

Universidad Técnica Federico Santa María
Departamento de Física
Valparaíso, Chile



“Acceptance Study of the CLAS12 Detector Using
the Forward Micromegas Tracker for Run Group E”

Bruno Benkel

Thesis Submitted for the Title of
Master in Science, Mention in Physics

Advisor: Hayk Hakobyan, PhD.
Correferent Professor: Raffaella De Vita, PhD.
Correferent Professor: William Brooks, PhD.

September 14th - 2023

Dedication

Dedicated to Opa, for instilling in me the passion for learning.

Acknowledgements

First, I want to thank my wife Eli for being the best. Without your support I would have surely collapse long before finishing this thesis. Continuing on family, I give a big thank you to my sister Paula. Your company (and patience towards my lack of housework) allowed for a great way to vent work-related stress. Finally, I also want to thank my Oma Hedi, my mother Isabel, and my father Walter for their unparalleled counsel and love.

Moving towards JLab, I want to express my deepest gratitude towards Raffaella De Vita. Working with you has truly been the greatest part of my physics career so far, and I hope to continue aiding you and the lab in whatever way I can. On the same line, I want to thank Veronique Ziegler and Maxime Defurne, for their counsel and aid during the FMT alignment work. I also want to thank Sebastian Kuhn, Mohammad Hattawy, and Yu-Chun Hung from RG-F, who helped me get started working with their data. Then, a thank you to Nathan Baltzell for his very abundant help with reconstruction and slow controls system. I also want to thank Stepan Stepanyan, Maurizio Ungaro, and Gagik Gavalian for showing me around the lab and making my stay there all-around better.

Regarding RG-E, I want to thank Hayk Hakobyan for assisting and pushing me to finish this thesis. I want to thank William Brooks as well, for his pinpoint questions and wise answers. Then my concert companion Taisiya Mineeva as well, for her dilligence in organizing the group and always providing interesting conversation subjects. From the group, I also want to thank my fellow students Esteban Molina, Antonio Radic, Claudio San Martín, and Matías Barria. While very distracting, sitting alongside you makes work so much more entertaining, and your help with ROOT is almost as invaluable as ChatGPT's. Then, I want to thank the target group, whose composition during my tenure was Sebastián Gálvez, Eduardo Valdivia, Jairo Gonzales, and Alonso Lepe.

Some final thank yous I couldn't fit anywhere else: To Alexandra Elbakyan, for being such a crucial actor in the democratisation of knowledge. To Ben Tatum, for being the coolest (and weirdest) friend one can hope for. In the same vein, to the other flyguys: Charlie, Corbin, and Ryan. You four are the best. To Nicole Benz and Loreto Romero, for their friendship and emotional support during (and outside of) work hours. To Oscar Castillo, for the long and fun conversations about both jazz and physics. And finally to Mark Clift from Galil, for providing so much help in the use of EPICS for the Galil motor.

Abstract

Abstract— Particle accelerators and detectors are the main source of data for High Energy Physics. In this context, they are usually massive machines with a great number of moving parts, most of which require work in calibration and maintenance to run in optimal conditions. Special care must be put into the software dedicated to this calibration, in addition to event reconstruction.

This thesis presents a calibration effort of this kind, done to the newest detector in the CLAS12 reconstruction chain — the Micromegas Vertex Tracker. The results obtained in this regard are favourable. The calibration proved successful, doubling vertex resolution, and providing harsher criteria for selecting useful particle tracks.

In addition, the thesis goes into detail about the improvements in reconstruction that come from this new detector, focusing on its acceptance, vertex resolution, and SIDIS variables. A full study is presented for the total acceptance in CLAS12, and the multiplicities of various types of particles are measured.

Keywords— Particle detectors; Jefferson Laboratories; CLAS12; Micromegas Detectors.

Resumen— Los aceleradores y detectores de partículas son la principal fuente de datos para la Física de Altas Energías. En este contexto, estos suelen ser máquinas masivas con un gran número de partes móviles, la mayoría de las cuales requieren un trabajo de calibración y mantenimiento para funcionar en condiciones óptimas. Hay que tener especial cuidado en el *software* dedicado a esta calibración, además del dedicado a la reconstrucción de eventos.

Esta tesis presenta un trabajo en calibración de este tipo, realizado al detector más nuevo de la cadena de reconstrucción de CLAS12 — el *Micromegas Vertex Tracker*. Los resultados obtenidos son favorables. La calibración resultó exitosa, duplicando la resolución de vértice y proporcionando criterios más duros para la selección de *tracks* de partículas útiles.

Además, la tesis profundiza en las mejoras en reconstrucción que aporta este nuevo detector, centrándose en su *acceptance*, la resolución de vértices y las variables de SIDIS. Se presenta un estudio completo de *acceptance* total de CLAS12, y se miden las multiplicidades de varios tipos de partículas.

Palabras Clave— Detectores de partículas; Jefferson Laboratories; CLAS12; Detectores Micromegas.

Glossary

BAND: Back Angle Neutron Detector
BMT: Barrel Micromegas Tracker
CCDB: Calibration and Conditions Database
CD: Central Detector
CEBAF: Continuous Electron Beam Accelerator Facility
CLAS12: CEBAF Large Acceptance Spectrometer for Operation at 12 GeV
CND: Central Neutron Detector
CS-Studio: Control System Studio
CTOF: Central Time-of-Flight
CVT: Central Vertex Tracker
DC: Drift Chambers
DIS: Deep Inelastic Scattering
DOCA: Distance Of Closest Approach
EB: Event Builder
EC: Electromagnetic Calorimeter
ECAL: Electromagnetic Calorimeters
ECIN: EC-Inner
ECOU: EC-Outer
EPICS: Experimental Physics and Industrial Control System
FD: Forward Detector
FMT: Forward Micromegas Tracker
FT: Forward Tagger
FTCal: Forward Tagger Calorimeter
FTOF: Forward Time-of-Flight
FTTrk: Forward Tagger Gas Tracker
FTHodo: Forward Tagger Hodoscope
GeV: Giga electronVolt
GUI: Graphical User Interface
HEP: High Energy Physics
HIPO: High-Performance Input Output File Format
HTCC: High Threshold Cherenkov Counter
IOC: Input / Output Controller
JLab: Jefferson Laboratories
linac: Linear Particle Accelerator
LTCC: Low Threshold Cherenkov Counter
MeV: Mega electronVolt
MM: Micro-Mesh Gaseous Structure (Micromegas)
mrad: milliradian
MVT: Micromegas Vertex Tracker
PID: Particle Identification
PCAL: Pre-shower Calorimeter
PDF: Parton Distribution Function
PMT: PhotoMultiplier Tube
pQCD: perturbative Quantum Chromodynamics
PV: Process Variable

QCD: Quantum Chromodynamics

RF: Radio-Frequency

RG-E: Run Group E

RG-F: Run Group F

RICH: Ring Imaging Cherenkov Detector

SVT: Silicon Vertex Tracker

SIDIS: Semi-Inclusive Deep Inelastic Scattering

T: Tesla

Contents

List of Figures	viii
1 Physics Motivation	1
1.1 Deep Inelastic Scattering	3
1.2 Semi-Inclusive Deep Inelastic Scattering	5
1.3 Hadronisation in the Nuclear Medium	6
2 Experiment	9
2.1 CEBAF	9
2.2 CLAS12	10
2.3 RG-E Experiment	26
3 FMT Alignment and Reconstruction	33
3.1 Forward Micromegas Tracker	33
3.2 FMT Alignment	36
3.3 FMT Reconstruction Work	39
3.4 Validation and Results	40
4 Data Analysis	45
4.1 <code>clas12-rge-analysis</code>	45
4.2 Cuts	50
4.3 Sampling Fraction	51
4.4 Acceptance Correction	53
5 Results and Conclusions	56
5.1 FMT Efficiency	56
5.2 Acceptance Correction Results	60
5.3 Study Results	66
5.4 Systematic Error Estimation	79
5.5 Conclusions	80
Appendices	81
A Reproducibility	81
B Fiducial Cuts	82
C FMT Layer Efficiency Error Estimation	84
D DIS plots in v_z bins	84
E RG-F Target Layout	84

List of Figures

1	The standard model	2
2	Deep Inelastic Scattering in Quantum Chromodynamics	3
3	Q^2 dependence of x Particle Distribution Function for the u quark	5
4	α_s dependence on Q^2	6
5	Continuous Electron Beam Accelerator Facility (CEBAF)	9
6	CEBAF Large Acceptance Spectrometer for 12 GeV (CLAS12)	10
7	High Threshold Cherenkov Counter (HTCC)	11
8	Drift Chambers (DC)	13
9	Low Threshold Cherenkov Counter (LTCC) Mirror System	13
10	Forward Time-of-Flight (FTOF) system	14
11	Forward Tagger (FT)	16
12	Central Vertex Tracker (CVT)	17
13	Central Time-of-Flight (CTOF) system	18
15	CVT momentum resolution vs. momentum	20
14	CLAS12 offline software reconstruction chain	20
16	DC momentum resolution vs. momentum	21
17	Particle going through the DC	22
18	Particle β vs. momentum for positively charged tracks	24
19	β distribution of neutrals	25
20	Double-target system	27
21	Run Group E (RG-E) CS-Studio main screen	31
22	Micromegas (MM) working principle	33
23	Micromegas Vertex Tracker (MVT) detector	34
24	Forward Micromegas Tracker (FMT) reconstruction summary	35
25	FMT detector geometry	36
26	Example FMT residuals plot	37
27	Examples of residuals goodness of fit plots	38
28	Residuals distribution improvement	39
29	DC vs. FMT z without geometric correction	41
30	z vs. θ for DC and FMT	42
31	DC vs. FMT z with geometry correction	42
32	FMT layers efficiency	43
33	Q^2 vs. ν comparison	52
34	Calorimeters E/p plots	53
35	Calorimeters p vs. E/p plots	54
36	v_z for DC and FMT for runs 12933 and 12016	57
37	v_z for DC and FMT, w/ and w/out the geometry cut	58
38	v_z , θ , and p FMT efficiencies for e^- , $e^-\pi^+$, and $e^-\pi^-$	61
39	ϕ efficiency and geometry cut study	62
40	e^- variables acceptance	64
41	θ study for negative particles	65
42	Hadronic variables acceptance	67
43	θ study for positive particles	68
44	Acceptance-corrected Q^2 separated in v_z bins	70
45	Acceptance-corrected ν separated in v_z bins	71

46	Acceptance-corrected z_h for $e^-\pi^+$ separated in v_z bins	72
47	Acceptance-corrected z_h for $e^-\pi^-$ separated in v_z bins	73
48	Acceptance-corrected p_T^2 for $e^-\pi^+$ separated in v_z bins	74
49	Acceptance-corrected p_T^2 for $e^-\pi^-$ separated in v_z bins	75
50	Acceptance-corrected ϕ_{PQ} for $e^-\pi^+$ separated in v_z bins	76
51	Acceptance-corrected ϕ_{PQ} for $e^-\pi^-$ separated in v_z bins	77
52	Statistics for e^- , $e^-\pi^+$, and $e^-\pi^-$ against v_z	78
53	ϕ vs. θ of e^- in p bins	82
54	ϕ vs. θ of π^+ and π^- , separated in p bins.	83
55	Q^2 separated in v_z bins	85
56	ν separated in v_z bins	86
57	z_h for $e^-\pi^+$ separated in v_z bins	87
58	z_h for $e^-\pi^-$ separated in v_z bins	88
59	p_T^2 for $e^-\pi^+$ separated in v_z bins	89
60	p_T^2 for $e^-\pi^-$ separated in v_z bins	90
61	ϕ_{PQ} for $e^-\pi^+$ separated in v_z bins	91
62	ϕ_{PQ} for $e^-\pi^-$ separated in v_z bins	92

1 Physics Motivation

Since ancient times, humanity has pondered the composition of matter. Philosophers from both East and West have often contemplated this question, and the modern view of this composition is based on the concept of the atom. The notion of the atom, or indivisible particle, was proposed by Democritus in the 6th century B.C. Despite its uncanny similarity to the modern concept of the atom, the model remained an abstract idea for more than two millennia.

In the grand scheme of things, it is only recently that we have been able to probe into the structure of matter and observe the atom. In 1897, J.J. Thomson discovered "corpuscles," or electrons, using a cathode ray tube. Based on this discovery, he proposed an atomic model consisting of a positively charged paste with lighter electrons floating inside. Then, in 1909, Ernest Rutherford put this model to the test in what became the first scattering experiment in history. By bombarding α -particles onto a thin gold foil, he proved that most of the atom's mass was concentrated in a small, positively charged nucleus at its center. He named the constituents of this nucleus protons.

Following that, Niels Bohr proposed the atomic planetary model in 1914. His theory precisely fit the experimental data for Hydrogen but did not apply to heavier atoms. This issue was resolved with the discovery of the neutron in 1932 by James Chadwick. This discovery made the masses of atoms consistent with available experimental data, marking the end of the era of classical particle physics. Then, in the 1950s and 1960s, a bewildering variety of particles were found in scattering experiments. The theory born from this "particle zoo" gave rise to the Standard Model, which explains these particles as combinations of a small number of fundamental particles. The model describes three of the four fundamental forces using force-mediating gauge bosons. Additionally, it describes 24 particles, which are the constituents of matter. Finally, it also includes one scalar boson, the Higgs boson, whose existence was proven in 2012 [Aad et al., 2012].

Half of these 24 particles are elementary constituents of hadrons. They were initially referred to as quarks by Murray Gell-Mann and George Zweig, and later as partons by Richard Feynman. Quarks are point-like spin-1/2 particles with a fraction of an electron's electric charge, and they are distinguished by "flavors". Both Gell-Mann and Zweig's constituent quark model and Feynman's parton model were later merged into the Quark-Parton model [Perkins, 2000].

Deep Inelastic Scattering (DIS) is the process used to investigate the interior of hadrons using leptons. The process is similar to Rutherford scattering and provided the first experimental evidence of quarks. DIS can be employed to delve even deeper into the structure of matter by utilising increasingly higher energies, thanks to Werner Heisenberg's uncertainty principle.

High-energy probes lead to asymptotic freedom, which is the property where the interactions between particles, such as quarks, become increasingly weak at shorter distances. This implies that inside hadrons, quarks mostly move as free, non-interacting particles. This allows for reliable calculation of event cross-sections in particle physics. Another important feature is colour confinement, which is the property of colour-charged particles that prevents their isolation. Colour charge is the Quantum Chromodynamics (QCD) equivalent of electric charge. There are

Standard Model of Elementary Particles

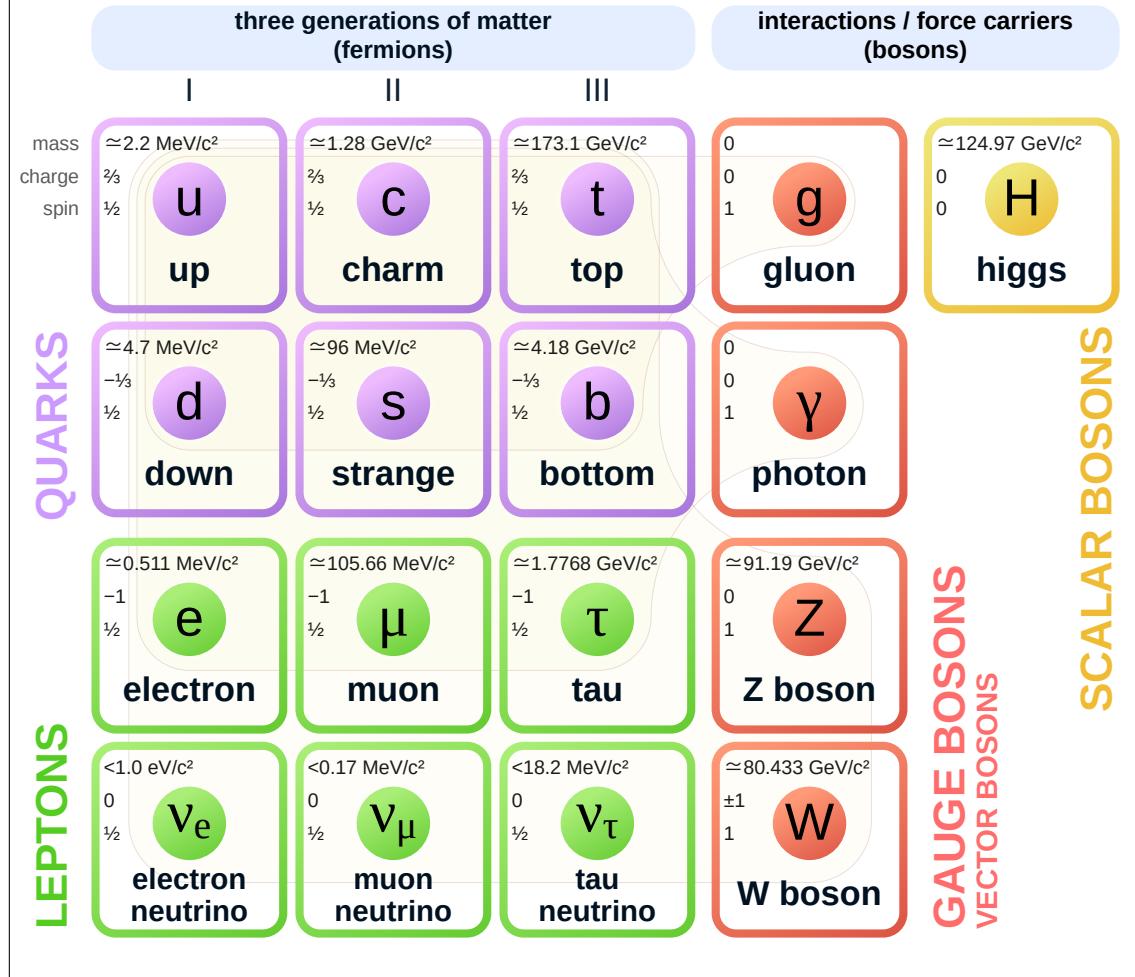


Figure 1: Fundamental particles in the Standard Model.

Source: Wikimedia Commons.

three colour charges and their corresponding anticharges. Quarks possess a single colour charge, while gluons, the force-mediators of the strong force, have a bi-colour charge. In a state of equilibrium, the strong force confines quarks to be in close proximity, forming quark-antiquark pairs or 3-quark triplets in such a way that the net colour charge is neutral.

QCD describes these properties and models the quark strong interaction. Similar to the electromagnetic force, the strength of the interaction is determined by the strong coupling constant α . However, unlike the electromagnetic force, this constant weakens as distances decrease. At a larger distance scale, or smaller resolution, confinement is expected to become visible. The α value increases to values close to 1, and the perturbative treatment of QCD is not viable. There is limited knowledge about the non-perturbative behaviour of QCD, and therefore, it mostly relies on phenomenological models at present. The aim of the Run Group E (RG-E) experiment is to gather new data on the hadronic structure in Semi-Inclusive Deep Inelastic

Scattering (SIDIS). Through this experiment, we hope to gain further insight into quark propagation and hadron formation.

1.1 Deep Inelastic Scattering

In its simplest description, DIS refers to the scattering of an electron off a quark inside a nucleon. Figure 2 displays the Feynman diagram illustrating DIS. The four-momentum of the nucleon is denoted by P , while that of the quark is represented by p . The initial and final four-momenta of the electron are given by k and k' , respectively. When k' is measured, the momentum transferred to the hadron system by the virtual photon is defined as $q = k - k'$. q is a spacelike vector conventionally denoted as $q^2 = -Q^2$.

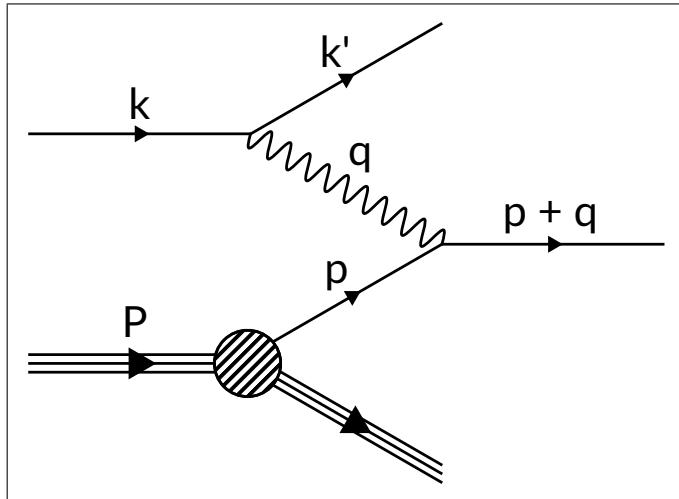


Figure 2: DIS in QCD. The diagram describes the stream of momentum in the scattering of a high energy electron off a quark. The quark's wave function is incorporated in the nucleon's wave function.

Source: Own elaboration, using Inkscape.

If Q^2 is sufficiently high, the quark is ejected from the nucleon. Soft processes, such as gluon emission and quark-antiquark pair production, then take place to neutralise the colour charge. This results in the transformation of the ejected quark into a jet of hadrons. The jet propagates in the direction of the transferred momentum from the electron.

1.1.1 DIS Cross Section

To approximate the cross section of electron-nucleon scattering, we will work in the center of mass reference frame. In this frame, the electron and nucleon are moving towards each other with sufficient energy, allowing us to neglect the nucleon's mass. As a result, the nucleon possesses nearly lightlike momentum along the collision axis. Consequently, the constituent quarks of the nucleon also have nearly lightlike momenta that are nearly collinear to the nucleon's momentum. Hence, as a first-order approximation, we can express the quark's momentum as

$$p = \xi P,$$

where ξ represents the longitudinal fraction of the quark's momentum, and thus $0 < \xi < 1$.

In the leading-order approximation, we can also disregard gluon emission and exchange during the collision. Therefore, the cross section of electron-nucleon scattering is equal to that of electron-quark scattering for a given ξ , multiplied by the probability that the nucleon contains a quark with a longitudinal momentum fraction of ξ , integrated over ξ .

This calculation encounters the issue that the probability of a nucleon containing a quark with a specific momentum cannot be computed within perturbative QCD. It relies on the soft processes that determine the nucleon's structure as a composite system of quarks and gluons. Consequently, we must consider this probability as an unknown function that needs to be measured in experiments.

Such probability functions are known as Parton Distribution Functions (PDFs). A PDF can be assigned to various types of quarks, antiquarks, and gluons, and is incorporated into the nucleon's wave function. For each parton f , its PDF is defined as

$$P_f = f_f(\xi) d\xi.$$

Therefore, the cross section for the inelastic scattering of an electron off a nucleon, within the leading-order approximation, can be expressed as

$$\sigma(e^-(k)p(P) \rightarrow e^-(k')X) = \int_0^1 d\xi \sum_f f_f(\xi) \cdot \sigma(e^-(k)q_f(\xi P) \rightarrow e^-(k') + q_f(p')),$$

where X denotes the final hadronic state. It is important to remember that this equation does not provide an exact QCD prediction but represents the first-term expansion of α_s [Halzen and Martin, 1991]. This approximation is known as the parton model.

1.1.2 The Parton Model

In the parton model, the cross section is given by

$$\frac{d^2\sigma}{dxdy} (e^- p \rightarrow e^- X) = \left(\sum_f x f_f(x, Q^2) Q_f^2 \right) \frac{2\pi\alpha s}{Q^4} (1 + (1 - y)^2), \quad (1)$$

where $s \equiv 2P \cdot k$, Q_f represents the charge of parton f , and x and y are the Bjorken variables defined as

$$x \equiv \frac{Q^2}{2P \cdot q}, \quad y \equiv \frac{2P \cdot q}{s}.$$

In the nucleon's rest frame, $y = q^0/k^0$, and it represents the energy transferred to the hadron by the incoming electron.

The PDFs in Equation (1) have a weak dependence on Q^2 due to gluon radiation. This leads to Bjorken scaling violation [Halzen and Martin, 1991]. When the structure functions are known for certain values of Q^2 , they can be evolved to other values using the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [Dokshitzer et al., 1991].

Figure 3 shows the predictions of the Altarelli-Parisi equations for the evolution of the PDFs with respect to Q^2 . Partons with large x tend to radiate and move

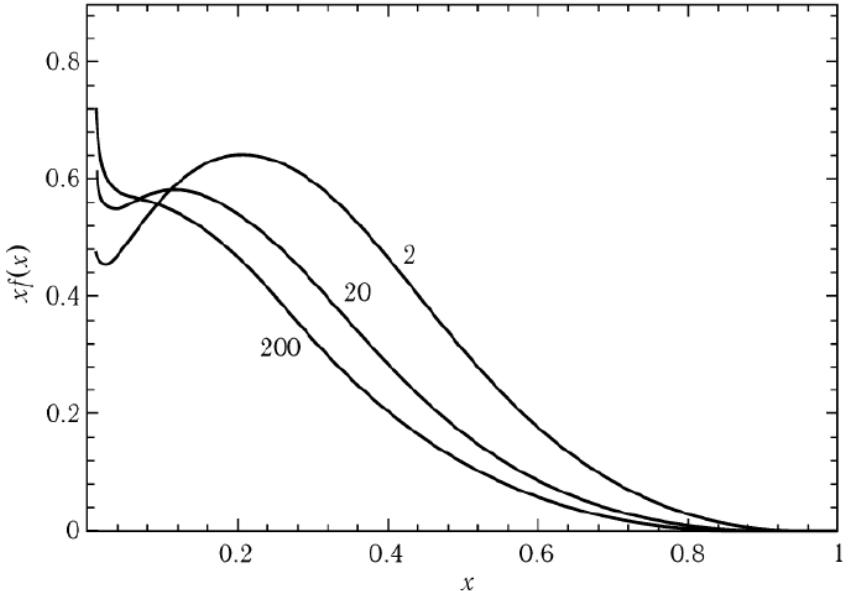


Figure 3: Parton distribution functions $xf_f(x)$ for the u quark at $Q = 2$, $Q = 20$, and $Q = 200$ GeV, showing the parton evolution effect according to the Altarelli-Parisi equations.

towards states with lower x values. Simultaneously, radiation generates new partons with low x values. As Q^2 increases, the parton distributions decrease for large x values while rapidly increasing for low x . At low Q^2 , the wavelength of the virtual photon is too large to probe the partons directly, resulting in probing the proton as a whole. The precise range of validity for the QCD-extended parton model is not known, but it is assumed to be applicable for $Q^2 > 1 \text{ GeV}^2$, corresponding to a spatial resolution of approximately 0.2 fm.

1.1.3 Strong Coupling Constant α_s

Measuring the experimental value of α_s is crucial for perturbative QCD calculations. To achieve this, the overall scale of renormalisation needs to be determined. Typically, the mass of the neutral Bose particle Z^0 , which is 91.19 GeV, is chosen as the scale. Furthermore, the renormalisation scheme should be fixed, which defines the coupling constant at a specific scale. The experimental results for α_s can be seen in Figure 4.

1.2 Semi-Inclusive Deep Inelastic Scattering

The previous section focused solely on the detection of the scattered electron in the inclusive reaction. However, in the case where one of the produced hadrons is identified, the event is referred to as semi-inclusive. The identification of this hadron provides information about the flavor of the quark that was struck. In SIDIS, it is possible to measure this flavor dependence. The identified hadrons in SIDIS are known as current fragments, and it is important to separate them in the analysis from the fragments originating from the target.

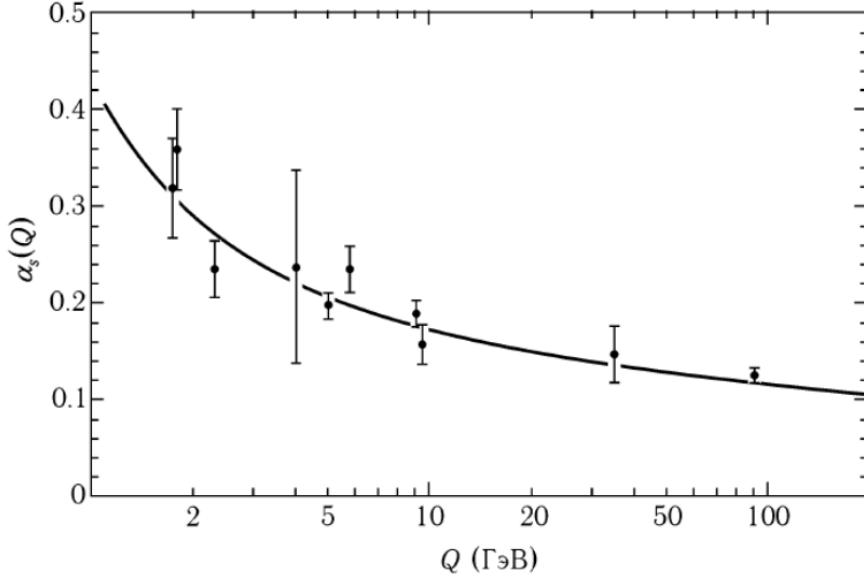


Figure 4: Experimentally measured dependence of α_s on Q . The measured values are compared with the theoretical predictions of renormalised evolution with an initial value of $\alpha_s(m_z) = 0.117$ GeV.

1.2.1 SIDIS Cross Section

In SIDIS, the fragmentation process corresponds to very low Q^2 and is not calculable in perturbative QCD (pQCD). It is parameterised by fragmentation functions $D_f^h(Q^2, z)$, which measure the probability that an f -flavoured quark fragments into an h -type hadron with a fraction z of the virtual photon energy ($E_h = z\nu$).

In the quark-parton model, the cross section for $eN \rightarrow ehX$ is assumed to be the differential cross section from Equation (1) multiplied by the fragmentation probability

$$\frac{d^3\sigma(eN \rightarrow ehX)}{dx dQ^2 dz} = \frac{d^2\sigma(eN \rightarrow eX)}{dx dQ^2} \cdot \frac{\sum_f e_f^2 q_f(x, Q^2) D_f^h(Q^2, z)}{\sum_f e_f^2 q_f(x, Q^2)}. \quad (2)$$

It is assumed that the quasi-free scattering process and the fragmentary process are independent in the cross section.

