) targets consisted of three disks of material, each approximately 1 g/cm^2 thick and each separated by a liquid helium layer (8.3 mm L_{He} for carbon, 3.5 mm for CH_2-CH_2). Each target has an aluminium cell window and beam window at the 'entry' point of the beam into the target cell, and one cell exit window at the exit point of the beam. A succession of windows follows the beams exit (bath, cavity, pumping tube, heat shield and vacuum), all of which are also made of aluminium. The total thickness of aluminium at either side of the target cell is around 0.15 mm, and when accounting for the multiple exit windows in succession 35 mm from the cell, this extends to a total aluminium thickness of around 0.43 mm. The shell, as in the NH_3 target, was poly-ether ether ketone (PEEK), and the total radiation length (expressed here as what percentage of each material's inherent radiation length is in the target) of each experimental target (non-empty) is made to be equal by design, at approximately 6-7%, with contributions from the liquid helium and aluminium windows amounting to $< 0.5\%$. Empty and optical targets both represent simply the PEEK shell with no material, the difference between them is that for the optical target, the LH_e bath was filled in order to model the background contribution from the cooling system. The NH_3 target consisted of beads of ammonia,

1-4 mm in diameter, packed into the PEEK shell, subsequently filled with liquid helium. The density of ammonia is assumed to be close to that of ammonia recorded at -80° C, which is 0.867 g/cm^3 [85].

Packing fraction

Since the target is made up of variable-sized beads packed into a cylindrical geometry, the packing fraction is non-trivial to calculate accurately. The packing fraction was initially assumed to be 60 %, however, the packing fraction can be calculated from information gleaned when taking data on CH_2 and other background targets. This is due to two reasons, one the physical parameters of the polyethylene target are well known and understood, and two it is possible to work under the assumption that due to the similar binding energies, and the proton to neutron ratio of carbon and nitrogen, the two targets can be likened to each other with a simple ratio factor of 7/6.

In equation form (taken from [86]), this looks as follows;

$$PF = \frac{6l_C\rho_C l_{CH}\rho_{CH}(n_A - n_{MT})}{2l_{CH}\rho_{CH}L\rho_A(n_{MT}-n_C)+9l_C\rho_C L\rho_A(n_{CH}-n_{MT})+l_C l_{CH}(n_{MT}-n_F)(9\rho_A\rho_C-2(\rho_A+3\rho_C)\rho_{CH})} \tag{4.5}$$

which simplifies down to two instances once all known target lengths and densities are substituted. The first scenario is when the full target is taken, i.e. the packing fraction of the ammonia relative to its surroundings, taking the full length of the liquid helium bathtub (5.86cm) into consideration. The second scenario takes the ammonia cell just on its own, which has a length of 5cm, and therefore the result from the first instance is multiplied by a factor of 5.86/5, or, 1.17 [86].

Scenario 1 looks as follows;

$$PF = 0.59713\frac{n_A - n_{MT}}{1.25055n_{CH} - 0.23688n_C - 0.013668n_F - n_{MT}} \tag{4.6}$$

which for the TCS final state counts, gives a packing fraction of 44%.

Scenario 2 looks as follows;

$$PF = 0.69983\frac{n_A - n_{MT}}{1.25055n_{CH} - 0.23688n_C - 0.013668n_F - n_{MT}} \tag{4.7}$$

which for the TCS final state counts, gives a packing fraction of 53%. There must also be an added acknowledgement that due to the fact that the target is oriented on its side, the beads will settle in a given orientation, making the packing fraction and density slightly non-uniform throughout the material. This is assumed to be a small contribution which will not affect the overall value of the dilution factor outside the scope of the quoted error.

The material radiation length of NH_3 is 40.87 g/cm^3, and taking into account the density, mul-

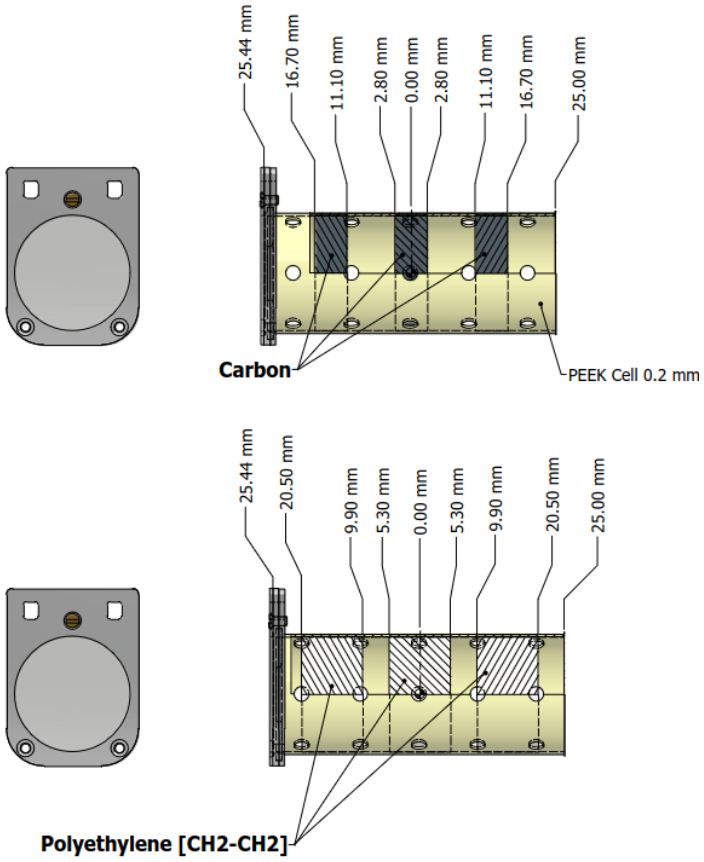


Figure 4.17: A schematic of the carbon and polyethylene targets used in the experiment [87].

tiplied by the packing fraction, and the targets' total length of 5 cm, the percentage radiation length is found as with the other experimental targets to be approximately 6-7%, with the liquid helium contributing approximately 0.05% and the aluminium windows at the edges of the target cell contributing approximately 1%.

The most accurate way to determine the value of n_{NH_3} from the nitrogen itself is to compare rates taken on CH_2-CH_2 and NH_3 targets as with the packing fraction calculation.

For TCS the final state upon which to calculate the dilution factor (and the packing fraction above) was established using the cuts shown in Table 4.5, as is used for the asymmetry extraction. The resulting counts on each target, normalised by Faraday Cup accumulated charge, can be seen in Figure 4.19.
Taking the packing fraction of the full target (Scenario 1), and including this in the calculation of the dilution factor in the following equation (a simplified form of the long-form DF equation in 4.3 but with the inclusion of the packing fraction terms);

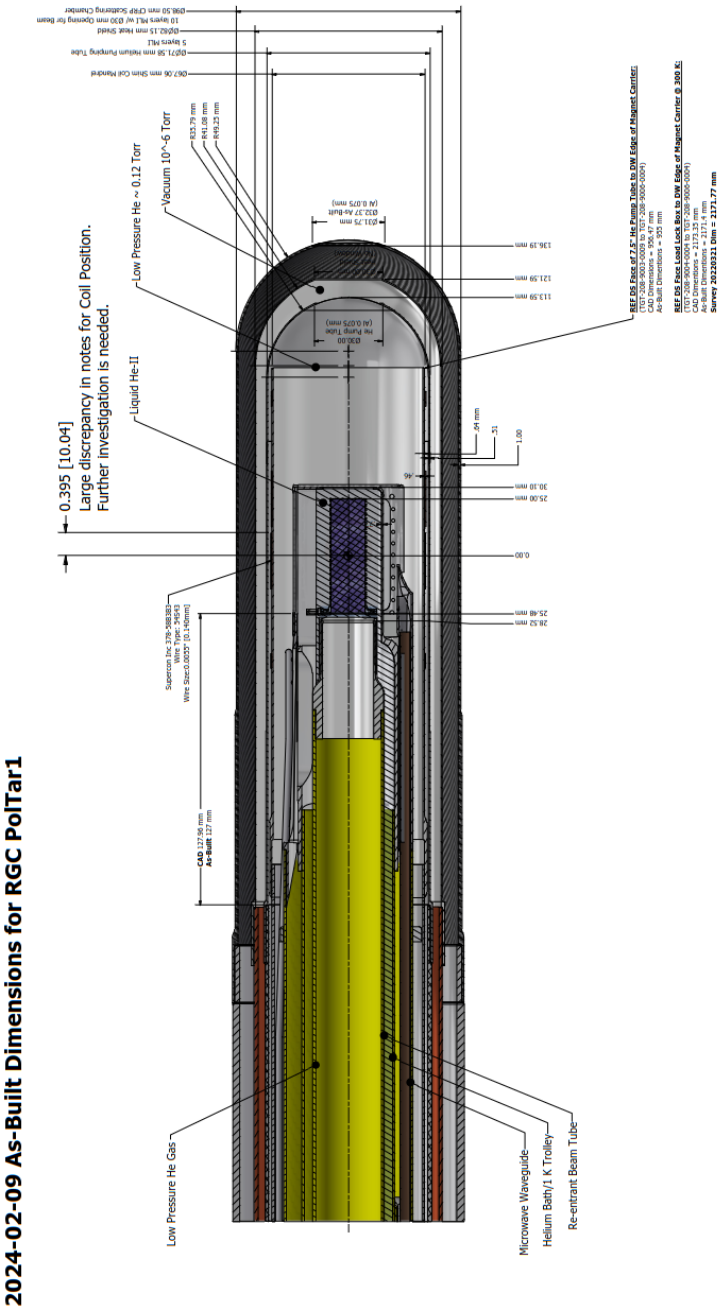


Figure 4.18: A schematic of the NH3 target used in the experiment, and the various positions of entrance and exit windows. Note the final position of the target material was at a z of -30 mm in the bath. All other thicknesses and lengths remained the same [88].

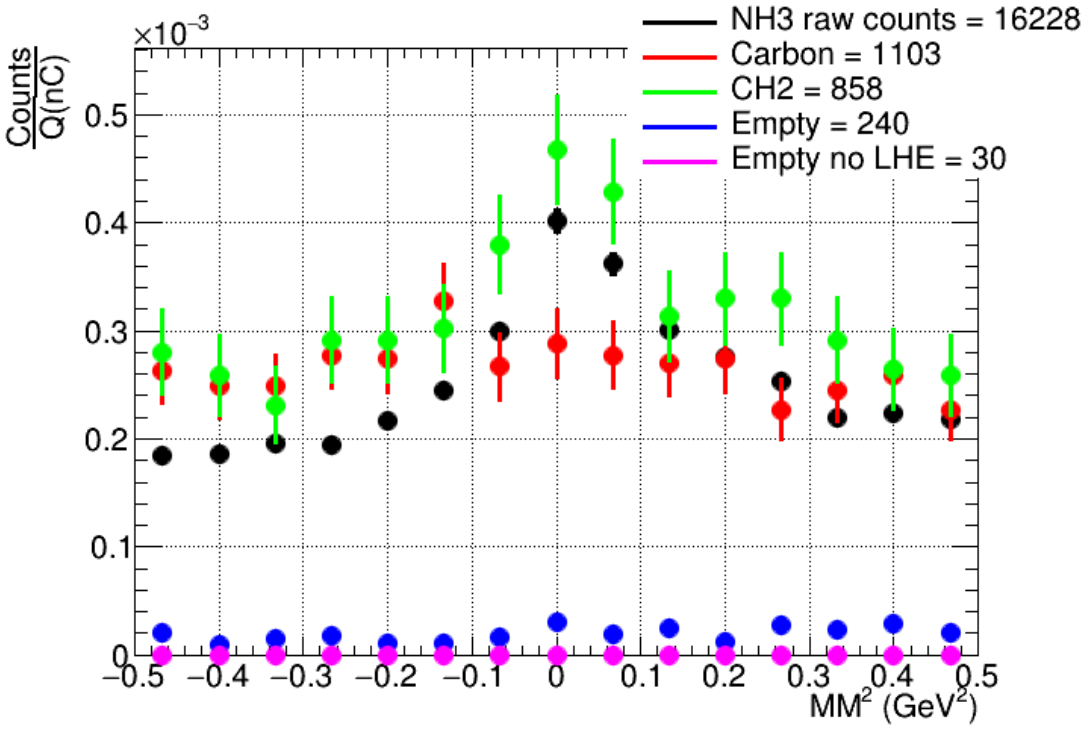


Figure 4.19: The counts taken on each target type in Summer 2022, normalised by FCup accumulated charge.

$$DF = 1.42574\frac{0.44}{n_A}(1.46891(n_{CH} - n_{MT}) + 0.43859(n_{MT} - n_F) - 1.25208(n_C - n_{MT})) \quad (4.8)$$

the result is 33%, which translates to a nuclear background contribution to the dataset of 67%. When taking the packing fraction of just the ammonia cell, of 0.52, the dilution factor becomes 44%. The correct method would be to take the first value as it is the closest representation of the physical target in beam, but it is interesting to note the large difference.
This method, while accurate analytically, has the downside of statistical availability. The counts on figure 4.19 show that there is a significant problem when considering the error on the two empty target runs (blue and magenta in the distribution). Although their contributions are relatively small compared to the contribution from carbon to the background, and the method proposed above takes the normalised counts into the full equation, the final statistical error on the dilution factor is large and would require much more data on empty targets to accurately quantify. Currently the error is taken by assuming Poisson statistics, and taking the square root of the raw counts, multiplied by the Faraday cup normalisation factor, but for a raw count of 30 events that pass cuts in the empty (no L_HE) set, a bin by bin error becomes very difficult to quantify on the same scale as the NH3 data, with 16228 raw counts.

A useful check on the dilution factor value obtained via the equation is to simply compare the NH3 and carbon runs directly, in the clearest signal region. Confidence in this comparison is awarded by looking at the CH_2-CH_2 counts (green) with respect the the NH_3 (black), they agree reasonably well within the previously established ratio factor of approximately 7/6. Looking at figure 4.19 the clear signal region is between a MM^2 of $\pm$ 0.1, and taking the ratio of NH_3 to carbon counts, following the simplified dilution factor form of $1 - \frac{C}{NH_3}$ gives approximately 0.3, which agrees with the analytical value obtained for the dilution factor.

4.1.3 Simulations

Simulations were performed using the unweighted **GRA**ce-based generator for **P**roton-**E**lectron collisions (GRAPE) [89]. This generator was the ideal candidate for simulating background processes in TCS as it contains Feynman diagrams that span the full kinematic space of the produced dilepton pair. In low invariant mass regions, QED Compton-type processes are dominant. At higher invariant masses the dominant background process is two-photon ($\gamma\gamma$) Bethe-Heitler, but there are also contributions from Z^0 production. In GRAPE, the calculation of the cross-section is based on any chosen subset of these processes so that an accurate picture of the background contributions within the invariant mass region of interest can be formed. A 10.6 GeV electron beam was required and a total of 1M events across the full invariant mass distribution was simulated according to the settings in Table 4.4.

Variable	Min	Max
Beam Energy (GeV)	0	10.6
Q^2 (GeV)	0	10
t (GeV^2)	0	7
Proton Theta (degrees)	0	120
Beam Electron Theta (degrees)	0	60
Final State Electron (degrees)	0	60
Final State Positron (degrees)	0	60

Table 4.4: Table of settings used in GRAPE simulations

TCSGen [90] is a weighted event generator, originally developed for the first TCS analysis in the era of the CLAS collaboration before its upgrade to CLAS12 [91]. This generator was used to simulate 1M elastic dilepton events across the full invariant mass range. A second set of events was also produced with a minimum invariant mass of 1.5 GeV, for more statistics in the TCS kinematic phase space of interest and to cut out large contributions from meson resonances below 1.5 GeV. TCSGen simulates pure TCS signal, pure Bethe-Heitler (BH) signal and the interference between TCS and BH. It is expected that in the full simulations, the cross-section contribution from pure TCS and the interference term will be small compared to that from Bethe Heitler, due to its dominance over the TCS signal.

The measured TCS cross-section, as discussed in Section 2.4.2 is, in essence, the sum of the pure TCS signal, the pure BH, the interference term, and any other sources of QED/Compton background. We can therefore see that taking the sum of events from both generators, appropriately normalised, will give a good representation of the free proton signal and dilepton background that can be expected from the dataset. Both simulations were passed through a GEANT4 simulation of the CLAS12 detector, with the same detector configurations, field settings and calibration constants as in the data. The simulated events were then passed through a ROOT macro to extract some base kinematics, without any exclusivity cuts, so that an understanding of how the final state looks could be established.

Figures 4.20a and 4.20b show the MM^2 of the scattered electron for GRAPE and TCSGen respectively. There is a clear, sharp peak at zero as expected, and it is noted that the weighting of TCSGen changes the scale when compared directly to GRAPE, this is dealt with in the normalisation stages in the following section. Figures 4.20c and 4.20d show t on a log scale. Figures 4.20e and 4.20f show the invariant mass of the decay lepton pair, where the key region of interest falls between 1.5 GeV and 3 GeV, therefore, TCSGen has poorer statistics in lower regions of invariant mass, out of the TCS signal phase space, but GRAPE performs well across the full region, due to the fact that it simulates background processes. Normalisation of the two generators

to the data will be explained in the next section.

Normalising Simulated events to data

Normalising the simulations to the existing data was done by calculating the integrated luminosity of the NH_3 data, based on the parameters listed as follows;

- Q = Charge accumulated in the Faraday cup
- e = Electron charge
- l = Length of target
- ρ = Density of target
- N_{NH_3} = Number of target atoms in the target
- N_A = Avogadro's number
- C = Conversion factor from cm^2 to pb^{-1}
- M_{NH_3} = Molar mass of ammonia

entered into the equation for integrated luminosity;

$$L_{INT} = N_{beam} \times n_{NH_3} = \frac{Q}{e} \times \frac{l \times \rho \times N_{NH_3} \times N_A \times C}{M_{NH_3}}. \tag{4.9}$$

The measured density of ammonia at temperatures of 1K differs slightly throughout the literature, for this study it is assumed to be $0.867\,g/cm^3$ which is the recorded value at -80°C, multiplied by a packing fraction of 44% as ascertained in Section 4.1.2. The molar mass of ammonia is 17.03 g/mol, and the conversion factor from cm^2 to pb^{-1} is 10^{36}. The number of target atoms in the target is taken as 3, as the measurement is done on a free proton. The length of the target is 5 cm, and the charge accumulated in the Faraday cup is taken from calculating the sum of the events recorded in $FCup^+$ and $FCup^-$, and has a value of 4.78 mC. Using all of these values, the integrated luminosity of the NH_3 data is calculated as $1478.96\,pb^{-1}$.
Normalising GRAPE is simple as it is an unweighted event generator, so each histogram is simply scaled by the normalisation factor as follows:

$$W_{GRAPE} = \frac{L_{INT}(= 1478.96) \times \sigma_{GRAPE}}{N_{gen}} \tag{4.10}$$

but for TCSGen it is a weighted event generator, so normalisation is done on an event-by-event basis, multiplying by the normalisation factor and then by an additional weight factor calculated using the flux and intrinsic weight of the generator as follows;

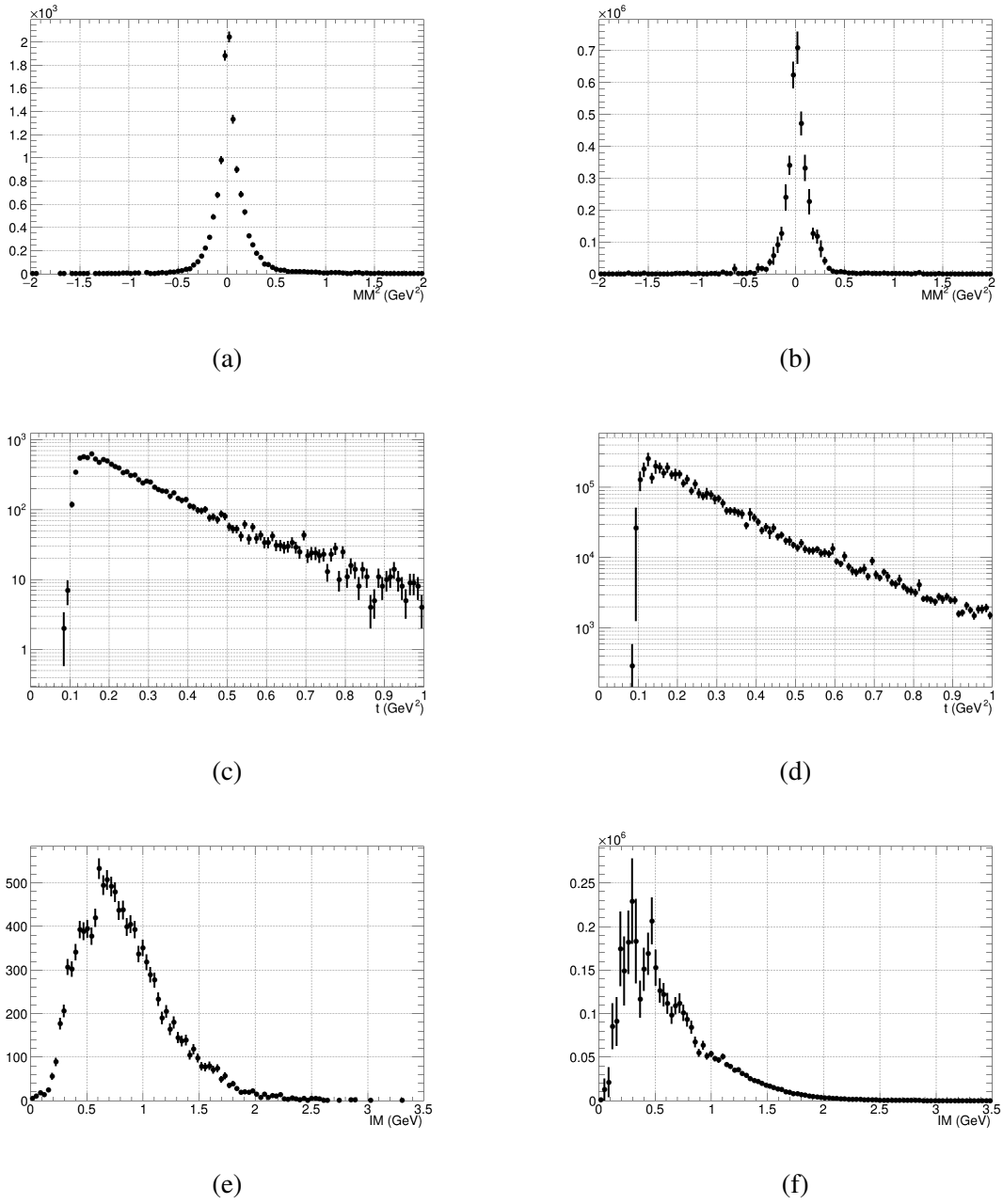


Figure 4.20:
Figures 4.20a and 4.20b show the counts in MM^2 for GRAPE and TCSGen respectively, Figures 4.20c and 4.20d show the counts in momentum transfer to the struck parton, t, and Figures 4.20a and 4.20f show the counts in invariant mass of the decay lepton pair (IM) all without any exclusivity cuts

$$W_{\text{TCSGen}} = L_{INT}(= 1478.96) \times \frac{p_{\text{beam}} \times e_{\text{beam}} \times \omega_{\text{gen}}}{N_{\text{gen}}} \tag{4.11}$$

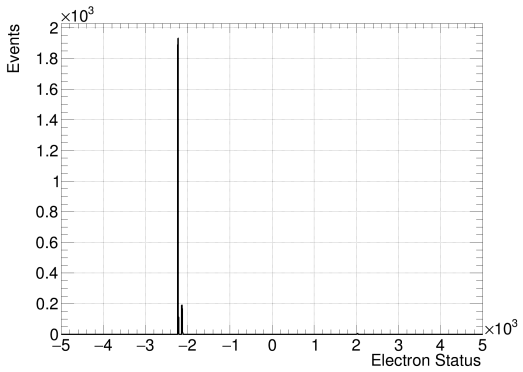
The normalisation is calculated such that the BH contribution represented by TCSGen takes into account only the contribution from the real photon flux, this is to avoid double counting the BH cross-section when summing TCSGen and GRAPE events together since GRAPE calculates cross-sections based on the quasi-real (virtual) photon flux. The real photon flux used in the TCSGen normalisation is calculated using [92];

$$\text{flux} = \frac{1}{2}\frac{d}{X_0}\frac{1}{E_\gamma}\frac{4}{3} - \frac{4}{3}\frac{E_\gamma}{E_{\text{beam}}} + \frac{E_\gamma^2}{E_{\text{beam}}^2} \tag{4.12}$$

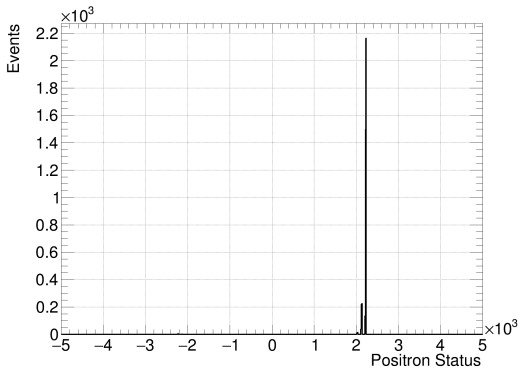
where d is the physical target length, X_0 is the radiation length of the target, E_γ is the energy of the photon, and E_{beam} is the energy of the beam. In a single proton target such as liquid hydrogen, the real photon flux contributes only about 10-15% when compared to the BH contribution modelled by GRAPE. Since this data was taken on a nuclear target, however, the contribution from the real photon flux is much larger, due to bremsstrahlung photons produced by the electron beam interacting with the nitrogen in the target.

4.2 Establishing Exclusivity Cuts

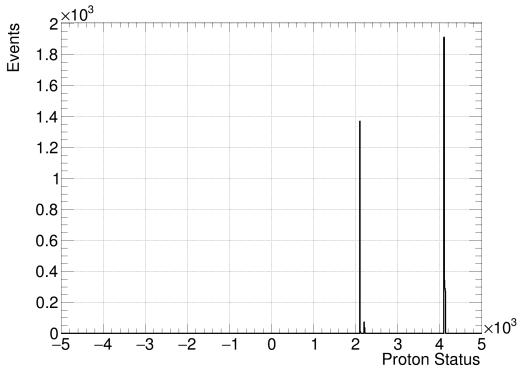
In this section, the exclusivity cuts for the final selection of the dataset will be discussed. As a reminder, the NH_3 dataset in this thesis consists of 157 runs, taken in the run of Summer 2022, and the final state the study seeks to identify is the dilepton pair e^+e^- and the recoil proton, leaving the scattered electron undetected, and reconstructed through the missing mass squared (MM^2). This dataset has been fully calibrated and processed for analysis. The first exclusivity cut in place is a cut on the 'Status' of the dilepton pair. The 'Status' refers to the detector subsystem the hit was registered in, and the cuts detailed ensure that the decay electron and decay positron are in the Forward Detector. This is around a Status of 2000, the Central Detector is at 4000 and the Forward Tagger is at 1000.



(a)



(b)



(c)

Figure 4.21: Status of each final state particle, 4.21a shows the electron in the FD, 4.21b shows the positron in the FD and 4.21c shows the proton in the FD and CD.

The absolute value of status was taken to ensure that the particle is selected even if it is a trigger particle (which is the first particle used to base the hit event on in the trigger) as for trigger particles, the value of Status is negative. The proton is studied in two different topologies, the forward and the central. Figure 4.21 shows the status of each final state particle. The electron and positron are shown to be in the forward detector (Status = $\approx$ 2000 in Figures 4.21a and 4.21b), note the electron is shown as a negative status due to it being the trigger particle. The final state proton shows two distinct peaks, one at around 2000, and one at around 4000, see Figure 4.21c, indicating that the final state can have a proton in either the forward or central detector regions.

The second exclusivity cut is a cut requiring the initial photon is quasi-real (very closely approximated to real, i.e a low Q^2). This requirement is in place to separate the TCS final state from a wealth of high Q^2 processes. To implement this cut, there are two methods. The first method involves calculating the momentum difference between the beam electron and the scattered elec-

tron, and placing a cut on this value being small, which requires a close-to real photon. The second method involves taking the ratio of the transverse momentum of the scattered electron to its three-momentum. This term is known as the transverse momentum fraction and is again expected to be very close to zero for quasi-real photon transfer. This is because the scattered electron is expected to be produced at very low angles, close to the beamline, or 'far forward', so by applying the cut on transverse momentum transfer it greatly reduces the electroproduction background.

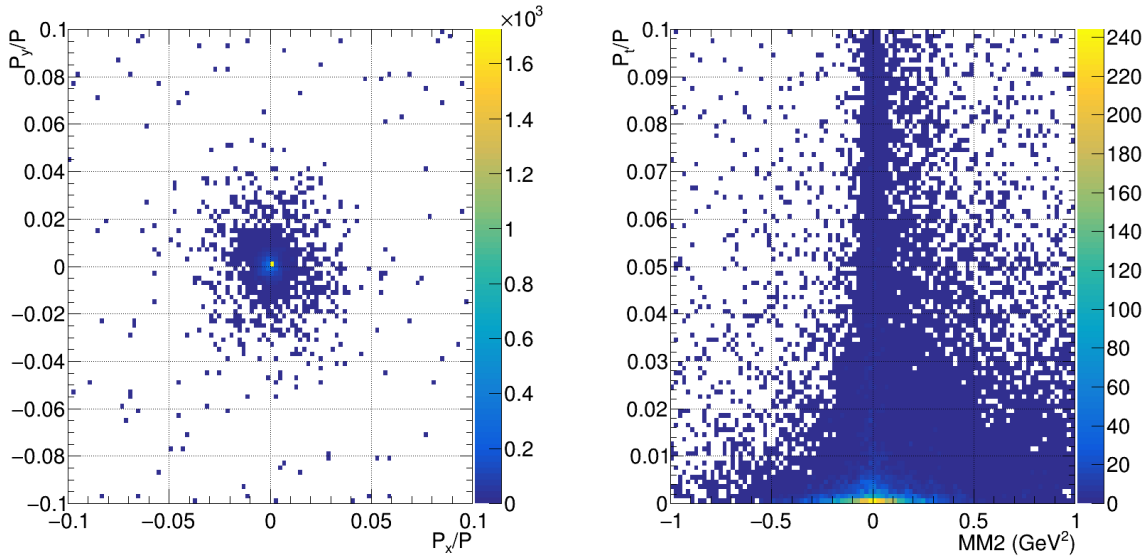


Figure 4.22: The transverse momentum fraction of the scattered electron, for GRAPE and TCSGen (left), and its comparison with MM^2 (right).