The hadron multiplicity per DIS event, denoted as $M_h(Q^2, z)$, is given by

$$M_h(Q^2, z) \equiv \frac{1}{\sigma} \frac{d^3\sigma(eN \rightarrow ehX)}{dQ^2 dz} = \frac{\int dx \sum_f e_f^2 q_f(x, Q^2) D_f^h(Q^2, z)}{\int dx \sum_f e_f^2 q_f(x, Q^2)},$$

where σ is the differential inclusive DIS cross section $\frac{d^2\sigma(eN \rightarrow eX)}{dx dQ^2}$.

1.3 Hadronisation in the Nuclear Medium

Hadronisation is the process in which quarks and gluons form hadrons. In the nuclear medium, this process is influenced by quark energy loss through two mediums: gluon radiation and multiple quark-nucleon scattering. Additionally, hadron-nucleon interactions also affect the process if hadronisation occurs inside the nucleus.

The primary experimental observable in hadronisation is the multiplicity of hadrons produced in a dense nucleus compared to a light one, such as deuterium. In the absence of attenuation from interactions with the medium, these multiplicities should be identical, resulting in a ratio of unity. This ratio is known as the attenuation ratio and is defined as

$$R_{\text{att}}^h(z, \nu) = \frac{\left(\frac{1}{\sigma} \frac{d^2\sigma(eN \rightarrow ehX)}{dz d\nu}\right)_A}{\left(\frac{1}{\sigma} \frac{d^2\sigma(eN \rightarrow ehX)}{dz d\nu}\right)_{^2H}},$$

where the derivative with respect to Q^2 is substituted with one with respect to ν . The ν and z dependence of this ratio can be used to study the nature of the hadron formation mechanism.

One type of observable that can be isolated is the characteristic times for the distinct stages of the hadronisation process. The existence of these stages is dictated by two fundamental properties in QCD: confinement and causality. Confinement implies that a coloured quark can only propagate for a limited distance. Causality dictates that the equilibrium colour field of a hadron cannot be formed instantaneously.

1.3.1 Virtual Photon Absorption

The first stage in the hadronisation process is the absorption of a virtual photon by a quark. This process occurs on a relatively short timescale, much less than 1 fm/c, and is governed by the wavelength of the virtual photon.

During this stage, the virtual photon interacts with a quark, transferring energy and momentum to the quark. The quark undergoes a transition to a higher energy state, which initiates the subsequent stages of hadronisation.

1.3.2 Production Time

Then, there must be a stage in which the coloured quark propagates as a quasi-free particle. During this second stage, there is gluon radiation with a differential spectrum given by pQCD as

$$d\omega^{q \rightarrow qg} = \frac{\alpha_s(k_\perp^2)}{4\pi} \frac{8}{3} \left[1 + \left(1 - \frac{k}{E}\right)^2 \right] \frac{dk}{k} \frac{dk_\perp^2}{k_\perp^2},$$

where E is the quark energy, k is the 4-momentum of the gluon, and k_\perp is its transverse momentum. The time associated with this stage is commonly referred to as the “production time” [Kopeliovich et al., 2004], and it represents the duration in which the quark is deconfined. The production time is a characteristic of the propagating quark and should be independent of the final hadron formed.

To a very good approximation, in deep inelastic kinematics and at $x > 0.1$, the struck quark absorbs all the energy of the virtual photon. Thus, the initial energy of the struck quark should be ν . This quantity is much larger than the quark’s mass (assuming an up or down quark), and thus we’ll ignore it in the treatment of this theory.

Conservation of energy then tells us that the final energy of the produced hadron should be no greater than ν . Additionally, the gluon radiation leads to a loss of energy, resulting in the hadron's energy being below ν . We can estimate this energy loss from the string model [Artru and Mennessier, 1974]. The main parameter in the string model is the string tension $\kappa \approx 1 \text{ GeV/fm}$. The growth of the string creates an energy loss governed by this κ . Therefore, we can estimate the rate of vacuum energy loss as

$$\frac{dE}{dx} \Big|_{\text{vacuum}} \approx 1 \text{ GeV/fm}.$$

Then, if

$$z_h = \frac{E_h}{\nu}, \quad (3)$$

is the fraction of the struck quark's energy retained by the hadron, then the term $\nu(1 - z_h)$ represents the energy loss through gluon radiation.

Thus, an estimate of the distance of gluon emission is

$$l_p = \frac{\nu(1 - z_h)}{\kappa},$$

and therefore the production time τ_p is l_p/c [Kopeliovich et al., 2004].

As an example, for a 10 GeV pion with $z_h = 0.6$, $l_p = 4 \text{ fm}$ – the production time is of the order of only a few fm for Jefferson Lab (JLab) energies.

1.3.3 Formation Time

In the final stage of the hadronisation process, the struck quark finds partner quarks to neutralise its colour, thus ending the gluon radiation. This stage is characterised by the evolution of the pre-hadron to an ordinary hadron and is commonly named the “formation time”. This is the time required to form the colour field of a hadron, and it should depend on the specific hadron being formed.

The formation time is not directly related to confinement, but rather it is a measure of the time required to form the non-perturbative colour field of the hadron, starting from a small colour-singlet object. This field generation time has a well-known analogue in Quantum Electrodynamics.

We can build a simple estimate for the formation time. To form a hadron of radius R from a point-like, single-colour singlet, the speed at which the field can arise (in its rest frame) is bound by the speed of light

$$\tau_{\text{formation}}^{\text{rest}} > \frac{R}{c},$$

which, when Lorentz-boosted into the lab frame, becomes

$$\tau_{\text{formation}}^{\text{lab}} > \frac{E}{m^*} \frac{R}{c},$$

where m^* is the mass of the propagating colour-singlet object. In principle, m^* ranges from the mass of two bare quarks to the fully formed hadron mass m_h . A lower limit for the formation time is given by setting $m^* = m_h$.

While this estimate is a classical calculation, a quantum mechanical analysis taking into account the gluon wavelength arrives at the same result, as is explained in [Dokshitzer et al., 1991].

2 Experiment

The Thomas Jefferson National Accelerator Facility is a High Energy Physics (HEP) laboratory located in Newport News, Virginia, USA. For simplicity and to follow convention, the laboratory will be called JLab hereafter. At the site, there is a recirculating linear electron accelerator named the Continuous Electron Beam Accelerator Facility (CEBAF). This accelerator is capable of delivering a 12 GeV electron beam to four experimental Halls simultaneously: Halls A, B, C, and D.

Different physics topics are studied in each of these halls. This thesis aims to provide a preparatory study for the RG-E experiment, which will take place in Hall B, where the CEBAF Large Acceptance Spectrometer for operation at 12 GeV (CLAS12) is located.

The first section provides a brief description of CEBAF at JLab. The second section then provides detailed information about CLAS12, including its central detector, forward detector, and offline event reconstruction. The third and final section discusses the RG-E experiment to be performed.

2.1 CEBAF

The CEBAF facility consists of a pair of 1.4-km-long antiparallel superconducting radio-frequency (RF) linear accelerators (linacs) constructed 8 meters below the

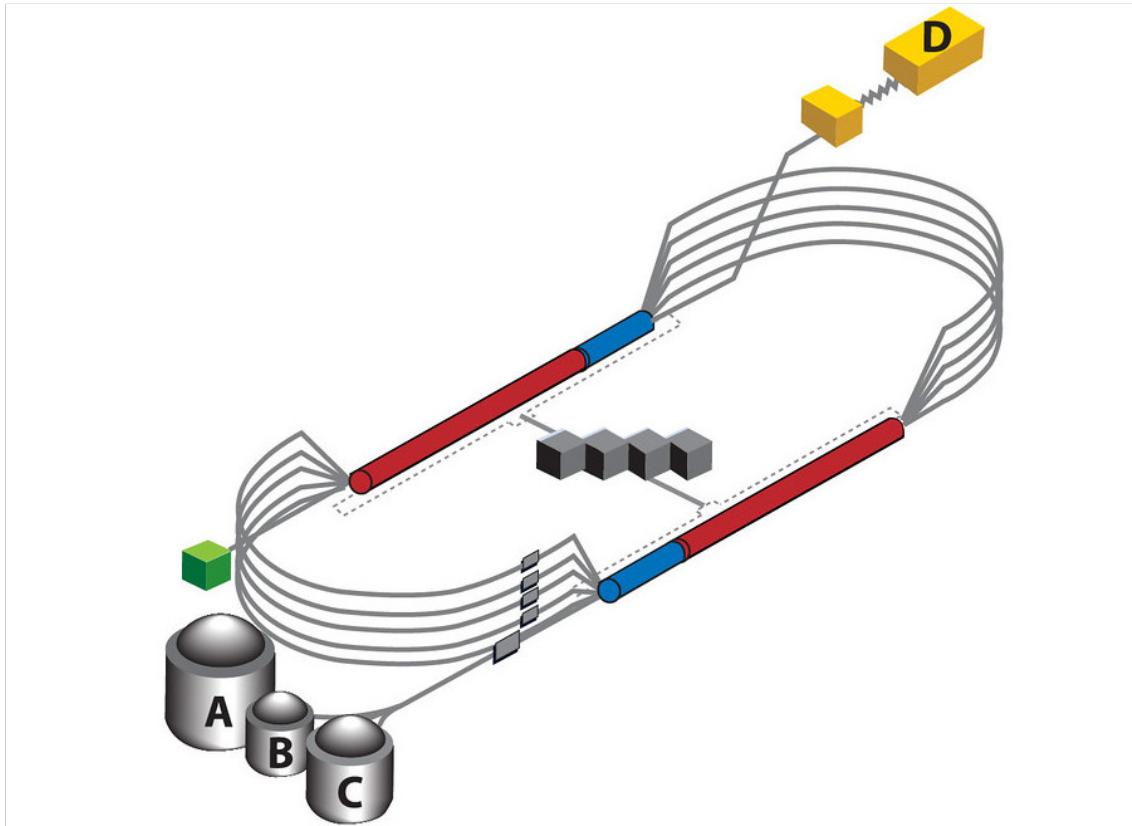


Figure 5: Simplified representation of CEBAF.

Source: jlab.org.

surface. The two accelerators are connected by two 180-degree arcs, each with a radius of 80 meters [Leemann et al., 2001]. A schematic diagram illustrating the design of CEBAF is provided in Figure 5.

The recirculating arcs are composed of five separate beamline sections, allowing the beam to traverse each linear particle accelerator (linac) up to five times. Within each linac, the energy gain of the beam ranges from 0.8 GeV to 1.2 GeV, resulting in a final beam energy of approximately 12 GeV. CEBAF is specifically designed for high-energy electron beam experiments aimed at studying the structure of mesons, nucleons, and nuclei [Rode and Team, 2010].

2.2 CLAS12

The main detector in Hall B is CLAS12, used to study electro-induced nuclear and hadronic reactions [Burkert et al., 2020]. The spectrometer provides efficient detection of charged and neutral particles over a large fraction of the full solid angle.

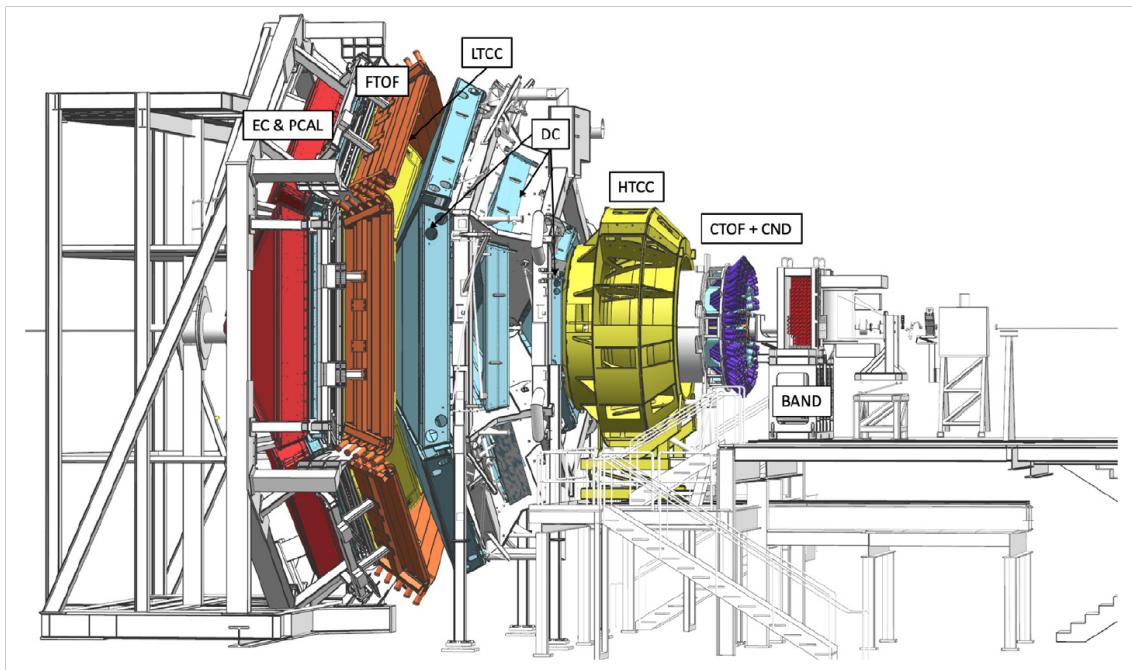


Figure 6: The CLAS12 detector in the Hall B beamline. The electron beam enters from the right and impinges on the target located in the center of the solenoid magnet shown at the right (upstream) end of CLAS12. The FD consists of the High Threshold Cherenkov Counter (HTCC) (yellow), the torus magnet (grey), the Drift Chambers (DC) tracking system (light blue), and another set of Cherenkov counters (hidden), time-of-flight scintillation counters (brown), and Electromagnetic Calorimeters (ECs) (red). Between the HTCC and the torus, the Forward Tagger (FT) is installed. The Central Detector (CD) consists of the Silicon Vertex Tracker (SVT) (hidden), which is surrounded by a Barrel Micromesh Tracker (BMT) (hidden), the Central Time-of-Flight (CTOF) system, and the Central Neutron Detector (CND) (blue). At the upstream end, a Back Angle Neutron Detector (BAND) (red) is installed.

Source: CLAS12 wiki.

CLAS12 is based on two superconducting magnets: a solenoid magnet and a 5 T torus magnet. The detector is divided into two parts: the Forward Detector (FD) and the Central Detector (CD). The FD, aided by the torus magnet, covers the forward polar range from 5° up to 35° , while the CD, aided by the solenoid magnet, covers the polar angles from 35° to 125° . Both detectors have full azimuthal coverage.

Trajectory reconstruction is performed using Drift Chambers (DC) in the forward direction, achieving a momentum resolution of less than 1%. In the central detector, trajectory reconstruction is done using a vertex tracker, resulting in a momentum resolution of less than 3%. Particle identification relies on Cherenkov counters, time-of-flight scintillators, and electromagnetic calorimeters [Burkert et al., 2020]. Fast triggering and high data-acquisition rates enable operation at a luminosity of $10^{35} \text{ cm}^{-2} \text{ s}^{-1}$ [Burkert et al., 2020].

A diagram of CLAS12 showing the position of each detector component is provided in Figure 6.

2.2.1 Forward Detector (FD)

High Threshold Cherenkov Counter (HTCC)

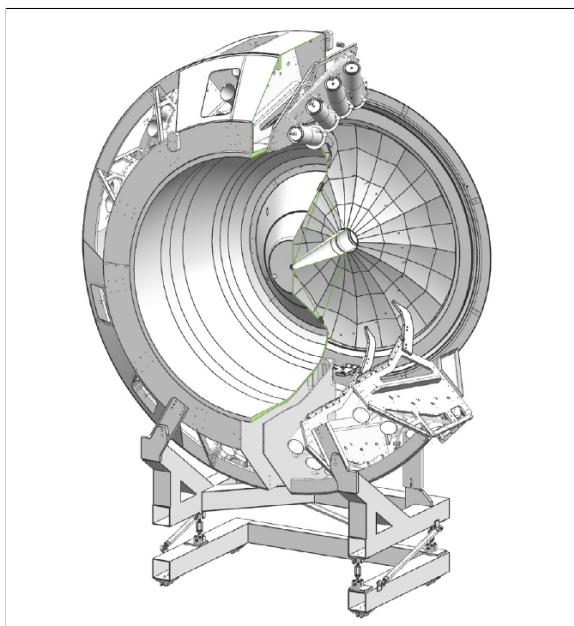


Figure 7: Render of the High Threshold Cherenkov Counter (HTCC). The container spans a diameter of about 4.5 m. The mirror is seen at the downstream end to the right. The Photomultiplier Tubes (PMTs) are mounted in 12 sectors and in groups of 4 at the outer perimeter of the container. Light collection uses additional Winston cones and 5-in PMTs with quartz windows.

Source: CLAS12 wiki.

The HTCC is specifically designed to separate electrons and positrons with momenta below 4.9 GeV from other ionizing particles. It achieves this through its capability for electron/positron identification, which provides high rejection of charged pions and low background noise. This is crucial for reliably identifying scattered electrons in an environment with a dense electromagnetic background.

The HTCC, positioned downstream of the target, is fitted in between magnets, upstream of the forward tracking detectors. It ensures full azimuthal coverage, meaning it can detect particles emitted from any direction around its circumference. In terms of the polar angle, it spans from 5° to 35° , covering a specific range of particle emission angles. Importantly, it has no blind areas in its complete solid angle coverage, meaning there are no regions where particles cannot be detected.

Operating in dry CO₂ gas at 1 atm pressure, the HTCC consists of a multi-focal mirror composed of 48 elliptical mirror facets. This mirror design enables the focusing of Cherenkov light

produced by charged particles passing through the detector. The focused light is then detected by 48 PMTs, with each PMT featuring a quartz window of 125 mm in diameter.

The PMTs are positioned within a magnetic field of up to $3.5 \cdot 10^{-3}$ T, which is oriented along the axes of the phototubes. To minimize the impact of the magnetic field on the PMTs, they are surrounded along their lengths by a multi-layer magnetic shield. This shield includes active compensation coils, which further help in shielding the PMTs from the effects of the magnetic field.

To minimize the effects of multiple scattering and its impact on the momentum analysis of charged tracks in the torus field, the HTCC mirror system is constructed using a backing structure made of low-density composite material. This choice of material helps to reduce the scattering of particles passing through the HTCC, thereby improving the accuracy of momentum measurements.

Since the HTCC is located in front of the momentum analyzing torus magnet, it is important to minimize the presence of materials in the path of charged particles, except for the radiator gas. This is done to prevent interactions and disturbances that could affect the accuracy of momentum analysis. The HTCC is designed with this consideration in mind.

In the HTCC, the density of solid material encountered by charged particles passing through its volume is approximately 135 mg/cm². This low-density configuration ensures that the material contribution to multiple scattering is minimised, allowing for more precise momentum measurements of charged tracks.

The HTCC also serves the purpose of generating a fast signal that can be used as a trigger for scattered electrons. This signal is utilised to identify and select scattered electrons for further analysis.

In conjunction with the energy deposited in the electromagnetic calorimeters, the HTCC plays a role in the identification of electrons with specific energies. By combining the information from the HTCC and the electromagnetic calorimeters, the experiment can accurately identify electrons of interest based on their energy deposition patterns.

A visual representation of the HTCC can be seen in Figure 7, which provides a cut view of the detector and its components.

Overall, the HTCC plays a crucial role in electron/positron identification by using a multi-focal mirror, PMTs, and a magnetic field setup. These components work together to ensure efficient detection and separation of electrons and positrons from other charged particles in a high-energy physics experiment environment.

Drift Chambers (DC)

The forward tracking system in the experiment consists of three independent drift chambers in each of the six sectors of the torus magnet. These drift chambers serve as the primary tracking detectors and are supported by the six coils of the torus magnet.

Each sector of the torus magnet contains a total of 36 layers in the drift chambers, with each layer having 112 sense wires. The sense wires are arranged in three regions, with each region comprising twelve layers. This configuration results in a total of 112 x 36 sense wires per sector.

The arrangement of the drift chambers around the torus coil can be visualised in Figure 8. The figure shows the positioning of the three regions of the drift chambers in each sector of the torus magnet.

In terms of location, the first region of the drift chambers is situated at the entrance to the torus magnetic field region. The second region is positioned inside the magnet, where the magnetic field is close to its maximum strength. Finally, the third region is located downstream of the torus magnet, in a low magnetic field space.

This arrangement of the drift chambers around the torus magnet provides independent and redundant tracking capabilities in each of the six torus sectors. It allows for precise reconstruction of charged particle trajectories and momentum measurements in the experiment.

Each of the three regions in the drift chambers consists of six "superlayers," where each superlayer comprises two layers. The wires in one layer are strung at a stereo angle of $+6^\circ$ relative to the sector midplane, while the wires in the other layer are strung at a stereo angle of -6° . This stereo configuration provides excellent resolution in the polar angle ($\Delta\theta < 2$ mrad) and good resolution in the azimuthal scattering angle ($\Delta\phi < 2$ mrad).

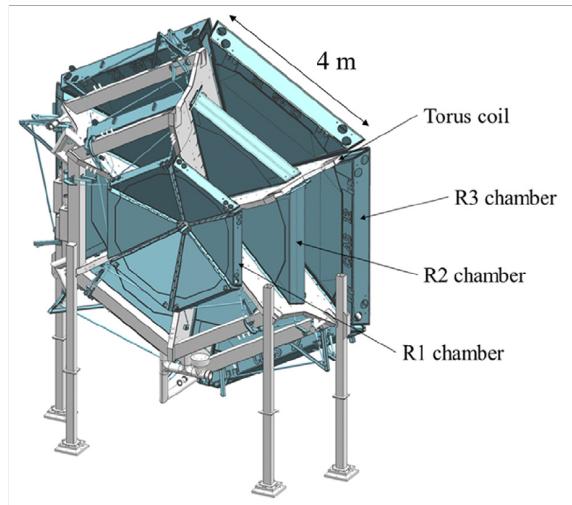


Figure 8: Drift Chambers (DC) render. Each of the DC regions are denoted as R1, R2, and R3 in the figure.

Source: CLAS12 wiki.

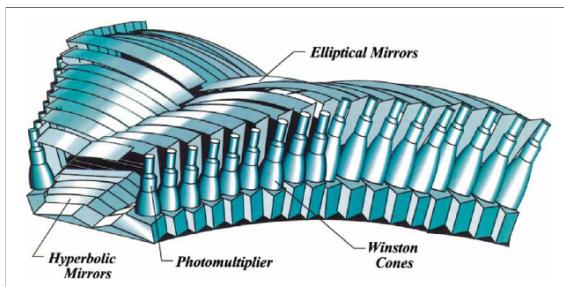


Figure 9: Layout and components of the optical mirror system within each LTCC box from the design model.

Source: CLAS12 wiki.

Low Threshold Cherenkov Counter (LTCC)

The LTCC system in CLAS12 is designed for the detection of charged pions and kaons in the momentum range of 3.5 to 9 GeV. It consists of boxes shaped like truncated pyramids, with four out of the six sectors of CLAS12 equipped with one LTCC box. Each LTCC box contains 108

lightweight mirrors with composite backing structures, 36 Winston light-collecting cones, 36 125-mm diameter PMTs, and 36 magnetic shields. The LTCC boxes are filled with heavy C4 F10 radiator gas.

As part of the 12 GeV upgrade, the LTCC detector from the previous CLAS experiment was refurbished to enhance its efficiency for charged pion and kaon detection. The enhancements included increasing the volume of the radiator gas, refurbishing the elliptical and hyperbolic mirrors with new coatings, and improving the sensitivity of the PMTs to Cherenkov light. The sensitivity improvement was achieved by coating the entrance windows of the PMTs with a wavelength-shifting material. This material absorbs ultraviolet (UV) light with wavelengths below 300 nm and re-emits two photons at a larger wavelength. These enhancements aim to improve the overall performance and detection efficiency of the LTCC system [Ungaro et al., 2020a]. A drawing from the design model of the LTCC can be seen in Figure 9.

Forward Time-of-Flight (FTOF)

The FTOF system is designed to measure the time-of-flight of charged particles emerging from the target during beam operation. It consists of six sectors of plastic scintillators with double-sided PMT readout. Each sector is divided into three arrays of counters separated into panels. Panel-1a has 23 counters, panel-1b has 62 counters, and panel-2 has 5 counters. The FTOF system is designed to provide excellent timing resolution for particle identification and good segmentation for flexible triggering options. The detectors cover a polar angle range from 5° to 45° , spanning 50% in azimuth at 5° and 90% at 45° . The lengths of the counters vary across the panels, ranging from 32.3 cm to 376.1 cm in panel-1a, from 17.3 cm to 407.9 cm in panel-1b, and from 371.3 cm to 426.2 cm in panel-2.

The timing resolution achieved in the FTOF system is 125 ps in panel-1a, 85 ps in panel-1b, and 155 ps in panel-2. This timing resolution allows for precise measurements of the particle's time-of-flight, which is crucial for particle identification purposes [Carman et al., 2020b]. A render of the FTOF detector can be seen in Figure 10.

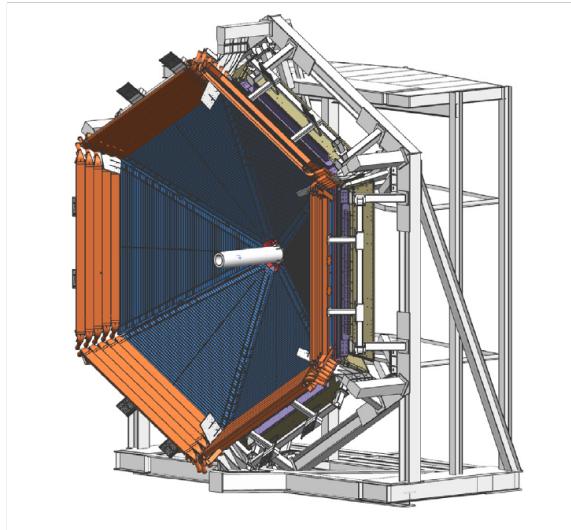


Figure 10: Render of the Forward Carriage with the FTOF system showing the panel-1b counters on the inside (dark blue), and the panel-2 counters on the outside (bronze). The panel-1a counters are located immediately downstream of the panel-1b counters and are not visible in the render. Part of the Pre-shower Calorimeter (PCAL) is visible downstream of the FTOF panels.

Source: CLAS12 wiki.

Ring Imaging Cherenkov Detector (RICH)

To improve particle identification in the momentum range of 3 – 8 GeV, a RICH detector was incorporated into one of the CLAS12 sectors, replacing the corresponding LTCC sector. The RICH detector enhances CLAS12’s capabilities in separating kaons from pions.

The RICH detector consists of aerogel radiators, visible light photon detectors, and a focusing mirror system. The focusing mirror system is designed to reduce the instrumented detection area to 1 m². Multi-anode PMTs are used as the photon detectors, providing the necessary spatial resolution and matching the aerogel Cherenkov light spectrum in the visible and near-UV region.

For forward scattered particles with momenta between 3 and 8 GeV and angles up to 13°, a proximity imaging method is employed. This method utilises a thin (2 cm) aerogel radiator for direct Cherenkov light detection.

For particles with larger incident angles between 13° and 25° and momenta ranging from 3 to 6 GeV, a different configuration is used. The Cherenkov light is produced by a thicker aerogel layer of 6 cm and focused by a spherical mirror. The light then undergoes two additional passes through thin radiator material and is reflected by planar mirrors before detection [Contalbrigo et al., 2020]. This setup enables efficient particle identification in this momentum and angular range.

Electromagnetic Calorimeters (ECAL)

The CLAS12 package incorporates the existing Electromagnetic Calorimeter (EC) from the CLAS detector and adds a new PCAL upstream of it. Together, they form the ECAL, which is primarily used for the identification and kinematical reconstruction of electrons, photons, and neutrons.

The ECAL consists of six modules, with the PCAL and EC divided into two parts each along the direction from the target. These parts are known as EC-inner (ECIN) and EC-outer (ECOU) and are read out separately. They provide longitudinal sampling of electromagnetic showers and also help improve particle identification through hadronic interactions.

Each module has a triangular shape and is composed of 54 layers of scintillators. The scintillators are 1 cm thick and segmented into strips that are 4.5 cm wide for PCAL and 10 cm wide for EC, fitted between 2.2-mm-thick lead sheets. The total thickness of the calorimeters corresponds to approximately 20.5 radiation lengths. The scintillator layers are grouped into three readout views, with 5/5/8 layers per view for PCAL/ECIN/ECOU, respectively. This arrangement allows for spatial resolutions of less than 2 cm for energy clusters.

To transmit the light signals from the scintillators, flexible optical fibres are used to route the light to the corresponding PMTs [Asryan et al., 2020].

Forward Tagger (FT)

The FT is an extension of the CLAS12 detector that allows for the detection of electrons and photons at very forward polar angles, specifically ranging from 2.5° to 4.5°. By detecting forward-scattered electrons, the FT enables electroproduction experiments at low photon virtuality Q^2 , providing a high-intensity, linearly polarized,

quasi-real photon beam with energy tagging. This setup is particularly suitable for hadron spectroscopy studies.

The FT consists of three main components: the FTCal (calorimeter), the FT-Trk (micro-strip gas tracker), and the FTHodo (hodoscope). The FTCal utilises 332 lead-tungstate (PbWO_4) crystals to identify electrons, measure the energy of electromagnetic showers, and provide fast trigger signals. The FTTrk, located in front of the FTCal, is responsible for measuring the scattering angles of charged particles. The FTHodo, a scintillator detector, assists in the separation of electrons and high-energy photons.

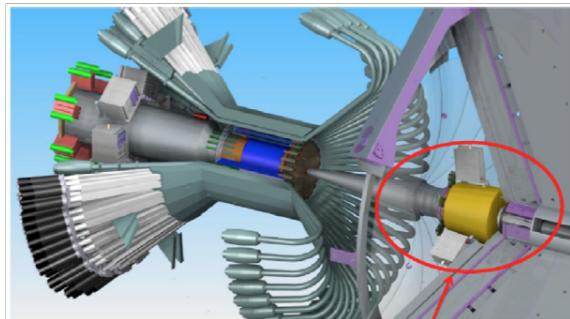


Figure 11: The Forward Tagger (FT) system circled downstream of the CD in front of the torus magnet warm bore entrance.

Source: CLAS12 wiki.

the FT detectors can be turned off, and additional shielding elements are installed in front of the FT, covering up to 4.5° of polar angle. This modified configuration, known as "FT-Off," reduces accidental background by one-third under the same beam conditions, enabling higher luminosity data acquisition with CLAS12 by mitigating background interference from the DC R1 chambers [Acker et al., 2020a].

A render of the FT can be seen in Figure 11.

2.2.2 Central Detector (CD)

In the CD, particles scattered from the target within the polar angle range of 35° to 125° are detected. The CD consists of various detectors that provide particle identification and tracking capabilities. Charged particles are detected in the Central Vertex Tracker (CVT) and the Central Time-of-Flight (CTOF) detector. Neutron detection is provided by the Central Neutron Detector (CND), which is located radially outside of the CVT and the CTOF. All detectors have full coverage in the azimuthal angle.

Central Vertex Tracker (CVT)

The CVT system is an integral part of the CD. It is primarily used for measuring the momentum and determining the vertex position of charged particles scattered from the target.

The CVT system is located inside the solenoid magnet, as depicted in Figure 12. It consists of two distinct detectors: the SVT and the Barrel Micromegas Tracker (BMT).

The SVT system is composed of three regions, each equipped with double-sided modules of silicon sensors. The regions have different numbers of modules: 10, 14, and 18, respectively. These silicon sensors are instrumented with a digital Application-Specific Integrated Circuit (ASIC) readout. The readout pitch, which refers to the distance between adjacent readout channels, is approximately 156 micrometers. The SVT system comprises 21,504 channels [Antonioli et al., 2020].

The BMT is composed of three layers of strips aligned along the beamline and three layers of circular readout strips around the beamline, totalling 15,000 readout elements. It significantly enhances momentum resolution and tracking efficiency. Each layer is divided azimuthally into three segments, providing 120° azimuthal coverage for each segment. The system is designed to operate at the full luminosity of $10^{35} \text{ cm}^{-2} \text{s}^{-1}$ [Acker et al., 2020b].

Central Time-of-Flight (CTOF)

The CTOF system is utilised for the purpose of charged particle identification by measuring their time-of-flight (TOF) in the momentum range of approximately 0.3 to 1.25 GeV. It consists of 48 plastic scintillators with double-sided PMT readout. The PMTs are connected to the scintillators via 1.0-meter-long upstream and 1.6-meter-long downstream focusing light guides.

These counters are arranged in a hermetic barrel configuration surrounding the target and the CVT, and they are aligned with the beam axis inside the 5 T solenoid magnet.

To ensure accurate measurements, the PMTs are positioned within a region of 0.1 T fringe field of the solenoid magnet and are enclosed within a triple-layer dynamic magnetic shield. This shield minimises the internal magnetic field near the PMT photocathode, achieving a field strength of less than 0.2 G. The CTOF system is designed to provide a time resolution of 80 ps, enabling precise charged particle identification in the CD [Carman et al., 2020a].

For a visual representation of the CTOF system, you can refer to Figure 13.

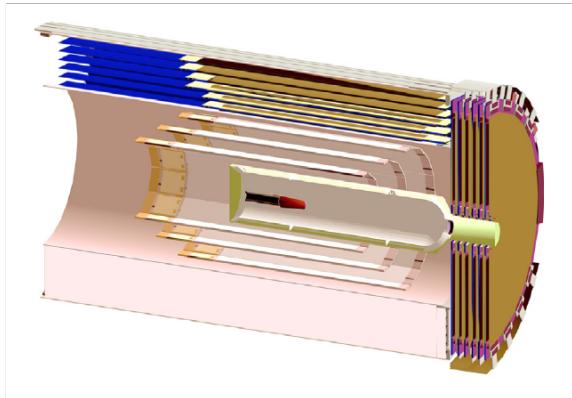


Figure 12: Render of the Central Vertex Tracker (CVT). From the inside, the figure shows the target cell and vacuum chamber, the three double layers of the SVT, followed by the six layers of the BMT. The beam enters from the left. The six Forward Micromegas Tracker (FMT) layers are shown at the downstream end at the right.

Source: CLAS12 wiki.

Central Neutron Detector (CND)

The CND is a component of the CD positioned radially outward of the CTOF system. Its primary function is to detect neutrons in the momentum range of 0.2 to 1.0 GeV by measuring their TOF from the target and the energy deposition in scintillator layers.

The CND is composed of three layers of scintillator paddles, with each layer containing 48 paddles. The paddles are coupled in pairs at the downstream end using semi-circular light guides. The signal generated in the scintillator paddles is read out at the upstream end by PMTs that are positioned outside the high magnetic field region of the solenoid magnet.

To transmit the scintillation light, the scintillators are connected to 1-meter-long bent light guides, which ensure efficient light propagation to the PMTs for signal detection and readout.

The combination of TOF measurements and energy deposition in the scintillator layers enables the CND to identify and detect neutrons within the specified momentum range in the CD [Chatagnon et al., 2020].

Back Angle Neutron Detector (BAND)

The CLAS12 spectrometer includes the BAND as a dedicated detector for neutron detection at backward angles. Positioned 3 metres upstream of the target, the BAND is designed to detect backward-scattered neutrons with momenta ranging from 0.25 to 0.7 GeV.

The BAND detector consists of 18 horizontal rows and five layers of scintillator bars. Each scintillator bar is equipped with PMT readout on both ends to measure the TOF of neutrons originating from the target. Additionally, there is an extra 1 cm scintillation layer specifically designed to veto charged particles, ensuring that only neutrons are detected.

Covering a polar angle range from 155° to 175° , the BAND detector achieves a design neutron detection efficiency of 35%. It also provides a momentum resolution of approximately 1.5%, allowing for precise measurements of the momentum of the detected neutrons [Segarra et al., 2020].

By utilising the BAND detector, the CLAS12 spectrometer is capable of detecting and characterising backward-scattered neutrons, providing valuable information for various physics studies and experiments conducted at CLAS12.

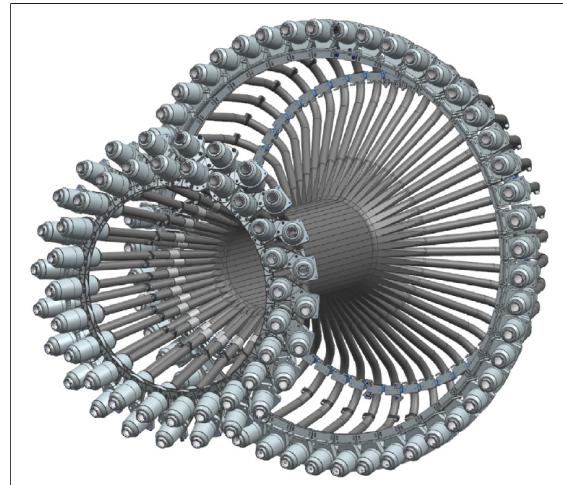


Figure 13: Render of the Central Time-of-Flight (CTOF). The render shows CTOF's 48 scintillator bars outfitted with light guides, PMTs, and magnetic shields at both ends of each counter.

Source: CLAS12 wiki.

2.2.3 Offline Reconstruction

The CLAS12 reconstruction and analysis process is facilitated by a data-stream processing framework called CLARA. CLARA adopts a service-oriented architecture, allowing the construction of software applications using micro-services that are connected via data-stream pipes [Gyurgyan et al., 2016].

In this framework, each service plays a specific role. It receives input data, processes it according to its functionality, and produces output data. The input and output data are organised in tabular structures known as “banks”, which are configured by the service developer to match the specific requirements of the service.

The services within CLARA form a data-flow path, where the output of one service becomes the input for the next service in the sequence. This design enables a flexible and versatile data processing application, as each service can be individually improved or replaced without necessitating structural changes to the framework.

To ensure consistency and modularity, the CLAS12 services are extensions of an abstract reconstruction engine. This engine provides common components such as initialisation and event processing methods, reducing the development complexity of individual micro-services and enforcing a uniform structure throughout the framework.

By leveraging the CLARA framework, the CLAS12 experiment benefits from a modular and adaptable data processing pipeline, allowing for efficient reconstruction and analysis of the collected data. The service-oriented architecture and data-stream processing approach contribute to the flexibility, scalability, and maintainability of the CLAS12 software framework.

The CLAS12 data reconstruction process involves data reader services that access decoded detector data stored in banks. Each entry in the bank represents a decoded detector hit and contains information such as sector, layer, component, order, and digitised data like signal charge, amplitude, time, or pedestal.

During the decoding stage, similar bank structures are created for various quantities required for event reconstruction, including hits, clusters, and tracks. Reconstruction algorithms specific to each CLAS12 subsystem fill these banks. The data persistence service appends and writes these banks to a file for later analysis.

The reconstruction algorithms are implemented as services that operate on input banks and produce output banks, which are then passed to subsequent algorithms in the reconstruction chain. The order in which the services are chained reflects the overall sequence of CLAS12 event reconstruction and the dependencies between subsystems.

The first step is the reconstruction of charged particle tracks in the Central and Forward Detector tracking systems, based on the recorded hit positions in the respective detectors. This process is known as “hit-based” tracking.

Simultaneously, hits recorded in other detectors are processed to reconstruct the energy and time of the associated particle interactions. The Event Builder (EB) service matches these reconstructed quantities with the tracks based on position and time information. Hits that are not matched to any track are retained as neutral particle candidates. The EB also determines the event “start time” and identifies the reconstructed particles.

Once the event start time is determined, a second iteration of forward tracking, known as “time-based” tracking, can be performed. This iteration incorpo-

rates the drift times in the Drift Chambers, providing improved tracking precision [Ziegler et al., 2020].

An overview of the composition of reconstruction application services, depicting the dependencies between the services, can be found in Figure 14.

Tracking

In the CLAS12 event reconstruction, charged particle tracking plays a crucial role and is divided into two main regions: the forward region and the central region.

In the forward region, charged particles are bent either inward or outward from the beamline by the torus magnet, depending on their charge. The magnetic field strength varies along the bending path, with an integral magnetic field ($\int B dl$) ranging from 2Tm at 5° to 0.5Tm at 40° . The forward tracking system responsible for tracking in this region consists of two components: the FMT and the DC.

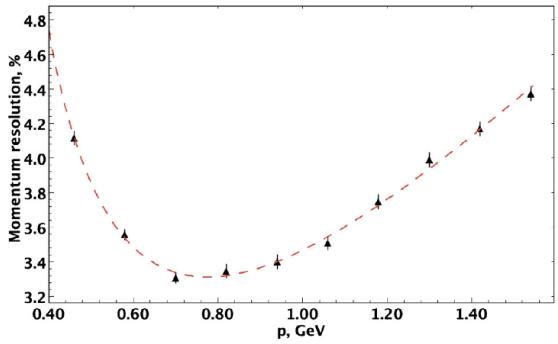


Figure 15: Momentum resolution vs. momentum of simulated protons in the CVT without background.

Source: [Ziegler et al., 2020].

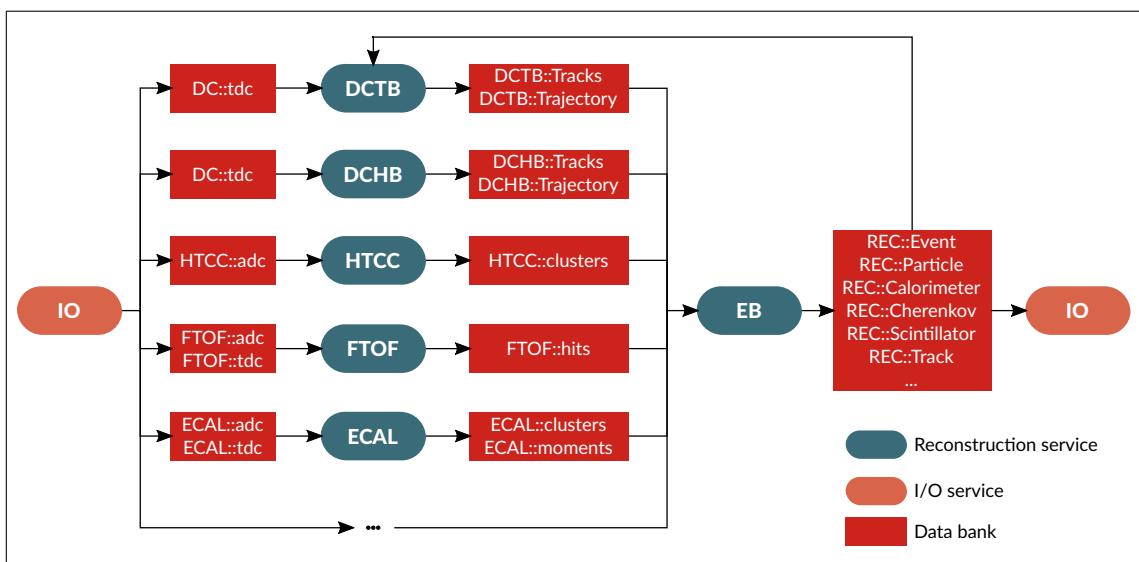


Figure 14: Graphical representation of the CLAS12 interdependencies between services and banks. The I/O service reads events from the input file and distributes them to the reconstruction services chain for processing. Each service reads the relevant banks, applies the reconstruction algorithm, and provides output banks that are passed to the next service in the chain. The Event Builder (EB) service is executed as last in the chain; it collects the relevant banks from all CLAS12 subsystems services and produces event, particle, and detector response banks that are written to the output file.

Source: Own elaboration, using Inkscape.

In the central region, charged tracks are curved into helices by the strong 5T solenoidal magnetic field. The central tracking system comprises the SVT and the BMT, which together form the CVT.

These tracking systems in both the forward and central regions use sophisticated detectors to measure the position of charged particle hits, allowing for the reconstruction of particle trajectories. By combining information from these detectors and utilising the magnetic field information, the tracking algorithms reconstruct the paths of charged particles, enabling precise determination of their momenta and vertices.

In both the forward and central tracking systems, track reconstruction involves two main steps: pattern recognition and track fitting. The first step is to identify hits, which are the recorded signals corresponding to a particle passing through a specific detector component. These hits are transformed from electronic signals into the position of the track within the geometry of the detector subsystem.

A hit is defined as a geometric object that represents a detector element. For example, in the central tracker, a hit can be represented by a line corresponding to a strip in the detector. These hit objects serve as the input for the pattern recognition algorithms.

Pattern recognition involves identifying clusters of hits and determining the spatial coordinates and corresponding uncertainties for the hits and clusters. During the pattern recognition stage, hits that are consistent with belonging to a trajectory, such as a particle track, are identified. This set of hits is then fitted to the expected trajectory, considering their uncertainties and incorporating knowledge of the detector material and the detailed magnetic field map. Figure 17 illustrates a particle passing through the DC.

The momentum resolutions in the central and forward trackers, as a function of momentum, are shown in Figures 15 and 16, respectively. The distributions are fitted with a function of the form $\sqrt{a + bx^2 + c/(1 + d/x^2)}$. In both distributions, the degradation of resolution at low momentum is attributed to multiple scattering effects. Furthermore, the resolution deteriorates as momentum increases beyond a minimum, primarily because of poorer track curvature resolution.

For central tracking, a simulated proton sample with momenta ranging from 0.5 to 2.5 GeV yields an average CVT reconstruction efficiency of 87.3%. A slight decrease in efficiency is observed for tracks with momenta below 600 MeV. The increased curvature of low transverse momentum (p_{\perp}) tracks leads to a rise in inefficiency due to acceptance effects. The primary source of inefficiency stems from the gaps between the sensitive volumes of the BMT and SVT detectors.

Regarding forward tracking, the momentum resolution in the DC is evaluated

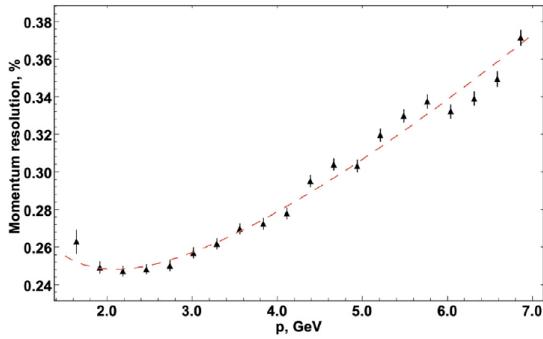


Figure 16: Momentum resolution vs. momentum in the DC evaluated using pions simulated at $\theta = 15^\circ \pm 5^\circ$ and at $\phi = 0 \pm 5^\circ$ without background.

Source: [Ziegler et al., 2020].

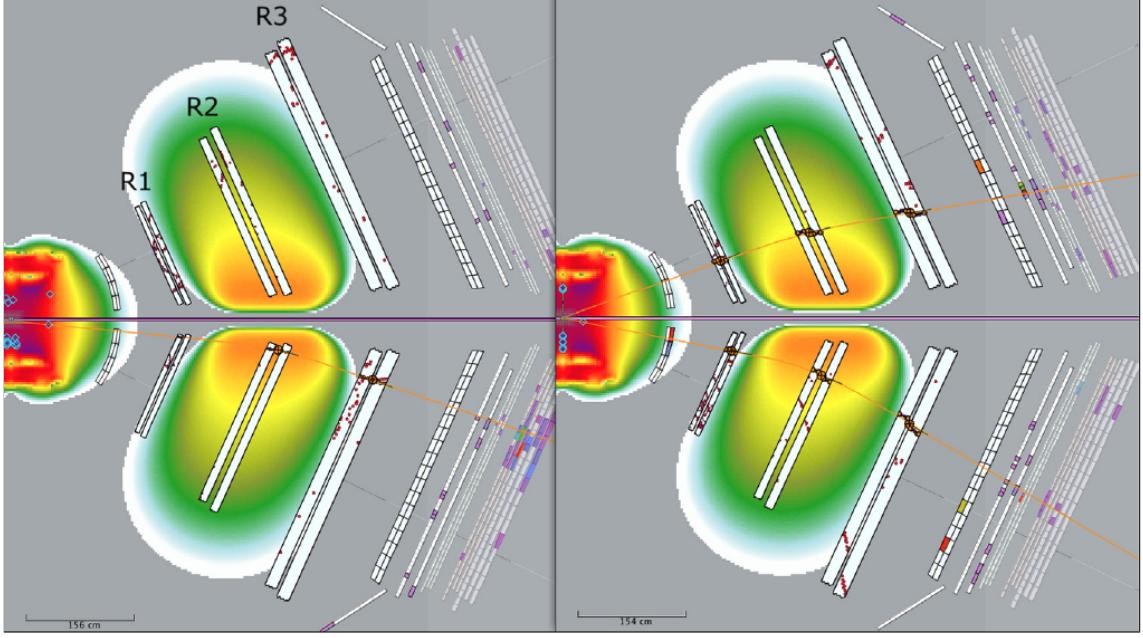


Figure 17: Views from CLAS12 Event Display (ced) of charged particle tracks in the DC showing cut-views to highlight different pairs of sectors of the CLAS12 Forward Detector. The coloured detector elements are the registered hits and the orange lines are the result of track reconstruction using the hits in the DC. The coloured areas about the detectors represent the regions of magnetic field from the torus and the solenoid. In these views the beam is incident from the left and the target is located in the middle of the solenoid (at the left edge of the image).

Source: ced render.

using tracks simulated at $\theta = 15^\circ \pm 5^\circ$ and $\phi = 0 \pm 5^\circ$. This range ensures that the majority of tracks fall within the sensitive volume. Moreover, the DC momentum resolution exhibits correlation with the polar angle since track curvature is determined by the magnetic field intensity, which is higher at lower angles in the torus field. These resolutions are obtained from a Monte Carlo sample that excludes out-of-time backgrounds or misalignments of the tracking volumes [Ziegler et al., 2020].

Particle Identification

The Particle Identification (PID) numbering scheme described here was initially introduced by the Particle Data Group (PDG) in 1988 [Yost et al., 1988]. Its purpose is to facilitate communication and data exchange between different generators, simulators, and analysis packages employed in particle physics. The system underwent subsequent revisions and adaptations in 1998 to allow for the systematic inclusion of undiscovered and hypothetical particles [Group et al., 1998]. The PID convention utilised in this thesis is based on the most up-to-date version available at the time of writing, as referenced from the 2020 Review of Particle Physics by the PDG [Group et al., 2020].

The EB is a crucial component within the reconstruction chain, serving multiple functions:

- It gathers information from upstream services.
- It correlates information from sub-detectors to form particles.
- It implements a general particle identification scheme.
- It organises the resulting information into a standardised and persistent data bank structure.

The EB service is executed twice using identical algorithms, first employing hit-based tracks and subsequently using time-based tracks. As mentioned previously, the results obtained from the hit-based EB are utilised to initialise time-based tracking.

In the definition of a reconstructed charged particle within CLAS12, the EB assumes that each reconstructed track in both the FD and the CD will be assigned an identification. The corresponding responses from the calorimeter, scintillator, and Cherenkov detectors are then associated with that particle based on geometric coincidences between the detector responses and the track. Matching criteria are established, which correspond to the resolution of each specific detector. The geometric matching process relies on the Distance of Closest Approach (DOCA) between the track and the detector response.

A similar procedure is employed for the creation of neutral particles, with the distinction that the seeding is presently performed using unassociated responses from the ECAL for the FD and the CND (or the BAND) for the CD, instead of using tracks.

A start time is assigned to the entire event and serves as the precise reference time for all time-based particle identification procedures. The determination of the start time relies on the optimal charged particle candidate in the FD with an associated timing response from the FTOF detector.

The EB assigns the start time based on the highest energy electron detected in the ECAL. If no electron is found in the ECAL, the EB then searches for a positron in the ECAL. In the absence of any lepton candidates, the next track in the priority list is a forward-going positive track, which is assumed to be a positive pion (π^+). Finally, if no forward-going positive track is identified, the EB searches for a forward-going negative track, assumed to be a negative pion (π^-). When searching for π^+ or π^- tracks, only the candidate with the highest momentum within each group is considered.

A parallel event start time is determined from the FT system to facilitate physics analyses and triggers specifically for events where the primary scattered electron is at very forward angles within the FT.

In such cases, all combinations of charged particles in both the FT and the FD are taken into account. The particle in the FT is assumed to be an electron, while all possible hadron mass hypotheses are considered for the FD tracks. The combination that exhibits the best time coincidence is selected, and the timing of the resulting FT electron is used to assign the start time.

A correction to the start time is subsequently applied using the RF signal from the accelerator, in conjunction with the reconstructed event vertex position. This correction effectively aligns the event start time with the most accurate measurement of the beam-bunch arrival time at the target.

The uncorrected measured vertex time of a particle, denoted as t_v , can be expressed as follows

$$t_v = t - \frac{P_L}{\beta c},$$

Here, t represents the measured time response (e.g., in a scintillator), P_L is the path length between the primary interaction vertex and the corresponding response, and βc denotes the speed of the particle.

Next, we calculate the time difference Δt_{RF} between t_v and the nearest beam bunch using the following formula

$$\Delta t_{RF} = t_v + \frac{z_0 - z_v}{c} - t_{RF} - \frac{N}{2f_{RF}},$$

In this equation, z_v represents the z -coordinate of the event vertex position, z_0 is the reference position calibration at the center of the target, and c denotes the speed of light in vacuum. t_{RF} and f_{RF} correspond to the measured and calibrated RF time and frequency of the accelerator. These values can either be 2.004 ns and 249.5 MHz, or 4.008 ns and 499 MHz, respectively. During the reconstruction process, these values are obtained from the Run Conditions Database.

Subsequently, the time can be further corrected to the nearest beam bunch using the equation

$$\Delta t'_{RF} = \text{mod} \left(\Delta t_{RF}, \frac{1}{f_{RF}} \right) - \frac{1}{2f_{RF}},$$

This correction allows for RF-correction to t_v . Thus, we obtain a final RF- and vertex-corrected start time for the event, defined as

$$t' = t_v - \Delta t'_{RF}.$$

The subsequent step involves a basic particle identification scheme designed to be flexible enough to accommodate various physics analyses while retaining essential information for future refinement of the criteria.

For charged particles, the identification process begins by utilising calorimetry and Cherenkov information to positively identify e^-/e^+ candidates in the FD. If the measured energy deposition aligns with the expected sampling fraction of the ECAL and the photoelectron response from the HTCC aligns with $\beta \sim 1$, the particle is assigned as an e^-

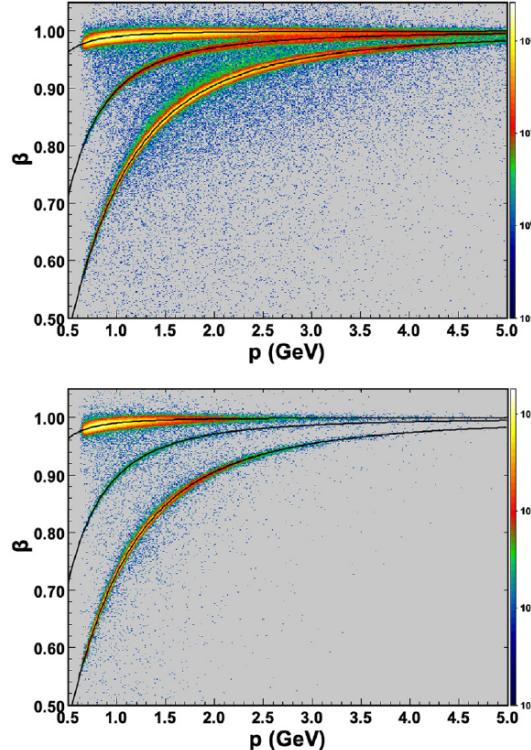


Figure 18: Particle β vs. momentum from simulation data for positively charged tracks with their start time from an electron in the FD (top plot) or in the FT (bottom plot).

Source: [Ziegler et al., 2020].

or e^+ based on the sign of curvature determined from forward tracking with the DC in the presence of the torus magnetic field.

The remaining charged particles are then assumed to be hadrons and are assigned an identity solely based on timing information. The p, K, π candidate with the smallest time residual is selected. This time residual is calculated as the difference between the measured flight time of the particle and the flight time computed for a specific mass hypothesis.

Figure 18 presents the β vs. momentum distributions for forward-going positively charged hadrons reconstructed using data from the FTOF and DC subsystems. The electron is reconstructed either in the FD (top) or in the FT (bottom). The computed curves for different mass hypotheses are superimposed on the distributions.

To identify neutral particles, the analysis assumes the presence of only neutrons and photons, distinguished solely by timing and topological information. In the FD, the identification is based on the ECAL, while in the CD, it relies on the CND. The reconstructed cluster positions of these detectors are used to calculate the particle's travel path from the event vertex, assuming a straight-line trajectory.

If the measured β value is close to 1, the particle is identified as a photon; otherwise, it is identified as a neutron. For photons in the FD, the momentum is determined from the deposited energy and the ECAL sampling fraction [Asryan et al., 2020]. For neutrons, the momentum is assigned based on the measured β , assuming the mass of a neutron.

Figure 19 illustrates an example of the reconstructed β values for neutral particles in the FD, demonstrating the separation between photons and neutrons.

A particle identification quality factor, represented as a signed- χ^2 or pull, is assigned based on the contributions from individual detector subsystem responses and their resolutions. For e^-/e^+ identification, the resolution-normalised distance from the expected ECAL sampling fraction is utilised, while for charged hadrons, the resolution-normalised time difference is employed. The resulting information is organised into standardised output bank structures for physics analysis. This includes the particle four-vectors, associated detector responses, and global event information such as beam RF and helicity details.

The accuracy of the currently implemented particle identification algorithm can be estimated by comparing the assigned particle identification with the true identification in Monte Carlo simulations. Table 1 presents the particle identification matrix for the FD (left) and CD (right). The values are derived from simulations in-

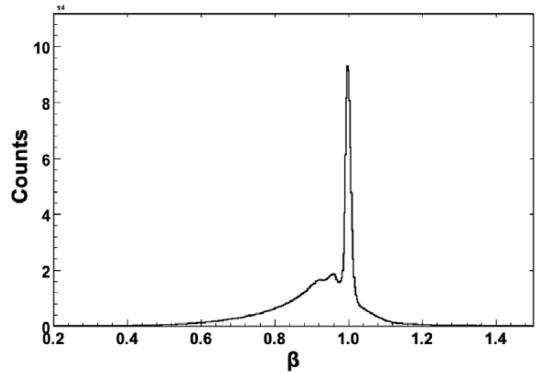


Figure 19: β distribution for neutral particles as measured by the ECAL from simulation data, showing a sharp peak at $\beta = 1$ from photons and a broader, slower distribution from neutrons.

Source: [Ziegler et al., 2020].

Forward Detector Truth							Central Detector Truth				
	e	π	K	p	n	γ		π	K	p	n
e	0.98							0.84	0.14	0.00	
π		0.93	0.10	0.00				0.11	0.80	0.01	
K		0.03	0.80	0.00				0.03	0.04	0.95	
p		0.03	0.02	0.98							0.11
n					0.66	0.01					
γ					0.14	0.95					

Table 1: Particle identification matrix for the FD (left matrix) and CD (right matrix). The FD matrix is based on simulated hadrons and photons with momentum between 1 and 2.5 GeV, and electrons up to 9 GeV. The CD matrix is based on simulated hadrons with momentum between 0.3 and 1.1 GeV. The diagonal elements are correctly identified, while the off-diagonal elements are misidentified. Detector inefficiencies are included.

volving electron-hadron or electron-photon pairs with hadron and photon momenta ranging from 1 to 2.5 GeV and electron momenta ranging from 1 to 9 GeV. The diagonal elements represent cases where the particle is correctly identified, while the off-diagonal elements represent cases of misidentification [Ziegler et al., 2020].

2.3 RG-E Experiment

The RG-E experiment, to be conducted in Hall B, focuses on measuring the hadronic multiplicity ratio between various nuclei and deuterium. To achieve this, a double-target system is being constructed. This system will enable a precise comparison between a deuterium target and heavy solid targets such as carbon, aluminium, copper, tin, lead, and others. The primary objective of the experiment is to enhance our understanding of hadronisation in the nuclear medium, colour transparency, and nuclear short-range correlations.