Looking at the TCSGen and GRAPE simulated events in Figure 4.22, it is clear that the transverse momentum fraction of the scattered electron is very close to zero. Looking at the data in Figure 4.23 the value is much more smeared out. This is due to a combination of a nuclear target, and the fact that the scattered electron is not directly detected in this channel, so resolution effects from all subsystems smear out the distribution. To allow for this smearing but also ensure that Pt/P is close to zero and encapsulates the clear central region shown by the simulated events, the cut is chosen to be $Pt/P < 0.05$.

The next exclusivity cut is on the value of the missing mass squared of the scattered electron, MM^2. This cut is made to ensure that the four-momentum of each detected final state particle, when summed, is equal to the four-momentum of the initial state particles, with a'missing' component pertaining to the mass of the scattered electron. Since the electron mass is around 0.511 MeV, on the GeV scale, it is expected that this value will be close to zero. A key thing

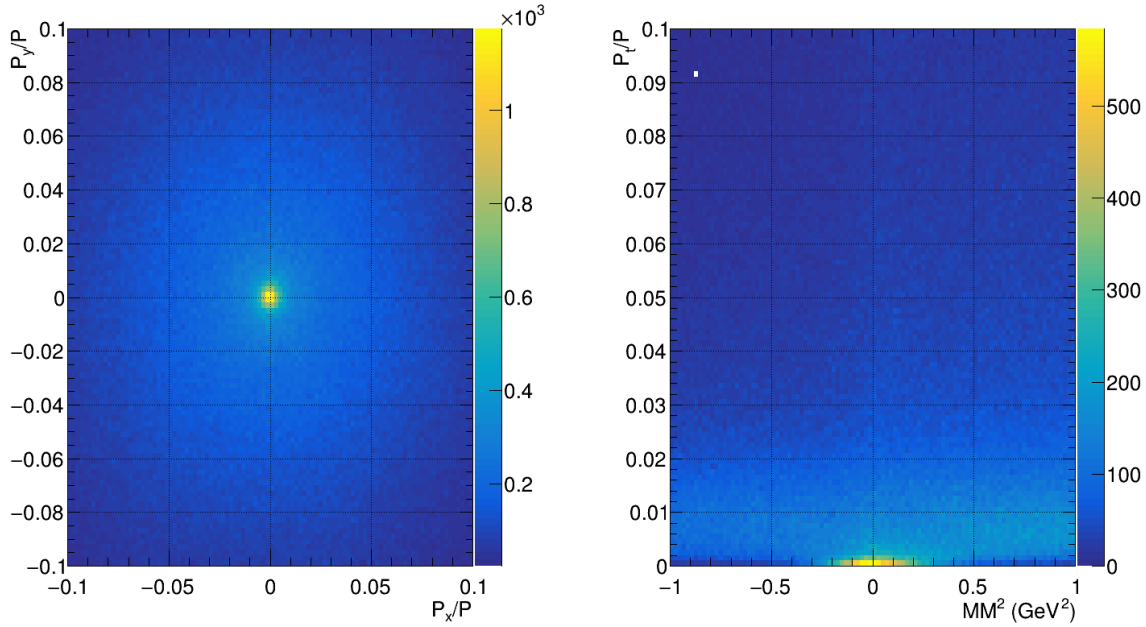


Figure 4.23: The transverse momentum fraction of the scattered electron, for NH_3 data (left), and its comparison with MM^2 (right).

to note about the calculation of MM^2 is that it includes initial and final state particles, making it sensitive to the effects of Fermi motion. Figure 4.24 shows the MM^2 without any exclusivity cuts (left) and with loose exclusivity cuts (right) applied. These loose exclusivity cuts are the previously established cuts on the dilepton pair being in the forward detector, exactly one of each particle in the final state and the cut on Pt/P < 0.05. On the left distribution, there is a clear peak at zero, with a large amount of background to the right. This background is attributed to channels which have a non-zero missing mass, as well as nuclear effects from the target itself. The right distribution, after the cuts, shows a reduction in both background and signal. The loss of the signal peak is attributed to losing those events that do not have a quasi real photon in the initial state, but still have a zero or close to zero missing mass value. It is clear looking at this distribution and referring back to Figures 4.20b and 4.20a in Section 4.1.3 that a sensible cut to remove background but that also accounts for detector resolution effects is around $-0.4\text{GeV}^2 < MM^2 < 0.4\text{GeV}^2$.

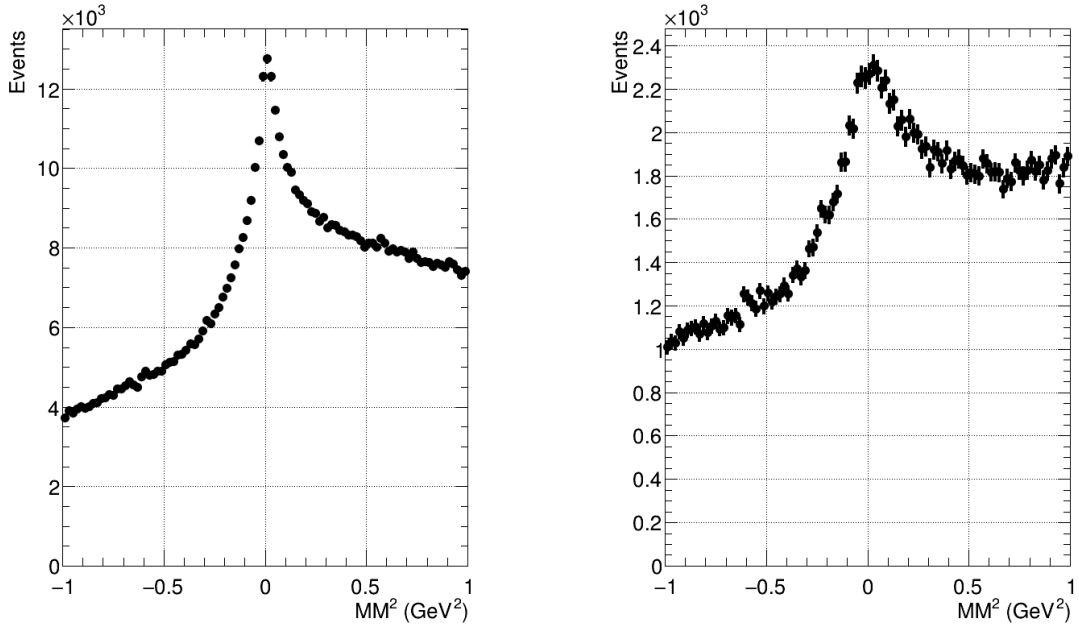


Figure 4.24: The Missing Mass squared of the scattered electron, for NH_3 data, with no exclusivity (left) and loose exclusivity cuts as defined in text applied (right).

The next exclusivity cut established was one on the invariant mass of the dilepton pair. This cut is required to remove events that pertain to meson production. There are a few peaks indicative of known meson resonances visible in Figure 4.25 particularly at around 770 MeV (0.77 GeV) there is a small bump associated with ω and ρ resonances with invariant masses 770 MeV and 782 MeV respectively. Just after the 1 GeV mark, there is a small peak corresponding to the invariant mass of the ϕ meson, at 1020 MeV. The peak at zero in Figure 4.25 is due to a combination of Dalitz decays ($\pi^0 \to \gamma e^+e^-$) and photon conversion to a lepton pair ($\gamma \to e^+e^-$) in the target material. This peak is magnified with respect to previous TCS results as we have a nitrogen target, hence more bremsstrahlung photons produced by the electron beam, and therefore more opportunities for decays to a lepton pair. Finally, a peak at around 3.1 GeV is due to the J/ψ meson, which is a well-known resonance in the dilepton channel. The cut on the invariant mass of the dilepton pair is chosen to be $1.5\,\text{GeV} < IM < 3\,\text{GeV}$, to remove the meson resonances. With a larger dataset, this cut could be tightened to a lower limit of 2 GeV, as further meson resonances exist between the ϕ meson and 2 GeV, namely the $\rho(1700)$, however this is not expected to have an impact on the asymmetries measured, and with the current statistics, a cut this tight would remove too many events, so for this measurement the cut is left at 1.5 GeV. From this point forward the loose exclusivity cuts now include the addition of the newly established invariant mass cut between 1.5 GeV and 3 GeV.

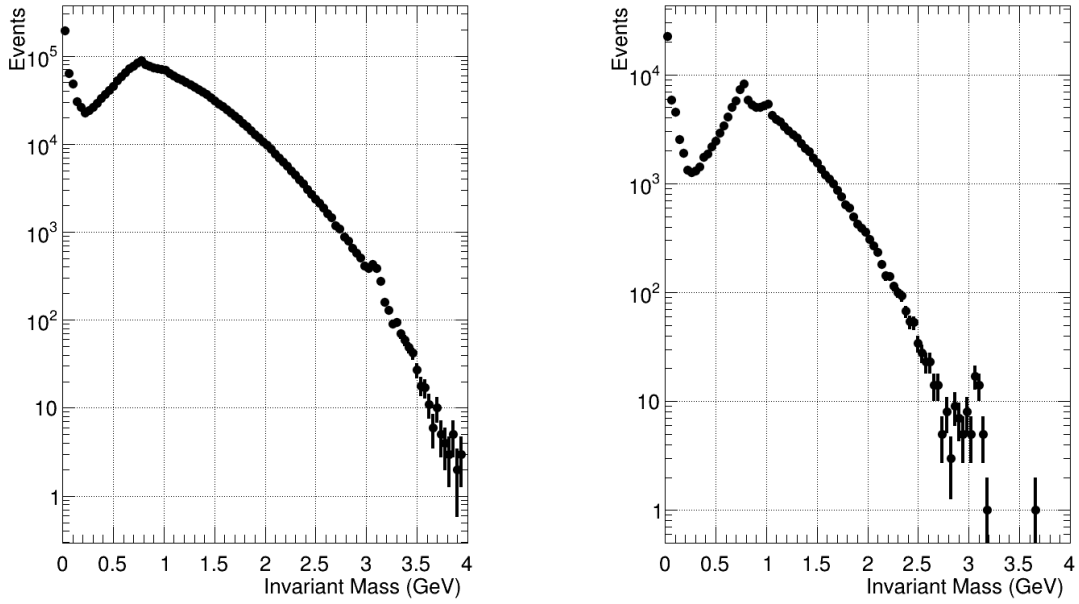


Figure 4.25: Invariant mass of the dilepton pair, for NH_3 data, shown on a log scale, with no exclusivity (left) and loose exclusivity cuts (right) applied.

Two more exclusivity cuts were placed on the events, one on the momentum transfer to the struck parton, t and the other on the energy of the incoming photon E_γ. Looking firstly at Figure 4.26, which shows the t distributions, most of the events fall in the low t region and the events above about $0.8\,\mathrm{GeV}^2$ show a steady fall towards zero. The upper limit on t ensures a hard scale to the scattering process, defined by $t \gg Q'^2$, while keeping to leading order in α_S and leading twist (the TCS amplitude is suppressed on the Q' scale by powers of (n-2), where n is the twist. At leading twist, n = 2 so there is no suppression and is the simplest case. [93], [94], [32]) and sacrificing as few statistics as possible.

A cut on the centre of mass energy of the γp system ($s = (p+q)^2 > 4\,\mathrm{GeV}$) ensures that the proton resonance region is cut out, while also ensuring the incoming photon is high energy, enough to probe a single parton inside the target nucleon. An additional requirement is made on the quasi-real photon energy E_γ to ensure the process is hard scale scattering. When looking at Figure 4.27, on the distribution without exclusivity cuts there is a large peak between around 1 GeV and 4 GeV. This peak is attributed to processes with a low incoming photon energy and hence a low dilepton invariant mass since when the cut on $IM > 1.5\,\mathrm{GeV}$ is applied, this structure disappears. It is clear looking at the right of Figure 4.27 that the previous cuts already set a data-driven cut range of $4\,\mathrm{GeV} < E_\gamma < 10.6\,\mathrm{GeV}$ due to the direct relationship between s and E_γ, and with the upper limit being set by the maximum beam energy provided by CEBAF to Hall B.

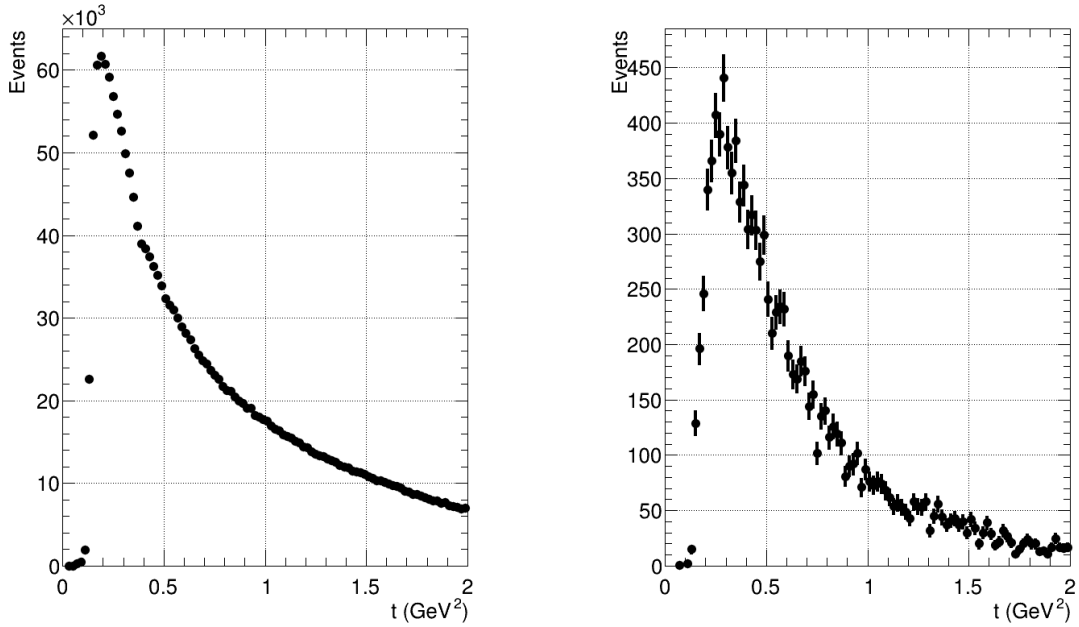


Figure 4.26: Momentum transfer to the struck parton, t, for NH_3 data, with no exclusivity (left) and loose exclusivity cuts (right) applied.

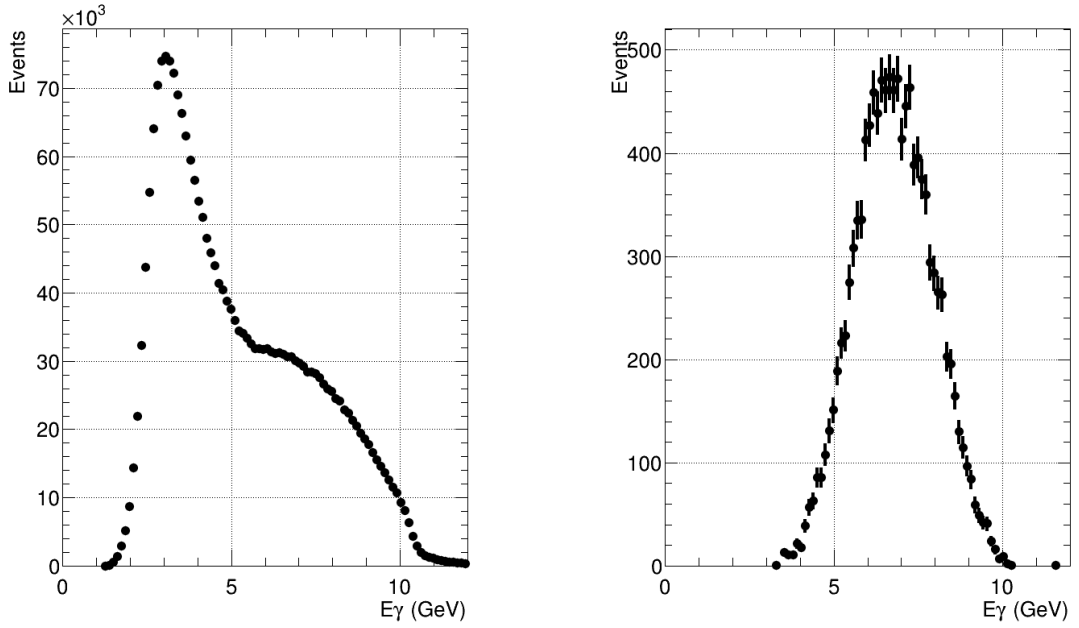


Figure 4.27: Energy of the incoming photon, E_γ, for NH_3 data, with no exclusivity (left) and loose exclusivity cuts (right) applied.

Collating all of the above studies, the exclusivity cuts for the final analysis are as follows;

Variable	Min	Max
MM^2 (GeV^2)	-0.4	0.4
-t (GeV^2)	0	0.8
IM (GeV)	1.5	3
E_γ (GeV)	4	10.6
Pt/P (GeV)	< 0.05	
|Final State Electron Status|	2000	
|Final State Positron Status|	2000	

Table 4.5: Table of exclusivity cuts that were used in the final analysis

Figures 4.28 - 4.30 show the momentum vs. θ distribution of each of the three final-state particles during the process of establishing full exclusivity, and 4.31 shows the same distributions after the full final state exclusivity has been applied.

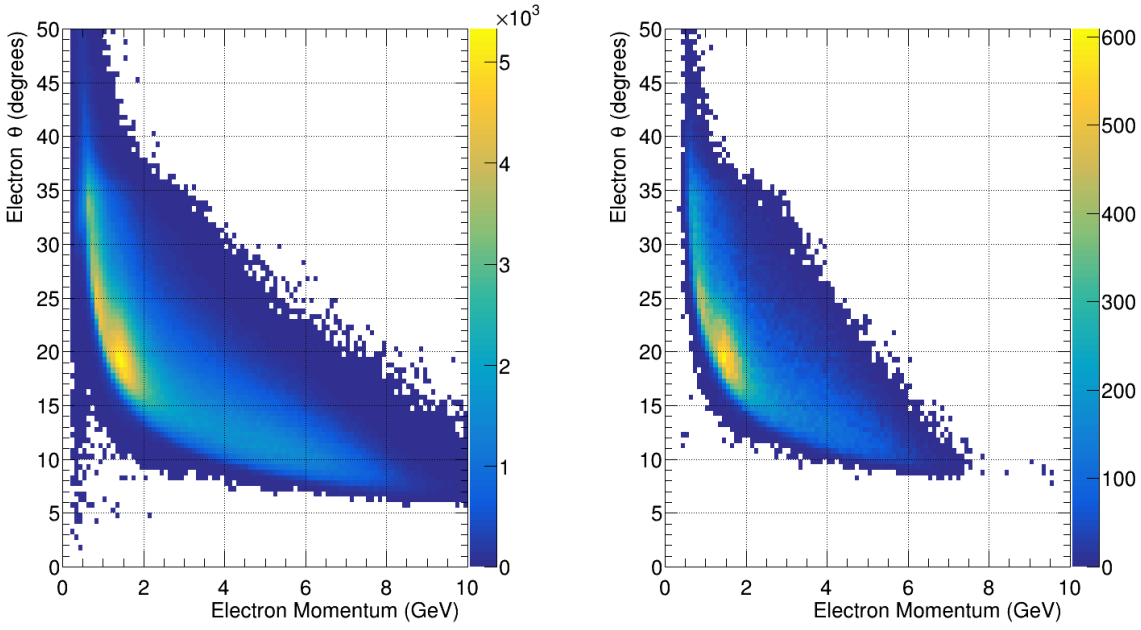


Figure 4.28: The momentum vs polar angle of the electron in the final state, on the left without any exclusivity, on the right with loose exclusivity cuts.

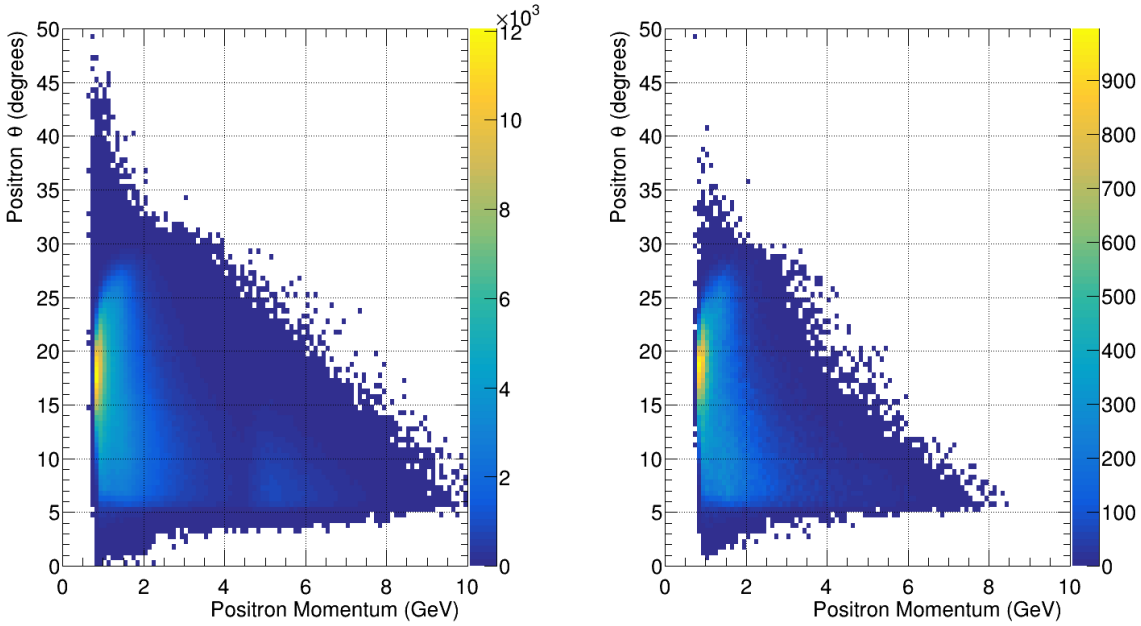


Figure 4.29: The momentum vs polar angle of the positron in the final state, on the left without any exclusivity, on the right with loose exclusivity cuts.

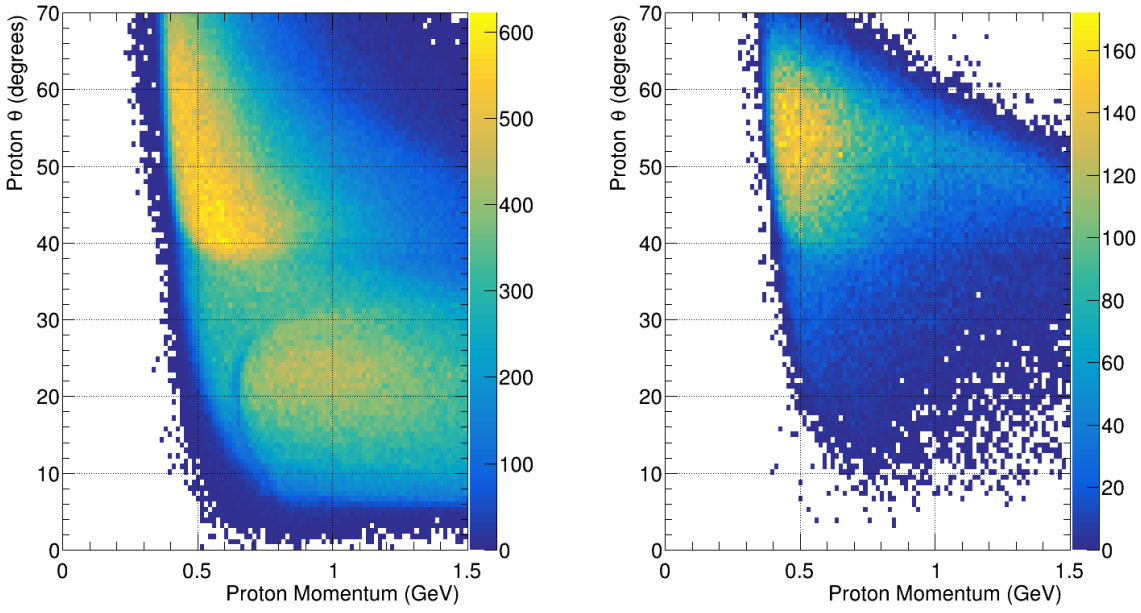
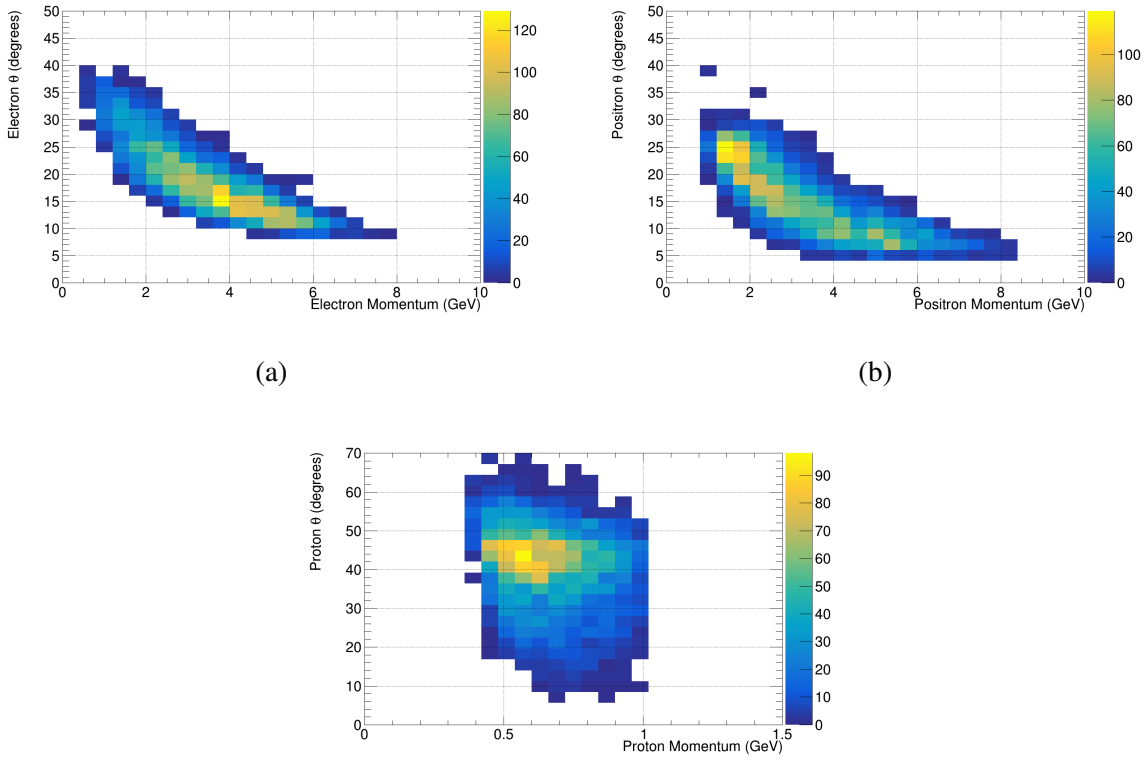


Figure 4.30: The momentum vs polar angle of the proton in the final state, on the left without any exclusivity, on the right with loose exclusivity cuts.



(a) (b)

(c)

Figure 4.31: Each of the three final state kinematic momentum vs theta distributions, after the full exclusivity cuts have been applied.

What is clear from these distributions is that the electron-positron pair are produced at low angles, with high momentum in the forward region, the location of the peak for each particle is opposite due to the opposing charge of each particle in the magnetic field. This supports the expectation that the final-state particles should be detected in the forward detector. Looking at the proton distribution, and recalling the CD/FD threshold lies at approximately 35° it is clear that the final state proton lies across this region, and scatters with a low momentum. This highlights the need for a possible two topology treatment of the final state asymmetries, one measurement in the FD and one in the CD. This can also be achieved by splitting the invariant mass cut into two regions, one for the forward proton (selecting a higher invariant mass) and one for the central proton (selecting a lower invariant mass). This relationship can be seen in Figure 4.32. The above treatment was attempted in the analysis, however, it was expected based on Figure 4.32 that the statistics would not be high enough in the FD to allow for asymmetry extraction in both regions separately.

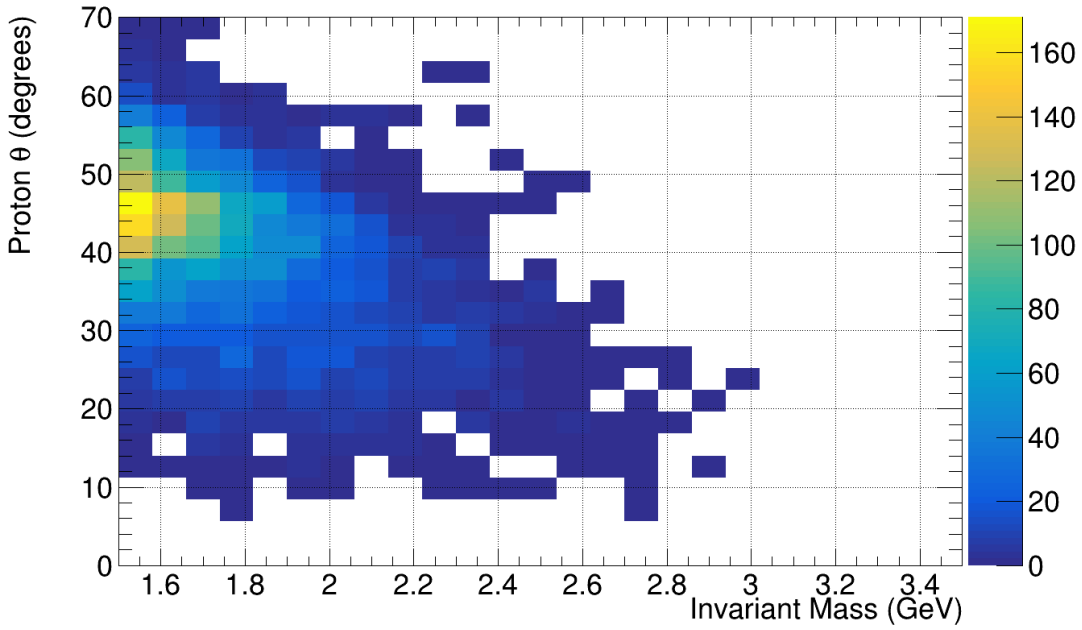


Figure 4.32: Relationship between final state $IM_{e^+e^-}$ and final state proton θ

Simulated Final State Kinematics and Comparison to data

After applying the cuts established for the dataset, including all fiducial cuts and radiative corrections, the simulated events were then normalised to the data using the integrated luminosity of the NH_3 data, as discussed in Section 4.1.3. The normalisation factors were then applied to each set of simulated events, and the final-state observables were extracted. The results can be seen in Figures 4.33 - 4.38b, where on the left is shown TCSGen and GRAPE separately, and on the right the combined distribution and counts. Note at this stage the TCSGen events used were produced with a minimum invariant mass of 1.5 GeV, to ensure a good representation of the TCS signal in the final-state observables. The GRAPE events remained across the full invariant mass range, to give a good representation of the background contributions in the whole phase space, this is the reason for the difference in statistical error.