During the data acquisition process, both the cryo-target (deuterium) and a solid target will be exposed to the electron beam simultaneously. To minimize differences in acceptance corrections between the targets, efforts are made to maintain a minimal distance between them. However, it is crucial to ensure that the targets can still be differentiated during the reconstruction process. The work behind this thesis contributes to the experiment by improving the offline reconstruction methods prior to the actual experiment, with the aim of reducing the distance between the targets. Further details on this topic can be found in Section 3.

Simultaneously positioning both targets in the beam allows for the cancellation of time-dependent systematic effects, thereby increasing the precision of the final results. These effects can include drifting gains and inefficient detector channels when measuring ratio-like observables. Since the target system will be placed in a vacuum, the switching between solid targets needs to be done remotely. It is essential to perform the switching quickly to maximise the beam time available for the experiment.

The double target system is designed to enable switching between up to five different solid targets. Previous experiments, such as EG2 conducted on the former

CLAS, have demonstrated that the design of the double target system offers significant advantages in reducing systematic uncertainties [Hakobyan et al., 2008]. The principle of the target system involves installing the solid targets on a carbon fibre band that slides on torlon rails. The movement of the band is accomplished using a piezo-motor, chosen for its insensitivity to magnetic fields. Figure 20 illustrates the proposed design of the double target system.

To ensure the proper functioning of the target, a series of tests have been conducted. These tests cover various aspects of the target's operation and performance. They include:

- Movement under vacuum and high magnetic field: The target system has been tested to ensure that it can operate smoothly and reliably under vacuum conditions and in the presence of high magnetic fields. This is important to verify that the target's movement and positioning are not affected by these conditions.
- Radiation hardness of selected materials: The materials used in the target system have been tested for their resistance to radiation. This is crucial because the target may be exposed to radiation during the experiment. By assessing the radiation hardness of the materials, it can be ensured that they will maintain their structural integrity and functionality throughout the experiment.
- Heat removal tests: The target system has undergone heat removal tests to ensure that it can effectively dissipate heat generated during operation. It is essential to keep the target temperature well below its melting point to prevent

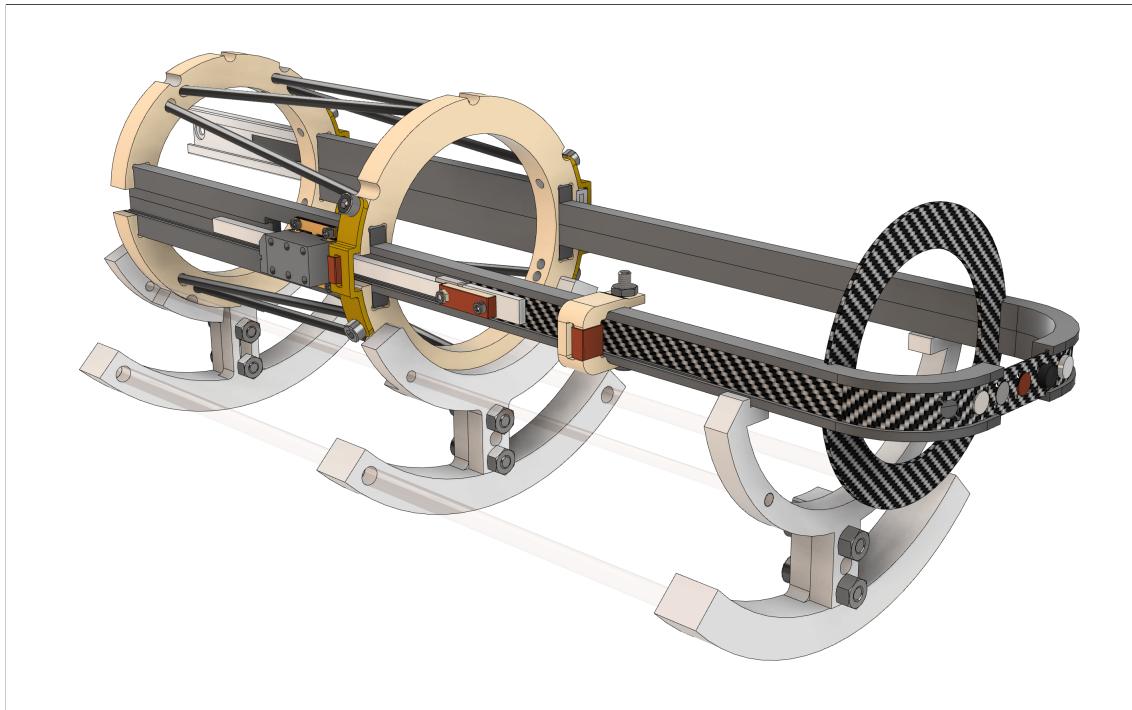


Figure 20: Double-target system CAD render.

Source: Render by Alonso Lepe, from the Double Target group.

any damage or degradation. By conducting heat removal tests, the target's cooling capabilities have been evaluated to ensure optimal performance and longevity.

These tests collectively aim to validate the target's operational robustness and reliability under the specific conditions of the experiment, thereby ensuring accurate and consistent data collection.

2.3.1 Slow Control System

A modern experimental physics experiment usually involves many moving parts and complex components that require constant monitoring and calibration. Additionally, continuous data collection is necessary. This task is accomplished through the use of slow control systems. These systems typically integrate the entire detector and experiment into one complex interface, facilitating easy access and maintenance.

The standard framework used in the context of HEP for achieving this is the Experimental Physics and Industrial Control System (EPICS). EPICS was developed at the Los Alamos National Laboratory to facilitate data acquisition and control for such experiments. This framework offers a distributed process control system that includes software communication, functional subsystems for data acquisition, supervisory control, closed-loop control, channel archiving, and alarm management [Dalesio et al., 1991].

Similar to many HEP experiments, the slow control system of CLAS12 is based on EPICS [Boyarinov et al., 2020].

To enable the integration of the RG-E target into the CLAS12 slow control system, the author developed an EPICS support module and Input/Output Controller (IOC). This module was not created from scratch since the motor developers had already developed a generic support module for Galil motors [Farnsworth, 2009]. The existing system proved to be capable of supporting the movement of the RG-E target system, requiring only the removal of unnecessary features and the addition of experiment-specific database variables.

The RG-E IOC, along with the complete set of EPICS support modules necessary to run it, can be found at

github.com/bleaktwig/rge-epics-support.

The EPICS module includes Process Variables (PVs) that have been added to facilitate the control of the RG-E target system. Below is a list of these PVs along with a brief description of each. The user-defined records database, where these PVs can be viewed and edited, is located at

`$EPICS_BASE/support/galil/3-6/db/galil_userdef_records.template`.

Analog Input (ai)

The normal use for this record type is to obtain an analog value from hardware and then convert it to engineering units [Stanley et al., 1998]. The record can also be used to write constants to be read from the database, such that they can be changed in runtime.

DMC01:A_curr_pos. Current position of the band. Displayed at the Graphical User Interface (GUI) and used for internal calculations.

DMC01:A_home. Position of the home. Displayed at the GUI and used for calculations.

DMC01:A_pos#. # is a number from 1 to 7. Positions of each of the seven targets. Displayed at the GUI and used for calculations.

DMC01:A_lowlimit. Position of the low limit. If DMC01:A_curr_pos is lesser than this value, a major alarm is fired.

DMC01:A_highlimit. Position of the high limit. If DMC01:A_curr_pos is greater than this value, a major alarm is fired.

DMC01:A_tolerance. Equivalence tolerance for the position of each target and the position of the band. It defines a valid range around the target position.

DMC01:COMMERR_STATUS. Variable that is true when there's a communication error with the controller and false otherwise. Used for triggering a communication alarm.

IOC01:SR_i_am_alive. Variable that is true when the IOC is up and running, false otherwise. Used for triggering a communication alarm.

Analog Output (ao)

The normal use for this record type is to output values to digital-analog converters. The desired output can be controlled by either an operator or a state program, or it can be fetched from another record [Stanley et al., 1998].

DMC01:A_go_home. Command to move the band to the home position, as defined in DMC01:A_home.

DMC01:A_go_pos#. Command to move the band to the position of target #, as defined in DMC01:A_pos#.

Calculation (calc)

The calculation or “Calc” record is used to perform algebraic, relational, and logical operations on values retrieved from other records. The result of its operations can then be accessed by another record so that it can be used [Stanley et al., 1998]. In the context of the RG-E target, each calculation returns a number from 0 to 11. This number represent the state the target is in, and is later used by a Select PV.

DMC01:A_at_pos#. Calculation that checks if the band position is equal to that of target # in DMC01:A_pos#, within the tolerance margin DMC01:A_tolerance. If it is, it returns #. Otherwise, it returns 0.

DMC01:A_at_home. Calculation that checks if the band position is equal to the home position in DMC01:A_home, within the tolerance margin defined by the tolerance. If it is, it returns 8. Otherwise, it returns 0.

DMC01:A_moving. Calculation that checks if the target is moving by checking the motor PV DMC01:A.MOVN. If it is, it returns 9. Otherwise, it returns 0.

DMC01:A_at_lowlimit. Calculation that checks if the band position is lesser than the low limit DMC01:A_lowlimit. If it is, it returns 10. Otherwise, it returns 0.

DMC01:A_at_highlimit. Calculation that checks if the band position is greater than the high limit DMC01:A_highlimit. If it is, it returns 11. Otherwise, it returns 0.

Select (sel)

The select record computes a value based on input obtained from up to 12 locations [Stanley et al., 1998]. By default, it is equal to the highest value among its input PVs.

DMC01:A_sel_tgtype. This record returns the highest value between the previously defined calculations. Thus, it associates the values returned to a state of the target. By convention, this PV assumes that *no more than one calc is greater than 0*. This assumption holds as long as DMC01:A_tolerance is not set higher than half the distance between targets and between the targets and the lower and higher limits.

Multi-Bit Binary Input (mbbi)

The normal use for the multi-bit binary input record is to read multiple bit inputs from hardware. The binary value represents a state from a range of up to 16 states. The multi-bit input record interfaces with devices that use more than one bit [Stanley et al., 1998].

DMC01:A_tgtype. This mbbi encodes the output of DMC01:A_sel_tgtype to a string and alarm level. The encoding is specified in Table 2.

CS-Studio

The GUI of CLAS12 EPICS is built on the Control System Studio (CS-Studio) toolkit. CS-Studio is specifically designed for monitoring and operating large-scale control systems and is based on the eclipse Rich Client Platform framework

Value	Name	Alarm Severity
0	Not Moving	Major
1-7	Target 1-7	No alarm
8	Home	No alarm
9	Moving	Minor
10	Low Limit	Major
11	High Limit	Major

Table 2: Names and alarm levels for the different values of the PV DMC01:A.tgtype.

[Kasemir et al., 2007]. To integrate the RG-E target system into CLAS12 EPICS, a CS-Studio screen was developed.

Below is a list of the requirements that the screen needed to fulfil. These requirements were derived from the specifications of the physics experiments as well as the needs of the electronics team responsible for the target system.

- Buttons to move the target band to the targets and a home location.
- A Button to stop the target band in case of emergency.
- A status check on the position of the target band.



Figure 21: RG-E CS-Studio main screen. The HOME and A, B, C, etc buttons move the target strip to the corresponding location, the STOP button is an emergency, instant stop of the target system, and the Motor Screens button allows the user to open additional motor screens.

Source: Own Elaboration, using CS-Studio.

- Alarm handling for the case when the band moves beyond low and high limits.
- Alarm handling for the IOC and communication problems.

The implemented screen, shown in Figure 21, incorporates these requirements. Along with the buttons, text displays indicate the position of each target within the band. Adjacent LED displays illuminate green when the band position aligns with the target position, within the tolerance specified in the database. Additionally, four LED displays are positioned to the right of the two IOC statuses, as well as the low and high limits. These LEDs indicate alarm conditions and work in conjunction with the CLAS12 Slow Control alarm system to alert the user of any issues.

Furthermore, in addition to the aforementioned screen, users can access other screens through the Motor Screens menu button. These screens are part of the motor EPICS support module and are retained for debugging purposes. They include a manual motor movement screen, a motor setup screen, a direct command line interface with the motor, and an amplifier configuration screen. All of these screens were developed by the Galil EPICS team [Farnsworth, 2009].

To enhance user experience, the GUI was designed using the Gruvbox colour palette. This palette is specifically chosen to ensure that colours are easily distinguishable, with sufficient contrast, while also being visually appealing and comfortable for the eyes. The Gruvbox colour palette can be found on GitHub at

github.com/morhetz/gruvbox.

Alarm System

To ensure the reliable operation of the CLAS12 detector, all EPICS-controlled subsystems within it are equipped with PVs that define alarm conditions. These alarms, along with their severity, associated subsystems, and pre-defined instructions on how to respond to them, are displayed in a centralised alarm system.

For each experiment conducted in Hall B involving a non-trivial target system, a specific set of alarms is required. In the case of the RG-E target, the list of implemented alarms and their corresponding PVs is provided in Table 3.

Name	Trigger	Severity
Band is not moving	DMC01:A_tgtype = 0	Major
Band is moving	DMC01:A_tgtype = 9	Minor
Band beyond low limit	DMC01:A_tgtype = 10	Major
Band beyond high limit	DMC01:A_tgtype = 11	Major
Controller comm. error	DMC01:COMMERR_STATUS = 1	Major
IOC comm. error	IOC01:SR_i_am_alive = 0	Major

Table 3: Alarm names, triggers, and severities for the RG-E target slow control system.

3 FMT Alignment and Reconstruction

In paper, the FMT offers a 3 to 10-fold increase in resolution when compared to DC [Aune et al., 2009]. Achieving this improvement requires work on the alignment and calibration of the detector, as well as the reconstruction from its data. This chapter focuses on the work carried out in this endeavour, specifically addressing alignment and reconstruction.

The first section provides a detailed description of Micromegas detectors in general and the FMT in particular. Subsequently, the second and third sections discuss the efforts made in alignment and reconstruction, respectively. Finally, the fourth and final section of the chapter elaborates on the resolution improvement resulting from this work.

3.1 Forward Micromegas Tracker

A Micro-Mesh Gaseous Detector or Micromegas (MM) is a gaseous particle detector that is derived from wire chambers. These types of detectors are commonly used in experimental physics for detecting ionising particles. The MM detector offers very precise temporal and spatial resolution, on the order of 100 nanoseconds and below 100 micrometers [Giomataris et al., 1996].

The detector operates by amplifying the charges created by ionisation in the gas volume. Its volume is divided into two parts by a metallic micro-mesh placed less than 150 micrometers away from the readout electrode or strip.

To illustrate this process, refer to Figure 22.

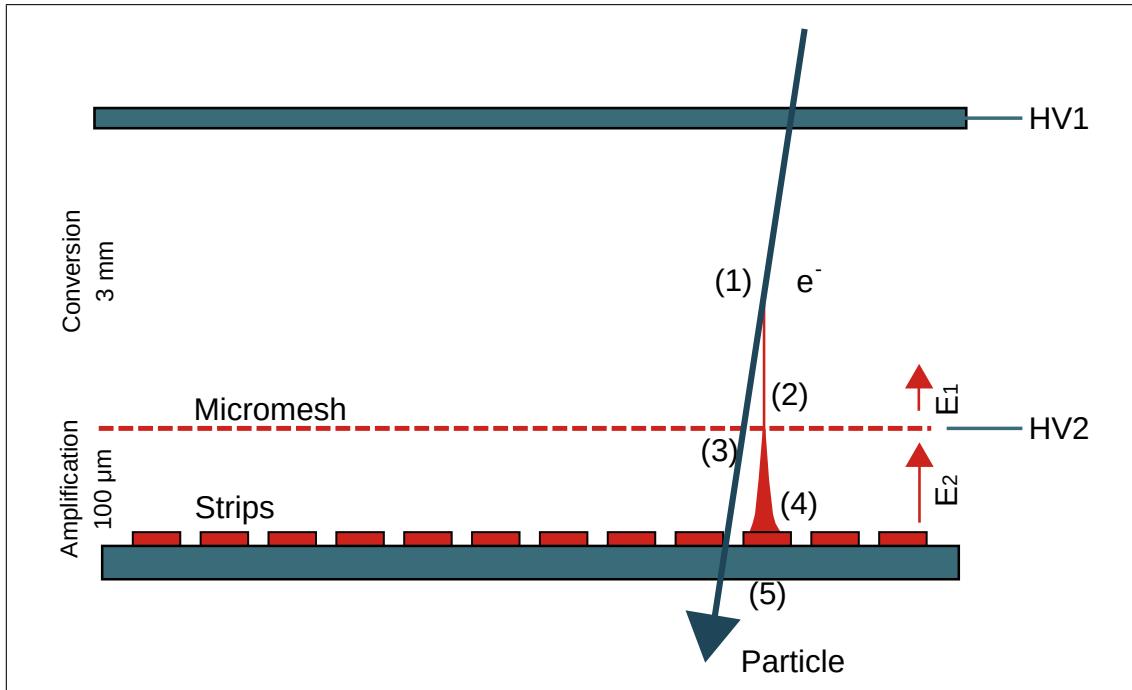


Figure 22: Micromegas (MM) working principle.

Source: Author's re-render of original illustration in [Giomataris et al., 1996], using Inkscape.

When a particle passes through the detector, it ionises the gas atoms by stripping an electron, resulting in an electron-ion pair (1). An electric field, denoted as E_1 , is applied to the gas, causing the electron to drift towards the amplification electrode (2) while the ion moves towards the cathode. As the electron crosses the mesh (3), it enters a strong electric field, denoted as E_2 , triggering an avalanche effect (4). This produces a significant signal at the readout strip (5), which can then be stored for reconstruction [Giomataris et al., 1996].

Inside the CLAS12 detector, a diverse set of tracking detectors is utilised to determine the positions and momenta of particles at different points along their trajectories. The proximity of these detectors to the particle source directly correlates to the precision achieved in determining the position and momentum at the interaction vertex, which refers to the point where the particle was generated. To optimise this precision, the Micromegas Vertex Tracker (MVT) is positioned as close as feasible to the target, as depicted in Figure 23.

Similar to the division of CLAS12 into the CD and FD, the MVT is also divided into two parts to maximise the angular coverage.

The first component is the BMT, which consists of 18 cylindrical detectors arranged in 6 layers. When combined with the SVT, this detector covers the angular region from 35 to 125 degrees, significantly enhancing the resolution of the polar

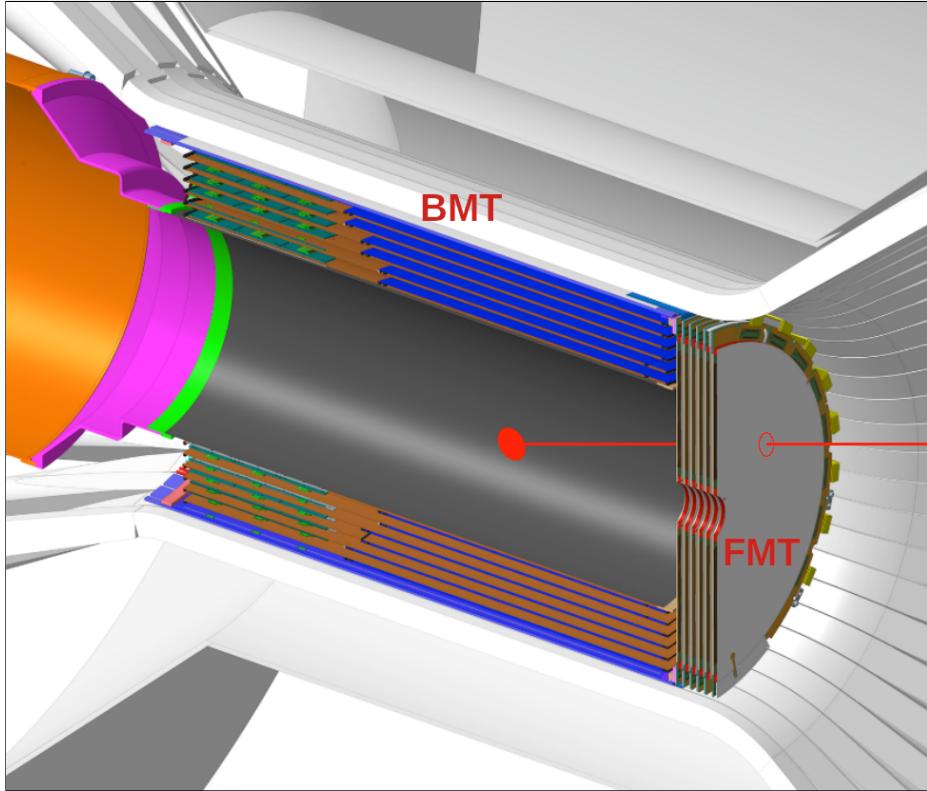


Figure 23: Micromegas Vertex Tracker (MVT) detector. The red dot denotes the $z = 0$ point in the beamline, the red line denotes an arbitrary track coming from that point, and the circumference in the FMT denotes where that track produces a signal in an FMT layer.

Source: CLAS12 wiki.

angle [Acker et al., 2020b].

Then, the FMT, which is made of six circular, flat detectors covering angles from 6 to 29°. In theory, it should improve the vertex resolution by a factor of 3 to 10 when compared to the DC [Aune et al., 2009]. While the original design of the FMT included six layers, the current implemented detector has only three layers installed. This is due to technical difficulties and concerns regarding its Lorentz angle [Konczykowski et al., 2010].

Each of the three FMT layers has 1024 readout strips, which follow a peculiar distribution, as can be seen in the image to the right of Figure 25. In addition, each layer's orientation differs by 60° to provide an accurate measurement in the xy-plane, as is shown in the image at the centre of the same Figure [Acker et al., 2020b].

The FMT is composed of six circular, flat detectors that cover angles ranging from 6 to 29 degrees. In its conception, it was anticipated that the FMT would improve the vertex resolution by a factor of 3 to 10 compared to the DC [Aune et al., 2009]. However, the currently implemented FMT deviates from the original design, with only three layers installed. This modification was necessitated by technical difficulties and concerns related to its Lorentz angle [Konczykowski et al., 2010].

Each of the three FMT layers is equipped with 1024 readout strips that exhibit a distinctive distribution, as depicted in the image on the right-hand side of Figure 25. Additionally, the orientation of each layer differs by 60 degrees to ensure accurate measurements in the xy-plane, as illustrated in the image at the center of the same Figure [Acker et al., 2020b].

3.1.1 FMT Track Reconstruction

Once a signal is detected on a readout strip and the data is stored, the information is extracted from it during offline reconstruction. The reconstruction process of the FMT closely resembles that of the DC.

To begin with, when a signal is detected in a strip, it undergoes digitisation, processing, and is transformed into an FMT Hit. A set of FMT hits is then processed using a Cluster Finding algorithm, where a Cluster is defined as a group of hits that likely originate from the same particle track. Groups of clusters from different layers are subjected to a DC Track Matching algorithm, where they are matched to DC tracks generated by the DC Reconstruction process. Subsequently, a Track Refitting algorithm is employed for each DC track using the data from the clusters, resulting

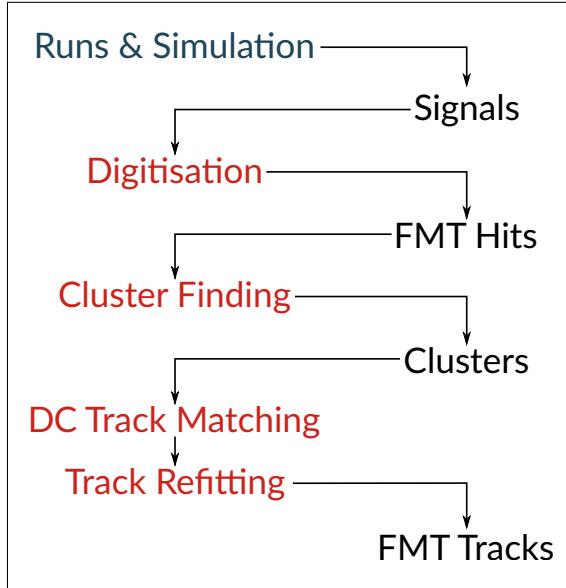


Figure 24: Forward Micromegas Tracker (FMT) reconstruction summary. Data taking is coloured blue, data in black, and processes in red.

Source: Own elaboration, using Inkscape.

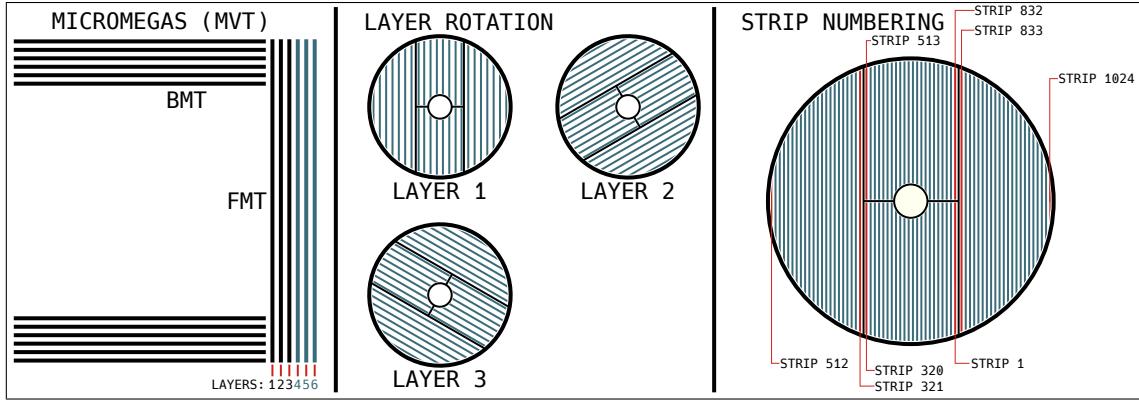


Figure 25: Forward Micromegas Tracker (FMT) detector geometry. The first picture shows the distribution of the Barrel Micromegas Tracker (BMT) and FMT layers, the second the different angle of each FMT layer, and the third the readout strip distribution of each FMT layer.

Source: Own elaboration, using Inkscape.

in updated tracks known as FMT tracks. An overview of this entire process is provided in Figure 24.

3.2 FMT Alignment

In an ideal scenario, the target and each detector would be installed precisely in their required positions. However, in the real world, there are inevitable misalignments in their placements. These misalignments must be accounted for and incorporated into the reconstruction process to ensure meaningful results.

Within the CLAS collaboration, the Calibration and Commissioning group is responsible for the alignment and calibration of each detector. The shifts and rotations necessary for alignment are included in the Calibration and Conditions Database (CCDB), which is then utilised during reconstruction.

The alignment work aimed to achieve three primary goals. Firstly, to provide FMT alignment tables to Run Group F (RG-F) for use in reconstruction. Secondly, to assess whether the resolution improvement obtained from the FMT justifies the additional material introduced into the CLAS12 detector. Lastly, to offer detailed information about these improvements, enabling Run Group E (RG-E) and other run groups to make informed decisions on whether to include the detector in their runs.

Alignment shifts can be performed in any of the three global axes: z , which is aligned with the beamline; x , which runs parallel to the ground; and y , which points upwards from the ground. Additionally, alignment rotations can be carried out around these axes. Specifically, for the purposes of this work, rotations around the z axis are referred to as ϕ rotation (roll), while rotations around the x and y axes are termed pitch and yaw, respectively.

To quantify misalignment, the DOCA between a reconstructed DC track and an FMT cluster is defined as a Residual. Due to the geometry of each layer (as depicted in Figure 25), only the residuals in the local y axis of a layer (perpendicular to the strips) can be measured. This implies that global z and ϕ alignment can be

performed independently for each layer. However, global x , global y , pitch, and yaw alignment must be carried out simultaneously for the entire detector.

To minimise residuals, they are plotted for a specific shift or rotation in the relevant axes. An example of such a plot is illustrated in Figure 26. Since the residuals are expected to follow a Gaussian distribution, a Gaussian fit is applied to them.

For z and ϕ alignment, the goodness of fit is heuristically evaluated by comparing the standard deviation (σ) of the Gaussian fits and selecting the shift with the smallest σ . For x , y , pitch, and yaw alignment, the goodness of fit is heuristically evaluated by choosing the fit with the mean closest to zero. It is important to consider a reasonable error margin when selecting the minima.

Examples of the distributions of goodness of fit for z and xy alignment can be observed in Figure 27.

3.2.1 Cuts

To mitigate background noise, cuts are applied to the DC tracks and FMT clusters. This process enhances data quality, ensuring more meaningful alignment results.

For DC tracks, the following cuts are implemented:

- $\text{track}.z < \text{layer}.z$: This cut removes tracks with a vertex z position further downstream than the FMT layer prior to swimming. Such occurrences result from reconstruction errors where the particle origin is outside the target.
- $|\text{track}.z - \text{layer}.z| < 0.05\text{cm}$: This eliminates tracks that are too far from the FMT layer after swimming, caused by bugs in the swimmer process that will be discussed in the subsequent section.
- $5\text{cm} < \sqrt{x^2 + y^2} < 25\text{cm}$: This removes tracks outside the active region of the layer.

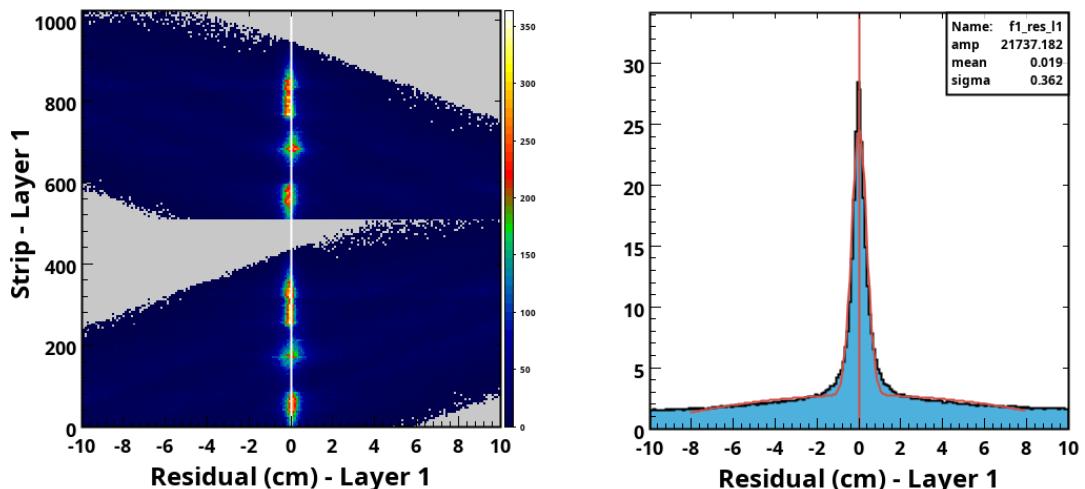


Figure 26: Example Forward Micromegas Tracker (FMT) residuals plot.