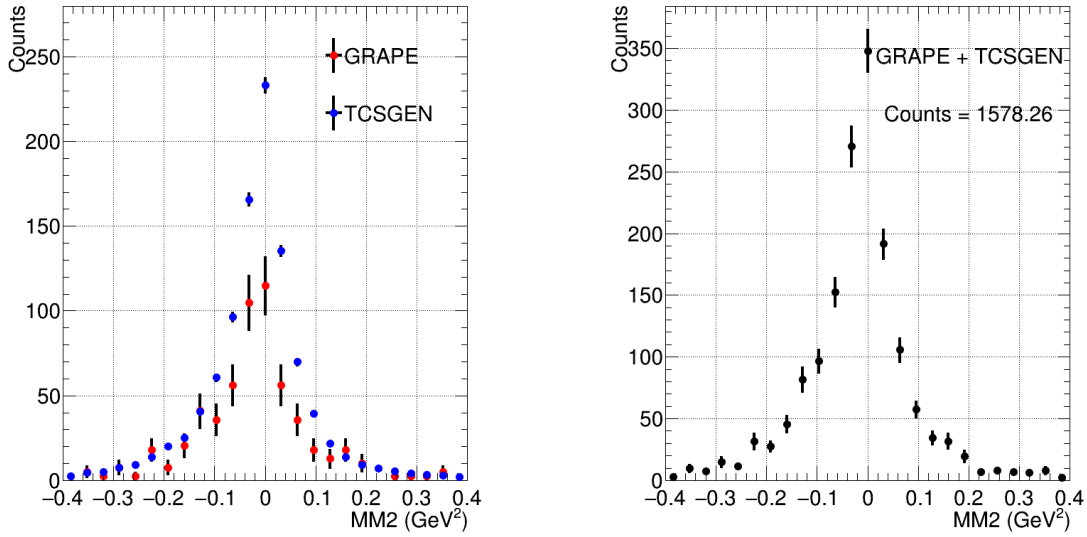


Figure 4.33: The simulated Missing Mass squared of the scattered electron in the final state, with all exclusivity cuts, on the left showing the split between TCSGen and GRAPE, and on the right the combined distribution.

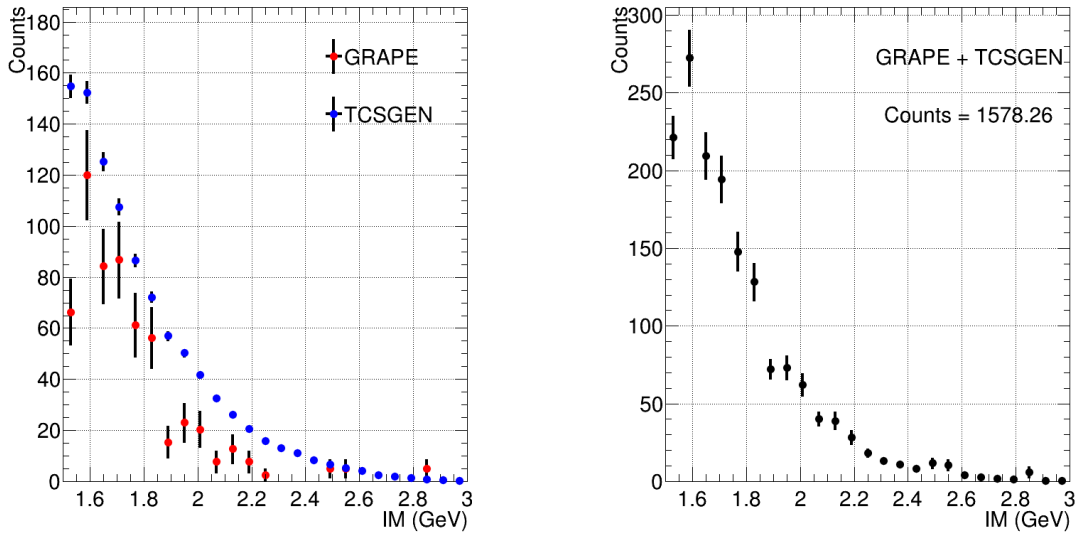


Figure 4.34: The simulated invariant mass of the dilepton pair in the final state, with all exclusivity cuts, on the left showing the split between TCSGen and GRAPE, and on the right the combined distribution.

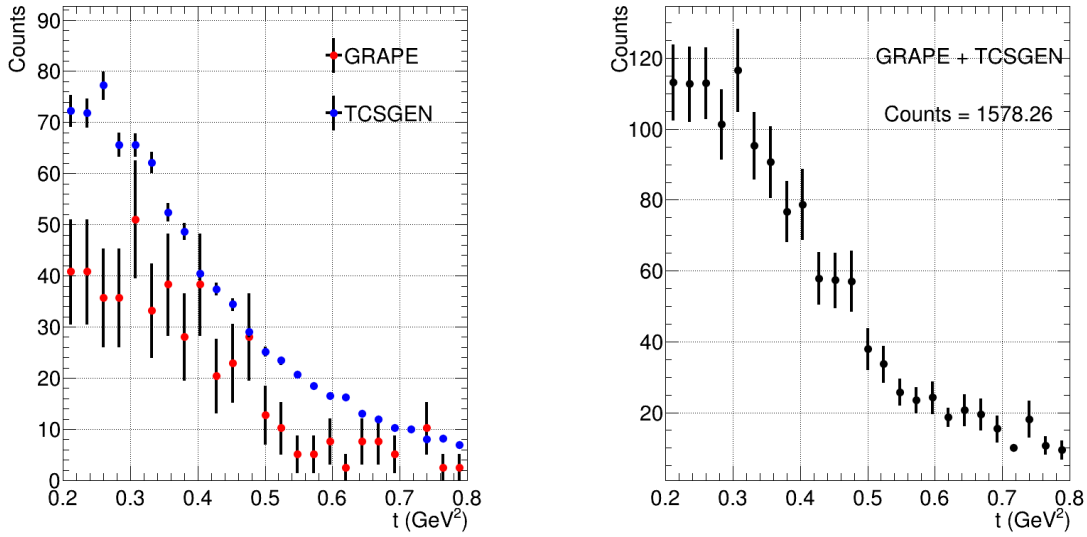
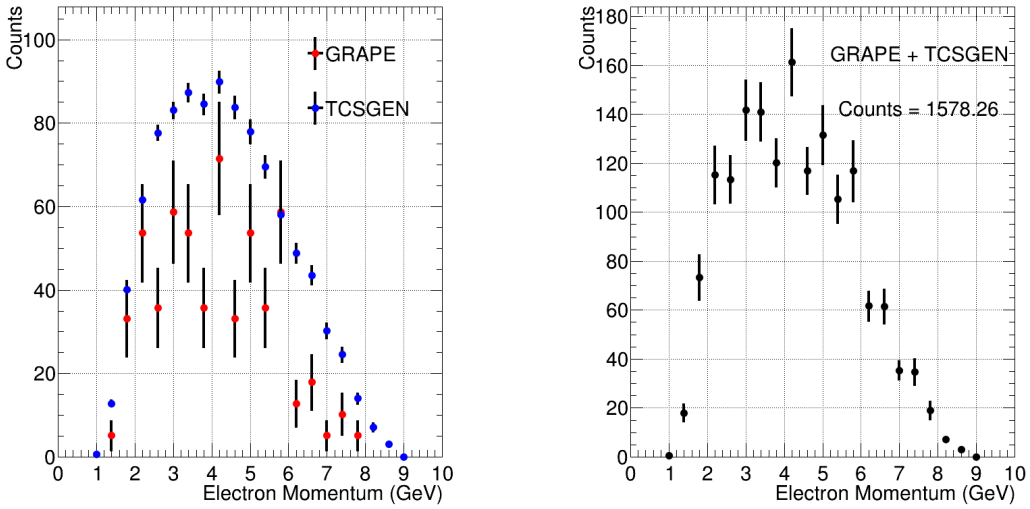
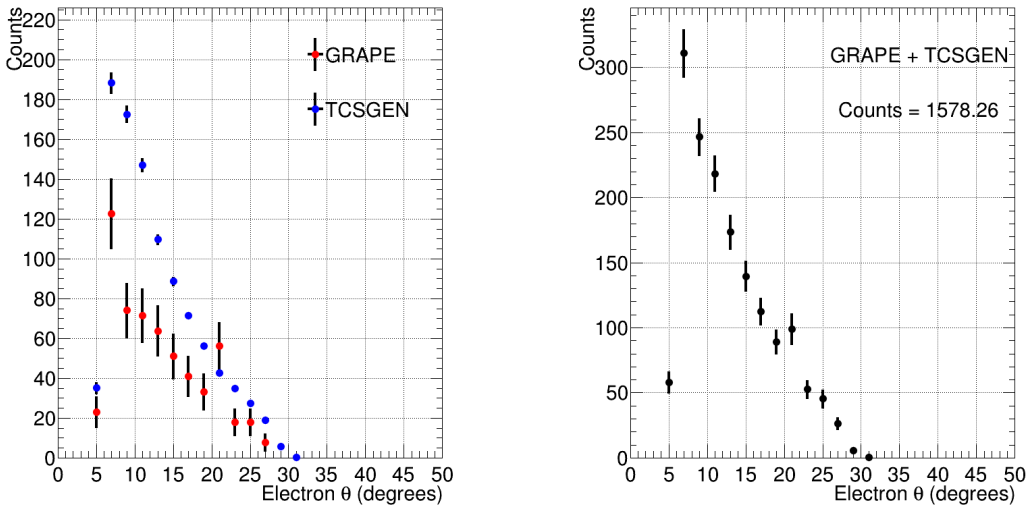


Figure 4.35: The simulated momentum transfer to the struck quark in the final state, with all exclusivity cuts, on the left showing the split between TCSGen and GRAPE, and on the right the combined distribution.

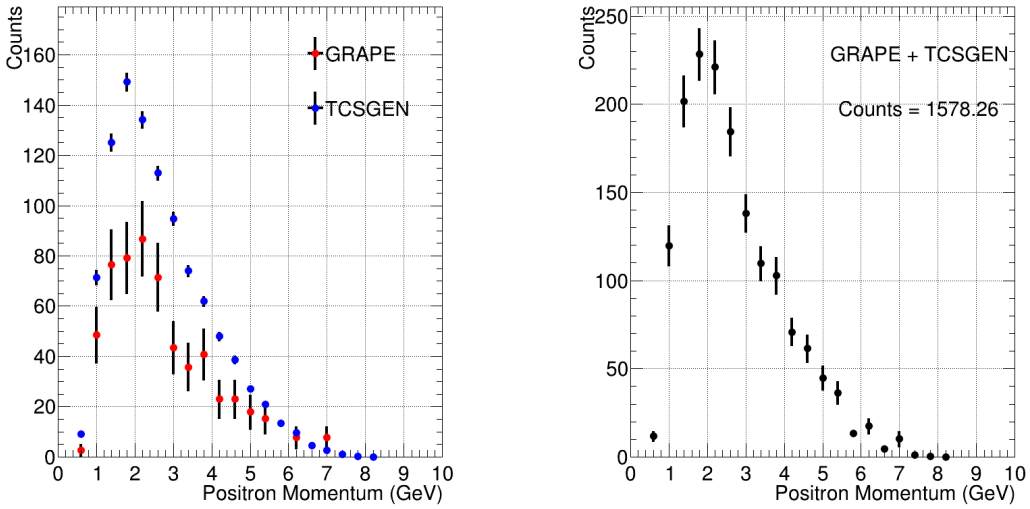


(a)

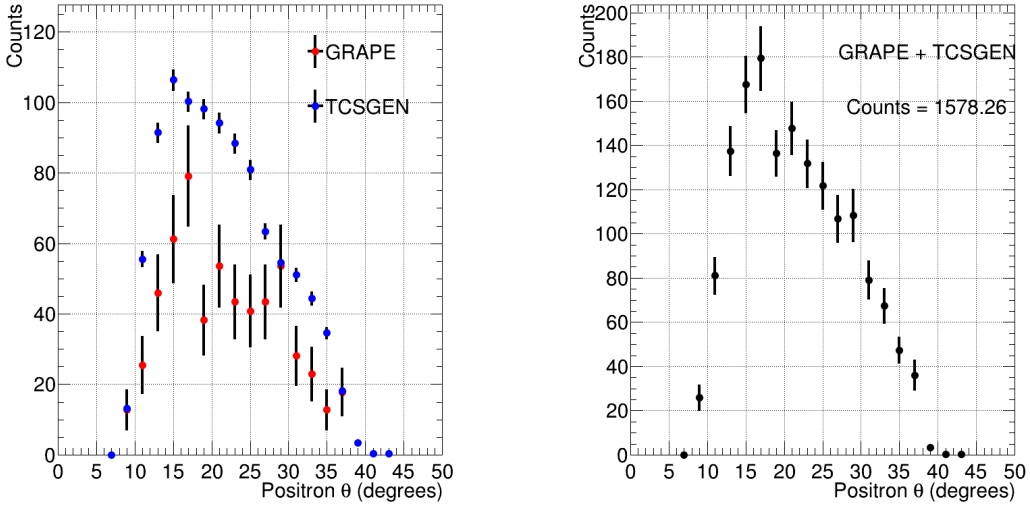


(b)

Figure 4.36: The simulated momentum and polar angle of the electron in the final state, with all exclusivity cuts, on the left showing the split between TCSGen and GRAPE, and on the right the combined distribution.

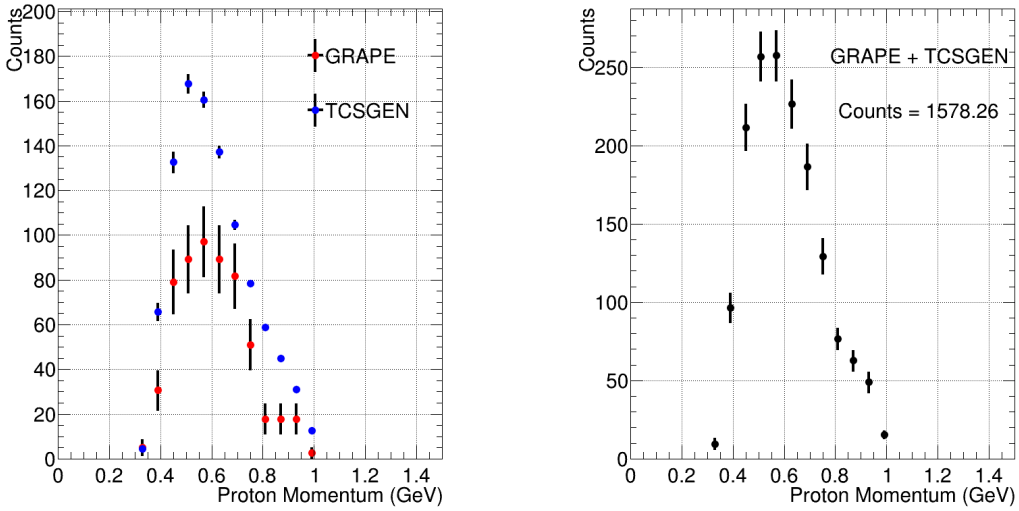


(a)

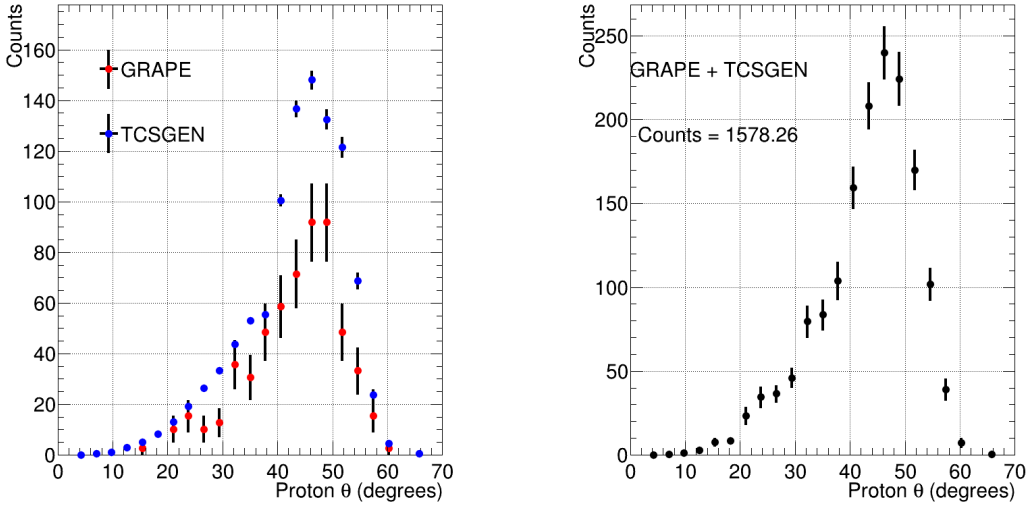


(b)

Figure 4.37: The simulated momentum and polar angle of the positron in the final state, with all exclusivity cuts, on the left showing the split between TCSGen and GRAPE, and on the right the combined distribution.



(a)

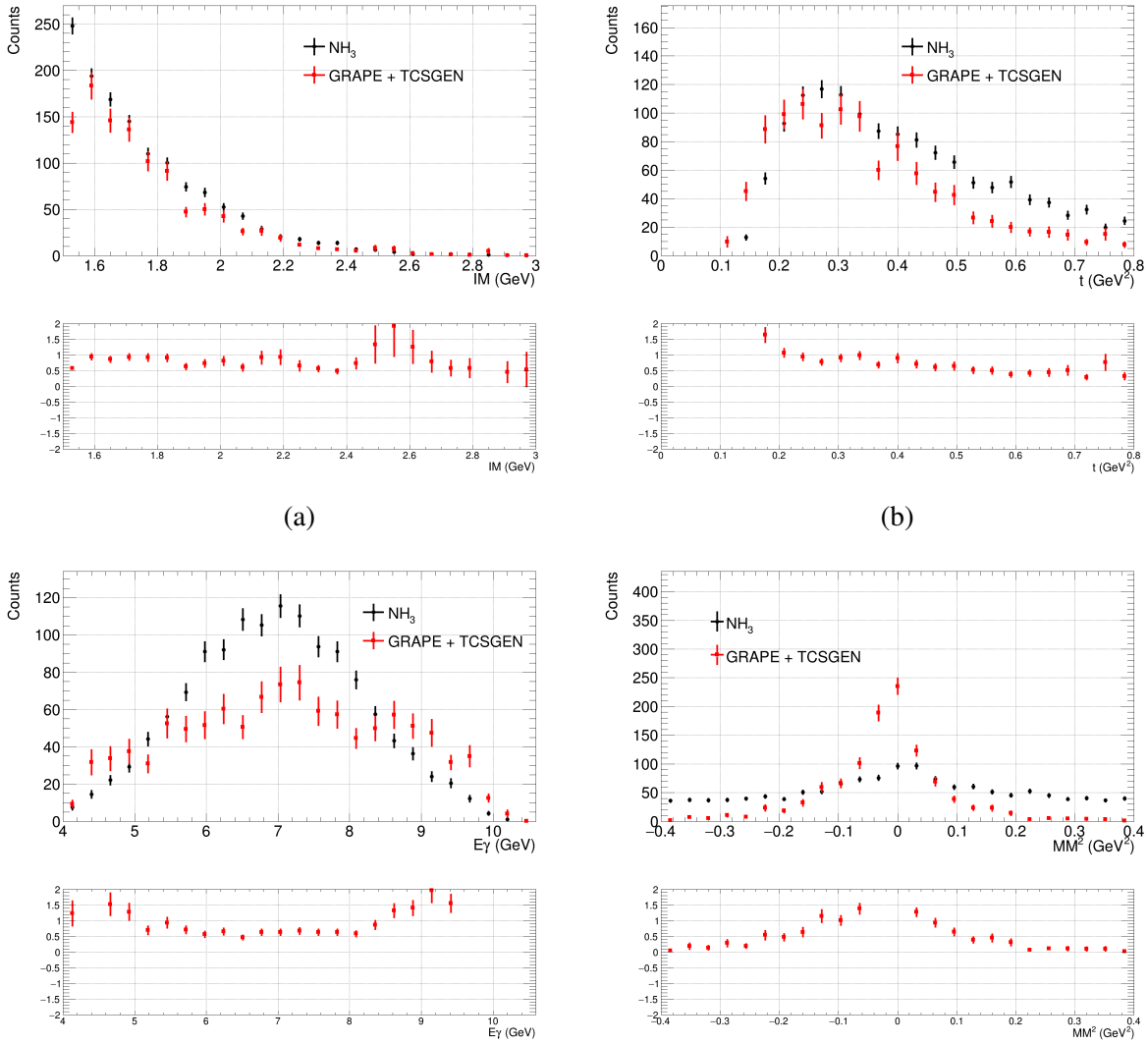


(b)

Figure 4.38: The simulated momentum and polar angle of the proton in the final state, with all exclusivity cuts, on the left showing the split between TCSGen and GRAPE, and on the right the combined distribution.

Using each of these distributions a comparison to the dataset was performed, initially by normalising the signal to the data as above and then scaling the data by the Dilution Factor calculated as 30% in Section 4.1.2. It is clear looking at the ratio of simulated events to data (below each corresponding figure), that there is fair agreement between simulation and data for t and IM, with an amplitude discrepancy for $E\gamma$ and a sharply deteriorating agreement when comparing

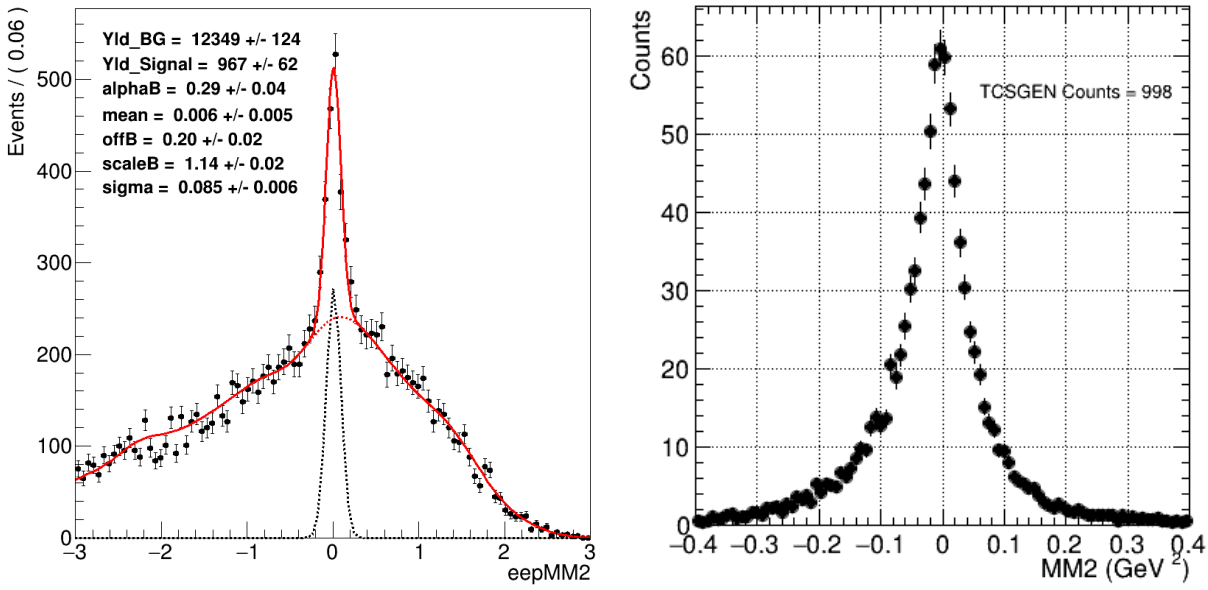
the MM^2 distributions. This is due to the fact that the simulations do not consider Fermi motion, which is clearly present in the data due to the nuclear target, therefore when calculating the MM^2 using information from the initial and final state particles, the resolution effects from the target material smear out the distribution, and the peak is diminished with respect to the simulation. This highlights the need for nuclear effects to be taken into account when comparing data kinematics with simulations on a free proton target, and that simply using a scaled dilution factor is not sufficient for comparing datasets where Fermi motion is present and not present. As described in Section 4.1.2, the sPlot technique can be used to compare the missing mass counts in the data to the simulation. To do this, an sPlot fit is performed to the dataset after all previously established cuts have been applied, again, with carbon data to model the background, see Figure 4.40. The counts in the signal peak after the subtraction of carbon + other background processes are shown to be 967 ± 67 events, which is comparable to the produced signal peak from the TCSGen simulation, with 998 counts. The variable $E\gamma$ is also calculated from initial and final state particles and therefore is also subject to detector resolution effects and Fermi motion effects which may serve to explain the discrepancy.



(a) (b)

(c) (d)

Figure 4.39: The data MM^2, IM and t, with all exclusivity cuts, shown in black scaled by the dilution factor and the sim shown in red.



(a) (b)

Figure 4.40: sPlot of the MM^2 (shown on the x-axis as eepMM2, measured in GeV^2) in the final state, with all exclusivity cuts, compared to the TCSGen signal simulation.

Looking at each of the final state particles there is a fair agreement for the proton momentum and theta, with a slight amplitude discrepancy that could be due to the error on the dilution factor. For electron and positron theta there appears to be some slight shift in the theta distributions in particular. This is likely attributed to a mixture of the inclusion of QED background processes from GRAPE, and the fact that both generators simulate dilepton events, however there are resonances between $M_{e^+e^-}$ = 1.5 GeV - 2 GeV that decay to pions which feature in the data and not in the simulation, thus slightly shifting the kinematics of the lepton pair when directly compared.

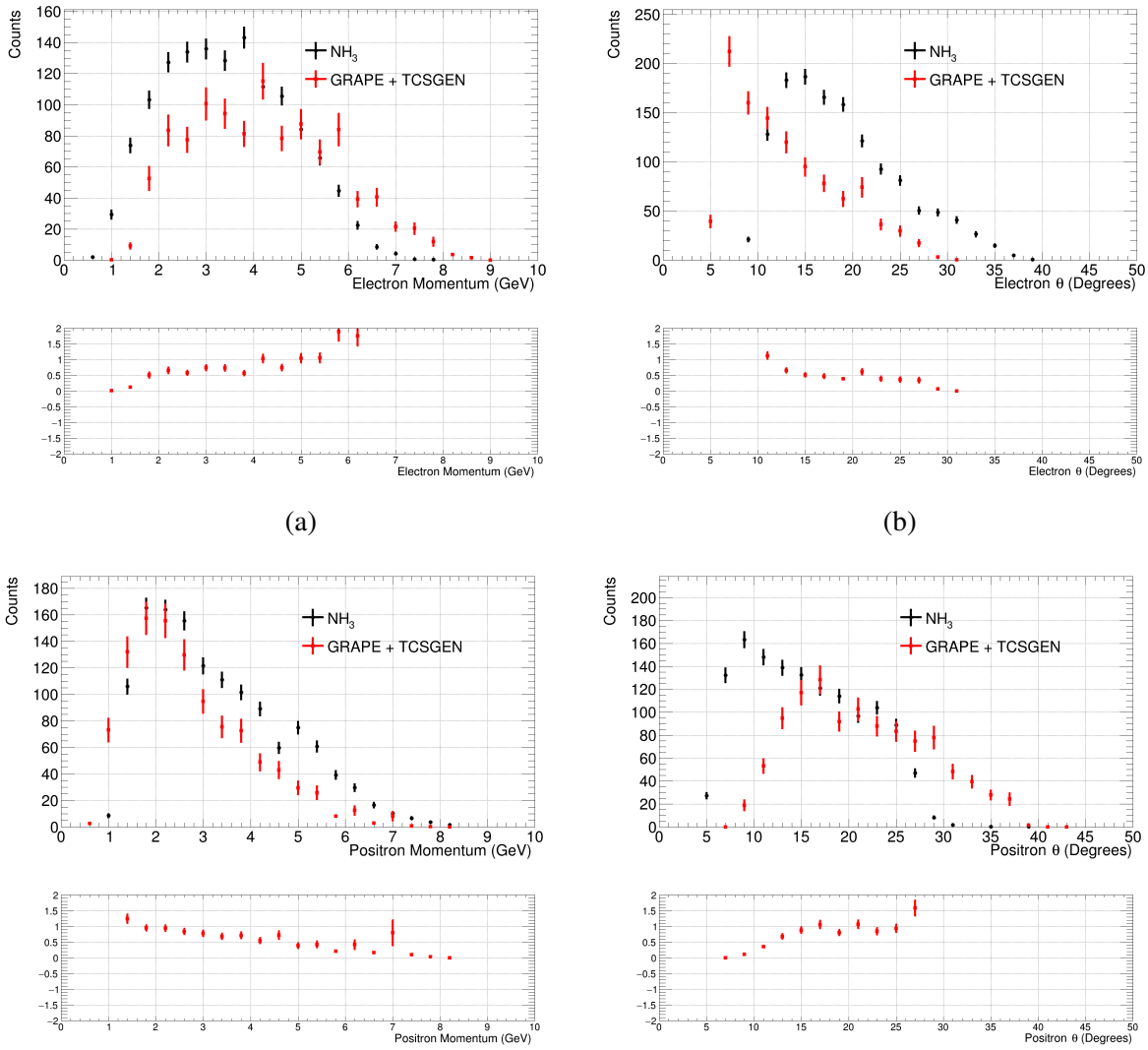
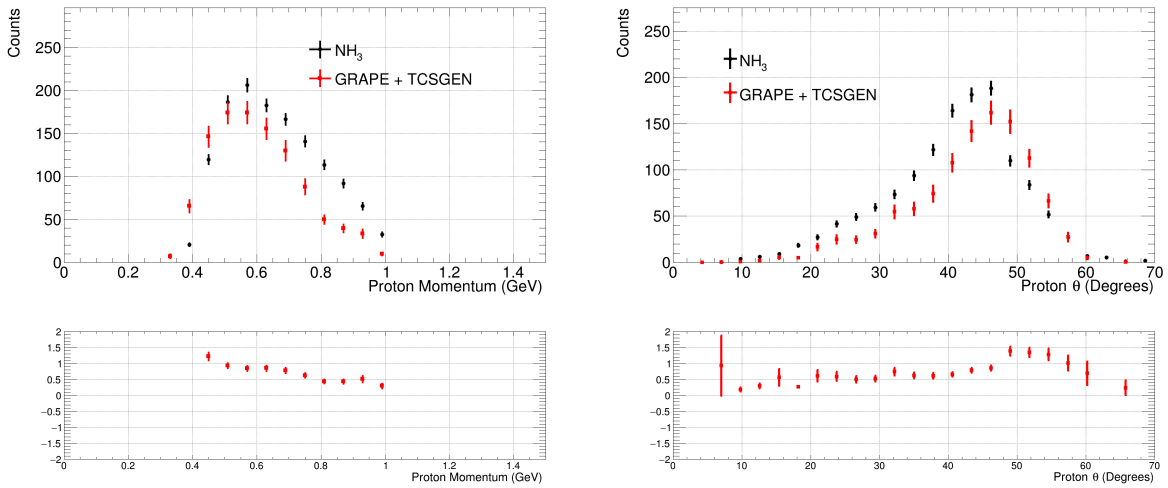


Figure 4.41: The data momentum and polar angle of the electron and positron in the final state, with all exclusivity cuts. The data shown in black is scaled by the dilution factor and the simulated events are shown in red.



(a) (b)

Figure 4.42: The data momentum and polar angle of the proton in the final state, with all exclusivity cuts. The data shown in black is scaled by the dilution factor and the simulated events are shown in red.

4.3 Extraction of the Observables

Now that the final-state phase space has been established and compared with the simulation, the final step is to extract the physics observables of interest from the data. Since TCS is largely dominated by BH, cross-section measurements are very difficult to extract, but since the BH contribution has a zero asymmetry, extracting asymmetries is much more feasible. As a reminder longitudinally polarised target gives access to three potential asymmetry observables, a Beam Spin Asymmetry (BSA), a Target Spin Asymmetry (TSA) and a Double Spin Asymmetry (DSA). The BSA is calculated as follows:

$$A_{\odot U} = \frac{P_t^-(N^{++} - N^{-+}) + P_t^+(N^{+-} - N^{--})}{Pb \times (P_t^-(N^{++} + N^{-+}) + P_t^+(N^{+-} + N^{--}))} \tag{4.13}$$

where $A_{\odot U}$ refers to circularly polarised photon and unpolarised target, P_t^+ and P_t^- represent the positive and negative target polarisation respectively, taken as a weighted average of each collected run.

The N^{ij} variables represent the number of counts in a kinematic histogram, usually binned in ϕ or t, with a beam helicity sign of i and a target polarisation sign of j. The counts in each of these histograms are normalised by the Faraday cup accumulated charge, where N^{++} for example is divided by the total accumulated charge of events with positive helicity and positive target polarisation in the Faraday cup. Finally, Pb is the beam polarisation, taken as 82.6% for the data-taking period. A detailed study of the variation in the beam polarisation across the run period and checking this average can be found in Appendix A.1.

The equation for the TSA is as follows:

$$A_{UL} = \frac{N^{++} + N^{-+} - N^{+-} - N^{--}}{Df \times (P_t^- (N^{++} + N^{-+}) + P_t^+ (N^{+-} + N^{--}))} \tag{4.14}$$

where A_{UL} refers to an unpolarised beam and a longitudinally polarised target and is much the same as the BSA, but with the addition of the dilution factor (D_f), which is calculated in Section 4.1.2 by taking the Faraday cup normalised carbon, empty target and LH_e and using the ratio of this to the NH_3 runs. This is to account for the target material aside from the polarised proton that we care for. Finally, the DSA is calculated as follows;

$$A_{\odot L} = \frac{N^{++} + N^{--} - N^{+-} - N^{-+}}{Pb \times Df \times (P_t^- (N^{++} + N^{-+}) + P_t^+ (N^{+-} + N^{--}))} \tag{4.15}$$

which uses the same variables as the prior two.

A study was performed to discern the variation in the dilution factor across each of the four final state kinematic bins in t. It is the expectation that the dilution factor should be independent of kinematic bin, and with this more well-established method of ascertaining the dilution factor and its error, the author was able to produce a value for each of the four final state t bins, firstly each value and its error is reported in table 4.6, and this is then plotted in Figure 4.43.

It is clear that there is very little notable variance in the dilution factor within the reported error, as expected, but the error bars are not insignificant and therefore more statistics would assure this conclusion. It is to be noted that the error on the dilution factor integrated over all t bins using the analytical method is +/-5%.

Table 4.6: Dilution factor values and associated errors for each t bin

tbin	Value	error
DF tbin1	0.340424	0.10992
DF tbin2	0.308554	0.0885508
DF tbin3	0.31214	0.103262
DF tbin4	0.38823	0.0831765

4.3.1 Calculating Target Polarisation

As outlined above, the target polarisation P_t is taken as a weighted average of each collected run. Throughout the experiment, the polarisation of the target was measured using NMR techniques (see Section 3.3.3), and these values are stored on a run-by-run basis. The issue with taking this result for P_t is that the physical structure of the target in the beamline (See Figure 3.16 in Section 3.3.3) has the NMR coils placed on top of the physical target material. This means the NMR is less sensitive to the polarisation of the target beyond its surface level, despite the centre of the

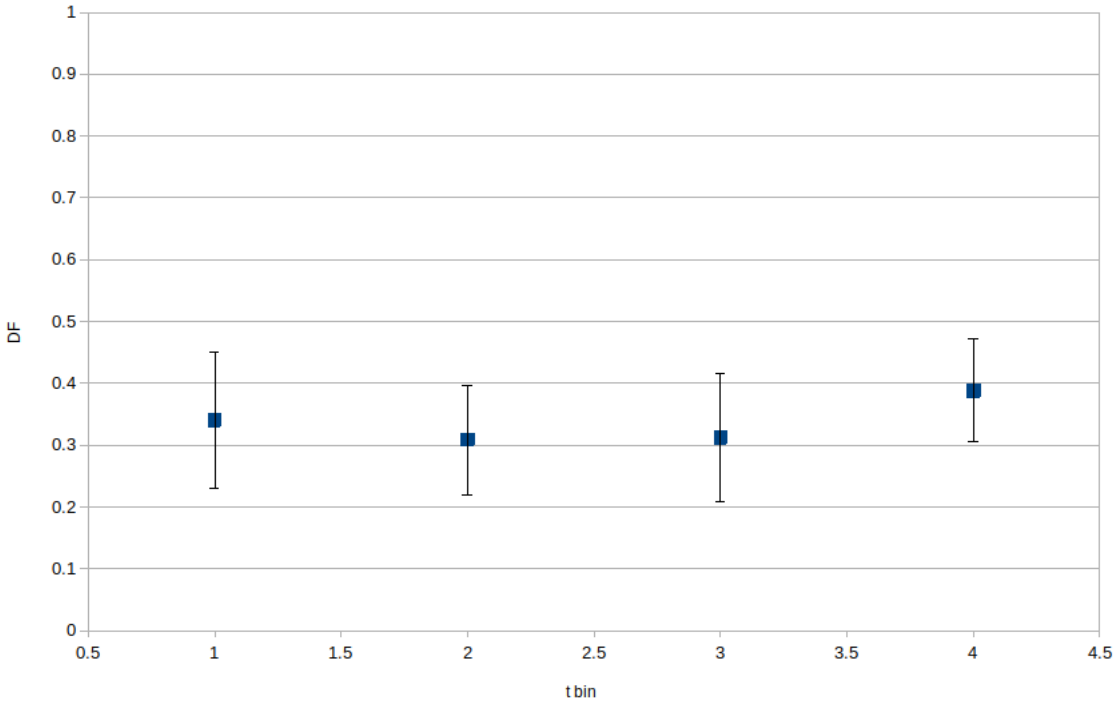


Figure 4.43: Dilution factor values and associated statistical error for each bin in t

target having the most direct beam exposure. There are also uncertainties associated with cell sizes and target misalignment [95]. This prompts the need for a secondary method to accurately determine a run by run target polarisation.

Following the method discussed in [96], the way this is achieved is by performing an analysis of quasi elastic $ep \rightarrow ep$ events and extracting their double spin asymmetry. This is then compared to a theoretical asymmetry;

$$A_{\odot L}^{\text{theoretical}} = \frac{2\tau(\frac{G_M}{G_E})[\tau\frac{M_p}{E_{\text{beam}}} + (1+\tau)\tan^2(\frac{\theta}{2})]}{1+(\frac{G_M}{G_E})^2\frac{\tau}{\varepsilon}} \tag{4.16}$$

calculated using the well established electromagnetic proton form factors $G_E(Q^2)$ and $G_M(Q^2)$ for a given set of Q^2 bins from [97]. The variables τ and ε are defined as $\tau = \frac{Q^2}{4m_N^2}$, where m_N is the nucleon mass, and $\varepsilon = [1+2(1+\tau)tan^2\frac{\theta}{2}]^{-1}$. The ratio of the measured DSA and theoretical DSA gives;

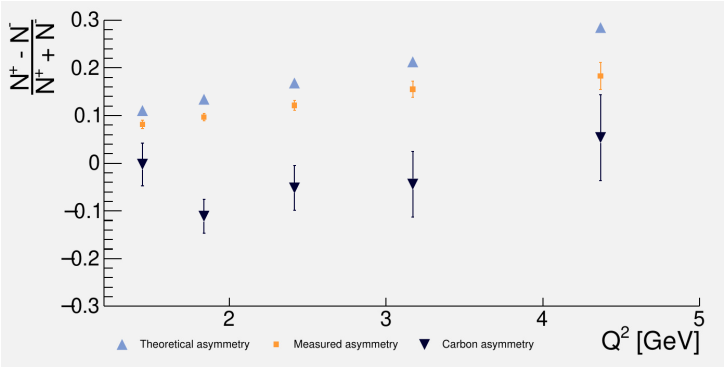
$$\frac{A_{\odot L}^{\text{measured}}}{A_{\odot L}^{\text{theoretical}}} = P_b P_t \tag{4.17}$$

and therefore this method can be used to extract P_t for known P_b. The author used code established to perform these calculations for analysts of the CLAS12 2022 run period, documented in [98]. The method first extracts the runs from a given run range, for this case the author split the 157 runs into 10 groups, creating a new group for every polarisation flip. The carbon runs taken across the full available run period were also provided, for the purpose of calculating the

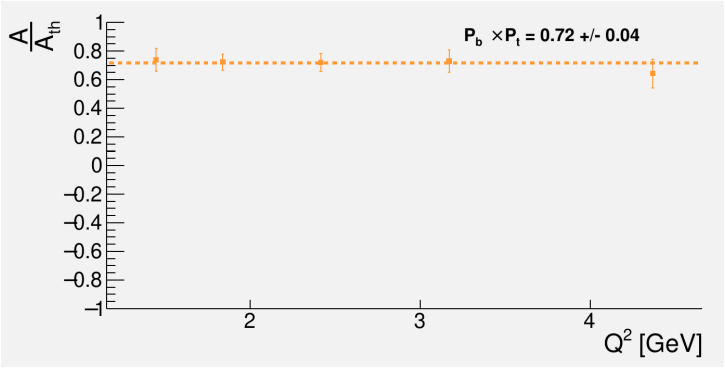
dilution factor for DSA extraction. Some exclusivity cuts are provided to ensure the events are quasi elastic, these are listed below [96] [98];

- $E_{e^-} > 9.5\,\text{GeV}$ and $E_{e^-} < 11.3\,\text{GeV}$
- $MM^2 < 0.04\,\text{GeV}^2$ and $MM^2 > -0.05\,\text{GeV}^2$
- $|\Delta\phi| > 178^\circ$ and $|\Delta\phi| < 181^\circ$
- $\delta p_p < 0.6\,\text{GeV}$ and $\delta p_p > -0.6\,\text{GeV}$
- $p_\perp < 0.4\,\text{GeV}$.

where $p_\perp$ is the total missing transverse momentum of the reaction and MM^2 here is the total missing mass of the reaction $ep \rightarrow ep$ which is assumed to be very close to zero. $|\Delta\phi|$ represents the difference between electron and proton azimuthal angles, or, a measure of the coplanarity. The cut on δp_p ensures that the reconstructed momenta agrees with proton momenta calculated via elastic kinematics. An example result for the lowest run range, runs 16137-16148, which have an average positive polarisation of 0.8 as recorded by online NMR [95], is shown in Figure 4.44. The extracted target polarisation using the elastic method is $0.72/0.826 \pm 0.04$. A similar set of plots is shown in Figure 4.45 for the highest negative run range, runs 16721-16766, which have an average negative polarisation of -0.55 as recorded by online NMR [95], the extracted target polarisation is $-0.66/0.826 \pm 0.04$.

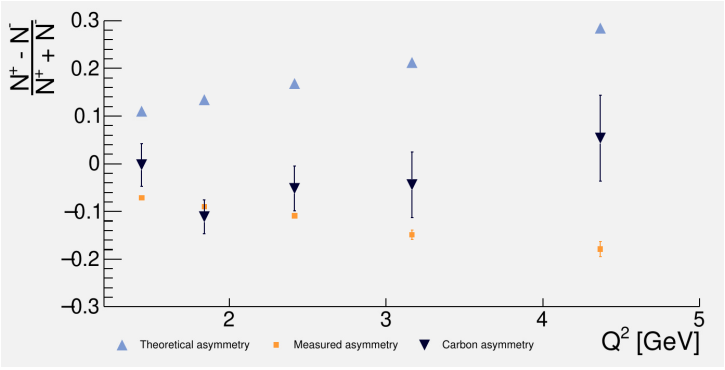


(a)

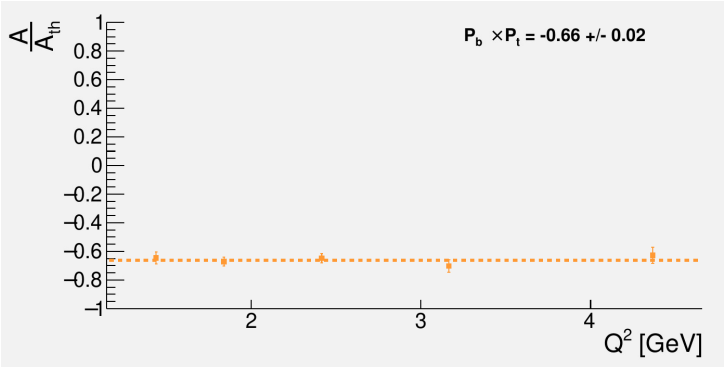


(b)

Figure 4.44: Figure 4.44a shows the measured (orange), theoretical (blue) and carbon (black) asymmetry in each Q^2 bin for a run range with positive target polarisation in the dataset, and for the same range, Figure 4.44b shows the ratio between measured and theoretical asymmetry giving $P_b P_t$.



(a)



(b)

Figure 4.45: Figure 4.45a shows the measured (orange), theoretical (blue) and carbon (black) asymmetry in each Q^2 bin for a run range with negative target polarisation in the dataset, and for the same range, Figure 4.44b shows the ratio between measured and theoretical asymmetry giving $P_b P_t$.

This analysis was performed for each of the 10 groups with results shown in Figure 4.46. It is clear that the positive and negative values have only slight variation across the run period, within established error bounds. Therefore, the average of these values is taken as the target polarisation for the run period, for positive runs this value is 0.871 and for negative runs this value is -0.796.

4.3.2 Error Formalism

The formalisms in the following sections are taken from the procedures outlined in [99].
With the functional form of the asymmetries defined as above, the error on each bin is calculated using the following:

$$\delta A_{BT} = \sqrt{\sum (\frac{\delta A_{ij}}{N^{ij}} \Delta N^{ij})^2} \tag{4.18}$$

for each beam target configuration B, T, where i, j are the helicity and target polarisation signs as usual. The first step is to calculate the partial derivatives for each beam target configuration, using each asymmetry equation. Looking first at the BSA in Equation 4.13, the partial derivatives

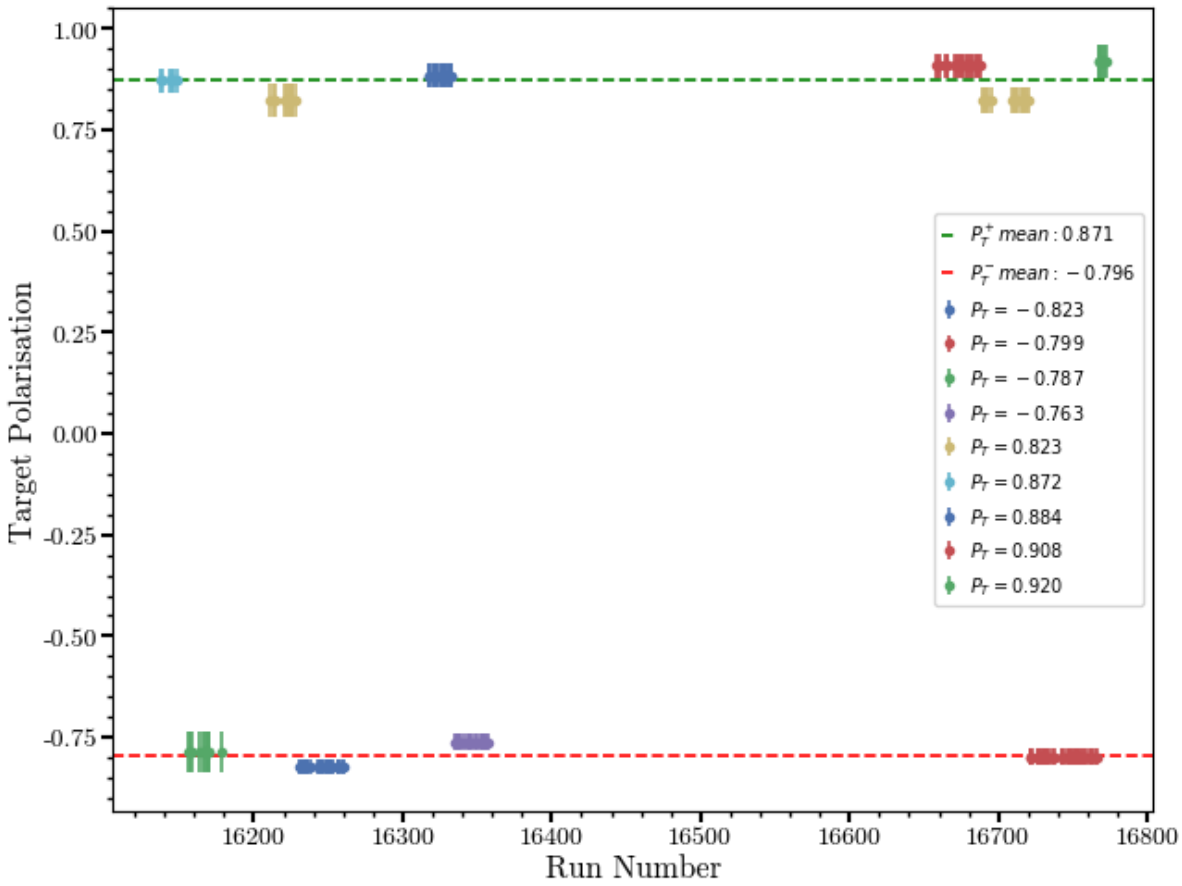


Figure 4.46: Target polarisation

are as follows;

$$\frac{\partial A_{\odot U}}{\partial N^{++}} = \frac{2P_t^-(N^{-+}P_t^- + N^{--}P_t^+)}{P_b[(N^{-+}+N^{++})P_t^- + (N^{--}+N^{+-})P_t^+]^2} \tag{4.19}$$

$$\frac{\partial A_{\odot U}}{\partial N^{--}} = \frac{-2P_t^+(N^{++}P_t^- + N^{+-}P_t^+)}{P_b[(N^{-+}+N^{++})P_t^- + (N^{--}+N^{+-})P_t^+]^2} \tag{4.20}$$

$$\frac{\partial A_{\odot U}}{\partial N^{-+}} = \frac{-2P_t^-(N^{++}P_t^- + N^{+-}P_t^+)}{P_b[(N^{-+}+N^{++})P_t^- + (N^{--}+N^{+-})P_t^+]^2} \tag{4.21}$$

$$\frac{\partial A_{\odot U}}{\partial N^{+-}} = \frac{2P_t^+(N^{-+}P_t^- + N^{--}P_t^+)}{P_b[(N^{-+}+N^{++})P_t^- + (N^{--}+N^{+-})P_t^+]^2} \tag{4.22}$$

and for the TSA;

$$\frac{\partial A_{UL}}{\partial N^{++}} = \frac{(N^{--}+N^{+-})(P_t^- + P_t^+)}{D_f[(N^{-+}+N^{++})P_t^- + (N^{--}+N^{+-})P_t^+]^2} \tag{4.23}$$

$$\frac{\partial A_{UL}}{\partial N^{--}} = \frac{-(N^{-+}+N^{++})(P_t^- + P_t^+)}{D_f[(N^{-+}+N^{++})P_t^- + (N^{--}+N^{+-})P_t^+]^2} \tag{4.24}$$

$$\frac{\partial A_{UL}}{\partial N^{-+}} = \frac{(N^{--}+N^{+-})(P_t^- + P_t^+)}{D_f[(N^{-+}+N^{++})P_t^- + (N^{--}+N^{+-})P_t^+]^2} \tag{4.25}$$

$$\frac{\partial A_{UL}}{\partial N^{+-}} = \frac{-(N^{-+}+N^{++})(P_t^- + P_t^+)}{D_f[(N^{-+}+N^{++})P_t^- + (N^{--}+N^{+-})P_t^+]^2}. \tag{4.26}$$

.

The next step is to calculate the error on the yield, namely ΔN^{ij}, this is calculated as follows;

$$\Delta N^{ij} = \frac{\sqrt{N^{ij}}}{FCup} \tag{4.27}$$

where $FCup$ refers to the accumulated charge in the Faraday cup to which the events are normalised. Errors on the target polarisation ($\pm 4\%$) beam polarisation ($\pm 0.2\%$) and dilution factor ($\pm 5\%$) are added in quadrature to Equation 4.18, and the total error is calculated for each bin in the kinematic histogram.

Chapter 5

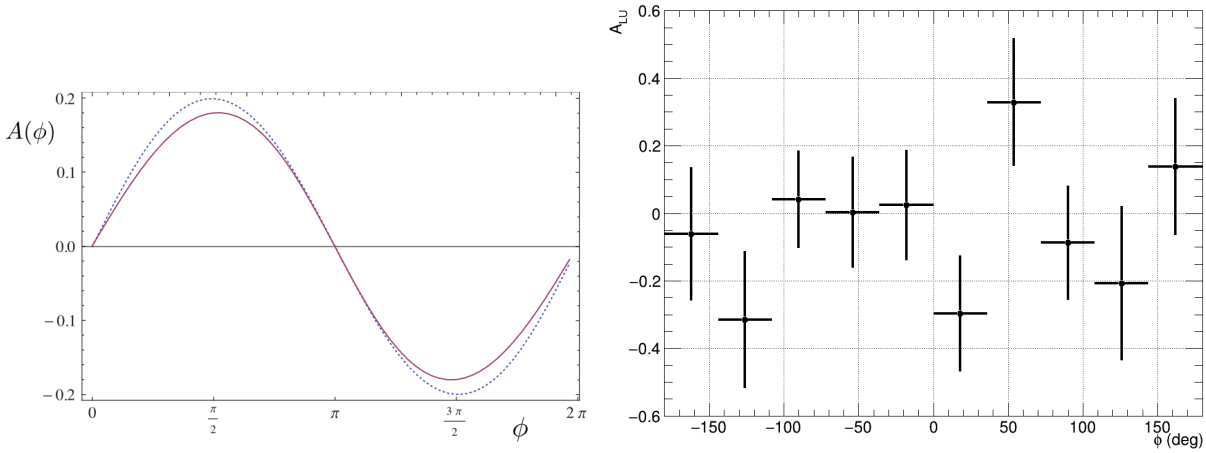
Results

Each BSA distribution is theorised to follow the functional form of $A + Bsin(\phi)$, where A and B are constants (A representing the offset of the fit from zero and B the amplitude of the fit), matching that of the previous TCS measurement in [42]. The TSA distributions are theorised to follow the functional form $A + Bcos(\phi)$, and are compared with the results from [49] in Figure 2.12.

The intention would ideally be to perform a 1:1 comparison with theoretical predictions. An attempt was made as follows, to compare with the BSA theory prediction in [48] the data is plotted in a central kinematic bin of $Q'^2 = 4\,\mathrm{GeV}^2$, $t = 0.1\,\mathrm{GeV}^2$ and $E_\gamma = 10\,\mathrm{GeV}$, the results are shown in Figure 5.1, note the theory result covers the ϕ range $[0, 2\pi]$ however the data is plotted between $\phi = [-\frac{\pi}{2}, \frac{\pi}{2}]$ so to directly compare the sign should be flipped.

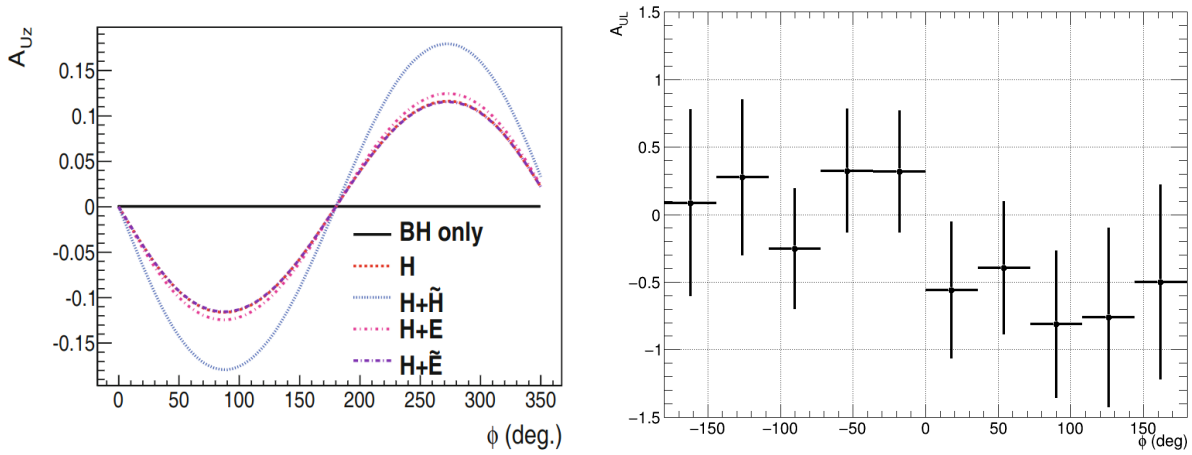
A comparison for the TSA was performed based on the binning described in [49], which stipulates the following kinematics: $Q'^2 = 7\,\mathrm{GeV}^2$, $t = 0.4\,\mathrm{GeV}^2$. The results are shown in Figure 5.2, to directly compare the sign should again be flipped due to the ϕ range. Since this investigation relies on enough statistics to be able to perform a multidimensionally binned extraction, a comparable result was unable to be obtained.

Looking now to maximising statistics in order to show a proof of principle for an asymmetry extraction, Figures 5.3 and 5.4 show the Beam Spin and Target Spin Asymmetry as a function of ϕ for $|MM^2| < 0.4\,\mathrm{GeV}^2$, $\frac{P_t}{P} < 0.05$ and $1.5\,\mathrm{GeV} < \mathrm{IM} < 3\,\mathrm{GeV}$ across four bins in t for the purposes of comparison with Figure 2.10 in Section 2.4.2. TSA results show consistency with theoretical predictions in two out of the four t bins, showing a flatter asymmetry for bin $0.25\,\mathrm{GeV}^2 < t < 0.34\,\mathrm{GeV}^2$ and bin $0.48\,\mathrm{GeV}^2 < t < 0.8\,\mathrm{GeV}^2$. The fit applied to bins $0.15\,\mathrm{GeV}^2 < t < 0.25\,\mathrm{GeV}^2$ and $0.34\,\mathrm{GeV}^2 < t < 0.48\,\mathrm{GeV}^2$ shows a comparable amplitude and shape with theory, however, the statistical errors are too high to claim full agreement. The BSA results show a flatter asymmetry when compared with the previous TCS measurement in [42]. This could be attributed to multiple factors, for one the general lack of statistics in the final state affects the ability to restrict tightly on the TCS phase space, therefore some background



(a) BSA in central kinematic bin theory. (b) BSA in central kinematic bin.

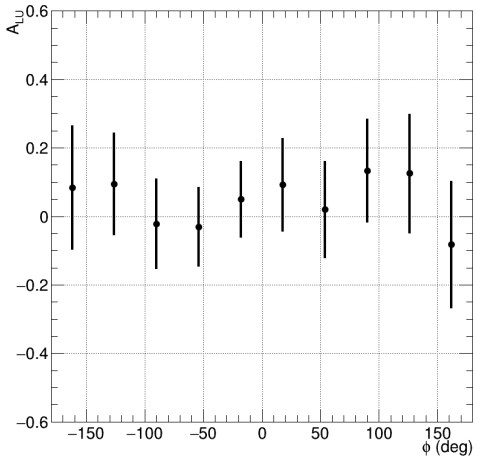
Figure 5.1: Comparison of BSA with theory.



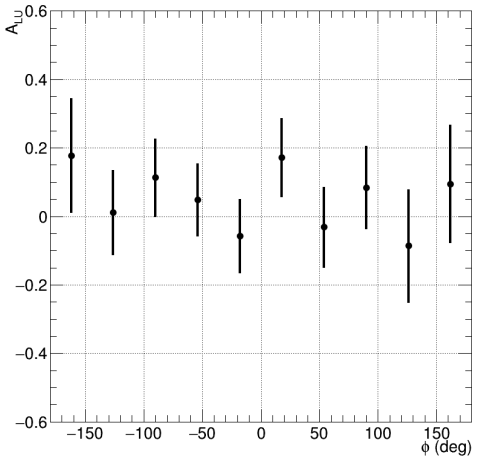
(a) TSA in central kinematic bin theory. (b) TSA in central kinematic bin.

Figure 5.2: Comparison of TSA with theory.