Source: Own elaboration, using CLAS12 alignment software.

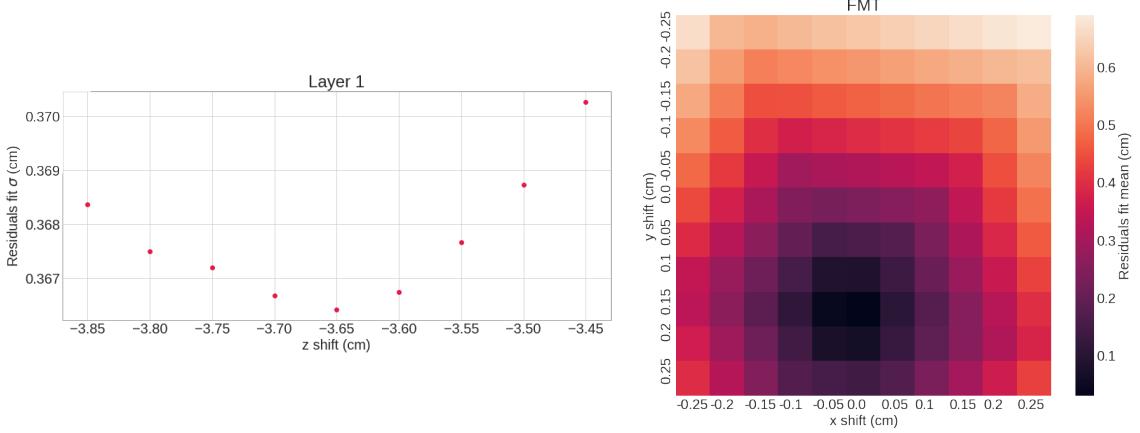


Figure 27: Examples of residuals goodness of fit plots.

Source: Own elaboration, using CLAS12 alignment software.

- $\theta < 66.42^\circ$: This excludes tracks with excessively high θ angles. When this occurs, a single particle affects multiple strips, rendering the detector's data less reliable for alignment purposes.

The implemented cuts applied to the FMT clusters are:

- $50\text{ns} < \text{Tmin} < 500\text{ns}$: This cut removes clusters with illogical values for Tmin , which represents the time of the first hit in the cluster.
- size $> 1 \& E > 100$: This eliminates small clusters with high energy, as they are generally considered to be of poor quality.
- size < 5 : This discards large clusters, as they are deemed to be less useful for analysis purposes.

3.2.2 Residuals Improvements

To validate the proposed alignment algorithm, it was applied to the data from RG-F, specifically Run 11983. The improvement in residuals is readily apparent when comparing the before and after alignment results, as depicted in Figure 28. It is important to note the difference in scale between the top and bottom plots, which further highlights the significant improvement achieved through the alignment process.

As depicted in Figure 28, the z and ϕ alignment significantly reduces the background, resulting in a higher concentration of residuals around the mean of the distribution. Moreover, the x and y alignment effectively aligns the mean of the distribution closer to zero, improving the overall alignment. However, for the pitch and yaw alignment, meaningful results could not be obtained. This can be attributed to the limited data available from the three layers, combined with the small rotations around the x and y axes, making it challenging to achieve precise alignment.

To determine the mean and standard deviation (σ) of the distribution, a Gaussian

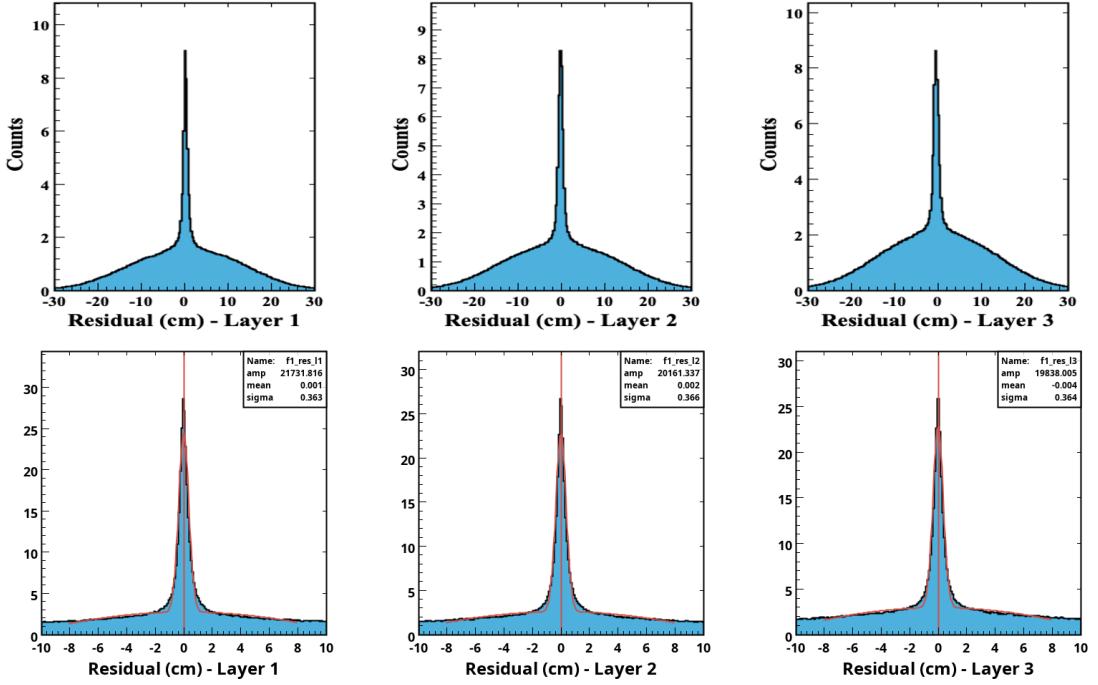


Figure 28: Residuals distribution before (upper image) and after (lower image) alignment.

Source: Own elaboration, using CLAS12 alignment software.

fit was applied. The parameters of the Gaussian fit are

$$\left(\text{amp} \cdot \text{gaus}(\mu, \sigma) \right) + \left(p_0 + p_1 \cdot x + p_2 \cdot x^2 \right)$$

gaussian + background

The results obtained are included in the CCDB at
clasweb.jlab.org/cgi-bin/ccdb/versions?table=/geometry/fmt/alignment

Alignment was successfully performed for the data from Run Group M, demonstrating that the alignment procedure is not specific to a particular run. This indicates the general applicability and effectiveness of the alignment procedure across different runs in the CLAS12 detector.

The impact of this alignment procedure on the resolution of the entire CLAS12 detector will be further investigated and discussed in the concluding subsection of the section.

The procedure described in this section is documented and shared publicly. It can be seen at

github.com/JeffersonLab/clas12alignment/tree/master/fmt

3.3 FMT Reconstruction Work

As the work on FMT alignment progressed, certain modifications were required in FMT reconstruction. These modifications primarily involved incorporating the alignment shifts determined during the alignment process into the reconstruction. Additionally, some fixes were made to address issues that were identified during the alignment work. These changes ensure that the reconstruction process takes into account the alignment information and addresses any related issues that were encountered.

First, the loading of shifts from the CCDB was included in the standard geometry class of the FMT reconstruction package. Then, standard methods to include the shifts in any frame of reference change were implemented. Finally, the code was studied in detail, and the shifts were added in all instances where they were required since the package originally didn't consider their application.

“Crossmaking” is the process of matching clusters in different layers to obtain an accurate 3D estimate of the position of a track [Ziegler et al., 2020]. This process was initially included in FMT reconstruction to facilitate the reconstruction for the six FMT layers. However, as mentioned before, only three FMT layers were installed for the RG-F run, and future runs may also use three layers due to concerns with the Lorentz angle when using six layers. To simplify the reconstruction process and make better use of the available number of layers, crossmaking was removed from the reconstruction.

Outside of FMT reconstruction, some minor changes were also required in the DC reconstruction package since some of its components depend on the FMT layers' positions. Additionally, the swim package diagnostic was updated as it failed to properly reconstruct the positions of tracks near the FMT layers.

A detailed list of the updates applied can be found in the following pull request to the `clas12-offline-software` repository

github.com/JeffersonLab/clas12-offline-software/pull/726.

3.4 Validation and Results

Just like the residuals validation, the results presented in this document are based on the application of this work to RG-F data. It is important to note that the RG-F target is approximately 55 centimetres long, which is much larger than the average CLAS12 target [Hattawy et al., 2019]. Specifically, the runs used for testing and validation are presented in Table 4, and the run used to obtain the data displayed in this section was 011983.

Run Number	Energy (MeV)	Current (nA)	Configuration	Target
011983	10389.4	50	Inbending	D2
012016	10389.4	250	Inbending	D2
012439	2186.4	15	Inbending	H2
012461	10196.6	20	Inbending	D2

Table 4: RG-F runs used for validation.

3.4.1 Cuts

Some additional cuts are applied to the tracks to obtain the plots presented in this section. These cuts are used to remove very poor tracks that would not be suitable for analysis regardless. The applied cuts are as follows:

- `abs(chi2pid) < 5`: This cut removes tracks that do not provide sufficient certainty regarding the particle's PID.
- `vz < fmtZ`: This cut removes tracks located further downstream than the FMT.
- `chi2/ndf < 15`: This cut excludes tracks with excessively high uncertainty.

3.4.2 Geometry Effect

To evaluate the enhancement in vertex resolution, we will compare the vertex positions of tracks that underwent only DC reconstruction with those that underwent both DC and FMT reconstruction. For convenience, we will refer to the former as DC tracks and the latter as FMT tracks. Considering that the z axis is aligned with the beamline, Figure 29 illustrates the z positions of the vertex for DC tracks versus FMT tracks.

To comprehend the plot in Figure 29, it is valuable to examine the RG-F target. The target consists of a large gas-filled chamber with a varying composition across different runs. The distance between the chamber windows measures 553.32 millimetres. Furthermore, it was observed that the upstream window of the target is positioned approximately 24 millimetres away from the beam window. All windows are constructed from aluminium and have a thickness of 15 micrometers. A detailed depiction of the target can be found in Addendum 1.

Based on Figure 29, it is evident that the FMT detector solely detects the upstream windows, completely overlooking the downstream one. This issue stems from a geometric constraint: the downstream window falls outside the active detection area of the FMT. This effect is clearly illustrated in Figure 30, where the θ angle is plotted against the vertex z coordinate. The two red lines in the plots represent the FMT's active area, and it is apparent that the downstream window lies outside this region, thereby explaining its absence.

To compensate for this geometric effect, we introduce an additional cut based on the plotted curves. The curves can be described by the following equations

$$c_1(z) = 57.29 \cdot \arctan\left(\frac{r_{\text{inner}}}{z_0 - z}\right), \quad c_2(z) = 57.29 \cdot \arctan\left(\frac{r_{\text{outer}}}{z_0 - z}\right). \quad (4)$$

Here, r_{inner} represents the radius of the hole at the center of the FMT, r_{outer} denotes the radius of the outer circumference of the FMT, and z_0 corresponds to the z position of

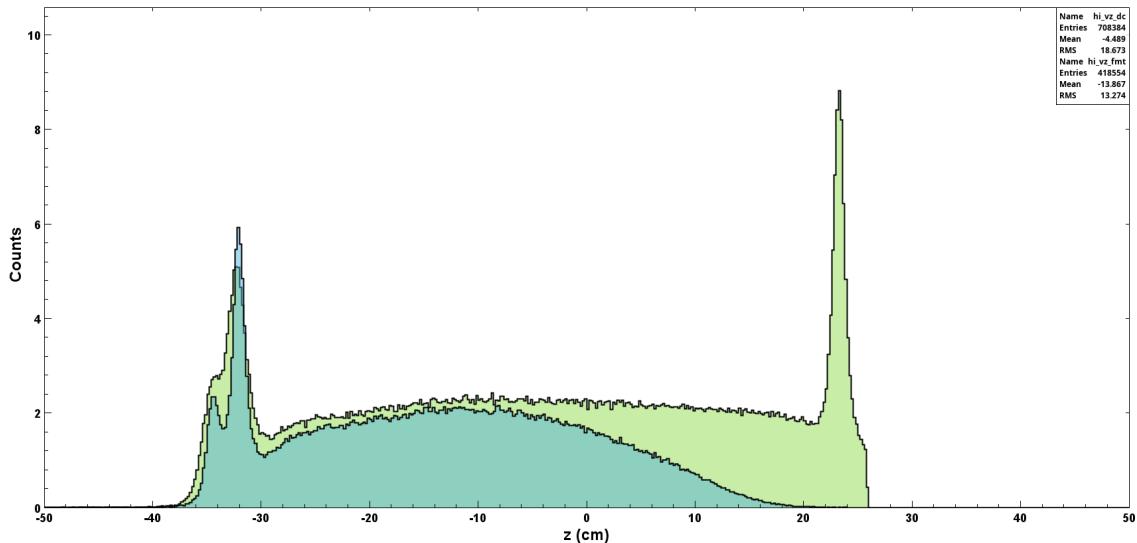


Figure 29: DC vs. FMT vertex z for electrons without any geometric correction. DC tracks are shown in green while FMT tracks are shown in blue. Note that the dark cyan colour comes from the overlap.

Source: Own elaboration, using the `fmtVertex.groovy` script in CLAS12 alignment software.

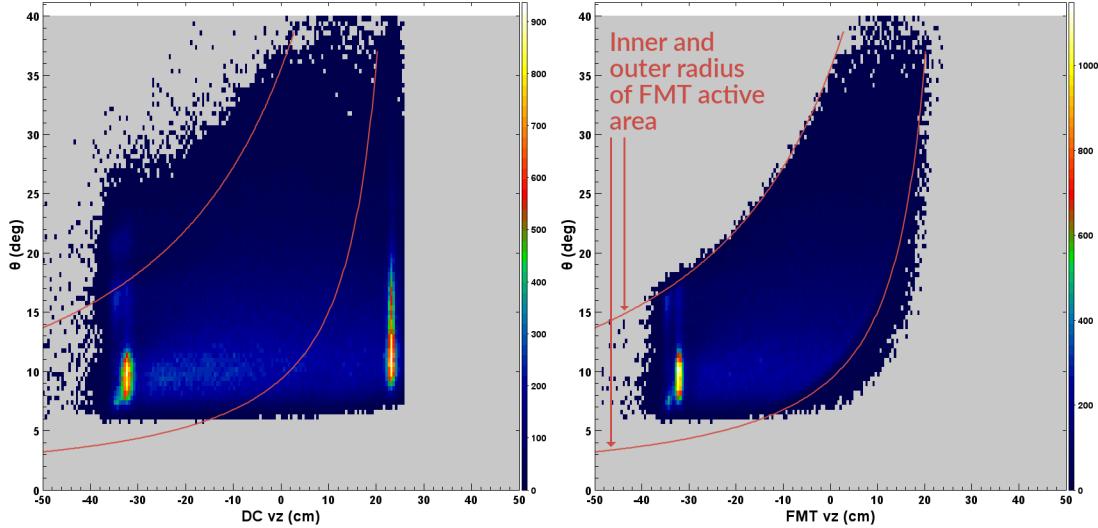


Figure 30: z vs. θ for DC and FMT for electrons without any geometry correction. FMT's active area are shown in red lines.

Source: Own elaboration, using the `fmtVertex.groovy` script in CLAS12 alignment software.

the first FMT layer plus the drift distance. All these parameters are obtained from the CCDB.

3.4.3 Vertex Resolution Improvement

Furthermore, an additional cut has been introduced. As of the FMT alignment work, beam alignment for RG-F data had not been carried out, resulting in a decrease in vertex resolution. To mitigate the impact of this alignment issue on reconstruction accuracy, we

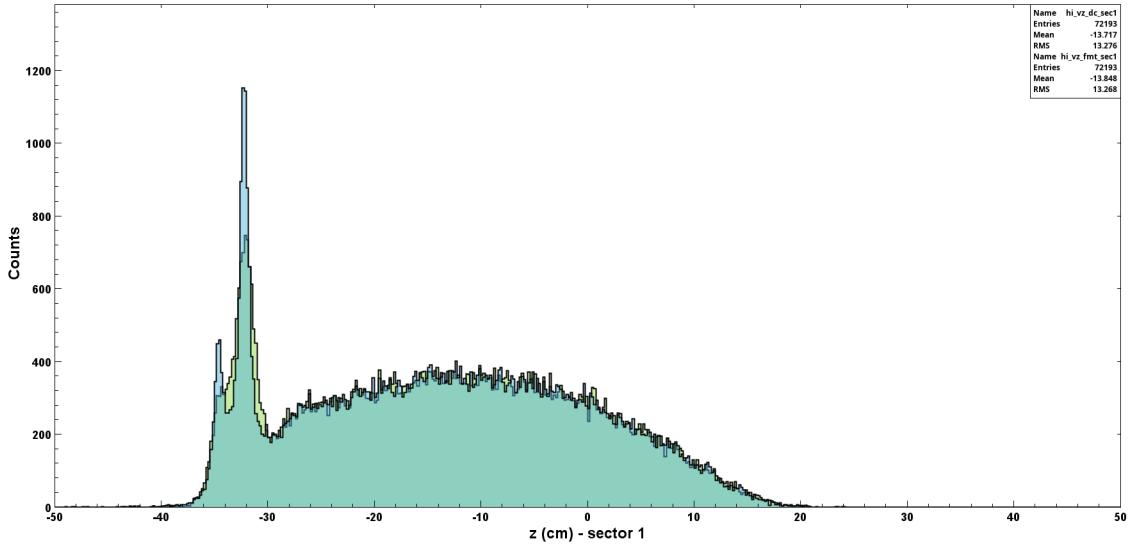


Figure 31: DC vs. FMT vertex z for electrons with a geometric correction. DC tracks are shown in green while FMT tracks are shown in blue. Data from only one CLAS12 sector was used to obtain this plot.

Source: Own elaboration, using the `fmtVertex.groovy` script in CLAS12 alignment software.

have implemented a cut to utilise only one sector of the detector.

The resolution plot, comparing DC and FMT tracks after applying all the previously mentioned cuts, is depicted in Figure 31. To evaluate the resolution for both DC and FMT tracks, we utilise a fit consisting of two Gaussian curves combined with a quadratic curve to account for the background. The fit is defined as follows

$$\text{amp}_1 \cdot \text{gaus}(z, z_{\max}, \sigma) + \text{amp}_2 \cdot \text{gaus}(z, z_{\max} - 2.4, \sigma) + p_1 + p_2 \cdot z + p_3 \cdot z^2,$$

where

- amp_1 represents the amplitude of the largest peak, and z_{\max} corresponds to its z position,
- amp_2 signifies the amplitude of the leftward peak, which has been measured to be at a position of 2.4 centimetres, and
- the remaining parameters, p_1 , p_2 , and p_3 , are obtained through the fitting process.

For electrons in run 011983 (low luminosity, 50 nA), the analysis yields a DC resolution of $\sigma_{\text{DC}} = 0.875$ cm and an FMT resolution of $\sigma_{\text{FMT}} = 0.387$ cm. This indicates a doubling of the resolution achieved with the inclusion of the FMT detector.

Similarly, for electrons in run 012016 (production luminosity, 250 nA), the analysis shows a DC resolution of $\sigma_{\text{DC}} = 1.009$ cm and an FMT resolution of $\sigma_{\text{FMT}} = 0.596$ cm.

3.4.4 Conclusions

Although the improvement in resolution is not as significant as initially anticipated for the detector, it remains an encouraging result. The enhanced resolution enables more precise measurements of the target position. As a practical implication, it allows for double targets to be positioned closer to each other, thereby benefiting the derived physics from experiments like the RG-E run.

The reason for the smaller-than-predicted improvement in resolution can be attributed to the initial projection, which assumed the presence of six FMT layers. The inclusion of six layers would provide additional positional data along the particle's track, thereby enhancing the accuracy of the vertex position measurement during the fitting process.

Furthermore, due to the limited number of layers and their close proximity in the FMT detector, it exhibits a small lever arm. As a result, the detector's contribution to the vertex

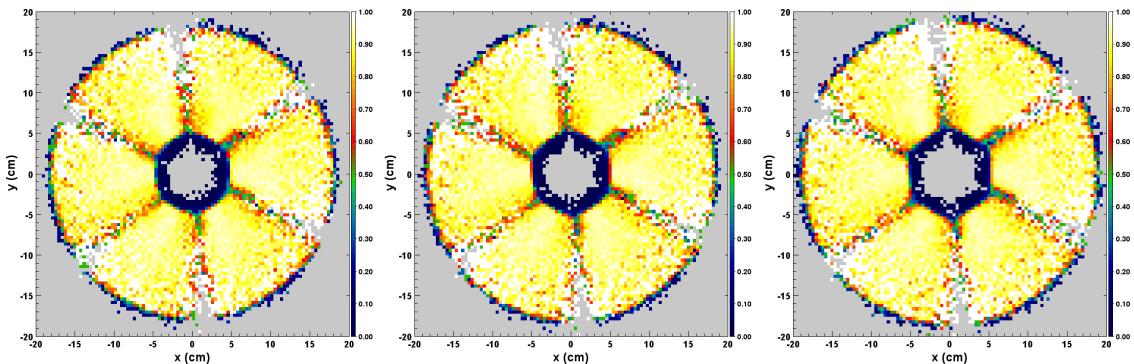


Figure 32: Efficiency of each FMT layer.

Source: Own elaboration, using the `fmtVertex.groovy` script in CLAS12 alignment software.

momentum resolution is not significant, as it does not provide sufficient additional data to accurately determine the track's momentum.

In order to gain a better understanding of the FMT detector, we conducted a brief study on its efficiency. The efficiency is defined as the ratio of the number of FMT tracks to the number of DC tracks, representing how many of the DC tracks were also detected by the FMT.

For the three-layer configuration, the efficiency is approximately 88.96%. Figure 32 illustrates the layer-by-layer efficiency, showing no anomalous geometric effects. The observed gaps in efficiency are solely a result of the CLAS12 detector's geometry, which is divided into six sectors.

4 Data Analysis

The first section describes a C program developed by the author to perform the analysis in this thesis. The second section discusses the DIS kinematics, which define the cuts made to the data. The third section provides detailed information about the approach used to include the sampling fraction in the analysis. Finally, the last section describes the simulations produced and the acceptance study conducted based on them.

4.1 clas12-rge-analysis

To perform the acceptance analysis reported in this thesis, the author developed an extensive C/C++ program that runs on the ROOT library. The purpose of this program is not only to conduct the analysis described in this thesis but also to be used for RG-E analysis in general. The program and its source code are freely available under the GNU LGPLv3 license and can be accessed on GitHub at: github.com/bleaktwig/clas12-rge-analysis. Issues and pull requests are welcomed, as they are crucial for maintaining the program and facilitating collaborative development in the repository.

After compiling the program using `make`, five executables are obtained in the `bin` directory. The following sections provide a description of each executable.

4.1.1 hipo2root

The CLAS12 reconstruction process utilises a custom file format called High-Performance Input Output (HIPO) format, developed by Gagik Gavalian [Chekanov et al., 2021]. The CLAS collaboration provides a set of tools that enable working with HIPO files, allowing users to read, write, and generate plots from these files. However, users at RG-E are generally more familiar with traditional ROOT files. Therefore, a conversion tool was developed.

The tool, called `hipo2root`, filters through a HIPO file's data and creates a ROOT file that includes pre-selected sections of storage known as banks. The selection of banks is based on data that is relevant to RG-E analysis, and users can easily add additional banks by modifying the program's source files. The output of the `hipo2root` executable is a ROOT file that contains the selected banks represented as trees, which are standard ROOT array-like variables.

Its manual entry is:

```
Usage: hipo2root [-hfn:w:] infile
* -h          : show this message and exit.
* -f          : set this to true to process FMT::Tracks bank. If this
                is set and FMT::Tracks bank is not present in the HIPO
                file , the program will crash .
* -n nevents : number of events.
* -w workdir : location where output root files are to be stored .
                Default is root_io .
* infile     : input HIPO file . Expected format is <text>run_no.hipo .
```

Convert a file from hipo to root format. This program only conserves the banks that are useful for RG-E analysis, as specified in the ‘lib/rge_hipo_bank.h’ file. It’s important for the input hipo file to specify the run number at the end of the filename (‘<text>run_no.hipo’), so that ‘hipo2root’ can get the beam energy from the run number.

Since simulation files don't have a run number, we use a convention for specifying the beam energy. For this files, the filename should be '>999XXX.hipo', where 'XXX' is the beam energy used in the simulation in [0.1*GeV].

4.1.2 extract_sf

To obtain the sampling fraction of the detector's calorimeters, certain analyses from CLAS12 are required. This is accomplished using the `extract_sf` executable, which utilises the momenta of particles and the corresponding deposited energy. The detailed methodology and the results obtained from this analysis are discussed in Section 4.3.

The manual entry of the program is:

```
Usage: extract_sf [-hn:w:d:] infile
* -h          : show this message and exit .
* -n nevents  : number of events
* -w workdir  : location where output root files are to be stored .
                  Default is root_io .
* -d datadir  : location where sampling fraction files are stored .
                  Default is data .
* infile      : input ROOT file . Expected file format:
<text>run_no.root .
```

Obtain the EC sampling fraction from an input file. An alternative to using this program is to fill the output file corresponding to the studied run (by default stored in the 'data' directory) with the data obtained from [CCDB] (https://clasweb.jlab.org/cgi-bin/ccdb/versions?table=calibration/eb/electron_sf). The function used to fit the data is

$$[0]*\text{Gaus}(x,[1],[2]) + [3]*x*x + [4]*x + [5]$$

where *[0]* is the amplitude of the Gaussian, *[1]* and *[2]* its mean and sigma, and *[3]*, *[4]*, and *[5]* the *p0*, *p1*, and *p2* used to fit the background.

The output of the program is the 'sf_params_<run_no>.txt', which contains a table with the sampling fractions and their errors. The table is formatted like the one at CCDB, as in

	sf0	sf1	sf2	sf3	sfs1	sfs2	sfs3	sfs4
sector 1	%11.8f							
sector 2	%11.8f							
sector 3	%11.8f							
sector 4	%11.8f							
sector 5	%11.8f							
sector 6	%11.8f							

4.1.3 acc_corr

The `acc_corr` executable is used to count the number of thrown and simulated events from two different ROOT files. Based on the program's input, it separates the data into appropriate 5-dimensional bins, counts the entries for all available particles in the

generated file, and exports the results into a text file. The results and plots presented in Section 5.2 were generated using the data obtained from this program.

The manual entry of the program is:

```
Usage: acc_corr [-hq:n:z:p:f:g:s:d:FD]
* -h          : show this message and exit.
* -q ...      : Q2 bins.
* -n ...      : nu bins.
* -z ...      : z_h bins.
* -p ...      : Pt2 bins.
* -f ...      : phi_PQ bins (in degrees).
* -g genfile  : generated events ROOT file.
* -s simfile  : simulated events ROOT file.
* -d datadir  : location where sampling fraction files are found.
               Default is data.
* -F          : flag to tell program to use FMT data instead of DC data
               from the simulation file.
* -D          : flag to tell program that generated events are in
               degrees instead of radians.
```

Get the 5-dimensional acceptance correction factors for *Q2*, *nu*, *z_h*, *Pt2*, and *phi_PQ*. For each optional argument, an array of doubles is expected. The first double will be the lower limit of the leftmost bin, the final double will be the upper limit of the rightmost bin, and all doubles between them will be the separators between each bin.

The output will be written to the ‘acc_corr.txt’ file, by default in the ‘data’ directory, which is formatted to make it easy to read by the ‘draw_plots’ program:

- * First line contains five integers; the size of each of the five binnings.
- * The next five lines are each of the binning schemes, in order *Q2*, *nu*, *z_h*, *Pt2*, and *phi_PQ*.
- * The following line contains one integer which is the size of the list of PIDs, followed by a line containing each of these PIDs.
- * Finally, a number of lines equal to the number of PIDs follows. Each line contains a list of the bins, ordered as ‘[Q2][nu][z_h][Pt2][phi_PQ]’.

4.1.4 make_ntuples

This executable operates on one or more files generated by hipo2root and produces a ROOT file that contains a set of ntuples relevant to the analysis. Furthermore, based on the specific requirements of this thesis’ analysis, the executable generates two sets of ntuples. Both sets have the same ntuple format, but the former utilises only DC tracking data while the latter incorporates both DC and FMT tracking data.

For each event, the program executes the following algorithm:

1. First, the program identifies the TOF of the trigger electron. The hits of the trigger electron in the FD scintillators and FD calorimeters are listed in order of priority based on the precision of each detector’s TOF measurement. The detectors are prioritised as follows: FTOF panel-1b, FTOF panel-1a, FTOF panel-2, PCAL, ECIN, and ECOU, as described in Section 2.2.1. Next, the program iterates over the list of hits, extracting the TOF value from the earliest hit in the most precise available layer.

2. Next, for each available track, two particle objects are instantiated. These objects contain the relevant data for the particle, including its vertex position, vertex momentum, charge, beta, and the CLAS12 sector through which it passed. The first object corresponds to the tracking data obtained from the DC, while the second object incorporates both the DC and the FMT tracking data. The assignment of the particle's PID will be carried out later in the process.
 3. The program computes and stores the particle's deposited energy in the calorimeters. This involves summing up the energy deposited by all the hits associated with the particle's track for each calorimeter.
 4. The program counts the number of produced photoelectrons in the HTCC and LTCC for the particle. Furthermore, the particle's TOF is computed using the same procedure as the one employed for the trigger electron's TOF, considering the hits in the detectors prioritised by their precision.
 5. The program assigns the Particle Identification (PID) to the particle. The process is very similar to the PID assignment in reconstruction, as described in Section 2.2.3. However, the assigned PID is not directly used in order to allow users to modify parameters and define new criteria for the PID assignment.
- Although this process typically yields the same results as reconstruction, there is a slight error in the PID assignment. This error is presented in Table 5. As shown in the table, some kaons and protons are misidentified as pions, but the degree of misidentification is not significant. Apart from that, all identifications are accurate.
6. Finally, two `ntuples` objects are created: one for the particle generated from the DC tracking data and another for the particle generated from both the DC and FMT data. These `ntuple` objects are then saved in an output file, which can be used directly for analysis or processed by the `draw_plots` program discussed in the next section.

Particle Identification Truth						
	e	π	K	p	n	γ
e	1.00					
π		1.00	0.09	0.02		
K			0.91			
p				0.98		
n					1.00	
γ						1.00

Table 5: Particle identification matrix for the FD. The rows show the PID assigned by reconstruction while the columns the one assigned by the `make_ntuples` program. The diagonal elements are correctly identified, while the off-diagonal elements are misidentified.

The manual entry of the program is:

```
Usage: make_ntuples [-hDf:cn:w:d:] infile
* -h : show this message and exit.
* -D : activate debug mode.
* -f fmtlyrs : define how many FMT layers should the track have hit.
              Options are 0 (tracked only by DC), 2, and 3. If set to
              something other than 0 and there is no FMT::Tracks bank
              in the input file , the program will crash. Default is
              0.
* -c : apply FMT geometry cut on data.
* -n nevents : number of events.
* -w workdir : location where output root files are to be stored.
              Default is root_io.
* -d datadir : location where sampling fraction files are. Default is
              data.
* infile : input ROOT file. Expected file format:
            <text>run_no.root .
```

Generate ntuples relevant to SIDIS analysis based on the reconstructed variables from CLAS12 data. The output of the program is the ‘ntuples_<run_no>.root’ file, which contains all relevant ntuples for RG-E analysis. This file can be studied directly in root or through the ‘draw_plots’ program.

4.1.5 draw_plots

The purpose of including the `draw_plots` executable is to save the user from rewriting similar code repeatedly to obtain plots. Using the `draw_plots` program is relatively straightforward: after running the program, the user is prompted with a series of questions to define various attributes related to the plots. This includes specifying any cuts and corrections to apply, setting up the binning, and selecting the variables to be plotted. Unless stated otherwise, the plots presented in Section 5 were generated using this executable.

The executable’s manual entry is:

```
Usage: draw_plots [-hp:cn:o:a:w:] infile
* -h : show this message and exit.
* -p pid : skip particle selection and draw plots for pid.
* -c : apply all cuts (general, geometry, and DIS) instead of
      asking which ones to apply while running.
* -n nentries : number of entries to process.
* -o outfile : output file name. Default is plots_<run_no>.root .
* -a accfile : apply acceptance correction using acc_filename.
* -A : get acceptance correction plots without applying
      acceptance correction. Requires -a to be set.
* -w workdir : location where output root files are to be stored.
              Default is root_io.
* infile : input file produced by make_ntuples .
```

Draw plots from a ROOT file built from ‘make_ntuples’. File should be named ‘<text>run_no.root’. This tool is built for those who don’t enjoy using root too much, and should be able to get most basic plots needed in SIDIS analysis.

4.2 Cuts

In order to enhance the relevance of particles for the analysis presented in this work, three types of cuts are applied:

- General cuts: These cuts are designed to exclude poorly reconstructed particles.
- Geometry cuts: These cuts define the region from which the reconstruction data is considered valid and useful for the analysis.
- DIS cuts: These cuts narrow down the analysis region to focus specifically on DIS.

By applying these cuts, the analysis can be focused on particles that meet the criteria for good reconstruction, fall within the desired geometry range, and are relevant to the DIS process.

4.2.1 General Cuts

For this analysis, only two cuts are considered as “general”. The first cut involves filtering out particles with PID values of 0 or 45. In CLAS12 reconstruction, these specific PID values are assigned to particles whose identification could not be successfully determined.

The second cut aims to exclude particles with imprecise tracking and is defined as follows

$$\frac{\chi^2}{\text{NDF}} < 15,$$

where χ^2 represents the final result from the χ^2 -test used in the Kalman filter fit of the tracking algorithm, as described in Section 2.2.3. The term NDF denotes the number of degrees of freedom associated with this same fit.

By applying these two cuts, particles with undetermined or uncertain identification (PID values of 0 or 45) and those with poor tracking precision (exceeding the specified χ^2/NDF threshold) are excluded from the analysis.

4.2.2 Geometry Cuts

Three geometry cuts are derived to constrain the reconstructed particle’s vertex position. The first two cuts ensure that the vertex is located along the beamline, while the third cut restricts it to the acceptance region of the FMT.

The first cut guarantees that the vertex is close to the beamline and is defined as

$$\sqrt{v_x^2 + v_y^2} < 4 \text{ cm},$$

where v_x and v_y represent the x and y coordinates of the vertex position, respectively.

The second cut ensures that the vertex originates from the target and is given by

$$-40 \text{ cm} < v_z < z_0 \text{ cm},$$

where v_z corresponds to the z coordinate of the vertex position, and z_0 represents the z position of the first FMT layer. For the RG-F Spring 2020 run, $z_0 = 26.12 \text{ cm}$.

The third and final cut ensures that the vertex falls within the FMT acceptance region, as defined in Section 3.4.2. It removes all particles whose v_z and θ values lie outside the region bounded by the two lines defined by Equation (4).

Fiducial cuts were not applied in this analysis. The reasoning behind this decision is explained in Appendix B, along with a brief explanation of how the cuts would be applied.

4.2.3 DIS Cuts

Three DIS cuts are applied to the scattered electron to restrict the phase space to that of DIS. If the trigger electron fails to pass these cuts, all particles in the event are disregarded.

The first cut is based on the invariant mass of the virtual photon, Q^2 , and is defined as

$$Q^2 > 1 \text{ GeV}^2,$$

where Q^2 is given by

$$Q^2 = 4E_b E' \sin^2(\theta_C/2). \quad (5)$$

Here, E_b represents the beam energy, E' is the energy of the scattered electron, and θ_C is the polar angle of the scattered electron. This cut ensures that the process falls within the DIS domain, as explained in Sections 1.1 and 1.1.2.

The second cut is imposed on the squared mass of the hadronic final state, W^2 , given by

$$W^2 > 4 \text{ GeV}^2.$$

Here, W^2 is defined as

$$W^2 = M^2 + 2M\nu - Q^2,$$

where M represents the mass of the nucleon, and ν is the energy transferred by the lepton probe. The details of the virtual photon energy ν and its effects on this analysis are explained in Section 1, and it is given by

$$\nu = E_b - E'. \quad (6)$$

This cut is applied to exclude nucleon resonances.

Lastly, an additional cut is imposed on the Bjorken-Y (Y_b) of the scattered electron, given by

$$Y_b < 0.85.$$

The Bjorken-Y ranges from 0 to 1 and is defined as

$$Y_b = \frac{\nu}{E_b}.$$

This cut effectively mitigates the influence of extreme radiative effects, which occur when a substantial portion of the incident electron's energy is transferred to the scattered electron.

The impact of these cuts on the Q^2 and ν of the scattered electron can be observed in plot 33.

4.3 Sampling Fraction

The energy deposited by electrons in the active area of the calorimeters is a fraction of their total energy, E_{tot} . This value is proportional to their momentum, P , for energies above a few hundred MeV. Heavier particles, due to their reduced penetration capabilities, tend to lose an amount of energy independent of their momentum. The electron sampling fraction measures the amount of energy lost depending on the momentum of a particle. This allows for both the measurement of the electron's energy and the differentiation of electrons from other particles [Wigmans, 2000].

To obtain the sampling fraction, the hits of each calorimeter by itself (PCAL, ECIN, and ECOU) are separated into arrays, with an additional array containing the union of the other three. Then, these arrays of hits are separated into 20 momentum bins. Each bin has a size of 0.4 GeV, starting at 1.0 GeV and ending at 9.0 GeV.

1-dimensional histograms are then created from the data in these arrays, measuring the deposited energy divided by the vertex momentum (E/p). A Gaussian fit plus a quadratic background is then applied, following the function described as

$$f(x) = p_0 g(x, \mu, \sigma) + p_1 x^2 + p_2 x + p_3, \quad \text{where } g(x, \mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left(-\frac{1}{2} \frac{(x - \mu)^2}{\sigma^2} \right),$$

where μ and σ represent the mean and standard deviation of the distribution, respectively. The fit is limited to the range between 0.15 and 0.30 for the expected E/p values for electrons based on theory.

Examples of these plots are shown in Figure 34. From the figure, it can be observed that there are not enough electrons in the extreme momentum ranges, such as from 1 to 1.4 GeV or from 8.6 to 9 GeV. Consequently, the sampling fraction fit, described in Equation (7), only considers data within the range of 1.4 to 8.6 GeV.

The mean of each of these fits is then extracted to serve as data points for a sampling fraction fit. A polynomial fit is employed since it effectively captures the shape of these

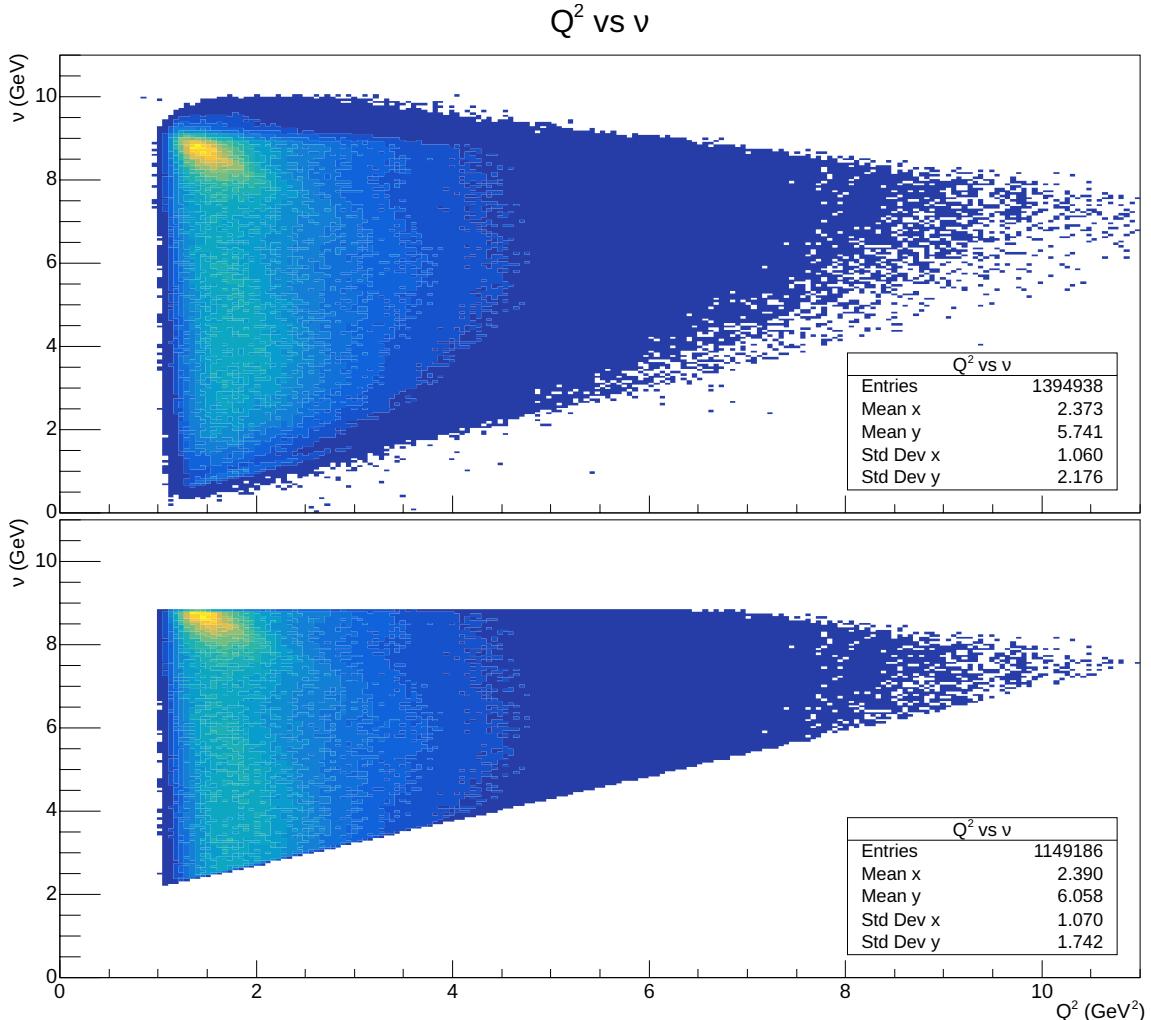


Figure 33: $e^- Q^2$ vs. ν before and after applying the $Q^2 > 1 \text{ GeV}^2$, $W^2 > 4 \text{ GeV}^2$, and $Y_b < 0.85$ cuts, run 12016.

Source: Own elaboration, using the clas12-rge-analysis software.

points and aligns with the reconstruction software. The fit is described as

$$f(x) = p_0 \cdot \left(p_1 + \frac{p_2}{x} + \frac{p_3}{x^2} \right). \quad (7)$$

The E/p distribution vs. p is depicted in Figure 35, together with this fit.

Finally, the parameters of the fit are saved in plain text files, following the convention used in the CCDB. These parameters can be utilised later for particle identification purposes, specifically for electrons and photons. Furthermore, they are employed to determine the energy of electrons and photons since, as mentioned previously, not all of their energy is deposited in the calorimeters.

4.4 Acceptance Correction

When discussing radiation detection, it is customary to distinguish between two types of efficiency: absolute efficiency and intrinsic detection efficiency. The former is defined as the fraction of events emitted by the source that are actually detected by the detector, expressed as

$$\xi_{\text{tot}} = \frac{\text{events registered}}{\text{events emitted by source}}.$$

This efficiency is influenced by the detector's geometry and the probability of an interaction occurring within the detector. The total efficiency is also referred to as the detector acceptance.

The total efficiency can be further decomposed into two components: the intrinsic efficiency, ξ_{int} , and the geometric efficiency, ξ_{geom} . The total efficiency is then given by

$$\xi_{\text{tot}} = \xi_{\text{int}} \cdot \xi_{\text{geom}}.$$

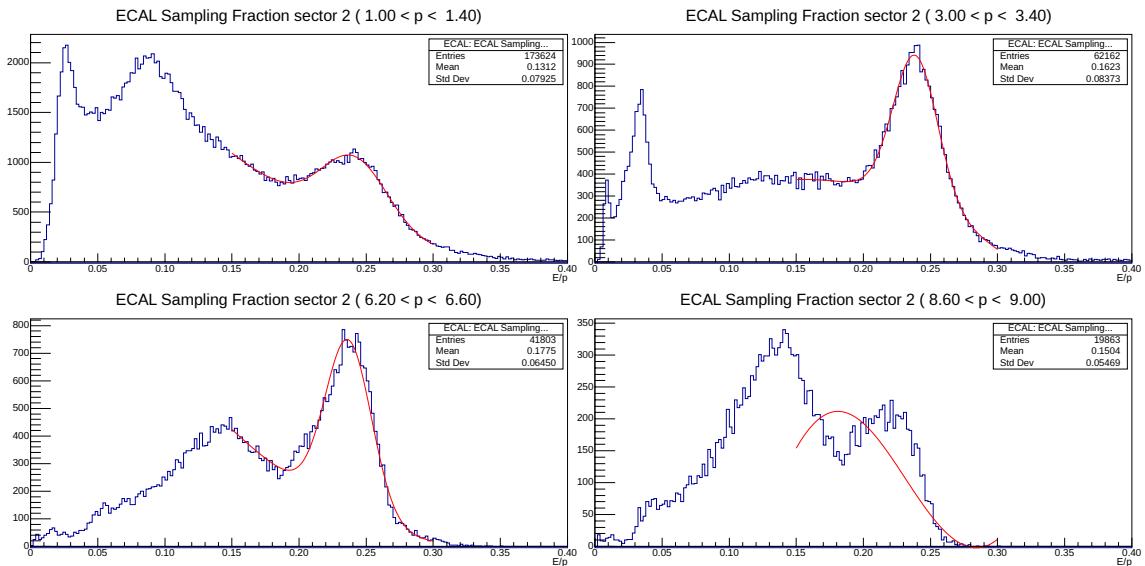


Figure 34: Four E/p plots describing the sum of the deposited energy per particle on all calorimeters (PCAL, ECIN, and ECOU). The particles' momentum is obtained from tracking and the event builder. As can be seen in the northwest and the southeast plots, the corner bins – 1.00 to 1.40 and 8.60 to 9.00 GeV respectively – are not very reliable.

Source: Own elaboration, using the clas12-rge-analysis software.

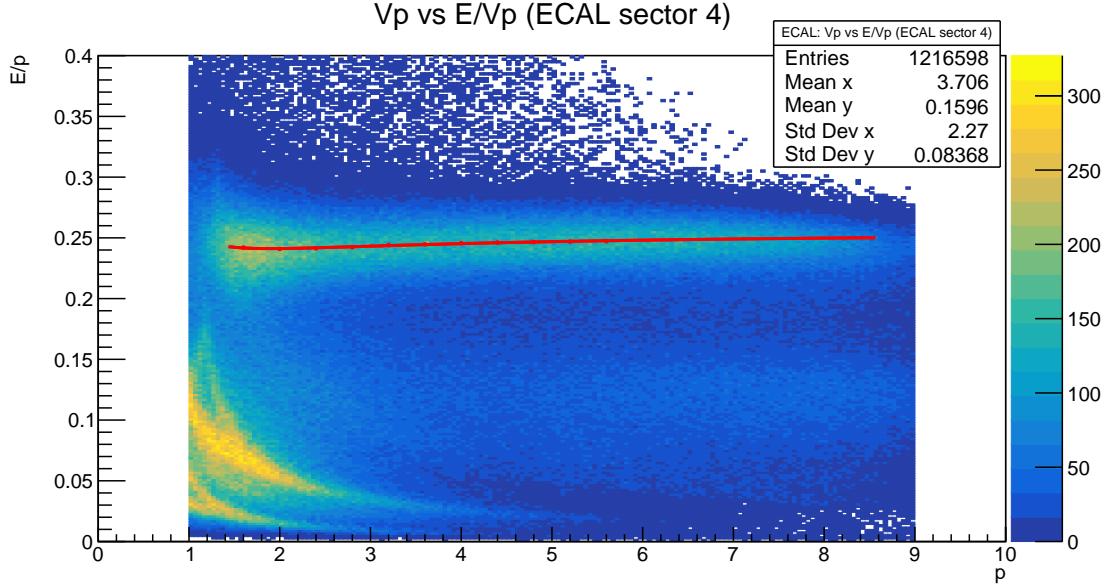


Figure 35: A 2D plot showing momentum p vs. deposited energy divided by momentum E/p . The particle's deposited energy on all calorimeters is measured. Its momentum is obtained from tracking and the event builder. The fit follows the deposited energy of electrons to find their sampling fraction.

Source: Own elaboration, using the clas12-rge-analysis software.

The intrinsic efficiency represents the fraction of events that actually reach and are detected by the detector

$$\xi_{\text{int}} = \frac{\text{events registered}}{\text{events impinging on detector}}.$$

This probability is dependent on the interaction cross-sections of the incident radiation with the detector medium. The intrinsic efficiency thus varies with the type of radiation, its energy, and the detector material [Leo, 1987].

Acceptance correction involves compensating for the total efficiency of the detector. To estimate this detector efficiency, a comparison is made between the total number of generated events, denoted as N_{thrown} , and the number of accepted events in a simulation of the detector, denoted as N_{simul} . This allows us to calculate an estimation of the detector efficiency, represented by $\tilde{\xi}_{\text{tot}}$, using

$$\tilde{\xi}_{\text{tot}} = \frac{N_{\text{simul}}}{N_{\text{thrown}}}.$$

Naturally, the value of $\tilde{\xi}_{\text{tot}}$ is influenced by the accuracy and reliability of the event generator and simulation programs employed in the study.

The acceptance of the detector exhibits variations across the phase space of the kinematic variables. Therefore, in order to achieve accurate acceptance correction, the ratio $\tilde{\xi}_{\text{tot}}$ needs to be divided into bins in a five-dimensional space. These bins correspond to the five variables under investigation: Q^2 , ν , z_h , p_T^2 , and ϕ_{PQ} . To simplify the analysis process and facilitate interpretation of the results, it is advantageous to have bins of the same size for each variable.

- For Q^2 , given by Equation (5), the lower edge of the binning is defined as 1 GeV^2 , considering that a DIS cut of $Q^2 > 1$ is applied, as described in Section 4.2.3. The upper edge is defined as 11 GeV^2 , and the bin size is set to 1 GeV^2 to ensure sufficient statistics per bin. This results in 10 bins for Q^2 .
- As shown in Figure 33, the minimum value of ν (described by Equation (6)) is 2 GeV due to the $W^2 > 4$ cut, and its maximum value is 9 GeV due to the $Y_b < 0.85$ cut. The bin size is set to 1 GeV , resulting in a total of 8 bins for ν .
- The fraction z_h of the virtual photon energy carried by the measured hadron is described in Section 1.3.2 and given by Equation (3). For this DIS study, we are interested in the region of this fraction ranging from 0 to 1. The bin size is defined as 0.1, resulting in 10 bins for z_h .
- p_T^2 represents the hadron's transverse momentum measured with respect to the virtual photon direction. For this experiment, very few hadrons are observed with $p_T^2 > 2 \text{ GeV}^2$, thus the binning range is defined from 0 to 2 GeV^2 . The bin size is set to 0.2 GeV^2 , resulting in 10 bins for p_T^2 .
- ϕ_{PQ} is the angle between the leptonic plane – the plane where the paths of the initial and scattered electrons lie – and the hadronic plane, which contains the virtual photon and the measured hadron. It is important to be cautious about the binning scheme for this variable, as some geometric effects can arise from the 6-sector geometry of the CLAS12 Forward Detector (FD), as described in Section 2.2.1. Hence, 13 bins are defined from -180° to 180° , resulting in a bin size of approximately 27.7° .

To calculate the acceptance, 10 million events were initially generated in deep inelastic kinematics using LEPTO, a Monte Carlo generator specifically designed for deep inelastic lepton-nucleon scattering [Ingelman et al., 1997]. Subsequently, these events were simulated under the experimental conditions of the RG-F experiment in CLAS12 using `gemc`, which is the standard tool for CLAS12 event simulation [Ungaro et al., 2020b]. The simulation took into account a torus field polarity of -1 and a solenoid field polarity of -0.745033 .

Finally, the simulated events were reconstructed using `coatjava`, which is the standard tool for CLAS12 event offline reconstruction [Ziegler et al., 2020]. Further details regarding the offline reconstruction process can be found in Section 2.2.3. The outcomes of the acceptance correction procedure are discussed in Section 5.2.

5 Results and Conclusions

The FMT efficiency is discussed in the first section, including a brief study of it. Then, the second section delves into the acceptance correction results, based on the methodology described in Section 4.4. Following that, the study results are discussed in detail, and the conclusions of the study follow shortly thereafter.

All Figures in this section that show results for data are from RG-F run 12016. For this run, a gas H₂ target was used. The beam energy was set to 10.4 GeV, and the luminosity to 250 nA. The solenoid field was set to an inbending polarity (-0.745033), and the torus to its regular polarity of -1 . The run has a total of 10,046,225 events.

5.1 FMT Efficiency

Compared to the alignment work described in Section 3, a low FMT is observed in this analysis. This is evident in Figure 36. Upon inspection, three causes can be attributed to this: the application of incorrect alignment constants, a geometry effect, and a general FMT offline reconstruction issue.

In this section, we extensively discussed the efficiency of the FMT detector, addressing alignment issues, as well as the impact of implementing a geometry cut. The key findings can be summarised as follows: despite applying corrections, the FMT efficiency for 3-layer tracks remained low, rendering them impractical for exclusive use. However, by switching to Spring 2020 data and applying the geometry cut, significant improvements were observed in the detection of trigger electrons and pions. Moreover, we thoroughly examined the correlation between efficiency and various variables, namely v_z , θ , ϕ , and p . The results were as anticipated: strong correlations were observed for v_z and θ , while no significant correlation was found for ϕ or p . These findings provide valuable insights into the performance and limitations of FMT, paving the way for an acceptance correction study and the following work.

5.1.1 Alignment Effect

The data from the RG-F experiment is divided based on the season in which the runs take place, namely Spring 2020 and Summer 2020. According to the run group's guidelines, it is recommended to use Summer data as it has undergone more calibration compared to the Spring data. However, the calibration work conducted so far does not include the FMT detector, resulting in a significant misalignment effect.

Through simple visual inspection, two distinct peaks can be clearly observed between $z = -36$ cm and $z = -30$ cm in Figure 29. The leftmost peak corresponds to the scattering chamber window, while the second peak corresponds to the RG-F target window. Both peaks appear flatter in the lines corresponding to run 12933 in Figure 36. As discussed in Section 3, this issue arises due to the lack of correction for FMT misalignments.

The simplest solution is to utilise Spring data. Although more calibration work has been performed on the Summer data, it mainly pertains to the CD, which is not used in this analysis. In the lines corresponding to run 12016 (Spring 2020) in Figure 36, we can see that both peaks are sharper. This indicates that the misalignment issue has been appropriately addressed in that particular run.

5.1.2 Geometry Effect

This problem has already been extensively discussed in Section 3.4.2. In summary, the FMT detector is located at approximately $z \approx 26$ cm, and it exhibits poor performance for targets located too close to it. The strength of this effect can be quantified by applying the geometry cut described by Equation (4) to both the DC and FMT tracks. Figure 37 illustrates the impact of this cut on both the DC and FMT tracks when applied to Figure 36. The effect of the cut on a v_z vs. θ plot can be observed in Figure 30.

Based on this cut and the FMT's position along the z -axis, subsequent plots will be confined to the range $-30\text{cm} < v_z < 20\text{cm}$. In addition, unless otherwise specified, the cut is applied for all subsequent plots.

5.1.3 Reconstruction Effect

Even after correcting for both the alignment and geometry issues, the FMT still exhibits lower efficiency compared to what was observed during alignment (compare Figure 37 with Figure 31). After conducting a thorough study, it was determined that this effect is not correlated with the run number, beam energy, or beam luminosity. Consequently, it can be concluded that the issue is not caused by hardware problems or run conditions.

Based on these findings, the logical conclusion is that the effect stems from a general issue in the FMT offline reconstruction for data. Identifying and rectifying this issue would require a larger project beyond the scope of this thesis and is therefore left as future work. For the purposes of this analysis, we will rely on a large number of events to minimize any statistical deficiencies.

5.1.4 Efficiency Study

With all these effects accounted for, we can proceed to study the efficiency in detail. First, if we define FMT efficiency as the percentage of DC tracks that get accepted by FMT, we

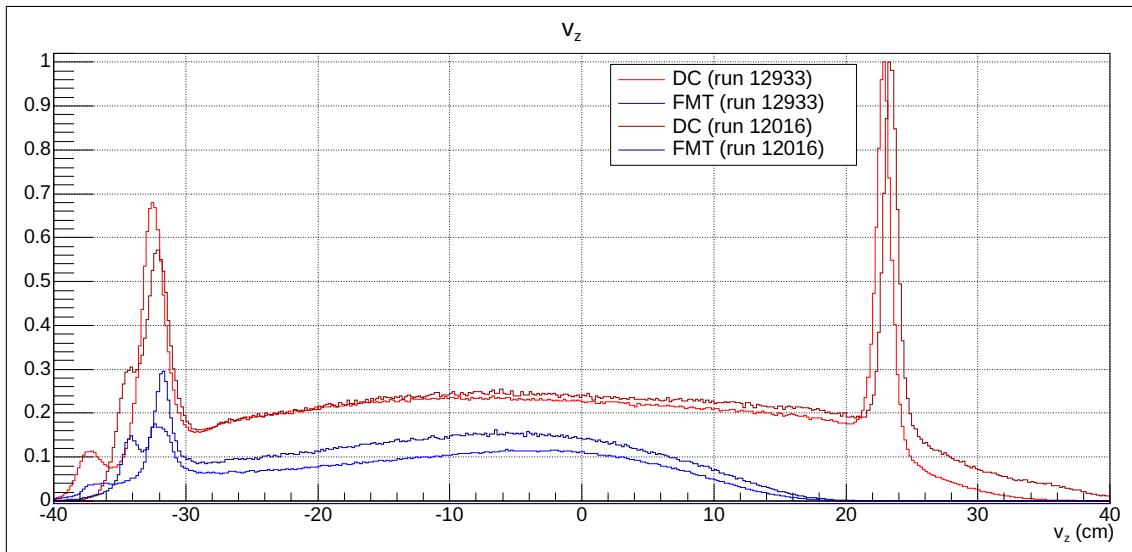


Figure 36: v_z for DC (in red) and FMT (in blue). Runs 12933 (Summer 2020) and 12016 (Spring 2020).

Source: Own elaboration, using the clas12-rge-analysis software.