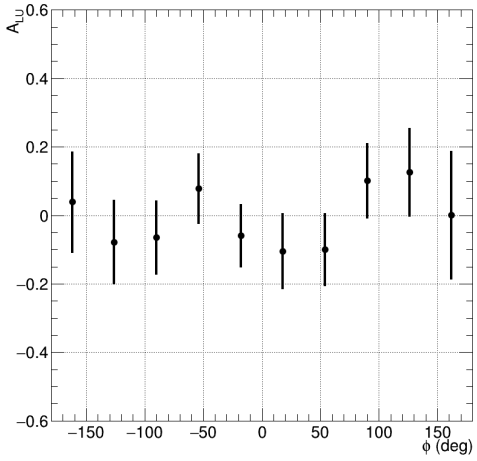
processes may be allowed in for example allowing the wider invariant mass cut of 1.5 GeV as opposed to 2 GeV. This could also be attributed to background sources from the nuclear target, as seen in figure 4.39d where the fermi motion effect is highlighted, this serves to smear out the missing mass distribution, therefore the resolution on this kinematic is much lower than if the target were only a free proton. There is a twofold effect this may have on the data, one, the cut on $MM^2 < 0.4\,\mathrm{GeV}^2$ may be too tight for the smeared out peak and some signal events could have been cut out, and two, the integration over kinematic bins in MM^2 and Q'^2 mean the result is dependent on good reconstruction of the 'missing' particle (the scattered electron), which is more difficult to achieve when the total final state reaction kinematics are smeared with nuclear effects. To a degree, the BSA should be less affected by this issue than the TSA since the nuclear background effects are cancelled out, however the $0.4\,\mathrm{GeV}^2$ cut may still be too tightly restricted to extract good statistics in the signal region, thus reducing the overall asymmetry.



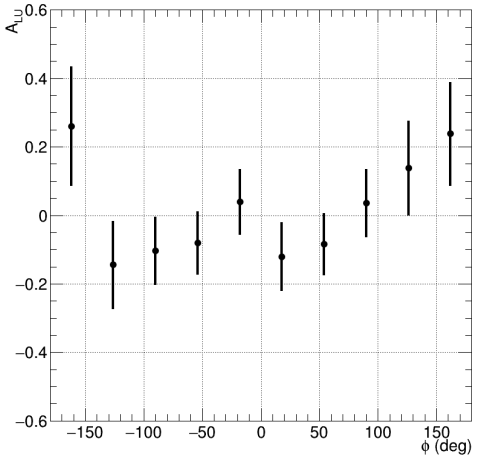
(a) $0.15\,\mathrm{GeV}^2 < t < 0.25\,\mathrm{GeV}^2$.



(b) $0.25\,\mathrm{GeV}^2 < t < 0.34\,\mathrm{GeV}^2$.

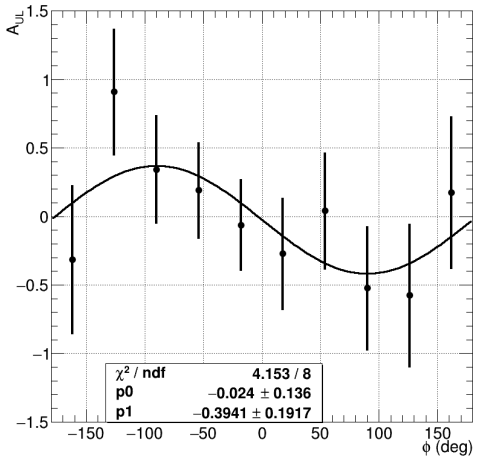


(c) $0.34\,\mathrm{GeV}^2 < t < 0.48\,\mathrm{GeV}^2$.

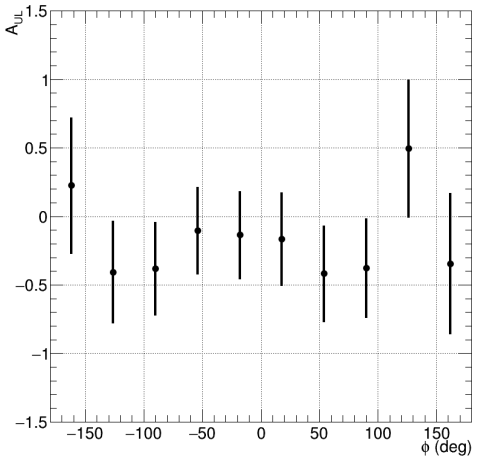


(d) $0.48\,\mathrm{GeV}^2 < t < 0.8\,\mathrm{GeV}^2$.

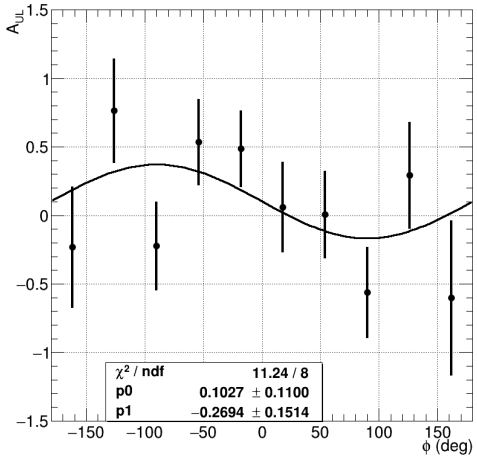
Figure 5.3: Beam Spin Asymmetry in each bin in t.



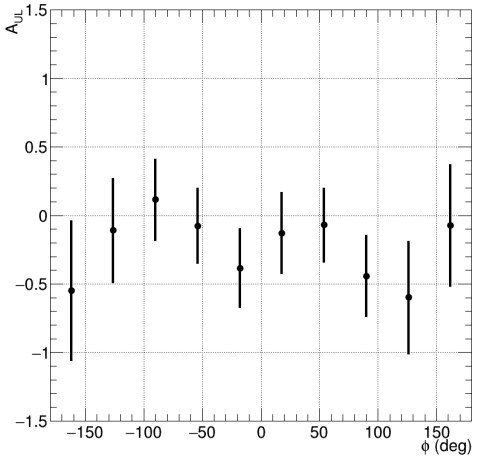
(a) $0.15\,\mathrm{GeV}^2 < t < 0.25\,\mathrm{GeV}^2$.



(b) $0.25\,\mathrm{GeV}^2 < t < 0.34\,\mathrm{GeV}^2$.



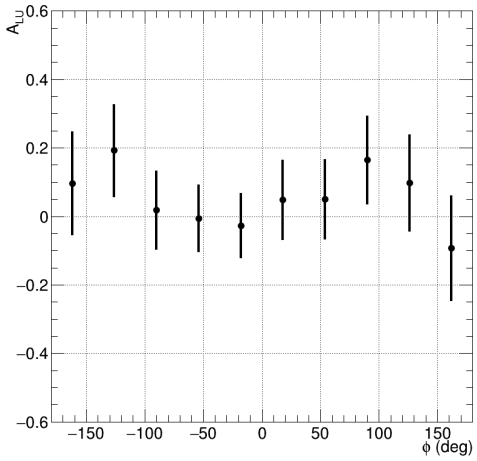
(c) $0.34\,\mathrm{GeV}^2 < t < 0.48\,\mathrm{GeV}^2$.



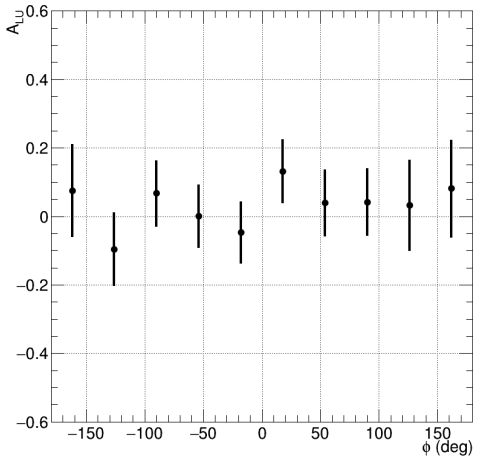
(d) $0.48\,\mathrm{GeV}^2 < t < 0.8\,\mathrm{GeV}^2$.

Figure 5.4: Target spin asymmetry in each bin in t.

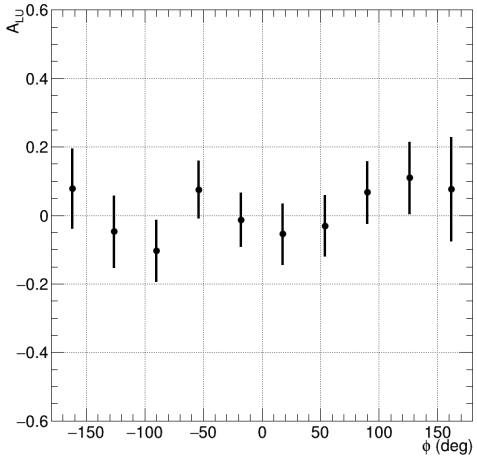
To investigate this, the Missing Mass squared cut was widened to $MM^2 < |0.7\,\mathrm{GeV}^2|$ and the resulting distributions are shown in Figures 5.5 and 5.6.



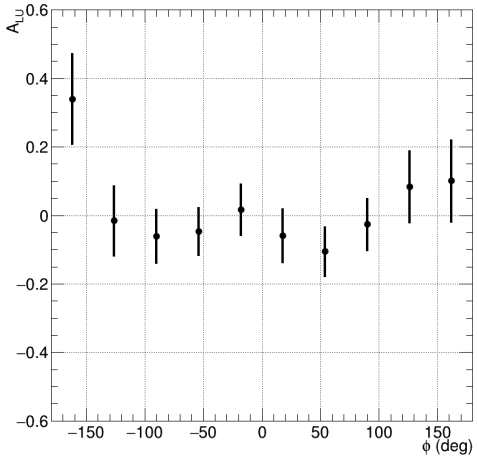
(a) $0.15\,\mathrm{GeV}^2 < t < 0.25\,\mathrm{GeV}^2$, $\mathrm{MM}^2 < |0.7\,\mathrm{GeV}^2|$.



(b) $0.25\,\mathrm{GeV}^2 < t < 0.34\,\mathrm{GeV}^2$, $\mathrm{MM}^2 < |0.7\,\mathrm{GeV}^2|$.

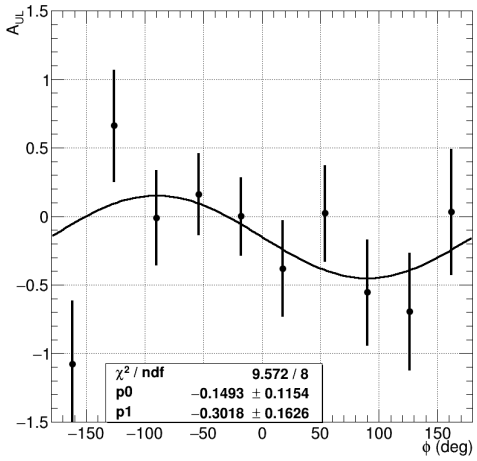


(c) $0.34\,\mathrm{GeV}^2 < t < 0.48\,\mathrm{GeV}^2$, $\mathrm{MM}^2 < |0.7\,\mathrm{GeV}^2|$.

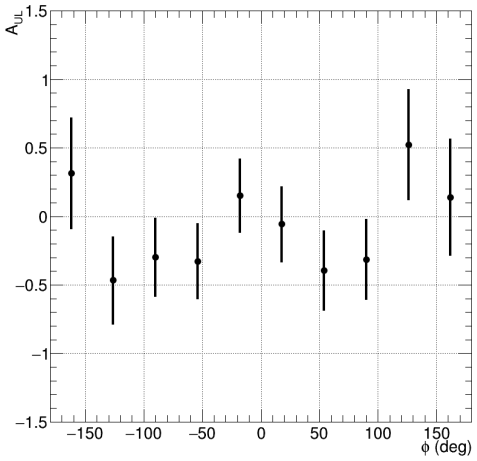


(d) $0.48\,\mathrm{GeV}^2 < t < 0.8\,\mathrm{GeV}^2$, $\mathrm{MM}^2 < |0.7\,\mathrm{GeV}^2|$.

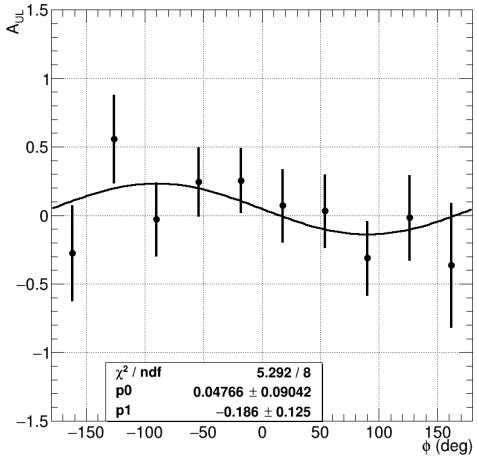
Figure 5.5: Beam spin asymmetry with widened Missing Mass Squared cut.



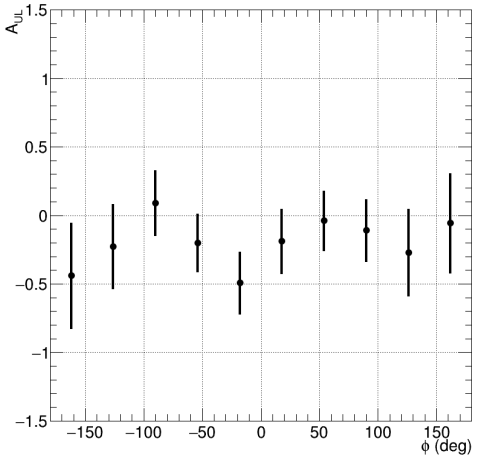
(a) $0.15\,\text{GeV}^2 < t < 0.25\,\text{GeV}^2$, $\text{MM}^2 < |0.7\,\text{GeV}^2|$.



(b) $0.25\,\text{GeV}^2 < t < 0.34\,\text{GeV}^2$, $\text{MM}^2 < |0.7\,\text{GeV}^2|$.



(c) $0.34\,\text{GeV}^2 < t < 0.48\,\text{GeV}^2$, $\text{MM}^2 < |0.7\,\text{GeV}^2|$.

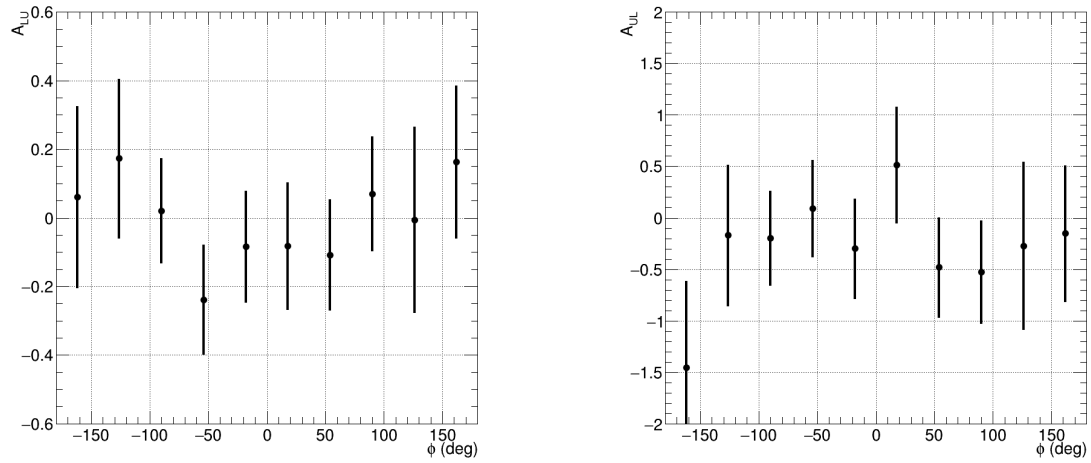


(d) $0.48\,\text{GeV}^2 < t < 0.8\,\text{GeV}^2$, $\text{MM}^2 < |0.7\,\text{GeV}^2|$.

Figure 5.6: Target spin asymmetry with widened Missing Mass Squared cut.

Unfortunately, increasing the statistics by widening this cut allows too much background which further dilutes the asymmetry, as evidenced by the flatter TSA distribution in Figure 5.6. The BSA results remain flat, specifically for the second and third t bins, the first t bin appears to form some sort of sinusoid, however, the bin is still far too statistically limited to have any confidence in the shape. In the case of the fourth t bin, there looks to be a couple of non-zero points amplified by the wider MM^2 cut, this is likely due to this t bin having the largest statistics at these kinematics. Since most of the TCS final state statistics are in the lower t bin as seen in Figure 4.39b, it is assumed that this bin contains the most influence from background processes,

that are being allowed in by the wider MM^2 cut, this is supported by applying a restriction on the Invariant Mass in this particular bin of $IM > 2\,\text{GeV}$, which brings some of those large offsets closer to the predicted amplitude, this again, however, comes at the cost of statistics and therefore no overall conclusion can be drawn from this bin.



(a) $0.48\,\text{GeV}^2 < t < 0.8\,\text{GeV}^2$, $\text{MM}^2 < |0.7\,\text{GeV}^2|$, IM > 2 GeV.

(b) $0.48\,\text{GeV}^2 < t < 0.8\,\text{GeV}^2$, $\text{MM}^2 < |0.7\,\text{GeV}^2|$, IM > 2 GeV.

Figure 5.7: Beam Spin Asymmetry (Left) and Target Spin Asymmetry (Right) with widened MM^2 and tightened IM cut for the fourth t bin (largest statistics at widened MM^2).

A systematic study was performed on the two extracted TSA asymmetries in Figures 5.4a and 5.4c, varying the MM^2 cut from its current 2σ width of $\pm 0.4\,\text{GeV}$ to 2.5σ ($\pm 0.525\,\text{GeV}$) and then finally down to 1.5σ ($\pm 0.315\,\text{GeV}$). The systematic uncertainties were calculated based on the following formula [45]:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(A_i - A_0)^2}{n}} \tag{5.1}$$

where A_i represents the asymmetry after the variation, A_0 represents the original asymmetry and n represents the number of variations performed.

Cut	TSA tbin1	TSA tbin3
1.5σ ($\pm 0.315\,\text{GeV}$)	-0.5602 ± 0.2086	-0.2935 ± 0.1646
2σ ($\pm 0.4\,\text{GeV}$)	-0.3941 ± 0.1917	-0.2694 ± 0.1514
2.5σ ($\pm 0.525\,\text{GeV}$)	-0.4925 ± 0.1784	-0.2326 ± 0.1383

Table 5.1: Table of systematic variation in MM^2

Each asymmetry from Table 5.1 was inserted into Equation 5.1 and the systematic uncertainty

for the two t bins was found to be;

$$\sigma_{TSA_{tbin1}} = 0.1365 \tag{5.2}$$

$$\sigma_{TSA_{tbin3}} = 0.0311 \tag{5.3}$$

Looking at the $\pm 0.2\%$ error in beam polarisation, a further systematic can be computed, taking the values of each asymmetry at a P_b of 82.6%, 82.4% and 82.8%, but there was no reportable change in asymmetry amplitude for this variation.
The error on the target polarisation was ± 0.04 for both positive and negative P_T, and the asymmetry fluctuation is reported in Table 5.2

Cut	TSA tbin1	TSA tbin3
$P_{T_+} = 0.911$, $P_{T_-} = 0.836$	-0.3761 $\pm$ 0.1829	-0.257 $\pm$ 0.145
$P_{T_+} = 0.871$, $P_{T_-} = 0.796$	-0.3941 $\pm$ 0.1917	-0.2694 $\pm$ 0.1514
$P_{T_+} = 0.831$, $P_{T_-} = 0.766$	-0.4114 $\pm$ 0.2001	-0.2812 $\pm$ 0.1581

Table 5.2: Table of systematic variation in P_T

Each asymmetry from Table 5.2 was again inserted into Equation 5.1 and the systematic uncertainty for the two t bins was found to be;

$$\sigma_{TSA_{tbin1}} = 0.0177 \tag{5.4}$$

$$\sigma_{TSA_{tbin3}} = 0.0121 \tag{5.5}$$

A final study was performed on varying the dilution factor $D_f \pm \Delta_{D_f}$, the results of which are shown in Table 5.3.

Cut	TSA tbin1	TSA tbin3
$D_f = 0.38$	-0.3422 $\pm$ 0.1665	-0.2339 $\pm$ 0.1315
$D_f = 0.33$	-0.3941 $\pm$ 0.1917	-0.2694 $\pm$ 0.1514
$D_f = 0.28$	-0.4645 $\pm$ 0.2259	-0.3175 $\pm$ 0.1785

Table 5.3: Table of systematic variation in D_f

These values give a resulting systematic uncertainty of;

$$\sigma_{TSA_{tbin1}} = 0.0618 \tag{5.6}$$

$$\sigma_{TSA_{tbin3}} = 0.0423 \tag{5.7}$$

It is clear from this study that the dominant source of systematic uncertainty are the exclusivity cuts on MM^2, which show a comparable magnitude of error to the bin-by-bin statistical error on each asymmetry point for TSA_{tbin1}, and are approximately $\frac{1}{5}$ of the statistical error on TSA_{tbin3}. The target polarisation systematic is around $\frac{1}{10}$ of the statistical error for each bin, the dilution factor is approximately $\frac{1}{4}$ of the statistical error and the beam polarisation contributes a negligible amount to the total systematic uncertainty.

Chapter 6

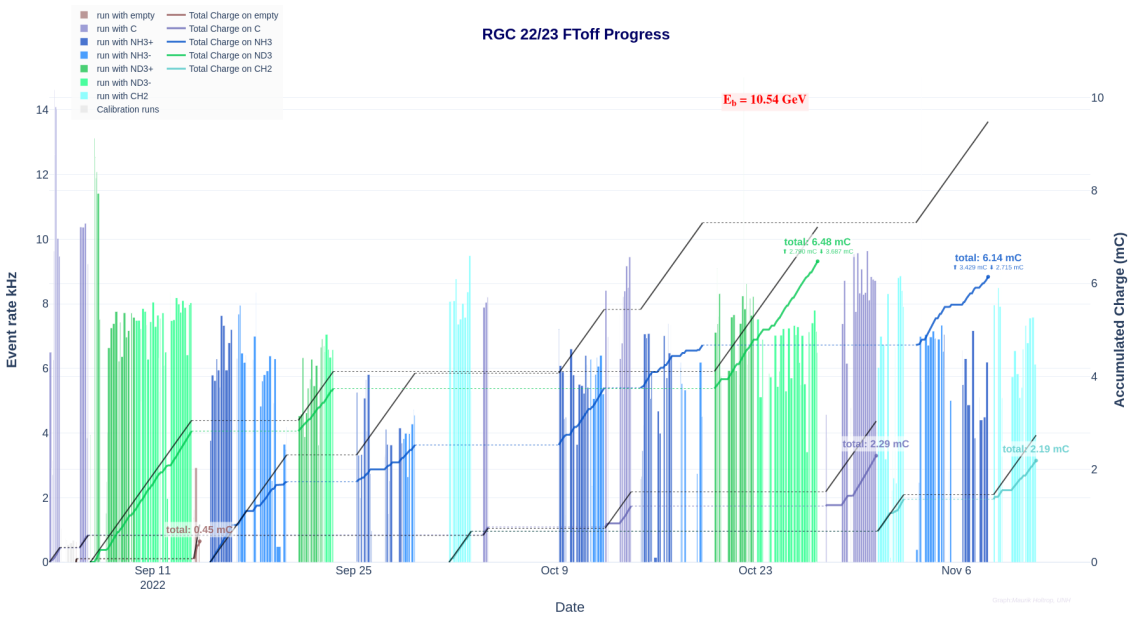
Future outlook for TCS

6.1 Increased Statistics

This thesis reports data taken to extract TCS on a longitudinally polarised NH_3 target with the 2022/2023 data taking run at Jefferson Lab. This concluding chapter will start with discussing some future prospects of this measurement, followed by the concluding remarks of the thesis.

The data used in this analysis included 157 runs, making up a total accumulated charge of approximately 4.78 mC. The dataset was from the run period taken across the summer months of 2022, however, the total data taking run in fact lasted between June 2022 and March 2023, therefore there are two additional run periods worth of data, Fall/Winter 2022, where the forward tagger was off, and Winter 2023. At the time of this thesis these additional run periods are still being processed for analysis. These two datasets provide an additional 8.28 mC, an extra two-thirds of the total available data. Analysing this data will serve to reinforce the results of this thesis, and allow for a more confident multidimensional binning of the final state observables. More statistics will also allow for a more high precision measurement of the asymmetries, restricting cuts for example in the invariant mass of the dilepton pair to $2\,\text{GeV} < \text{IM}_{e^+e^-} < 3\,\text{GeV}$, cutting out any potential contributions from the meson resonances between 1.5 GeV and 2 GeV. It would also be interesting, with higher statistics, to observe if any asymmetry in fact lies between this region, and whether this contributes to the TCS asymmetry in any way. More statistics would also serve to give a better chance of extracting a Double Spin Asymmetry, which will allow access to $H, \tilde{H}, \tilde{E}$ [49].

The errors on the asymmetry can be extrapolated to include the additional data from all three run periods, and to also include additional Program Advisory Committee (PAC) days that were awarded to replace those lost during a power supply failure that disrupted the Fall 2022 run period. This extrapolation is shown in Figure 6.2. Note that this is purely an extrapolation of error bars, not a prediction based on actual data. The binned distributions from Chapter 5



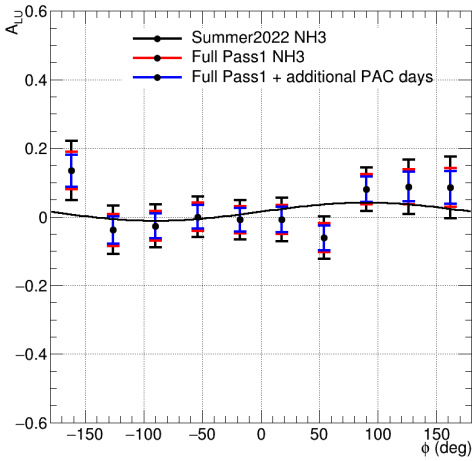
(a)

(b)

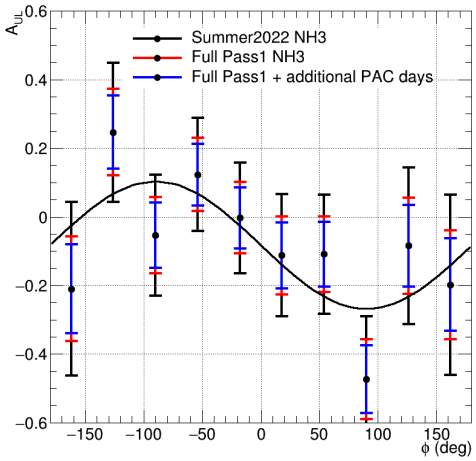
Figure 6.1: The accumulated charge across the two remaining run periods of the experiment.

are shown in Figures 6.3 to 6.7 and it is clear that error bars are significantly reduced with the increase in statistics, and the prediction is that for the most part this increase in statistics serves to make the shape of the asymmetries more discernible. It is to be noted that the blue error bars, representing the increase in statistics awarded by the extra PAC experiment time, assumes an additional parameter, known as the 'duty factor' which accounts for the time the beam is actually running, which is approximately 0.5, therefore the actual increase in statistics including this factor is a conservative estimate. In a significant number of the plots shown however, the

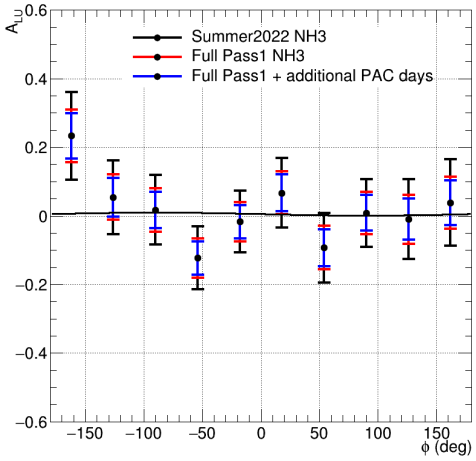
decrease in error shown by both the blue and red error bars, where red represents the additional data from the Fall/Winter 2022 and Winter 2023 run periods, is enough to bring the statistical error down to a point where the asymmetry shape is much more clearly defined. The additional data also serves as a proof of principle for a potential multidimentionally binned extraction, as seen by the extrapolation of Figure 5.7 in Figure 6.7 which will allow for a more detailed study of the TCS asymmetry. Figure 6.2 also presents the opportunity to separate contributions from protons in the forward and central detectors, which prior to the increase in data would have been too statistically limited to extract any sort of meaningful distribution. This separation can be used in investigations of potential contributions from the invariant mass region between 1.5 GeV and 2 GeV, since this is where the majority of central detector protons will fall, and see if any of these dilute or affect the asymmetry in a different way in a nuclear target with background dilution coming from scattering events occurring on bound protons in the nitrogen, compared to a free proton target, where they were shown to have little to no effect in [42].



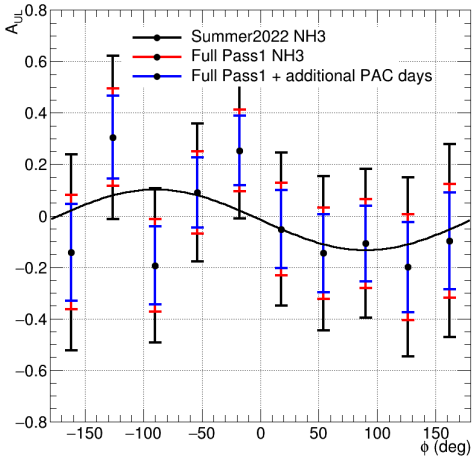
(a) BSA with no restriction on the proton.



(b) TSA with no restriction on the proton.

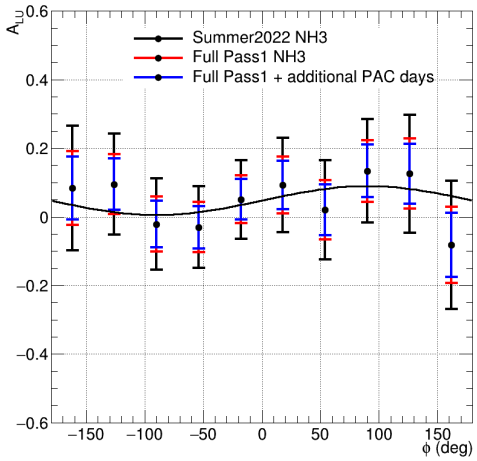


(c) BSA with the proton restricted to the FD.

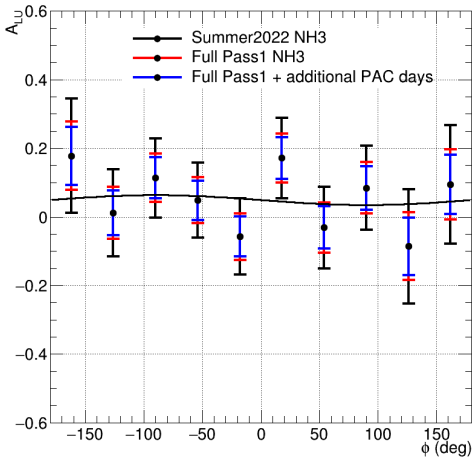


(d) TSA with the proton restricted to the FD.