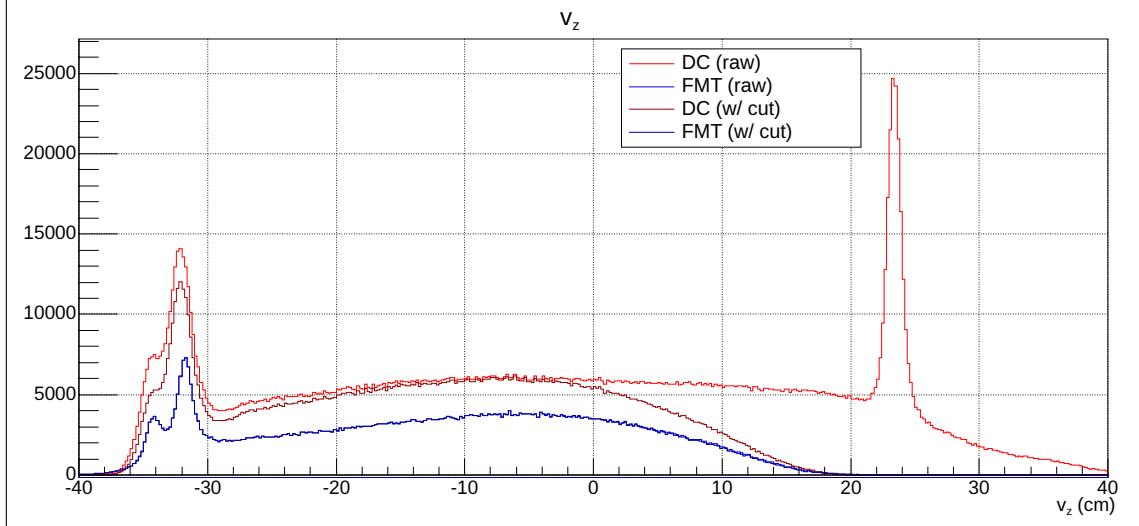


Figure 37: v_z for DC without the geometry cut (in red), with it (in dark red), for FMT without it (in blue), and with it (in dark blue). Spring 2020 data, run 12016. The effect of the cut is very clear on DC tracks, yet it almost doesn't affect FMT tracks.

Source: Own elaboration, using the clas12-rge-analysis software.

get the results in Table 6 for runs 12933 (Summer 2020) and 12016 (Spring 2020), as well as the simulation described in Section 4.4.

To estimate the efficiency of one FMT layer and the associated errors for the two types of tracks, we define $P(L_n)$ as the probability of layer n detecting a particle, with $1 \leq n \leq 3$. Assuming that all layers have the same efficiency, denoted as E_1 ,

$$P(L_1) = P(L_2) = P(L_3) = E_1,$$

the efficiency E_3 for 3-layer tracks can be obtained using the probabilities $P(L_n)$ as follows

$$\begin{aligned} E_3 &= P(L_1)P(L_2)P(L_3) \\ E_3 &= E_{1(3)}^3, \end{aligned} \tag{8}$$

where $E_{1(3)}$ is the 1-layer efficiency estimated from 3-layer tracks.

Run 12933			Run 12016		
	raw	w/ cut	raw	w/ cut	
e^-	2 layers	25.1 ± 1.5	37.5 ± 0.7	32.7 ± 2.5	53.7 ± 0.8
	3 layers	5.6 ± 2.7	8.5 ± 1.7	9.9 ± 6.3	16.4 ± 3.6
$e^- \pi^+$	2 layers	6.5 ± 0.2	13.7 ± 0.9	11.1 ± 0.2	28.0 ± 1.3
	3 layers	0.3 ± 0.1	0.7 ± 0.4	1.0 ± 0.1	2.7 ± 1.4
$e^- \pi^-$	2 layers	5.6 ± 0.0	14.2 ± 1.0	8.9 ± 0.4	29.5 ± 1.4
	3 layers	0.3 ± 0.0	0.7 ± 0.5	0.9 ± 0.3	2.9 ± 1.6

Table 6: Results of the FMT efficiency study performed in %.

For 2-layer tracks, the efficiency E_2 can be obtained as

$$\begin{aligned}
E_2 &= P(L_1)P(L_2)(1 - P(L_3)) \\
&\quad + P(L_2)P(L_3)(1 - P(L_1)) \\
&\quad + P(L_3)P(L_1)(1 - P(L_2)) \\
&\quad + P(L_1)P(L_2)P(L_3) \\
E_2 &= 3E_{1(2)}^2 \left(1 - E_{1(2)}\right) + E_{1(3)}^3 \\
E_2 &= 3E_{1(2)}^2 \cdot \left(1 - E_{1(2)}\right) + E_3,
\end{aligned} \tag{9}$$

where $E_{1(2)}$ is the 1-layer efficiency estimated from 2-layer tracks.

From (8), we can estimate $E_{1(3)}$ as

$$E_{1(3)} = \sqrt[3]{E_3}. \tag{10}$$

In contrast, $E_{1(2)}$ cannot be obtained explicitly from (9), but it can be estimated numerically.

Using Equations (9) and (10), we can estimate the weighted average efficiency \bar{E}_1 as

$$\bar{E}_1 = \frac{4E_{1(2)} + E_{1(3)}}{5},$$

where the weights are assigned based on the number of ways 2 and 3-layer tracks can be obtained, 4 and 1, respectively.

From \bar{E}_1 , we can estimate the errors on $E_{1(2)}$ and $E_{1(3)}$ as

$$\begin{aligned}
\delta(E_{1(2)}) &= |\bar{E}_1 - E_{1(2)}|, \\
\delta(E_{1(3)}) &= |\bar{E}_1 - E_{1(3)}|.
\end{aligned}$$

To propagate these errors to the efficiencies E_2 and E_3 , we use the variance formula

$$\delta(f(x)) = \frac{\partial f(x)}{\partial x} \cdot \delta(x),$$

where $\delta(E_2)$, obtained from Equation (9), is

$$\begin{aligned}
\delta(E_2) &= \frac{\partial}{\partial E_{1(2)}} \left(3E_{1(2)}^2 - 3E_{1(2)}^3 + E_{1(3)}^3\right) \cdot \delta(E_{1(2)}) \\
\delta(E_2) &= \left(6E_{1(2)} - 9E_{1(2)}^2\right) \cdot \delta(E_{1(2)})
\end{aligned}$$

and $\delta(E_3)$, obtained from Equation (8), is

$$\begin{aligned}
\delta(E_3) &= \frac{\partial}{\partial E_{1(3)}} \left(E_{1(3)}^3\right) \cdot \delta(E_{1(3)}) \\
\delta(E_3) &= 3E_{1(3)}^2 \cdot \delta(E_{1(3)}).
\end{aligned}$$

A Python script was written to calculate E_2 and E_3 from each measurement shown in Table 6. The script is included in Appendix C. The corresponding errors were obtained this way, and are included in the table.

The table illustrates the positive impact of switching to Spring 2020 data and applying the geometry cut. The switch results in a 30.1% increase in detected trigger electrons, as well as a 70.8% increase for positive pions and a 58.9% increase for negative pions.

Furthermore, by applying the geometry cut based on v_z and θ , an additional 64.2% increase in trigger electrons is achieved, resulting in a total increase of 113.9%. The pion yield experiences a substantial enhancement, with positive pions increasing by 150.2% and negative pions increasing by 231.5%, resulting in a total increase of 330.8% and 426.8%, respectively.

Next, we need to ensure that we are not introducing a systematic error by applying these corrections. To achieve this, we must investigate the effect of the geometry cut on different detected variables for electrons (e^-), positive pions ($e^-\pi^+$), and negative pions ($e^-\pi^-$). Based on the definition of the cut, we anticipate a strong correlation between efficiency and v_z and θ , and at most a weak correlation with ϕ and p .

Our prediction is validated by Figures 38a and 38b. The efficiency displays a pronounced dependence on both v_z and θ for all three particles under study. It is worth noting that the dependence is more pronounced for electrons compared to pions. This outcome aligns with expectations, as the only accepted pions are those detected in events where an electron is also detected. Thus, the “final” pion efficiency incorporates the combined detector efficiencies for electrons and pions.

Next, Figure 38c confirms our expectation that there is no strong correlation between efficiency and momentum.

Last, we examine the ϕ efficiency, as depicted in Figure 39a. Our prediction holds true for both positive and negative pions, as the efficiency demonstrates no dependence on the value of ϕ . However, in the case of electrons, upon initial inspection, the sharp valleys appear less prominent after applying the geometry cut.

Nevertheless, this is merely a visual effect. When we tally the number of electrons detected by the DC and FMT, the valleys become more pronounced in the DC, while they remain the same in the FMT. Consequently, the increase in FMT efficiency becomes more pronounced in these valleys, as demonstrated in Figure 39b.

As can be seen in Table 6 and Figures 38 and 39, the pion efficiency for 3-layer tracks does not exceed 3%. Consequently, it is not practical to work exclusively with FMT tracks that traverse all three FMT layers in this study. For the remainder of the document, 2 and 3-layer tracks will be considered together, and will be collectively referred to as FMT tracks.

5.2 Acceptance Correction Results

DC and FMT acceptances were obtained by following the procedure described in Section 4.1.3. In this section, we will study the acceptance percentage for each electron and hadronic variable. These numbers represent the percentage of thrown particles that were detected by the DC and by 2 or 3 FMT layers in addition to DC in the GEMC simulation.

Since they represent the acceptance percentage, all plots display a calculated division as follows

$$y_{\text{acc}} = \frac{y_r}{y_t}, \quad (11)$$

where y_{acc} is the percentage of accepted particles, y_r is the number of reconstructed particles by the studied detector in the GEMC simulation, and y_t is the number of particles thrown by LEPTO.

Then, to propagate the error of y_r (e_d) and y_t (e_t) to y_{acc} (e_{acc}), we have

$$\begin{aligned} e_{\text{acc}} &= \delta \left(\frac{y_r}{y_t} \right) \\ &= y_{\text{acc}} \cdot \sqrt{\left(\frac{e_r}{y_r} \right)^2 + \left(\frac{e_t}{y_t} \right)^2}, \end{aligned}$$

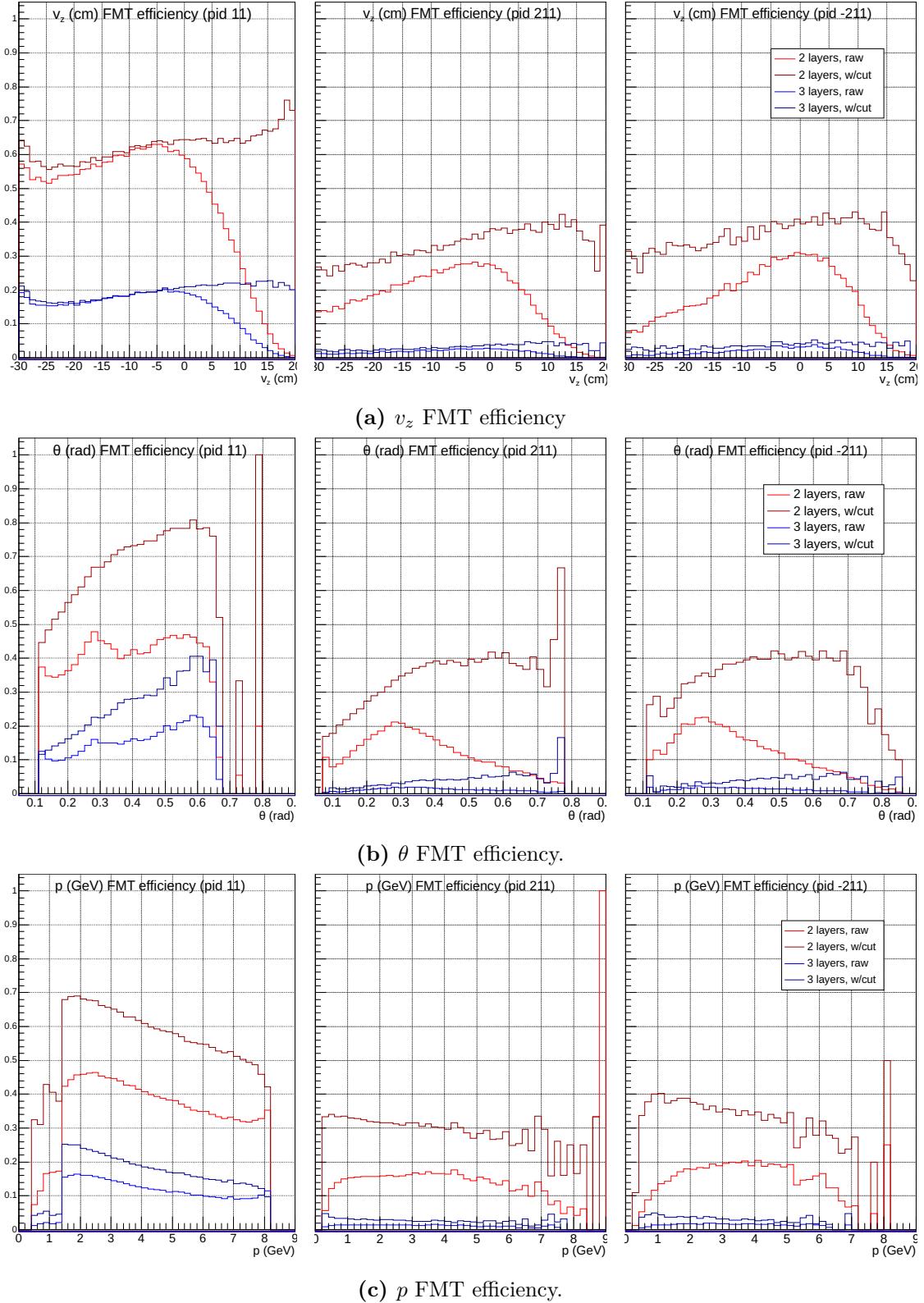
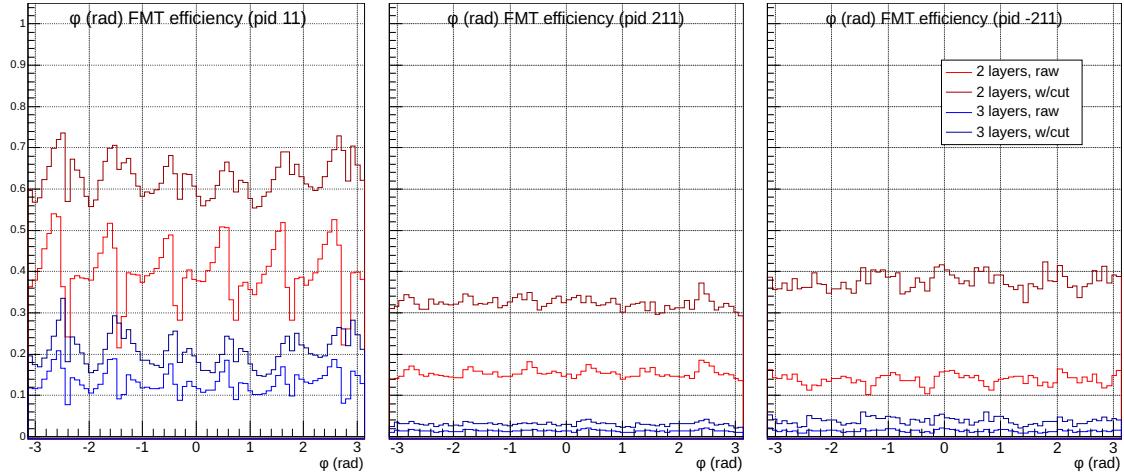
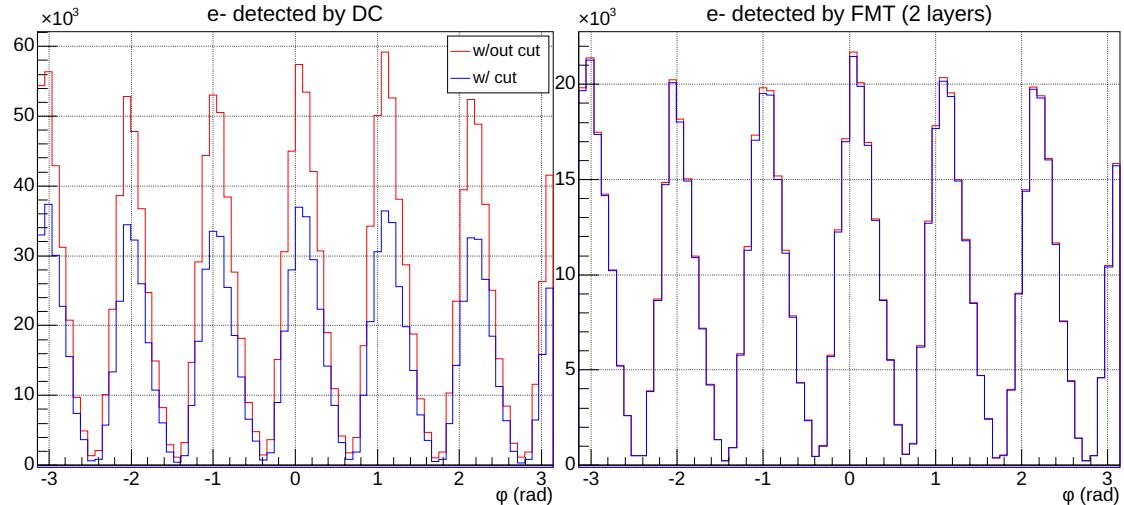


Figure 38: v_z , θ , and p FMT efficiencies for e^- , $e^-\pi^+$, and $e^-\pi^-$. FMT efficiency is defined as the percentage of DC tracks that are detected by 2 FMT layers. Run 12016.

Source: Own elaboration, using the clas12-rge-analysis software.



(a) ϕ FMT efficiency for e^- , $e^- \pi^+$, and $e^- \pi^-$. FMT efficiency is defined as the percentage of DC tracks that are detected by 2 FMT layers.



(b) Number of electrons detected in terms of ϕ for DC and FMT.

Figure 39: ϕ FMT efficiency and number of e^- detected for DC and FMT. Run 12016.

Source: Own elaboration, using the clas12-rge-analysis software.

since y_t is the number of trials, we can assume $e_t = 0$, and thus

$$= y_{\text{acc}} \cdot \frac{e_r}{y_r}. \quad (12)$$

From the Central Limit Theorem, assuming a normal distribution, the variance e_r^2 is given by

$$e_r^2 = y_t \cdot y_{\text{acc}}(1 - y_{\text{acc}}).$$

Replacing this in Equation (12), we obtain

$$e_{\text{acc}} = y_{\text{acc}} \cdot \frac{\sqrt{y_t y_{\text{acc}}(1 - y_{\text{acc}})}}{y_r}.$$

Now, replacing y_{acc} with its definition from Equation (11), we arrive at the final error expression

$$e_{\text{acc}} = \sqrt{\frac{y_{\text{acc}}(1 - y_{\text{acc}})}{y_t}}. \quad (13)$$

5.2.1 Electron Variables

The acceptance of the scattered electron variables Q^2 and ν is presented in Figure 40. Each one is shown in an integrated kinematical region for the other variable.

In Equation (5), it can be observed that Q^2 has a quadratic dependence on the scattering angle θ_C of the scattered electrons. Hence, it is important to understand the θ acceptance of CLAS12 in order to distinguish the geometric effect related to θ from the inherent Q^2 acceptance of the FD. Figure 41 illustrates this dependence for negative particles.

The triangular shape of each DC sector, combined with the inbending tracks resulting from the negative solenoid field, leads to a significantly low acceptance at low θ angles. When integrating across ϕ , this results in a low θ efficiency for $\theta \lesssim 0.15$ radians. Referring back to the Q^2 plot in Figure 40a, the decrease in Q^2 acceptance between 1 and 4 GeV^2 can be clearly attributed to this geometric effect, which is purely of a geometric nature.

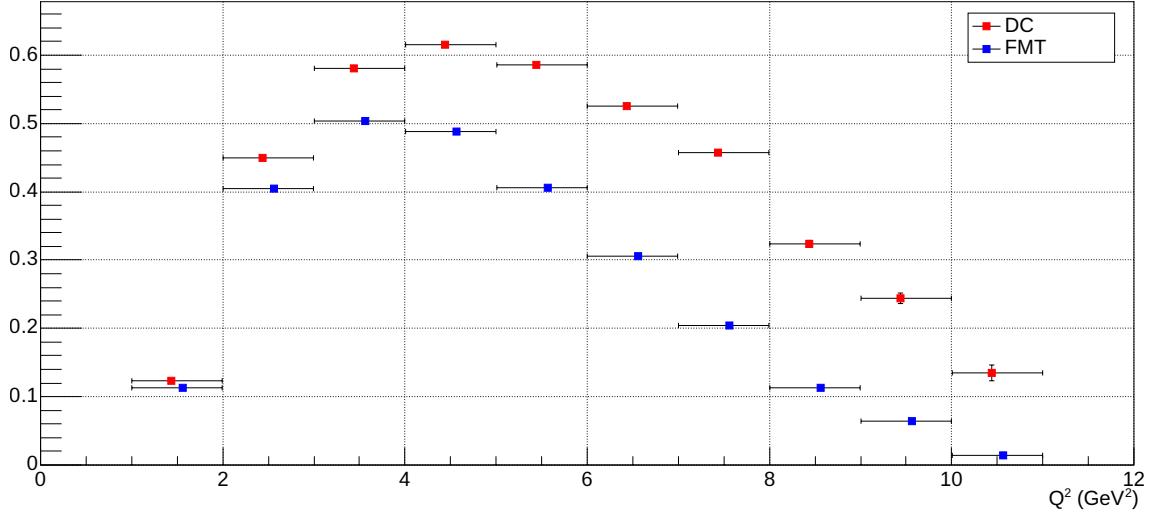
On the other hand, since ν does not exhibit a direct correlation with θ_C (as indicated by Equation (6)), we can assume that the acceptance observed in Figure 40b is intrinsic to the detector.

5.2.2 Hadronic Variables

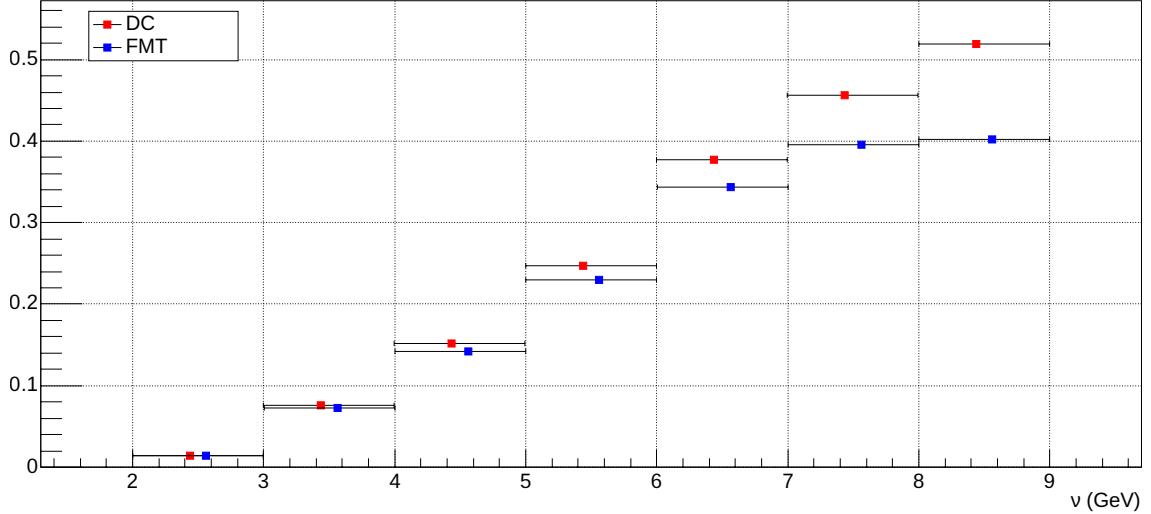
The acceptance of the hadronic variables z_h , p_T^2 , and ϕ_{PQ} for $e^- \pi^+$ and $e^- \pi^-$ is presented in Figure 42. Each plot shows the acceptance in an integrated kinematical region for all electron variables and other hadronic variables.

It can be observed from the figures that the acceptance for the hadronic variables is lower compared to that of the electron variables. This is expected since the detection of the hadronic variables requires both the trigger electron and at least one hadron to be detected by the CLAS12 detector. This effect is also reflected in the entries for $e^- \pi^+$ and $e^- \pi^-$ in the efficiency Table 6.

Furthermore, the acceptance for the hadronic variables is shown to be approximately half for $e^- \pi^-$ compared to $e^- \pi^+$. This difference is related to the θ acceptance of CLAS12, which is influenced by the non-trivial magnetic field of the solenoid. As explained in Section 5.2.1, during the majority of run 12016, the solenoid field was configured to a polarity of -0.75 . This causes negative particle tracks to bend inward towards the low- θ region of the Drift Chambers (DC), while positive particle tracks bend outward towards the high- θ region.



(a) Q^2 acceptance.



(b) ν acceptance.

Figure 40: e^- variables acceptance. ν is integrated in 40a, and Q^2 is integrated in 40b. The bin markers are slightly shifted in x to improve legibility.

Source: Own elaboration, using the clas12-rge-analysis software.

The ϕ and θ acceptances of the detector for negative particles can be seen in Figure 41, whereas those for positive particles are shown in Figure 43. These plots clearly demonstrate that the DC is generally more receptive to positive particles. Consequently, there is a higher detection efficiency for π^+ compared to π^- , which explains the differences in the acceptance of the hadronic variables between $e^-\pi^+$ and $e^-\pi^-$.

The variable z_h does not depend on the θ angle of either the hadron or the scattered electron, as can be seen in Equation (3). Therefore, the observed z_h acceptance ratios in Figures 42a and 42b are solely related to the intrinsic z_h acceptance of CLAS12.

Regarding p_T^2 , initially described in Section 4.4, it is defined as $p_T^2 = p_h^2 \cdot (1 - \cos^2(\theta_{PQ}))$, where p_h represents the momentum of the hadron in the lab frame, and θ_{PQ} is the angle between the virtual interaction photon and the detected hadron. Both p_h and θ_{PQ} have a non-trivial dependence on the θ angles of both the scattered electron and the detected hadron. The variable p_T^2 is presented in Figures 42c and 42d.

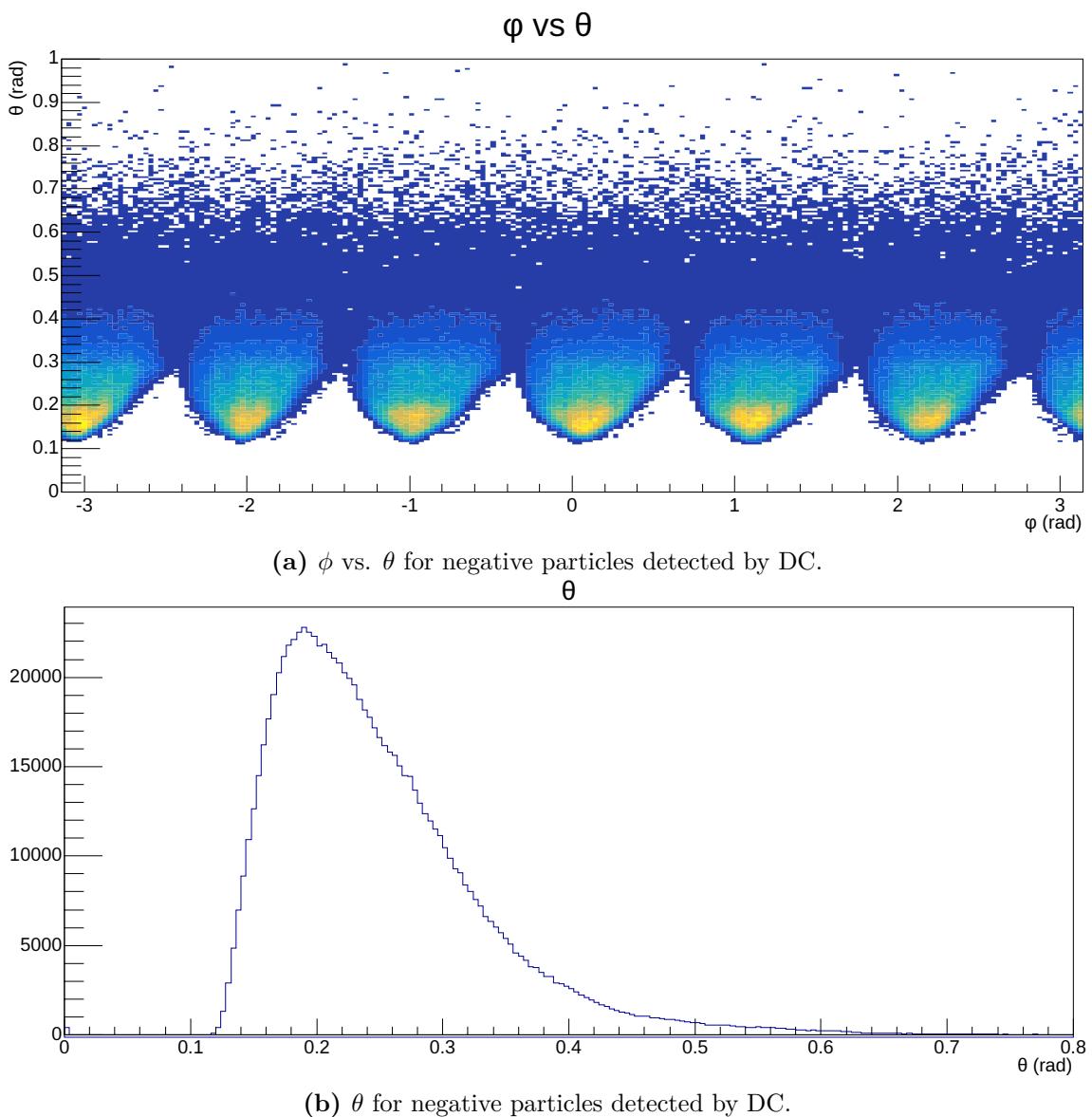


Figure 41: θ study for negative particles. Simulated RG-F target.

Source: Own elaboration, using the clas12-rge-analysis software.

In addition to its shape, an increase in both value and uncertainty can be observed for $p_T^2 \gtrsim 1.4$ in both the DC and FMT acceptance plots. This increase is purely due to statistical uncertainty, as very few particles with these values of p_T^2 are accepted by CLAS12 in the simulated data. These uncertainties will subsequently be inherited by the plots obtained in Section 5.3 of the thesis.