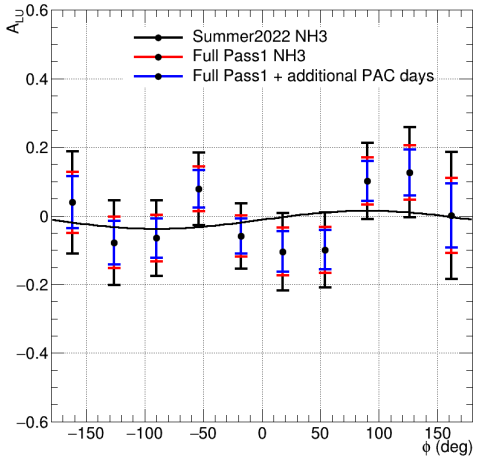
Figure 6.2: BSA (Left) and TSA (right).



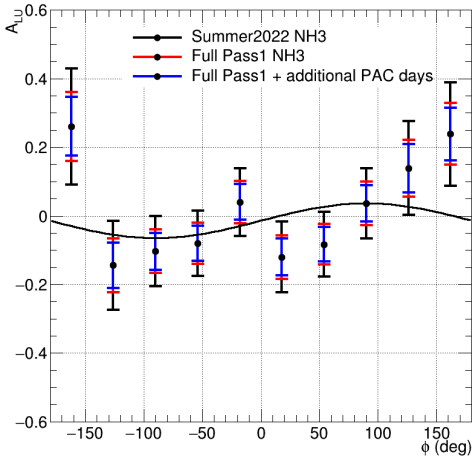
(a) $0.15\,\mathrm{GeV}^2 < t < 0.25\,\mathrm{GeV}^2$.



(b) $0.25\,\mathrm{GeV}^2 < t < 0.34\,\mathrm{GeV}^2$.

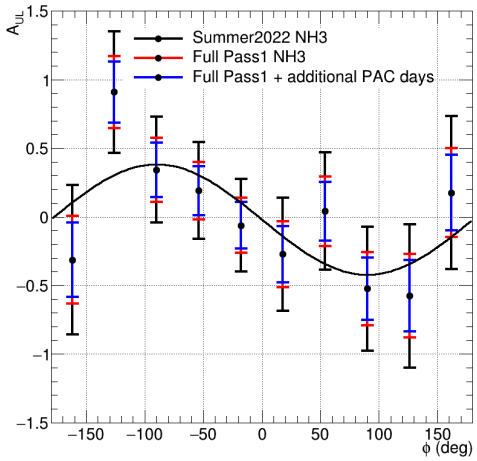


(c) $0.34\,\mathrm{GeV}^2 < t < 0.48\,\mathrm{GeV}^2$.

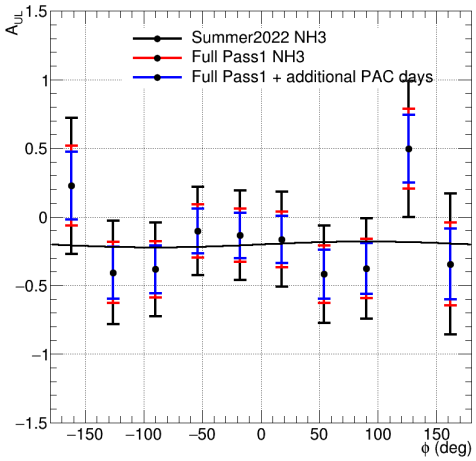


(d) $0.48\,\mathrm{GeV}^2 < t < 0.8\,\mathrm{GeV}^2$.

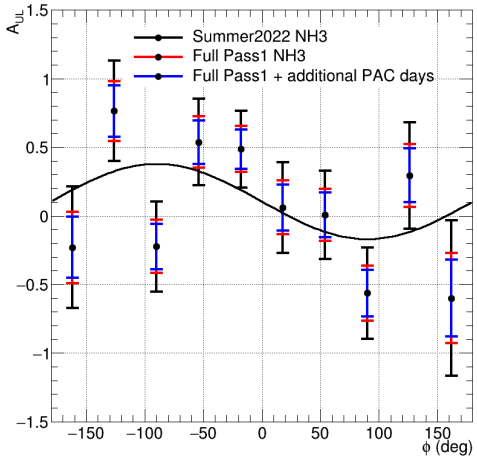
Figure 6.3: BSA in each bin in t.



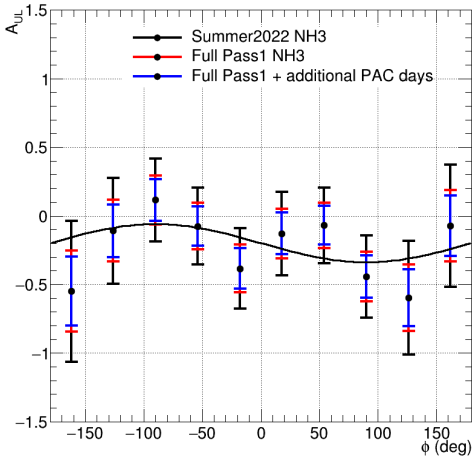
(a) $0.15\,\mathrm{GeV}^2 < t < 0.25\,\mathrm{GeV}^2$.



(b) $0.25\,\mathrm{GeV}^2 < t < 0.34\,\mathrm{GeV}^2$.

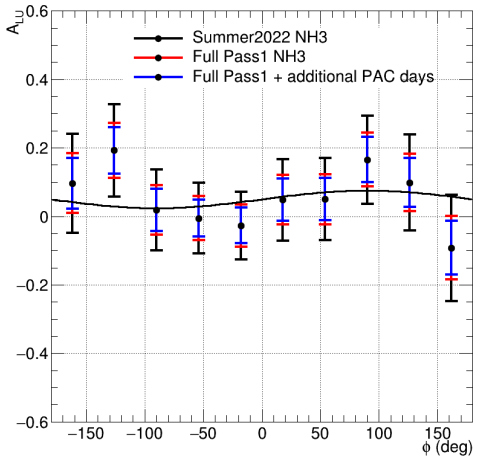


(c) $0.34\,\mathrm{GeV}^2 < t < 0.48\,\mathrm{GeV}^2$.

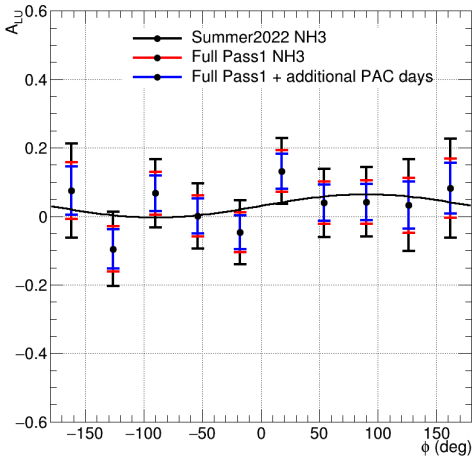


(d) $0.48\,\mathrm{GeV}^2 < t < 0.8\,\mathrm{GeV}^2$.

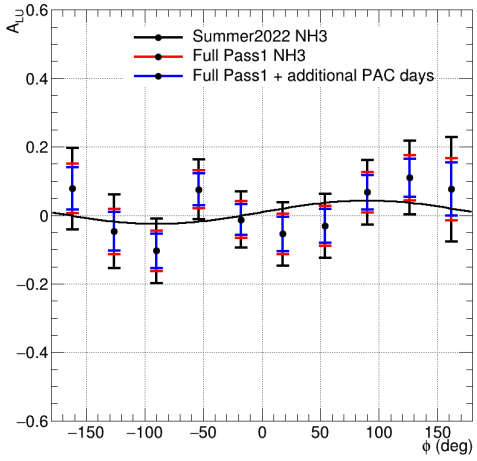
Figure 6.4: TSA in each bin in t.



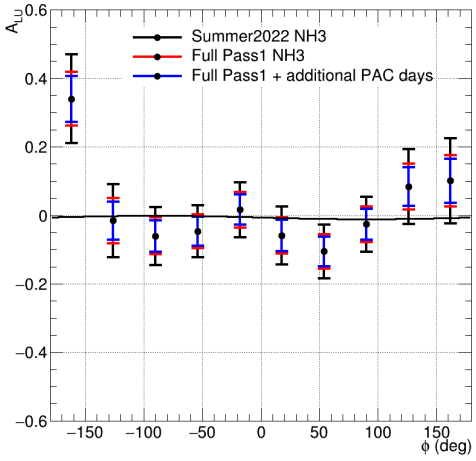
(a) $0.15\,\mathrm{GeV}^2 < t < 0.25\,\mathrm{GeV}^2$, $\mathrm{MM}^2 < |0.7\,\mathrm{GeV}^2|$.



(b) $0.25\,\mathrm{GeV}^2 < t < 0.34\,\mathrm{GeV}^2$, $\mathrm{MM}^2 < |0.7\,\mathrm{GeV}^2|$.

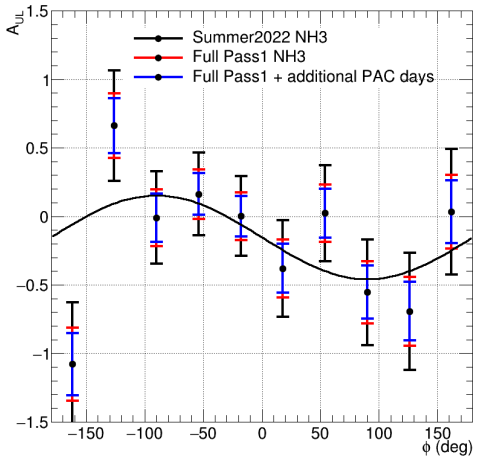


(c) $0.34\,\mathrm{GeV}^2 < t < 0.48\,\mathrm{GeV}^2$, $\mathrm{MM}^2 < |0.7\,\mathrm{GeV}^2|$.

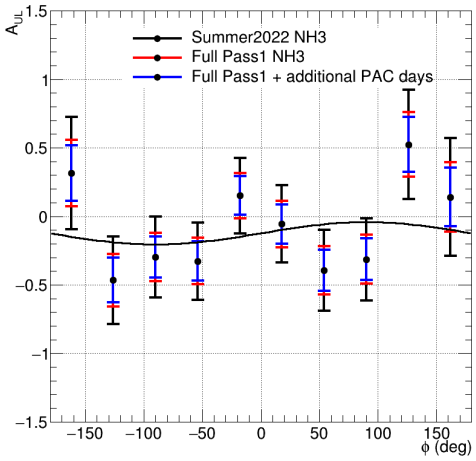


(d) $0.48\,\mathrm{GeV}^2 < t < 0.8\,\mathrm{GeV}^2$, $\mathrm{MM}^2 < |0.7\,\mathrm{GeV}^2|$.

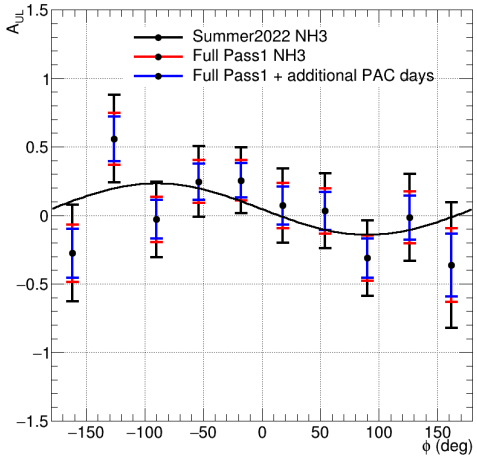
Figure 6.5: BSA in each bin in t and with a widened MM^2 cut



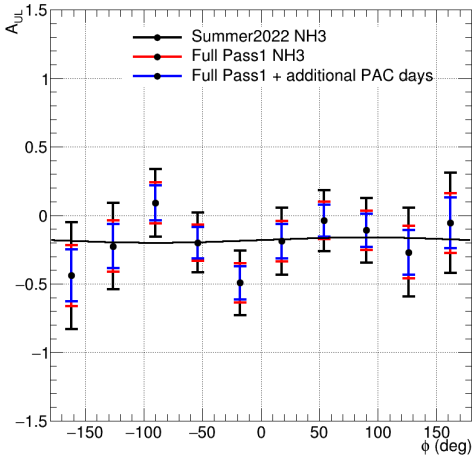
(a) $0.15\,\mathrm{GeV}^2 < t < 0.25\,\mathrm{GeV}^2$, $\mathrm{MM}^2 < |0.7\,\mathrm{GeV}^2|$.



(b) $0.25\,\mathrm{GeV}^2 < t < 0.34\,\mathrm{GeV}^2$, $\mathrm{MM}^2 < |0.7\,\mathrm{GeV}^2|$.

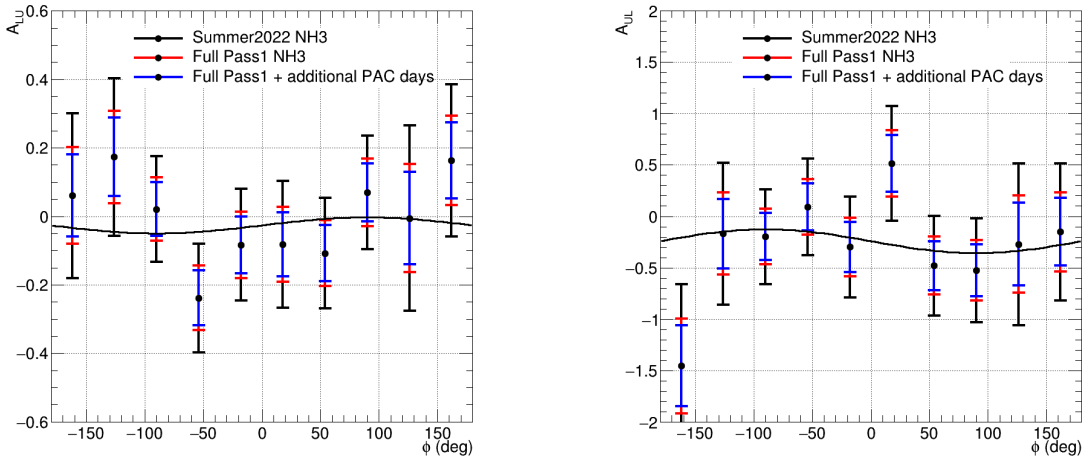


(c) $0.34\,\mathrm{GeV}^2 < t < 0.48\,\mathrm{GeV}^2$, $\mathrm{MM}^2 < |0.7\,\mathrm{GeV}^2|$.



(d) $0.48\,\mathrm{GeV}^2 < t < 0.8\,\mathrm{GeV}^2$, $\mathrm{MM}^2 < |0.7\,\mathrm{GeV}^2|$.

Figure 6.6: TSA in each bin in t and with a widened MM^2 cut.



(a) $0.48 < t < 0.8$, $MM^2 < |0.7|$, $IM > 2$ GeV.

(b) $0.48 < t < 0.8$, $MM^2 < |0.7|$, $IM > 2$ GeV.

Figure 6.7: BSA and TSA in the final bin in t, with a widened MM^2 cut and an additional tightened cut on Invariant mass.

6.2 TCS at the Electron-Ion Collider

Part of the additional work in this thesis served to explore the possibility of measuring the TCS cross-section at the Electron-Ion Collider, which is an upcoming high luminosity collider facility which will collide polarised electron beams with polarised protons and a range of polarised light ions. The work done by the author focused on developing feasibility studies for a potential detector design for the first interaction region of the EIC.

6.2.1 EIC Yellow Report

Initially, a toy Monte Carlo event generator was created, at the time referred to as eic-pi0-toy-MC, by Pawel Snadjer and Kemal Tezgin with the author contributing some TCS-specific kinematics. The generator was then used to simulate pure TCS and pure Bethe Heitler in a collider configuration, as opposed to a fixed target at JLab. The events were simulated at two different beam energies, 5×41 GeV (proton beam energy × electron beam energy) and 18×275 GeV. The following plots are the result of an analysis performed by the author, of the generated TCS events produced by eic-pi0-toy-MC, with particle momenta smeared using the package eic-smear, which is used to mimic realistic detector resolutions [100]. The following studies were also featured in the EIC Yellow Report, which was a document made to detail the physics outlook, early detector concepts, and opportunities available at the EIC, prior to its approval [3]. Figure 6.8 shows the distributions of the squared invariant mass of the lepton pair ($Q'^2 = e^+ + e^-$

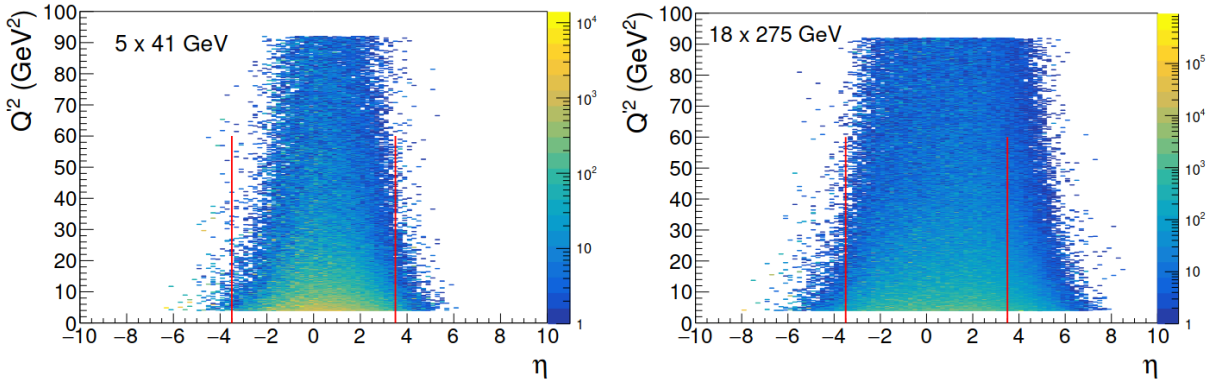


Figure 6.8: Distributions showing Q'^2 versus η for beam energies 5×41 GeV (left) and 18×275 (right). The red lines are indicative of the nominal acceptance region of the detector.

where e^+, e^- represent the four-momentum of the dilepton pair in the TCS final state), versus the pseudorapidity η, which represents the angle of detection relative to the beam axis;

$$\eta = -\ln(\tan(\frac{\theta}{2}))$$

where θ is measured in radians [101]. As θ goes to zero, pseudorapidity tends to infinity. It is often preferable to use pseudorapidity over θ due to its Lorentz invariance. The left plot shows the

distribution of collision energy 5×41 GeV, and the right shows the distribution for 18×275 GeV. The red lines marked on the distributions represent the nominal acceptance range of the detector at the time of the Yellow Report. It is clear that at lower collision energies (left) there is a small loss of lepton pair events for high values of η (very forward leptons) and at largest collision energies (right), it is again noted that very forward (high η) electrons are missed, however, this time these missed events span a much larger range in Q'^2, with a significant portion of events in the lower Q'^2 region where the TCS signal is most cleanly isolated ($2.25\,\text{GeV} < Q'^2 < 9\,\text{GeV}$). It was learned through this investigation that there must be a trade-off between overall yield and physical space required for detector components.

Figure 6.9 shows distributions of scattered proton transverse momentum versus η, again the left distribution corresponds to 5×41 GeV and the right to 18×275 GeV. Both distributions show that the proton scatters with a low transverse momentum, at low angles (high η), meaning that it remains fairly close to the beam line. Detection of the scattered proton is a crucial requirement for the reconstruction of t as discussed in Section 2.3, thus the detector setup at the EIC must account for detection of the final state particles close to the beam axis. To achieve this, Roman Pots could be used, which are silicon tracking stations installed close to the beam, inside the beam-pipe [102].

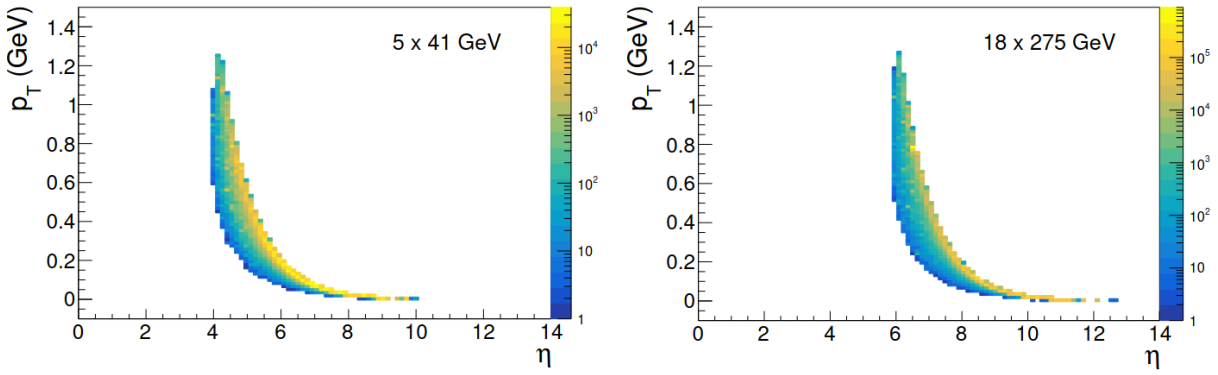


Figure 6.9: Distributions showing the transverse four-momentum of the scattered proton versus η for beam energies 5×41 GeV (left) and 18×275 GeV (right).

The natural progression of ascertaining the behaviour of the scattered proton is to ensure that its detection is, in fact, the best method of calculating t. To do this, the resolution on t, i.e Δt = Generated t - Reconstructed t was calculated. This was done in two ways, firstly, using the momentum information from the real and virtual photon, calculated via $(q - q')^2$ where, as a reminder, q represents beam electron - scattered electron (i.e. the four-momentum of the real photon) and q' represents the four-momentum of the produced virtual photon, calculated by adding the four momenta of the produced lepton pair. The results of this for beam energies 5×41 GeV are shown in the bottom plot of Figure 6.10. The purple distribution shows the full event yield, while the dark blue distribution shows the yield for which the proton was also recon-

structed. This distribution shows a wide sigma, meaning that the resolution on t ascertained from photon information is fairly poor. Another point to note is that for higher energies, reconstruction using photon information is not possible due to the fact that in the quasi-real regime, and at these collider kinematics, the scattered electron is backward, thus the quasi-real photon cannot be reconstructed without far backward, which were not present in this toy simulation. Moving on to the top two distributions of Figure 6.10 the resolution on t is shown again, however, this time calculated using momentum information from the initial and scattered proton, $(p' - p)^2$ for beam energies 5×41 GeV (left) and 18×275 GeV (right). It is immediately evident from the definition of the peak compared to the previous method, that the resolution is better defined with proton information, rather than via the photon. This again emphasises the necessity of detecting the scattered proton at collider kinematics as well as fixed target.

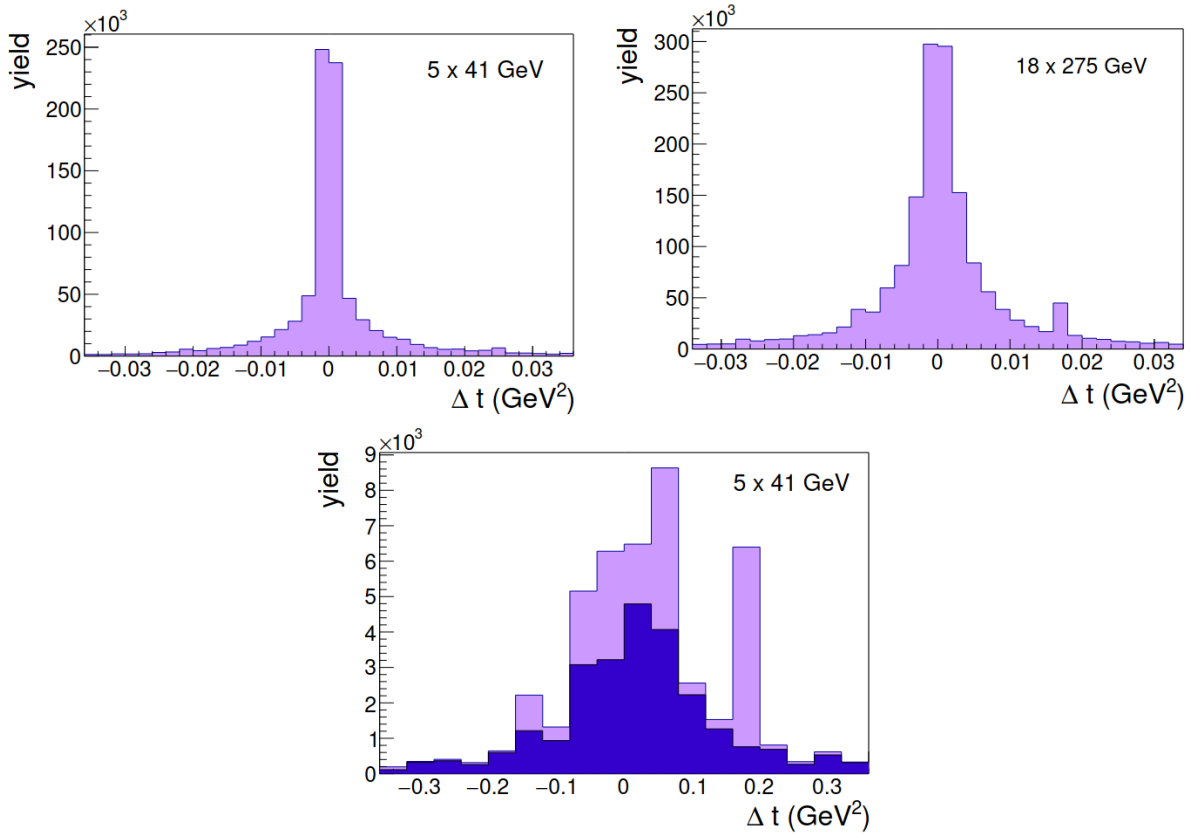


Figure 6.10: Top; Resolution on t using proton information ($t = (p' - p)^2$) at energy setting 5×41 GeV (Left) and 18×275 GeV (right). Bottom; Resolution on t using photon information ($t = (q - q')^2$), the dark blue region represents events where a proton was reconstructed in the final state [3].

A final study performed looked into the merits of detecting the scattered electron directly, as opposed to calculating it as the 'missing' component of the final state momentum reconstruction. The resulting plots are shown in Figure 6.11 across a range of η again for beam energies 5×41 GeV (left) and 18×275 GeV (right). The distribution in cyan represents the number of total scattered electron events, the distribution in yellow represents the number of events that

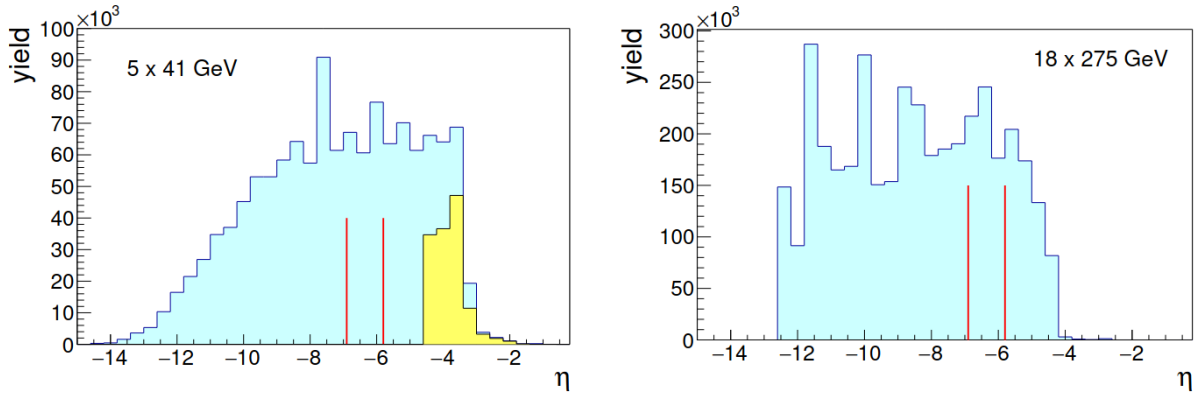


Figure 6.11: Distributions showing the total scattered electron events (cyan), indicating those picked up within the acceptance range of the detector (yellow) and highlighting the position of a low Q^2 tagger (red lines) for beam energies 5×41 GeV (left) and 18×275 GeV (right).

fall within the nominal acceptance range of the detector, and the red lines mark out the proposed region of a low Q^2 tagger at the time of the Yellow Report (2020). From these distributions, it can be discerned that for low energies, the events that fall into the detection region make up only a small fraction of the total number of events, and for high energies, this fraction is hardly even visible. This indicates that the best method of obtaining information on the scattered electron is via reconstruction through missing mass and momentum, much like is done for the JLab analysis, and, while the low Q^2 tagger will serve to increase statistics somewhat, it is not a necessity for this result. The next stage of the analysis was to introduce beam smearing and Bethe-Heitler interference to the generated events, and perform tests with a full simulation of the detector system, to ascertain a more realistic estimate of the physics capabilities of the EIC, and inform the design of its detectors.

6.2.2 The ECCE Detector

Following the Yellow Report [3] the Department of Energy (DoE) called for detector design proposals for the first interaction region of the EIC. The EIC Comprehensive Chromodynamics Experiment (ECCE) detector was an answer to this call, and was chosen as the reference detector in March 2022 [103]. The detector is now known by the name ePIC after the merging of the ECCE and ATHENA collaborative groups. For the purposes of this study, which was conducted by the author prior to the merge, and to avoid confusion with the EpIC generator (discussed in Section 6.2.3), it will continue to be referred to as ECCE.

Far Forward Instrumentation

As already established in this thesis, the detection of the dilepton pair and proton are crucial for extracting TCS final state kinematics. In the Yellow Report [3] studies showed that, as expected, at both low and high energy beam settings of the EIC, the proton scatters with a

low transverse momentum, at low polar angles or high pseudorapidity. To this end, the ECCE detector configuration includes a far forward region, which consists of a series of detectors aimed at detecting the scattered proton. The detector systems used to pick up the proton are the B0 and the Roman pots. The Roman Pots are detectors housed in cylindrical vessels, close to the beam line, and their main purpose was initially to measure the cross-section of the beam in the collider. They can also be used to detect high energy, charged particles, close to the beamline, making them a perfect candidate for detection of the scattered proton [102]. The B0 system is used to measure charged particles and tag neutral particles in the far forward region, with coverage in η at a more forward range than that of the Roman Pots and will aid in further detection of far forward protons. [104].

Central Detector

The Central Detector (CD) in the ECCE configuration proposed re-using the BaBar superconducting solenoid, which will provide a 1.5 T field [105]. Within this central region there are three components, the barrel, which covers the central pseudorapidity region $-1 < \eta < 1$, and forward and backward endcaps, covering $1 < \eta < 3.5$ and $-1 < \eta < -3.5$ respectively. Tracking in the ECCE CD is achieved using multiple technologies, the Monolithic Active Pixel Sensor (MAPS) based silicone vertex system, the μRWELL gas tracking subsystem and an AC-LGAD (Low-Gain Avalanche Diode) for Time Of Flight (TOF) measurements [106].

6.2.3 Simulated Studies

The ECCE detector configuration was simulated using Fun4All software [107] which is a GEANT4 based model of each subsystem. This software used to conduct a feasibility test of measuring TCS at the EIC. The first stage in this feasibility study was to generate a series of TCS Monte Carlo events, that could be passed through the detector simulation to ascertain detector performance for measurement and reconstruction of TCS observables. The generator used for this purpose was the EpIC generator [108], which was developed from the toy generator in Section 6.2.1. EpIC is a Monte Carlo event generator which utilises existing GPD models from the Partonic Tomography Of Nucleon Software (PARTONS) [109] framework and is used to generate cross-sections for a given set of kinematics. This framework is used with the mini FOAM library, which is a general-purpose MC event simulator integrated with ROOT, and together they generate random events in phase space. The algorithm behind mFOAM has a two stage process, exploration and generation. In the exploration stage, the shape of the distribution function is "explored" using MC methods to divide the integration domain into a series of cells, or "foam". The generation stage takes a random cell, and a point in this cell is generated and weighted, this goes on to be the MC ‘event’ [110]. In terms of what the user sees, EpIC takes in an input .xml file, within which experimental parameters are defined, for example, helicity,

polarisation, event number, etc. and it generates the four vectors of all the particles as the output.

The generator configuration used for the TCS events was as follows;

- Event number for each energy setting $= 1,000,000$ to give a good number of statistics. Cross-section scaling factors 0.34 nb for energy 5×41 GeV and 0.97 nb for energy 18×275 GeV were taken into account using additional factors at the final plotting stage.
- FOAM nCells $= 15,000$ defining the maximum number of 'foam' cells in the integration domain.
- FOAM nSamples $= 3,000$ defining the number of Monte Carlo events to be used in the cell exploration stage.
- FOAM nBins $= 3,000$ defining the number of bins in the edge-histogram in the cell exploration stage.
- Decay mode $= e^-$ indicating an electron-positron decay channel
- Helicity $= -1$ both helicity settings were generated but due to time constraints, only the negative helicity setting was analysed.

The kinematics in EpIC were set to either match or slightly extend upon the original kinematics studied in the Yellow Report, and are detailed below;

- $0\,GeV^2 < t < 1\,GeV^2$
- $2\,GeV^2 < Q'^2 < 20\,GeV^2$ to ensure a hard scale for the scattering and to minimise background from the low resonance region.
- $0.01 < y < 0.95$
- $0 < \phi < 2\pi$ to obtain a full lab frame azimuthal (ϕ) angular coverage.
- $0 < \phi_S < 2\pi$, (where ϕ_S represents the angle between the leptonic plane, see Figure 2.8c, and the transverse component of the polarisation of the target nucleon), to obtain a full ϕ_S angular coverage.
- $\frac{\pi}{6} < \theta < \frac{5\pi}{6}$,

The final state particles detected in the Fun4All simulation of the ECCE detector were;

- Scattered proton using truth (generated) momenta matched to hits in Roman Pots and B0 with cuts on both to separate out the beam.

- Decay lepton pair: Truth momenta associated with hits in the Endcap Electromagnetic Calorimeter (EEMC), Forward Electromagnetic Calorimeter (FEMC) and BECAL and separated by PID taken from calorimeter cluster information.

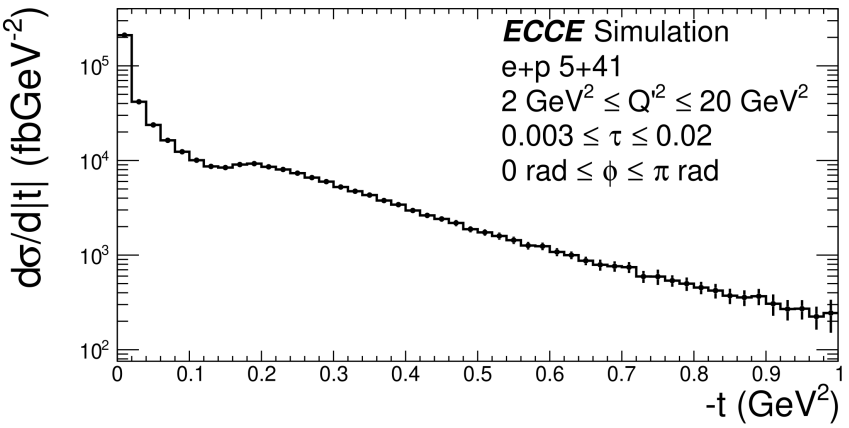
- Scattered electron: Calculated as opposed to being detected. This is due to the original findings of the YR study, which indicated that the e' would be difficult to directly detect, without implementation of a low Q^2 tagger, and that it is instead better to use the momenta of other final state particles compared to initial beam momenta to calculate it [3]. In more recent versions of Fun4All, there has been an integration of a low Q^2 tagger, which would mean that a missing mass study could be performed with the scattered proton as the 'missing' particle, as it is within detector resolution to calculate this at high energies, however, this will be a low statistics measurement.

The following shows some key plots produced by the author that were featured in the ECCE design proposal and in the subsequent Exclusive Diffractive and Tagging publication, where a collaborative effort of physicists working on exclusive, diffractive and tagging physics channels published their feasibility studies using the ECCE design concept [111]. One of the key parameters t, was calculated as $t = (p' - p)^2$ after the findings of the Yellow Report indicated that the use of photon momenta to calculate this led to poorer resolution, see Figure 6.10. The plot in Fig. 6.12a and 6.12b show the projected cross-section for t, averaged over Q'^2, τ and ϕ. The cross-section was normalised with an integrated luminosity factor $10\,\text{fb}^{-1}$. The full equation to calculate the differential cross-section is;

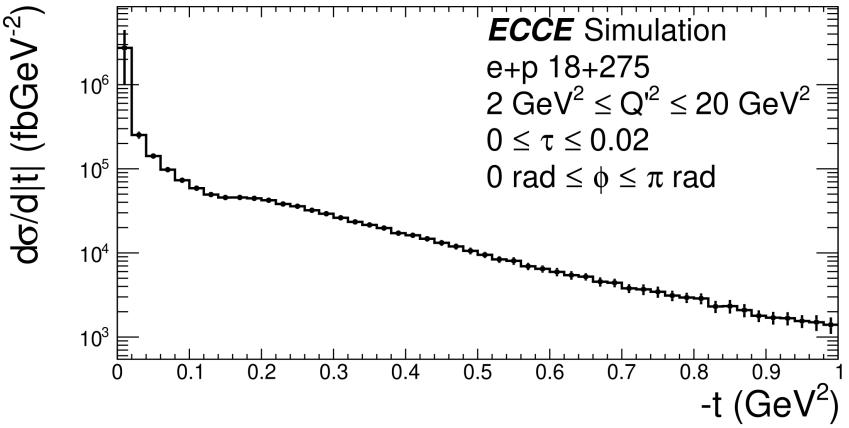
$$\frac{d^4\sigma}{dt} = \frac{1}{\mathscr{L} \cdot Acc_{bin} \cdot \Delta\Omega}(N \pm \sqrt{N}), \tag{6.1}$$

where $\mathscr{L}$ is the integrated luminosity and in this case is equal to $10\,\text{fb}^{-1}$, representing a year of operation at the EIC. Acc_{bin} represents the bin acceptance, calculated via Reconstructed data/Generated data for each bin. $\Delta\Omega$ is the multidimensional bin width, calculated as as $\Delta\Omega = \Delta Q^2 \Delta\ x_B \Delta t \Delta\phi_h$, using information from each bin width and range. Finally, N is the number of counts in the bin. The cross-section plots align with expectations in literature, for a t cross-section measurement of TCS to be around 1-10 nb [32], see Figure 6.13. The distributions show this for both energy settings. This gave confidence in the ability of ECCE to detect the scattered proton with high enough efficiency to reconstruct this variable across its full range.
Figures 6.14 show some preliminary pseudorapidity plots that describe the acceptance of the scattered proton in each forward detector region. It appears from these plots that for the low energy setting, 5×41 GeV, many counts fall within $4.2 \leq \eta \leq 6.2$, which indicates the B0 is the key subsystem for the reconstruction of the scattered proton at these energies. We also see a substantial number of counts within the region $6.2 \leq \eta \leq 8$, within the acceptance of the Roman Pots. Looking at the higher energy setting, we see no hits at all in the η region covered by the

B0, instead all counts are in the Roman Pots region. This emphasises the requirement of this system for 18×275 GeV energies.



(a) $5\times$ 41 GeV.



(b) $18\times$ 275 GeV.

Figure 6.12: TCS Differential cross-section versus the momentum transfer to the struck parton t reconstructed using the beam and scattered protons $t = (p' - p)^2$.

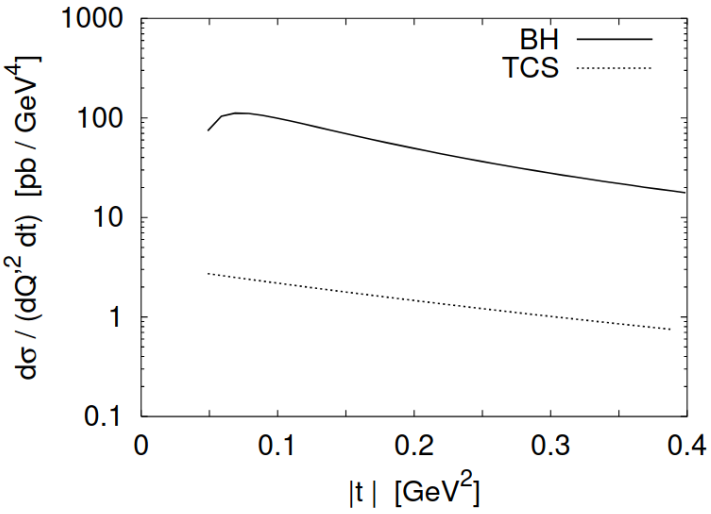
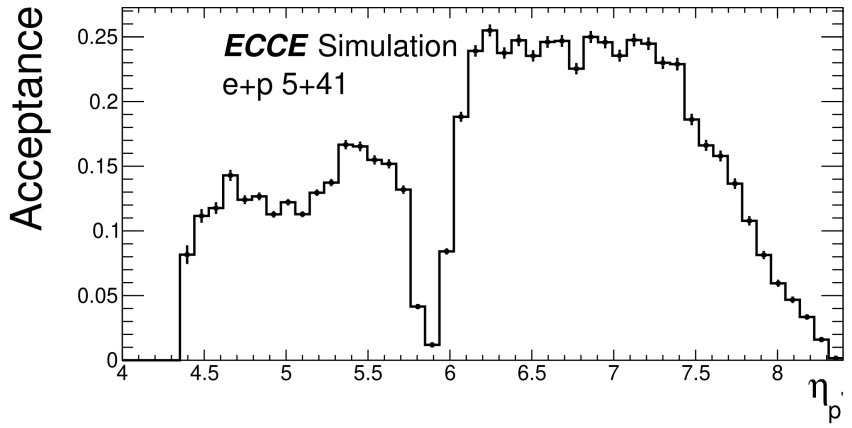
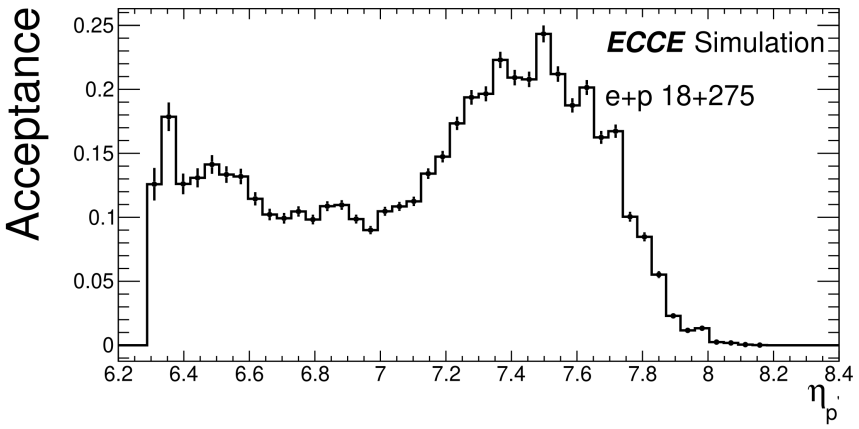


Figure 6.13: Theoretical TCS cross section (dashed) as compared with BH cross section (solid). Taken from [32].



(a)



(b)

Figure 6.14: Top: 5×41 GeV acceptance vs pseudorapidity (η) of the scattered proton from TCS events. Bottom: 18×275 GeV acceptance vs pseudorapidity (η) of the scattered proton. Note acceptance is given as a value where 1 corresponds to 100%

As a further element to the study, the author generated 50,000 muon TCS events to allow exploration into the di-muon channel and pion background contamination. The muon channel would be advantageous over the electron channel, as it is difficult to separate the scattered electron from background electroproduction processes from the electron produced after the decay of the virtual photon in the TCS process. With the di-muon channel this separation is not required, however, this would require a dedicated μ detector system and therefore would be more difficult to measure. The distributions for this were featured in an as yet unpublished study related to AI techniques in background processes [112].

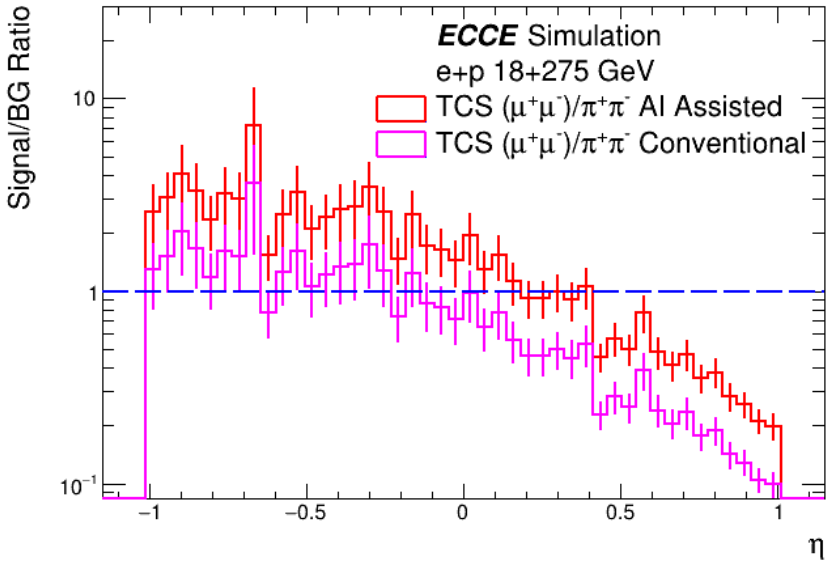


Figure 6.15: Signal to background ratio for muon TCS events using AI assisted tracking versus conventional tracking [112].

The conclusion drawn from this was that the ECCE detector was a suitable design for a potential measurement of TCS at the EIC. It showed the capability of reconstruction of a full range of t with a fair acceptance for each variable at energies 5×41 GeV, and although acceptance showed a slight decline at higher energy settings, this could be supplemented with further statistics, a low Q^2 tagger with which to detect the scattered electron, and improved RP configurations for better scattered proton detection. Since the publication of the ECCE proposal in 2022 many improvements have been made to the simulation and reconstruction algorithms in the software, so it is expected that these preliminary results could show much improvement in a future study.

Chapter 7

Conclusion

To summarise the work done in this thesis, a novel extraction of a Beam Spin asymmetry and Target Spin asymmetry was performed on a longitudinally polarised, NH_3 target, with the CLAS12 Detector at Jefferson Lab. The author played an active role in both the data taking and processing since this run occurred during the time period of this PhD. The slight downside to this however is that processing algorithms and reconstruction software are constantly being improved and revisited, with the goal of providing better final state reconstruction and tracking of particles through CLAS12, and therefore the dataset the results are extracted on at this stage is relatively fresh, what is referred to in the collaboration as a 'Pass1', and is also incomplete, as there are two more periods worth of data taking still to be processed. Since TCS is a statistically hungry channel, this affects the overall quality of the extracted observables, meaning one must favour maximising statistics over taking a precision measurement. This means that in the final extracted observable there will still be sources of background contamination, the key area of this will be from the nuclear target. A good way to minimise this is, once all of the data is available, is to recalculate the dilution factor with a higher degree of precision from the increased statistics, and improve the exclusivity cuts. This will not only decrease the errors in the extraction of the TSA, but will also serve to provide a better understanding of the total contribution from each element of the nuclear background (nitrogen specifically, foils, liquid helium etc). A secondary way to improve understanding of the final state and improve confidence in the final result is to perform a full simulation of TCS events with the addition of a nuclear target. The TCSGen generator is currently undergoing implementation of a Helium target, and there are plans to extend this to allow simulations on the deuteron, and potentially this can be extended to more nuclear targets. This will allow effective comparison of the data and expected distributions, to fully quantify the nuclear backgrounds affecting the reconstruction of the TCS final state.

Appendix A

The following details a study conducted by the author of the beam polarisation throughout the full run period.

A.1 Beam Polarisation

Firstly, the recorded beam polarisation obtained from each Møller run, along with its associated error, for the full set of data, accounting for Summer, Fall and Winter periods, is reported in Figure A.1.

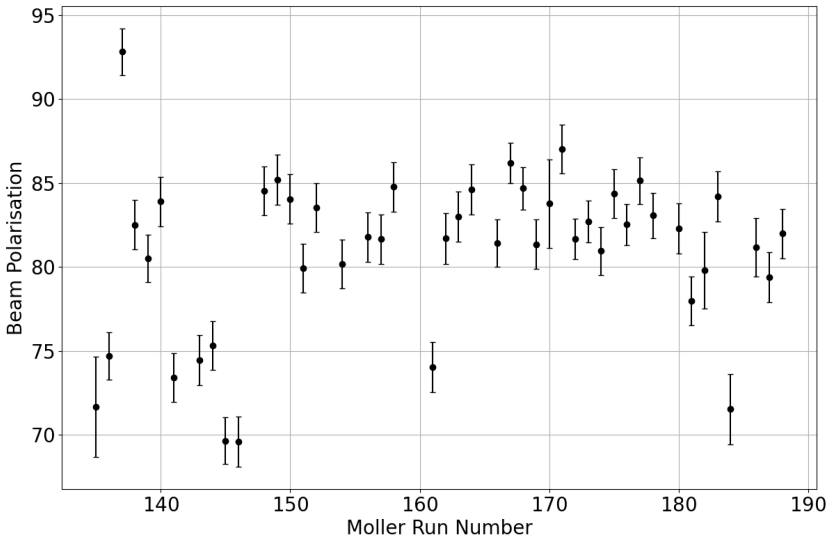


Figure A.1: Beam Polarisation as a function of Møller run number for all run periods of RGC (Summer, Fall and Winter). Errors marked on the plot are the absolute errors on the Beam Polarisation specific to each run.

Next, these Møller runs were filtered by the conditions listed in Table A.1, which details the requirements for a Møller run to be considered 'good'. The Beam Charge Asymmetry values indicate if there is an issue independent of the Møller system, i.e. related to the accelerator systems and the Accidental/Coincidence ratio (a measure of signal-to-background ratio on the

Variable	Condition
2C21 BPM Current	$\approx$ 15nA
Beam Charge Asymmetry	$< 0.2\%$ (typical $\approx 0.1\%$)
Accidental/Coincidence Ratio	< 0.1
Final Beam Polarisation Error	$< 1.5\%$ (absolute)

Table A.1: Table of conditions for a good Møller run.

coincident electron pair measured by the two polarimeter detectors, typically the background rates are less than 10 % of the true coincidence rate [53]) is a beam quality test used to ensure there is limited bleed through from other halls, and the left and right Møller PMTs are set at the correct voltages.

Figure A.2 shows the Beam Charge Asymmetry (BCA) of all Møller runs, with bands added to designate the acceptance threshold of 0.2 % marked in red, and a tighter set of bands used to indicate the typical expected values of 0.1 % marked in blue.

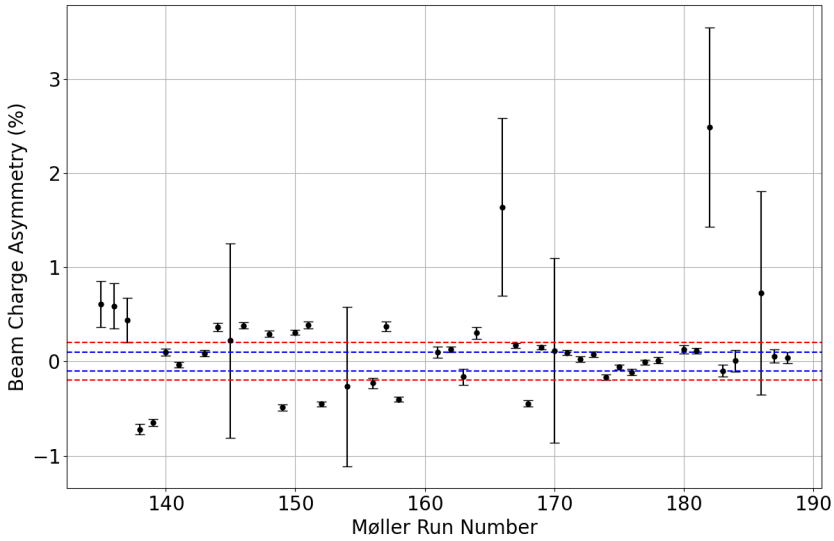
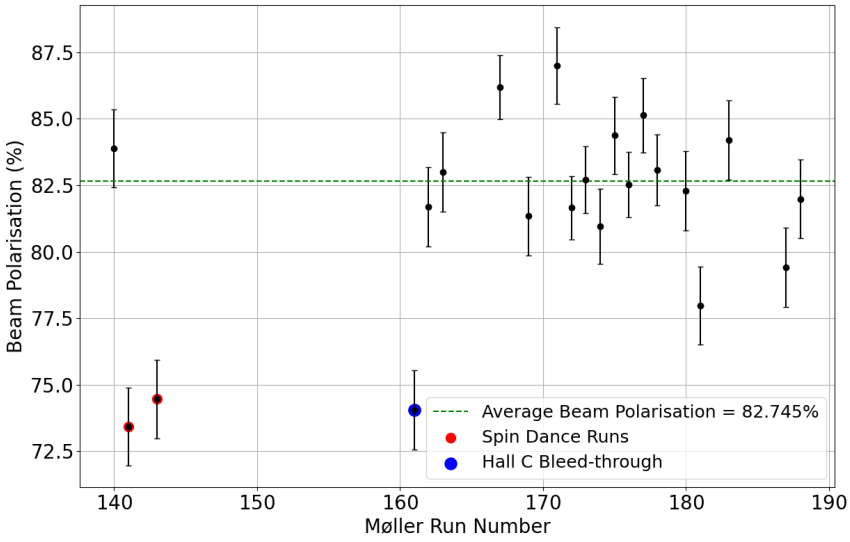


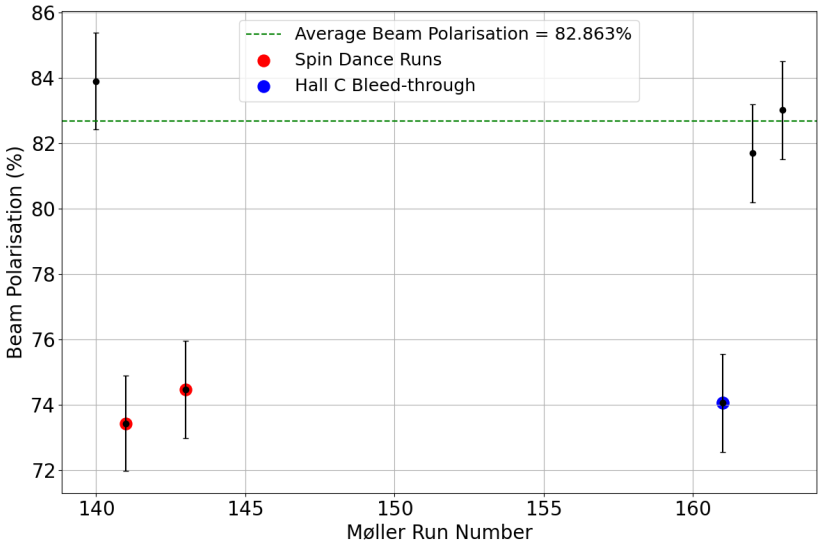
Figure A.2: Beam charge asymmetry as a function of run number for the Møller data. Red dashed lines mark the $\pm 0.2\%$ threshold required to designate a 'good' Møller Run, and the dashed blue lines mark the $\pm$ 0.1% threshold.

Applying the restriction of BCA $< 0.2\%$ to the Møller runs, the Beam Polarisation as a function of run number is shown in Figure A.3a. The runs from just the Summer 2022 period (this thesis) that meet this condition, are shown in Figure A.3b.

Since the Møller runs are taken during the experiment, there can be a number of issues with the reported values of BCA, A/C ratio, and error, which can be due to the beam quality, the Møller system, or simply human error in recording the values (stopping a Møller acquisition run to early or late, taking a run during unstable beam etc.). For this reason, it was advised to widen



(a)

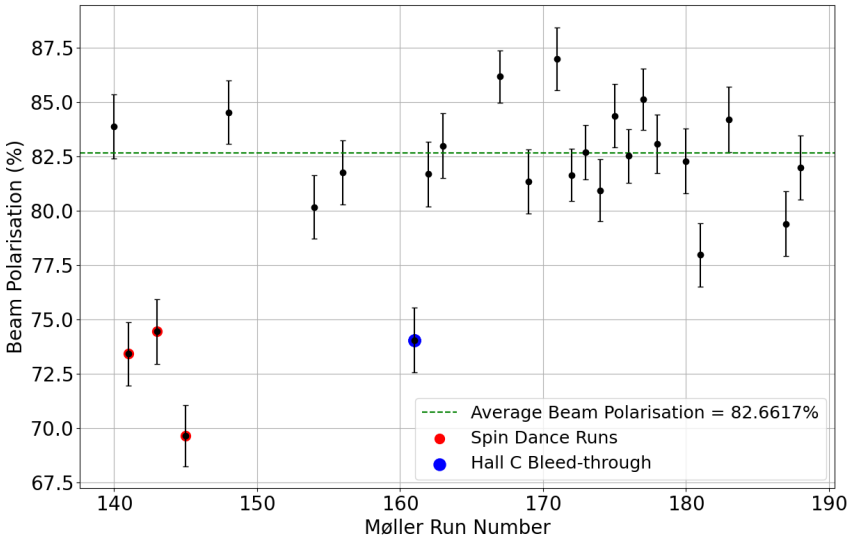


(b)

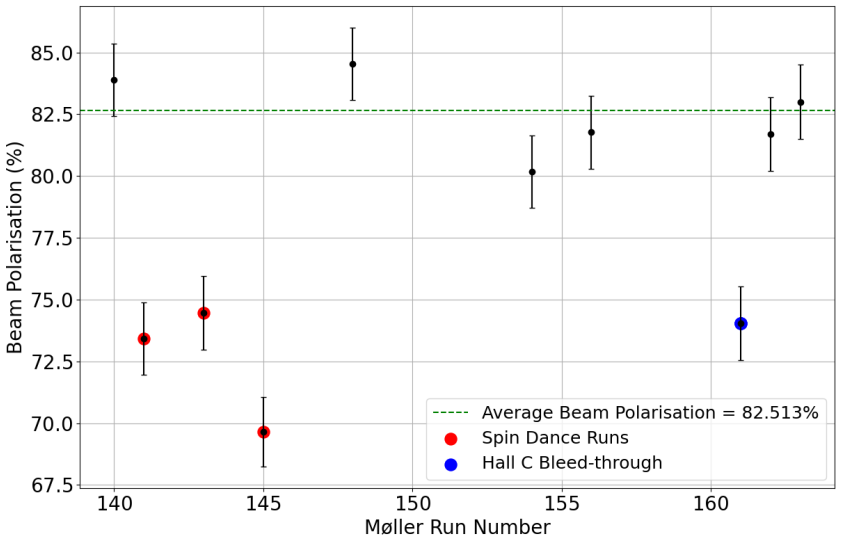
Figure A.3: Figure A.3a shows the Beam Polarisation as a function of run number for the full RGC run period after the condition of Beam Charge Asymmetry < ± 0.2 shown in Table A.1. Runs taken during a spin dance are shown with a red marker, and runs where Hall-C had excessive beam bleed-through are shown in blue. The average polarisation after bad runs are filtered out is shown in green. Figure A.3b shows the same, but restricted to only the Summer 2022 run period.

the BCA condition to < ± 0.3 % and Pbeam error < ± 1.5 % to ensure that the Møller runs are not being filtered out due to these issues. The Beam Polarisation as a function of run number after applying this condition is shown in Figure A.4a, and for just the Summer 2022 period in

Figure A.4b. There are now more data points available to confirm the beam polarisation is stable throughout the full run period and has a value of approximately 82.7 % ±0.2 %.



(a)

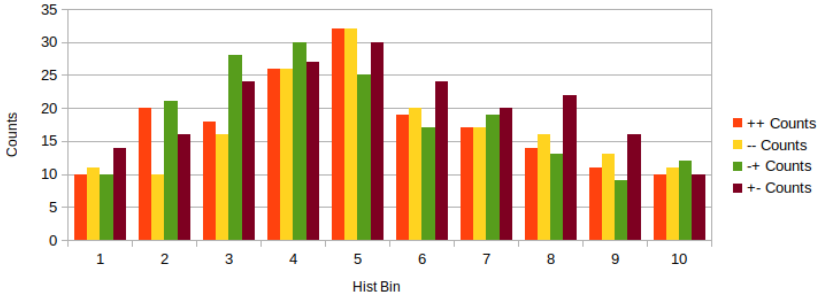


(b)

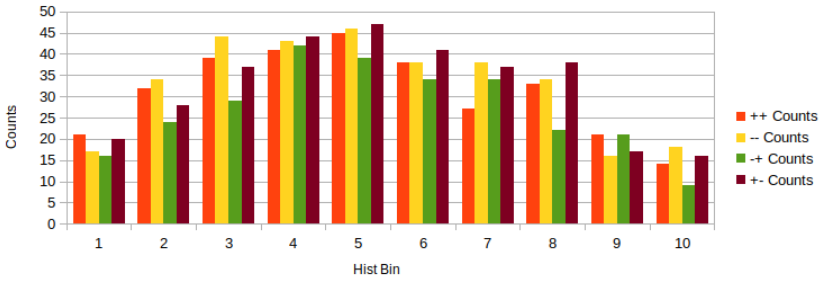
Figure A.4: Figure A.4a shows the Beam Polarisation as a function of run number for the full RGC run period after the condition of Beam Charge Asymmetry widened to $< \pm 0.3$ shown in Table A.1. Runs taken during a spin dance are shown with a red marker, and runs where Hall-C had excessive beam bleed-through are shown in blue. The average polarisation after bad runs are filtered out is shown in green. Figure A.4b shows the same, but restricted to only the Summer 2022 run period.

A.2 Rebinned Asymmetry

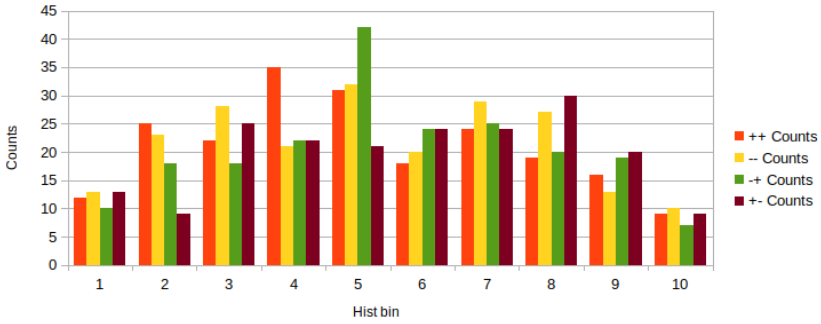
This study looked to observe the distribution of events in each helicity and polarisation configuration that went into the calculation of the final observable. The distribution of counts is shown in figure A.5, for each of the four final state t bins. The original reasoning for a 10-bin split was to directly compare with the previously published results, as the final state statistics were similar. Consolidating these counts into six bins rather than 10 should serve to reduce the size of the error bars, and potentially reveal a clearer shape to the asymmetry. The results of this for the two tbins that a result was able to be extracted in, from the thesis, are shown in figure A.6 and the extrapolated figures in A.7. Looking at the fit parameters, the thesis showed for the first tbin $p1 = -0.3941 \pm 0.1917$ and here $p1 = -0.379 \pm 0.193$. For the third tbin the thesis showed $p1 = -0.2694 \pm 0.1514$ and here $p1 = -0.2561 \pm 0.1521$ so in neither instance has there been a real noticeable change in the preliminary TSA extraction.



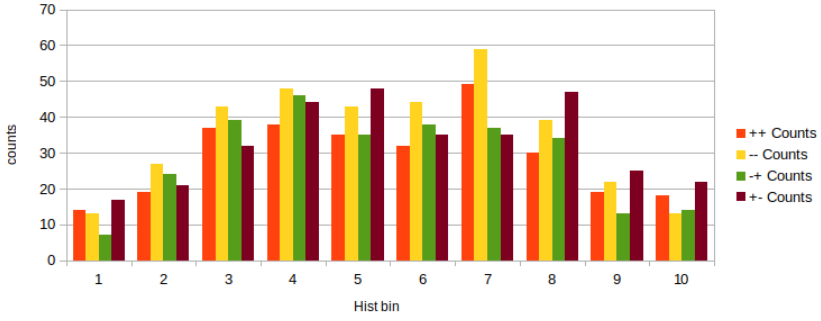
(a) tbin1



(b) tbin2



(c) tbin3



(d) tbin4

Figure A.5: Distributions showing the counts in each N^{ij} bin, split into the four final state t bins

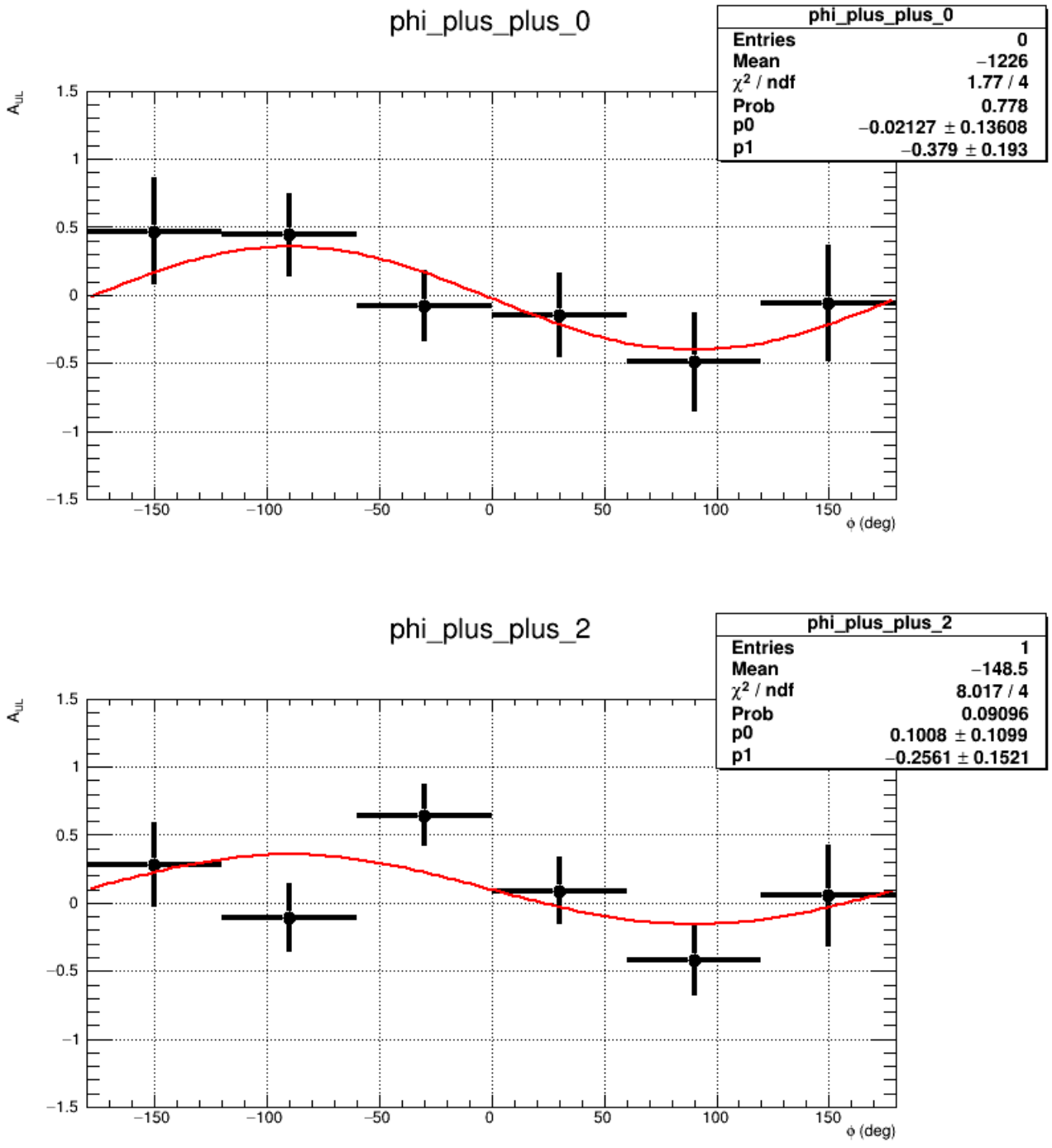
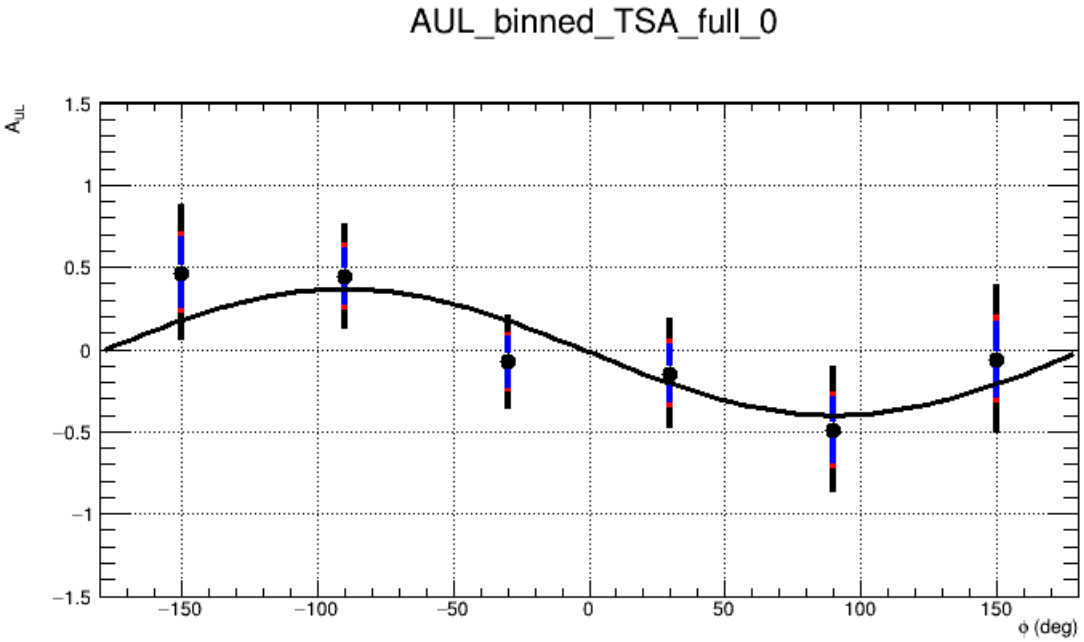


Figure A.6: TSA 6 bins



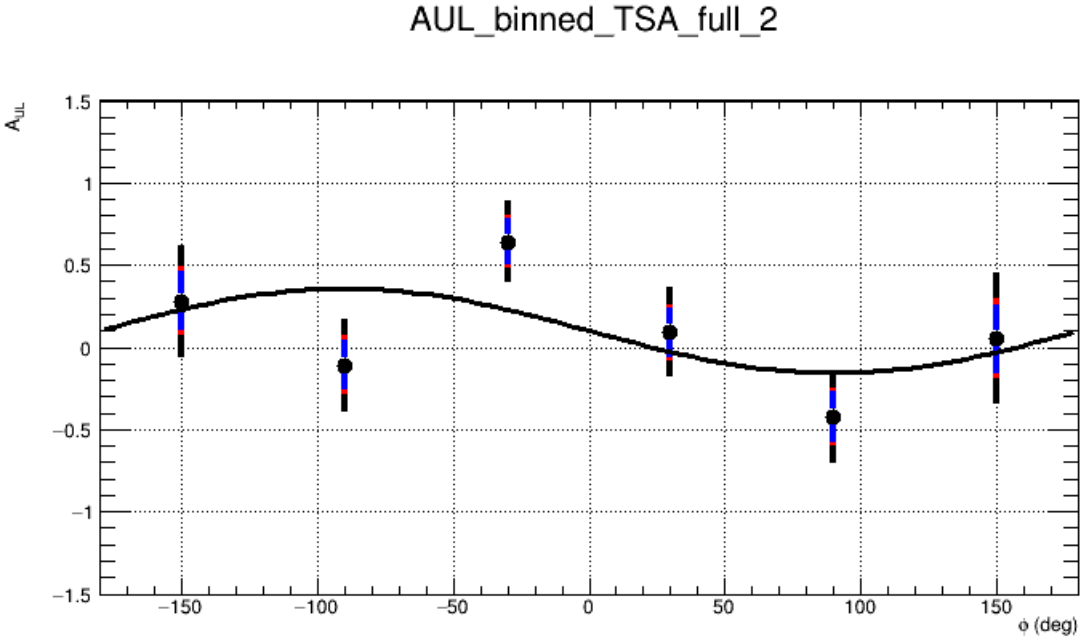


Figure A.7: TSA extrapolated

Bibliography

[1] R. L. Workman et al. Review of Particle Physics. *PTEP*, 2022:083C01, 2022.

[2] A. Accardi et al. Electron Ion Collider: The Next QCD Frontier. *Eur. Phys. J.*, A52(9):268, 2016.

[3] R. Abdul Khalek, A. Accardi, et al. Science Requirements and Detector Concepts for the Electron-Ion Collider. *Nuclear Physics A*, 1026:122447, oct 2022.

[4] L. Adamczyk et al. Precision Measurement of the Longitudinal Double-spin Asymmetry for Inclusive Jet Production in Polarized Proton Collisions at $\sqrt{s} = 200$ GeV. *Phys. Rev. Lett.*, 115(9):092002, 2015.

[5] C. Alexandrou, S. Bacchio, M. Constantinou, J. Finkenrath, K. Hadjiyiannakou, K. Jansen, G. Koutsou, H. Panagopoulos, and G. Spanoudes and. Complete flavor decomposition of the spin and momentum fraction of the proton using lattice QCD simulations at physical pion mass. *Physical Review D*, 101(9), may 2020.

[6] Daniel de Florian, Rodolfo Sassot, et al. Extraction of spin-dependent parton densities and their uncertainties. *Physical Review D*, 80(3), aug 2009.

[7] Astrid Morreale. Nuclear physics at the energy frontier: recent heavy ion results from the perspective of the Electron Ion Collider. *Universe*, 5(5):98, 2019.

[8] Arts at MIT and Jefferson Lab. Visualising the Proton. `https://www.youtube.com/watch?v=e2FrALuacZ4&t=11s`, Oct. 2022.

[9] H. Moutarde, P. Sznajder, and J. Wagner. Border and skewness functions from a leading order fit to DVCS data. *The European Physical Journal C*, 78(11):890, Nov 2018.

[10] Elke-Caroline Aschenauer, Salvatore Fazio, Jason Lee, Heikki Mäntysaari, B Page, B Schenke, T Ullrich, R Venugopalan, and P Zurita. The electron-ion collider: Assessing the energy dependence of key measurements. *Reports on Progress in Physics*, 82, 01 2019.

[11] Cédric Mezrag. An Introductory Lecture on Generalised Parton Distributions. *Few Body Syst.*, 63(3):62, 2022.

[12] David Griffiths. *Introduction to elementary particles*. John Wiley & Sons, 2020.

[13] M Thompson. *Modern Particle Physics*. Cambridge University Press, University Printing House, Cambridge, CB28BS, United Kingdom, 2013.

[14] E. Jones and R. Childers. *College Physics*. Addison-Wesley, Massachusetts, United States, 1990.

[15] H. Young, R. Freedman, and A. Ford. *University Physics*. Pearson, New York, United States, 2012.

[16] R. C. Walker, B. W. Filippone, et al. Measurements of the proton elastic form factors for $1 \leq Q^2 \leq 3$ $(\mathrm{GeV}/c)^2$ at SLAC. *Phys. Rev. D*, 49:5671–5689, Jun 1994.

[17] E. D. Bloom, D. H. Coward, H. DeStaebler, J. Drees, G. Miller, L. W. Mo, R. E. Taylor, M. Breidenbach, J. I. Friedman, G. C. Hartmann, and H. W. Kendall. High-energy inelastic $e-p$ scattering at 6° and 10°. *Phys. Rev. Lett.*, 23:930–934, Oct 1969.

[18] M. Breidenbach, J. I. Friedman, H. W. Kendall, E. D. Bloom, D. H. Coward, H. DeStaebler, J. Drees, L. W. Mo, and R. E. Taylor. Observed behavior of highly inelastic electron-proton scattering. *Phys. Rev. Lett.*, 23:935–939, Oct 1969.

[19] Johannes Blümlein. The theory of deeply inelastic scattering. *Progress in Particle and Nuclear Physics*, 69:28–84, March 2013.

[20] Prof.M. Thompson. Particle physics. https://www.hep.phy.cam.ac.uk/ thomson/lectures/partIIIparticles/Handout6_2009.pdf, accessed: 24/09/2024, 2009.

[21] F. Halzen and A. D. Martin. *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. John Wiley and Sons, New York, United States, 1984.

[22] G. M. Prosperi, M. Raciti, and C. Simolo. On the running coupling constant in QCD. *Progress in Particle and Nuclear Physics*, 58:387–438, 7 2006.

[23] Alexandre Deur, Stanley J. Brodsky, and Guy F. de Teramond. The QCD Running Coupling. *Nucl. Phys.*, 90:1, 2016.

[24] Particle Data Group and Workman, R L. and others. Review of Particle Physics. *Progress of Theoretical and Experimental Physics*, 2022(8):083C01, 08 2022.

[25] V. D. Burkert, L. Elouadrhiri, F. X. Girod, C. Lorce, P. Schweitzer, and P. E. Shanahan. Colloquium: Gravitational form factors of the proton, 2024.

[26] Maxim V. Polyakov and Peter Schweitzer. Forces inside hadrons: Pressure, surface tension, mechanical radius, and all that. *International Journal of Modern Physics A*, 33(26):1830025, September 2018.

[27] V. D. Burkert, L. Elouadrhiri, and F. X. Girod. The pressure distribution inside the proton. *Nature*, 557(7705):396–399, May 2018.

[28] Ringaile Placakyte. Parton Distribution Functions. In *Proceedings, 31st International Conference on Physics in collisions (PIC 2011): Vancouver, Canada, August 28-September 1, 2011*, 2011.

[29] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *European Physical Journal C*, 63:189–285, 1 2009.

[30] M. Defurne et al. A glimpse of gluons through deeply virtual compton scattering on the proton. *Nature Communications*, 8(1408), 2017.

[31] Satvir Kaur and Harleen Dahiya. Study of kaon structure using the light-cone quark model. *Phys. Rev. D*, 100:074008, Oct 2019.

[32] E. R. Berger, M. Diehl, and B. Pire. Timelike Compton scattering: Exclusive photoproduction of lepton pairs. *European Physical Journal C*, 23(4):675–689, Apr 2002.

[33] Markus Diehl. Introduction to GPDs and TMDs. *The European Physical Journal A*, 52(6), Jun 2016.

[34] A.V. Belitsky and A.V. Radyushkin. Unraveling hadron structure with generalized parton distributions. *Physics Reports*, 418(1–6):1–387, October 2005.

[35] M Diehl. Generalized parton distributions. *Physics Reports*, 388:41–277, 2003.

[36] Michel Guidal, Hervé Moutarde, and Marc Vanderhaeghen. Generalized Parton Distributions in the valence region from Deeply Virtual Compton Scattering. *Rept. Prog. Phys.*, 76:066202, 2013.

[37] A. V. Radyushkin. Double distributions and evolution equations. *Physical Review D*, 59(1), December 1998.

[38] D. Müller, D. Robaschik, et al. Wave Functions, Evolution Equations and Evolution Kernels from Light-Ray Operators of QCD. *Fortschritte der Physik/Progress of Physics*, 42(2):101–141, 1994.

[39] M. Vanderhaeghen, P. A. M. Guichon, and M. Guidal. Hard Electroproduction of Photons and Mesons on the Nucleon. *Phys. Rev. Lett.*, 80:5064–5067, Jun 1998.

[40] S. V. Goloskokov and P. Kroll. The target asymmetry in hard vector-meson electroproduction and parton angular momenta. *The European Physical Journal C*, 59(4):809–819, December 2008.

[41] T. Gehrmann and M. Stratmann. The Bethe-Heitler process in polarized photon - nucleon interactions. *Phys. Rev.*, D56:5839–5843, 1997.

[42] P. Chatagnon et al. First Measurement of Timelike Compton Scattering. *Phys.Rev.Lett.*, 127:262501, Dec 2021.

[43] T. Horn, Y. Illieva, F. J. Klein, P. Nadel-Turonski, R. Paremuzyan, and S. Stepanyan. Timelike Compton Scattering. *AIP Conference Proceedings*, 1374(1):542–545, 10 2011.

[44] S. Chen et al. Measurement of deeply virtual compton scattering with a polarized-proton target. *Physical Review Letters*, 97(7), August 2006.

[45] S. Pisano et al. Single and double spin asymmetries for deeply virtual compton scattering measured with clas and a longitudinally polarized proton target. *Physical Review D*, 91(5), March 2015.

[46] Angela Biselli. Studies of gpds at jefferson lab: results and future experiments. *Journal of Physics: Conference Series*, 938:012003, 12 2017.

[47] CLAS Collaboration, G. Christiaens, M. Defurne, D. Sokhan, et al. First clas12 measurement of dvcs beam-spin asymmetries in the extended valence region, 2022.

[48] H. Moutarde, B. Pire, et al. Timelike and spacelike deeply virtual Compton scattering at next-to-leading order. *Phys. Rev. D*, 87(5):054029, 2013.

[49] Marie Boër, Michel Guidal, and Marc Vanderhaeghen. Single and double polarization observables in timelike Compton scattering off proton, 2015.

[50] Reza Kazimi. Simultaneous Four-hall Operation for 12 GeV CEBAF. In *4th International Particle Accelerator Conference*, page THPFI091, 2013.

[51] Andrea Celentano. The BDX experiment at Jefferson Laboratory. *EPJ Web of Conferences*, 96:01006, 01 2015.

[52] V.D. Burkert, L. Elouadrhiri, et al. The CLAS12 Spectrometer at Jefferson Laboratory. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 959:163419, 2020.

[53] N. Baltzell, V.D. Burkert, J. Carvajal, N. Dashyan, R. De Vita, L. Elouadrhiri, G. Kharashvili, A. Kim, R. Paremuzyan, B.A. Raue, Y.G. Sharabian, S. Stepanyan, M. Tiefenback, M. Ungaro, and K. Wild. The CLAS12 beamline and its performance. *Nuclear*

Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 959:163421, 2020.

[54] Stepan Stepanyan. The HPS Beamline Manual, July 2021.

[55] N Baltzell and S Stepanyan. CLAS12 Møller Operations Manual-v1.7, November 2023.

[56] R. Fair, N. Baltzell, et al. The CLAS12 superconducting magnets. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 962:163578, 5 2020.

[57] M.D. Mestayer, K. Adhikari, et al. The CLAS12 drift chamber system. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 959:163518, 2020.

[58] D.S. Carman, L. Clark, R. De Vita, et al. The CLAS12 Forward Time-of-Flight system. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 960:163629, 2020.

[59] G. Asryan, Sh. Chandavar, et al. The CLAS12 forward electromagnetic calorimeter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 959:163425, 2020.

[60] Y.G. Sharabian, V.D. Burkert, et al. The CLAS12 high threshold Cherenkov counter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 968:163824, 2020.

[61] M Ungaro, D Anderson, G Asryan, M A Antonioli, P Bonneau, V D Burkert, S Christo, M Cook, B Duran, B Eng, A Hoebel, D Insley, J Jacobs, S Joosten, M Leffel, T Lemon, M Mcmullen, Z E Meziani, R Miller, P B Rojas, Y G Sharabian, and A Yegneswaran. The CLAS12 Low Threshold Cherenkov detector. *Nuclear Inst. and Methods in Physics Research*, 957:163420, 2020.

[62] M. Contalbrigo, L. Barion, et al. The large-area hybrid-optics CLAS12 RICH: First years of data-taking. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1057:168758, 2023.

[63] A. Acker et al. The CLAS12 Forward Tagger. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 959:163475, 2020.

[64] D.S. Carman, G. Asryan, et al. The CLAS12 Central Time-of-Flight system. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 960:163626, 2020.

[65] M.A. Antonioli, N. Baltzell, et al. The CLAS12 Silicon Vertex Tracker. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 962:163701, 2020.

[66] A. Acker, D. Attié, et al. The CLAS12 Micromegas Vertex Tracker. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 957:163423, 3 2020.

[67] S.J. Paul, A. Peck, et al. Alignment of the clas12 central hybrid tracker with a kalman filter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1049:168032, 2023.

[68] P. Chatagnon, J. Bettane, et al. The CLAS12 Central Neutron Detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 959:163441, 2020.

[69] B. Raydo, S. Boyarinov, et al. The CLAS12 Trigger System. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 960:163529, 2020.

[70] S. Boyarinov, B. Raydo, et al. The CLAS12 Data Acquisition System. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 966:163698, 2020.

[71] V. Ziegler, N. A. Baltzell, et al. The CLAS12 software framework and event reconstruction. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 959:163472, 4 2020.

[72] Pierre Chatagnon. *Nucleon Structure studies with CLAS12 at Jefferson Lab: Timelike Compton Scattering and the Central Neutron Detector*. PhD thesis, Paris-Saclay, Orsay, France, October 2020.

[73] P. Pandey, S. Kuhn, et al. Longitudinal Solid Polarized Target for CLAS12. *Journal of Nepal Physical Society*, 8(2):23–30, Dec. 2022.

[74] N.Pilleux on behalf of the CLAS Collaboration. RG-C end of run and first look at physics. `https://indico.ijclab.in2p3.fr/event/9131/contributions/28860/attachments/20757/28909/RGC_end_of_run_report.pdf`, accessed:22/02/2024.

[75] J Brock. Raster Target Run Group C, June 2022.

[76] D.Glazier. chanser. https://github.com/dglazier/chanser, accessed:16/05/2022.

[77] D.Glazier and others. JeffersonLab/CLAS12Root: Data Analysis Tools for hipo4 data format. https://github.com/dglazier/clas12root, accessed:13/06/2022.

[78] Rene Brun, Fons Rademakers, et al. root-project/root: v6.18/02, June 2020.

[79] D.Glazier. chanser documentation. https://chanser.readthedocs.io/en/latest/combitorials.html, accessed:13/06/2024.

[80] H. Avakian et al. "clas12 rg-a - analysis note overview and procedures phase i towards sidis clas12 first publications". https://www.jlab.org/Hall-B/shifts/admin/paper_reviews/2020/RGA_Analysis_Overview_and_Procedures_Nov_4_2020-6245173-2020-12-09-v3.pdf, accessed:20/06/24.

[81] S. Kuhn. Dilution Factor Determination for RG-C. https://userweb.jlab.org/ kuhn/RGC/RG-C_Dilution.pdf, accessed:04/09/23.

[82] Sucheta Shrikant Jawalkar. *Measurement of Single and Double Spin Asymmetries in $\vec{p}(e,e'\pi^{\pm,0})X$ Semi-Inclusive Deep-Inelastic Scattering*. PhD thesis, The College of William and Mary, Williamsburg, Virginia, January 2012.

[83] D.Glazier. brufit. https://github.com/dglazier/brufit, accessed:26/01/2023.

[84] M. Pivk and F.R. Le Diberder. : A statistical tool to unfold data distributions. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 555(1–2):356–369, December 2005.

[85] Alexander Blum. On crystalline character of transparent solid ammonia. *Radiation Effects*, 24(4):277–279, 1975.

[86] S.E. Kuhn. Dilution Factor Determination for RGC. Note, Old Dominion University, Norfolk, Virginia, October 22024.

[87] C Keith. Carbon and Polyethylene Targets for Run Group C, June 2022.

[88] J Brock. TGT APOLLO Cross Section, March 2021.

[89] Tetsuo Abe. GRAPE dilepton (Version1.1): A Generator for dilepton production in e p collisions. *Comput. Phys. Commun.*, 136:126–147, 2001.

[90] Paremuzyan. R et al. TCSGen- Generator for Timelike Compton Scattering Processes. https://github.com/JeffersonLab/TCSGen/tree/master, accessed:21/06/2024.

[91] R. Paremuzyan. *Timelike Compton Scattering. PhD thesis*. PhD thesis, Yerevan Physics Institute, 2010.

[92] Paremuzyan, R. and others. TCSGen. https://github.com/JeffersonLab/TCSGen/blob/master/src/K

[93] Braun, Vladimir M. Higher twists. *EPJ Web Conf.*, 274:01012, 2022.

[94] David J. Gross and S. B. Treiman. Light-cone structure of current commutators in the gluon-quark model. *Phys. Rev. D*, 4:1059–1072, Aug 1971.

[95] Pushpa Pandey. *Longitudinal solid polarized target for CLAS12 and study of spin structure of nucleons*. PhD thesis, 2024.

[96] X. Zheng, K. P. Adhikari, et al. Measurement of target and double-spin asymmetries for the $\vec{e}\vec{p}\to e\pi^+(n)$ reaction in the nucleon resonance region at low Q^2. *Phys. Rev. C*, 94:045206, Oct 2016.

[97] J. Arrington. Implications of the discrepancy between proton form factor measurements. *Phys. Rev. C*, 69:022201, Feb 2004.

[98] N.Pilleux and C.Dilks. PbPt calculation for RGC. https://github.com/N-Plx/shared_for_RGC,accessed:10/08/2024.

[99] Biselli. A et al. Measurements of single- and double-spin asymmetries for deeply virtual Compton scattering with a polarized electron beam and a longitudinally polarized proton target, 2014.

[100] BNL EIC task force. eic-smear. https://github.com/eic/eic-smear, 2020. Accessed: 2020-11-15.

[101] Elna Gerchtein. Pseudorapidity. https://scienceworld.wolfram.com/physics/Pseudorapidity.html, accessed: 2021-22-04", 2007.

[102] A. Adare et al. Concept for an Electron Ion Collider (EIC) detector built around the BaBar solenoid. 2 2014.

[103] Brookhaven National Laboratory. Department of Energy Selects Site for Electron-Ion Collider. `https://www.bnl.gov/newsroom/news.php?a=116998`, 2019. Accessed: 2020-01-12.

[104] Tanja Horn on behalf of the ECCE Consortium, CUA and JLab. ECCE Detector. `https://indico.bnl.gov/event/13614/contributions/58010/attachments/38779/64096/ECCE_Detector-nb.pdf`, accessed 2022-01-24. EIC Detector Proposal Advisory Panel Meeting.

[105] ECCE Consortium. Notes on the use of the BaBar solenoid in ECCE, 2021. ecce-note-det-2021-01.

[106] ECCE Consortium. ECCE Tracking System, 2021. ecce-note-det-2021-01.

[107] C Pinkenburg. Fun4All - the EIC Software Group Website. https://eic.github.io/software/fun4all.html.

[108] P Sznajder, K Tezgin, et al. EpIC. https://drf-gitlab.cea.fr, 2021.

[109] B. Berthou et al. Partons: Partonic tomography of nucleon software. *The European Physical Journal C*, 78(6):478, Jun 2018.

[110] S. Jadach and P. Sawicki. mFOAM-1.02: A compact version of the cellular event generator FOAM. *Computer Physics Communications*, 177(5):441–458, sep 2007.

[111] A. Bylinkin et al. Detector requirements and simulation results for the eic exclusive, diffractive and tagging physics program using the ecce detector concept. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1052:168238, July 2023.

[112] ECCE Consortium. Deep Learning-based Muon Identification for the ECCE Detector. 2021.