The variable ϕ_{PQ} , as described in Section 4.4, represents the angle between the leptonic and hadronic planes. It has a highly non-trivial relationship with the θ and ϕ angles of both the scattered electron and the detected hadron. Despite this complexity, it is known that the ϕ_{PQ} acceptance generally follows a similar shape as observed in Figures 42e and 42f. This behaviour has been observed and documented in DIS analysis publications such as [Morán et al., 2022] and others.

5.3 Study Results

To select a region in v_z to place the RG-E target, two criteria were considered: a phase space study and a statistics study. In the phase space study, the RG-F target gas region ($-30 < v_z < 20$ cm) was divided into ten 5 cm bins, and the phase space of each kinematic variable was analysed in each bin. The results of this study are presented in Section 5.3.1. For the statistics study, the region chosen in the phase space study was further divided into 1 cm bins. The study aimed to determine which 7 cm region within the chosen range had the largest statistics. The results of this study are presented in Section 5.3.2.

The total statistical error on the acceptance-corrected result, denoted as e_{corr} , needs to consider both the statistical error of the measurements (e_{meas}) and the acceptance correction (e_{acc}). The error of the measurements, e_{meas} , is purely statistical in nature and is given by

$$e_{\text{meas}} = \frac{\delta y_{\text{meas}}}{y_{\text{meas}}},$$

The acceptance correction error, e_{acc} , was derived in Equation (13).

Since e_{meas} arises purely from experimental data and e_{acc} is purely from simulation, they are considered to be completely uncorrelated. Therefore, the total statistical error of the acceptance-corrected result can be estimated as the quadratic addition of the two errors, i.e.,

$$e_{\text{corr}} = \sqrt{e_{\text{meas}}^2 + e_{\text{acc}}^2}.$$

5.3.1 Phase Space Study

Considering the objectives of this DIS study, it is advantageous to maximise the phase space of each kinematic variable of study. This approach broadens the scope of investigation, increases sensitivity to detect rare phenomena, and facilitates the testing of theoretical predictions for future studies using the double target system. Therefore, the first criterion for selecting a v_z region for the target is to find a region that provides the maximum range of kinematic variables.

The resulting plots show the acceptance-corrected DIS variables separated into v_z bins. Figures 44 and 45 display the distributions of the electron variables Q^2 and ν , respectively. The z_h distributions for $e^- \pi^+$ and $e^- \pi^-$ can be observed in Figures 46 and 47, respectively. Figures 48 and 49 show the distributions of p_T^2 for $e^- \pi^+$ and $e^- \pi^-$, respectively. Finally, Figures 50 and 51 present the distributions of ϕ_{PQ} for $e^- \pi^+$ and $e^- \pi^-$, respectively. These plots provide insights into the dependence of each DIS variable on the v_z coordinate. For these same distributions without acceptance correction, please see Appendix D.

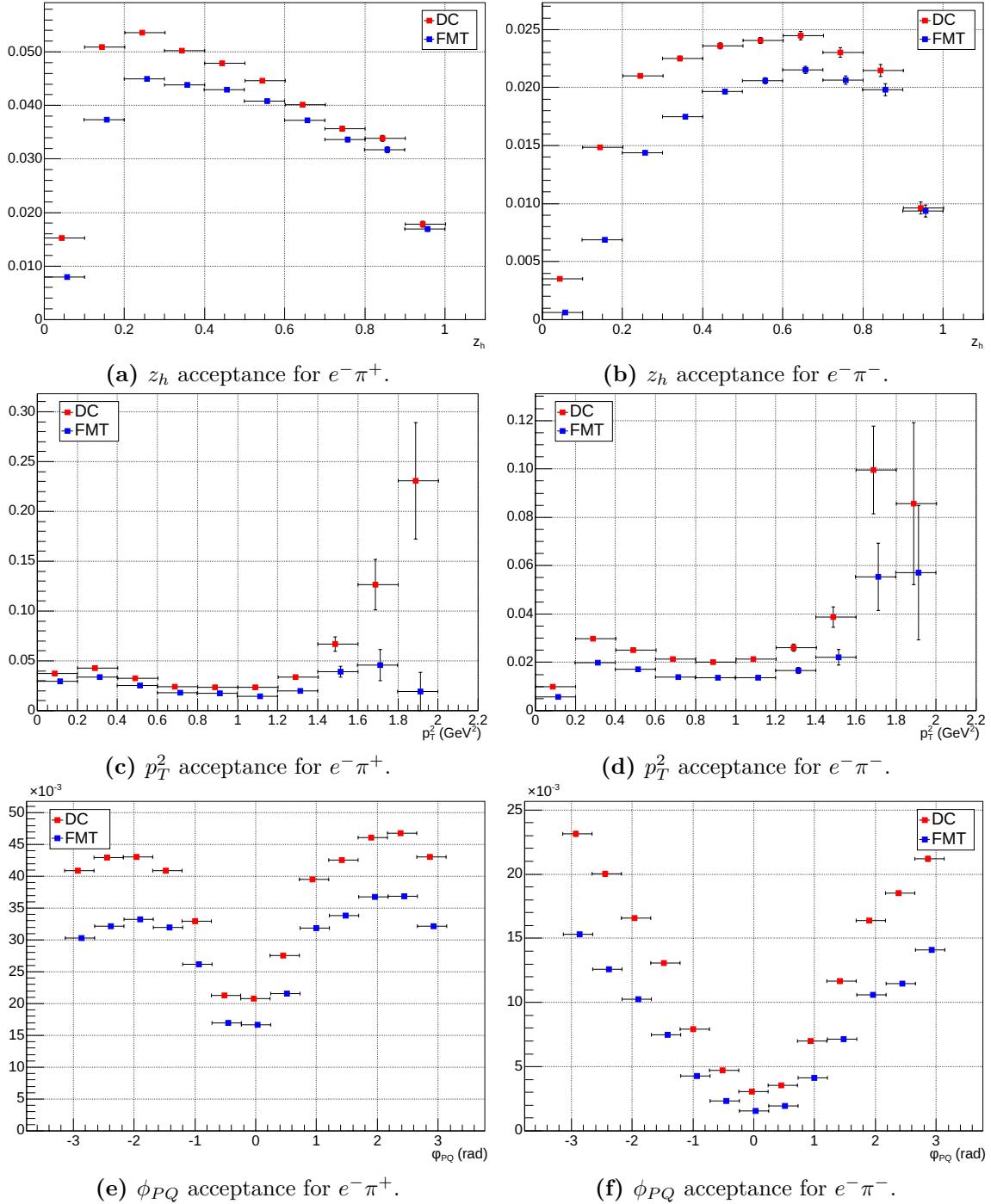
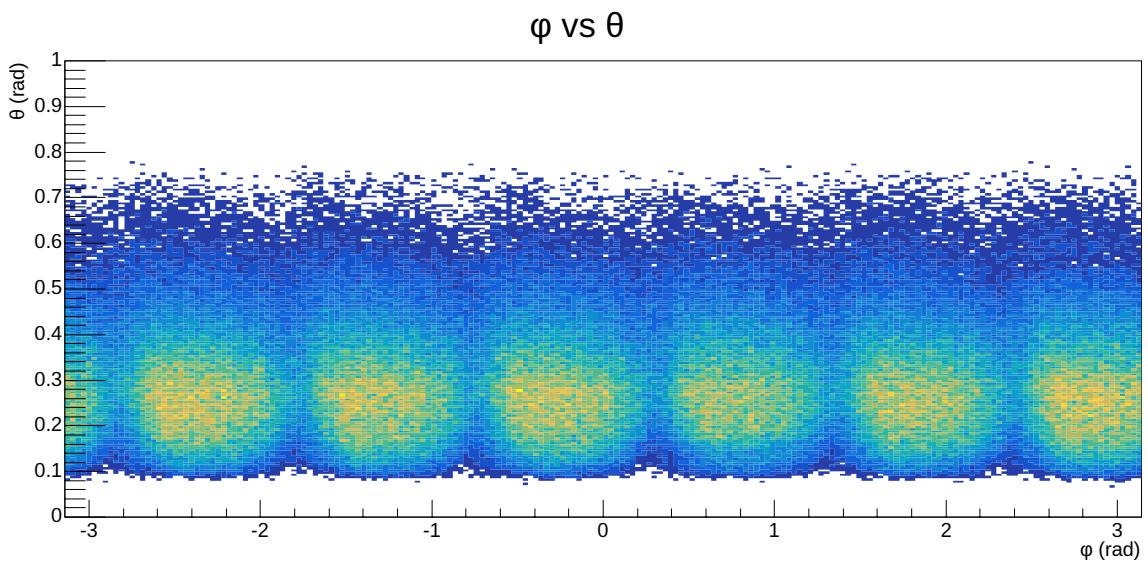
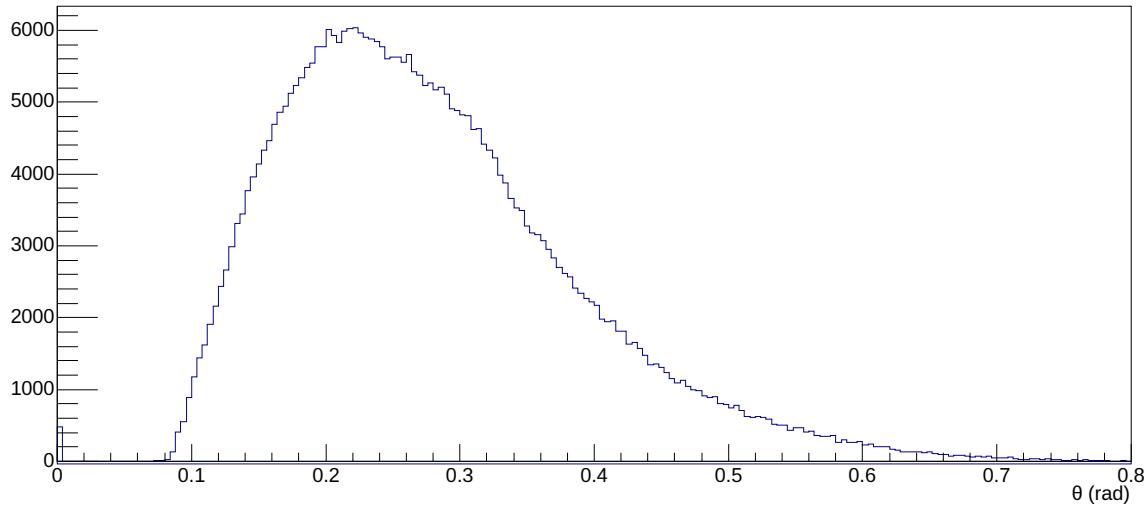


Figure 42: z_h , p_T^2 , and ϕ_{PQ} acceptances for $e^- \pi^+$ and $e^- \pi^-$. All electron and other hadronic variables are integrated in all Figures. The bin markers are slightly shifted in x to improve legibility.

Source: Own elaboration, using the clas12-rge-analysis software.



(a) ϕ vs. θ for positive particles detected by DC.
 θ (rad)



(b) θ for positive particles detected by DC.

Figure 43: θ study for positive particles. Simulated RG-F target.

Source: Own elaboration, using the clas12-rge-analysis software.

In the study of Q^2 , as shown in Figure 44, the higher end of the variable's phase space is limited for $v_z < -5$ cm, with the effect becoming more pronounced as we move further upstream. This effect can be understood by considering the compounded effect of the θ efficiency for negative particles (as seen in Figure 41) and the limited acceptance region of FMT (described by Equation (4) and illustrated in Figure 30).

The higher end of θ becomes limited for lower v_z values. Based on the objective of maximising the phase space of each variable, this suggests setting the minimum v_z for the RG-E target near -5 cm. Additionally, it is noted that the variable exhibits an unusual shape for $10 \text{ cm} < v_z < 20 \text{ cm}$, likely due to the cut in low θ angles in that region, which is another consequence of the FMT acceptance region.

In the study of ν , as seen in Figure 45, it was previously observed that ν has no direct correlation with the scattering angle θ_C (Section 5.2). Therefore, no significant effect on the phase space of ν is observed for $v_z < -5$ cm, unlike Q^2 . However, a loss is observed in the lower end of the phase space for $v_z = 10$ cm and downstream. Based on this effect, it is reasonable to keep v_z below approximately 10 cm to preserve the largest possible phase space of ν .

In the study of z_h , despite its lack of direct correlation with the electron's and pion's θ , clear differences are observed across different v_z bins, as shown in Figures 46 and 47. However, this can be explained by its inverse correlation with ν (as described in Equation (3)). Similar to ν , the extreme phase space loss is primarily observed for $v_z > 10$ cm, and therefore, no additional severe restrictions on the v_z region are imposed beyond those defined based on the studies of Q^2 and ν .

In the study of p_T^2 , as depicted in Figures 48 and 49, large statistical fluctuations are observed for $p_T^2 > 1.4 \text{ GeV}^2$, consistent with the prediction in Section 5.2.2. Studying the phase space of the variable, a cutoff at high p_T^2 values is observed for $v_z < -5$ cm and $v_z > 15$ cm, similar to what was seen for Q^2 . Based on this observation, no additional restrictions are imposed on the v_z region under study.

Regarding the study of ϕ_{PQ} , as shown in Figures 50 and 51, no easily discernible loss is observed in the phase space of ϕ_{PQ} as we vary v_z . While there are significant changes in the shape of the variable distribution across different v_z bins, conducting a detailed shape study is beyond the scope of this thesis, as ample information is already provided by the other DIS variables.

5.3.2 Statistics Study

After obtaining the v_z region with the maximum range of kinematic variables, our second criterion is to maximise statistics inside this region. With phase space already maximised, this allows us to increase the statistical precision of future studies, enabling a thorough exploration of the parameter space.

Based on the results of the different phase space studies, we'll limit the scope of this statistics study to the region $-5 \text{ cm} < v_z < 10 \text{ cm}$, the result of which can be observed in Figure 52. The objective of this study is to find the 7 cm region in v_z with the highest statistics. The choice of 7 cm in particular comes from the length of the double target, where the liquid target has a length of 3 cm, the separation between the liquid and solid target is 4 cm, and the solid target is of negligible width for the scope of this study.

First, we look at e^- statistics, seen in Figure 52a. It is trivial to note that, in the studied range, the more upstream we look, the more statistics are seen. Therefore, the best region for the placement of the target is from -5 to 2 cm. This result is only reinforced by the results observed in $e^- \pi^+$ and $e^- \pi^-$ statistics, which can be seen in Figures 52b and 52c.

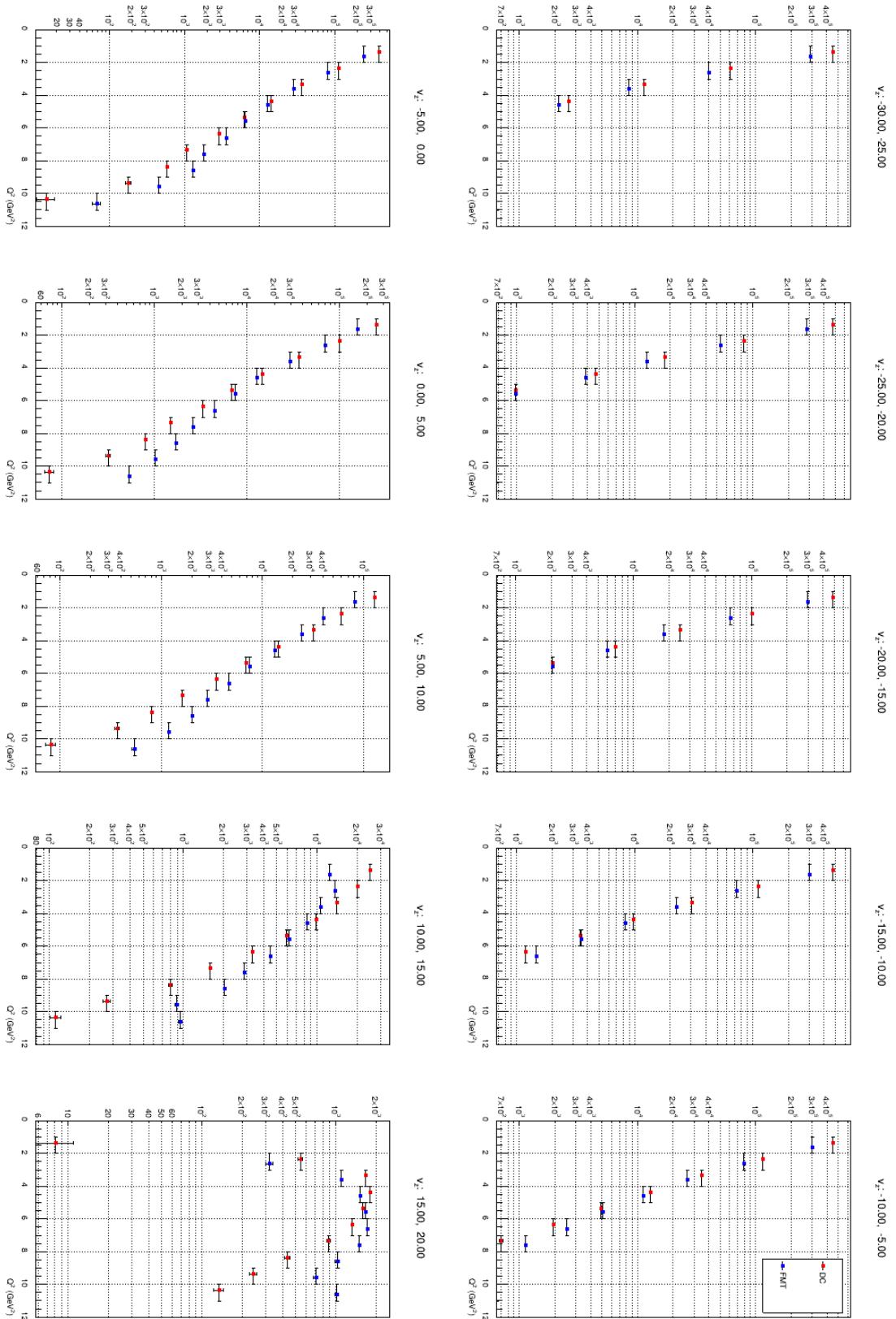


Figure 44: Acceptance-corrected Q^2 detected by DC and FMT, separated in v_z bins. Run 12016. The bin markers are slightly shifted in x to improve legibility.

Source: Own elaboration, using the clas12-rge-analysis software.

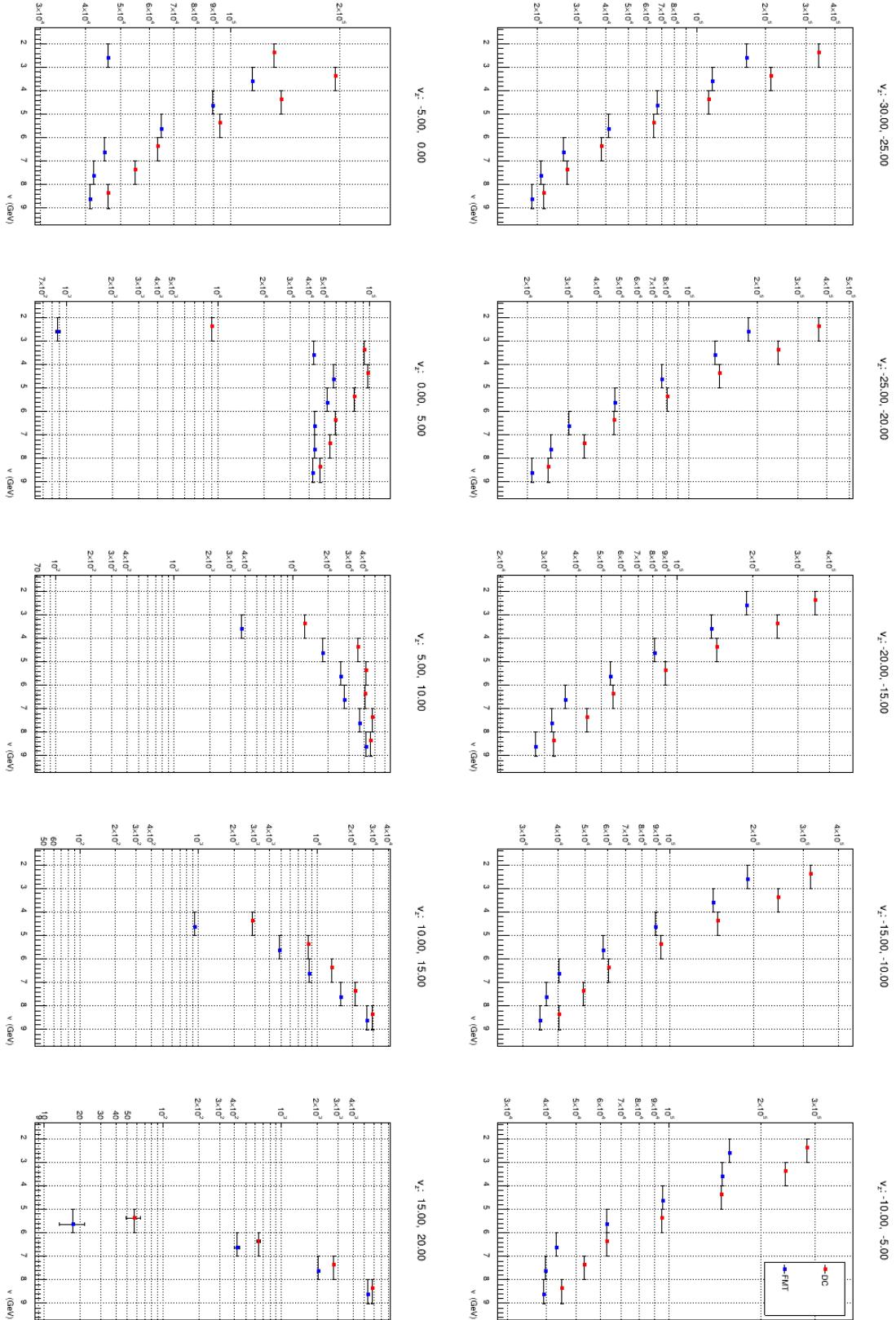


Figure 45: Acceptance-corrected ν detected by DC and FMT, separated in v_z bins. Run 12016. The bin markers are slightly shifted in x to improve legibility.

Source: Own elaboration, using the clas12-rge-analysis software.

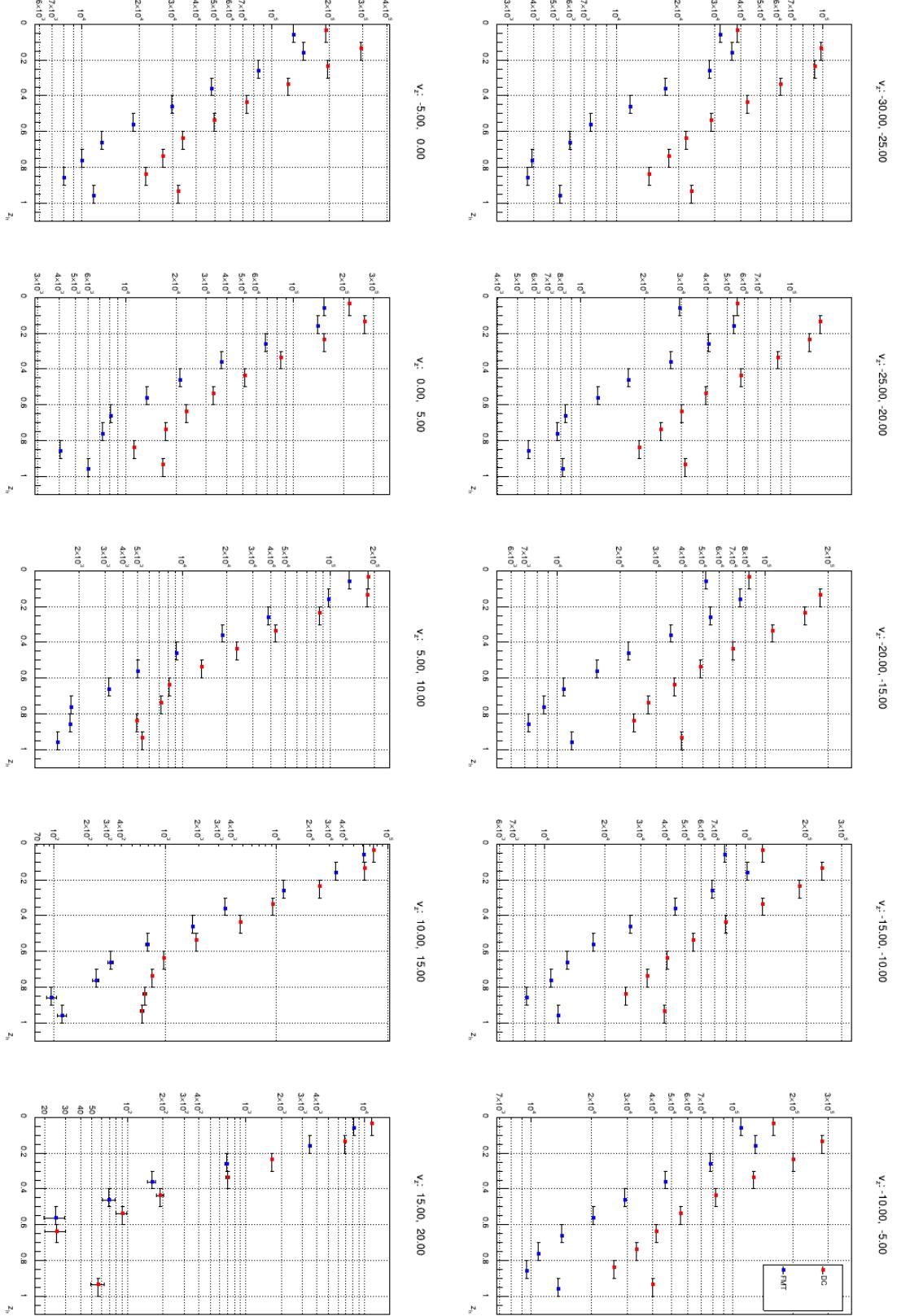


Figure 46: Acceptance-corrected z_h for $e^- \pi^+$ detected by DC and FMT, separated in v_z bins. Run 12016. The bin markers are slightly shifted in x to improve legibility.

Source: Own elaboration, using the clas12-rge-analysis software.

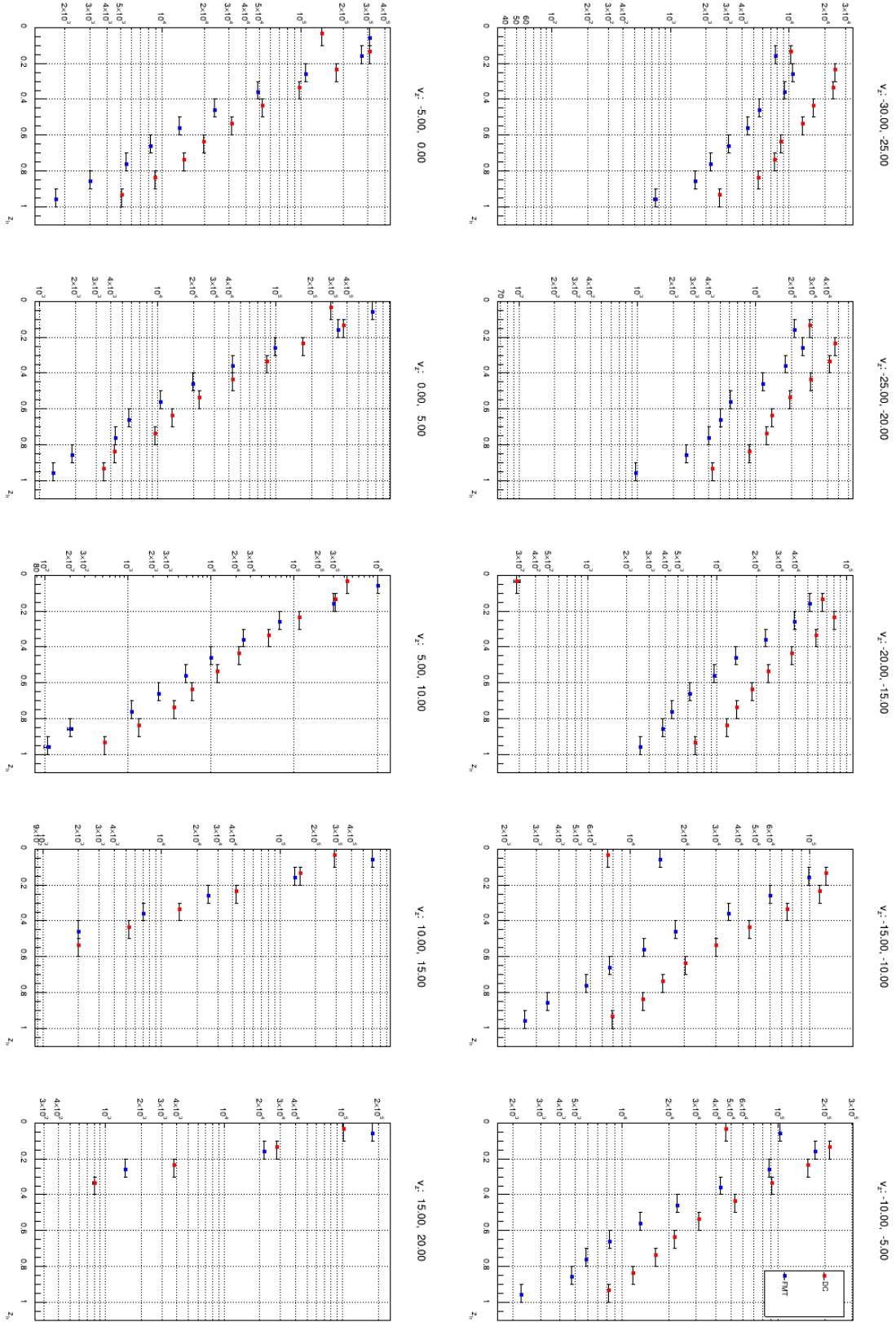


Figure 47: Acceptance-corrected z_h for $e^- \pi^-$ detected by DC and FMT, separated in v_z bins. Run 12016. The bin markers are slightly shifted in x to improve legibility.

Source: Own elaboration, using the clas12-rge-analysis software.

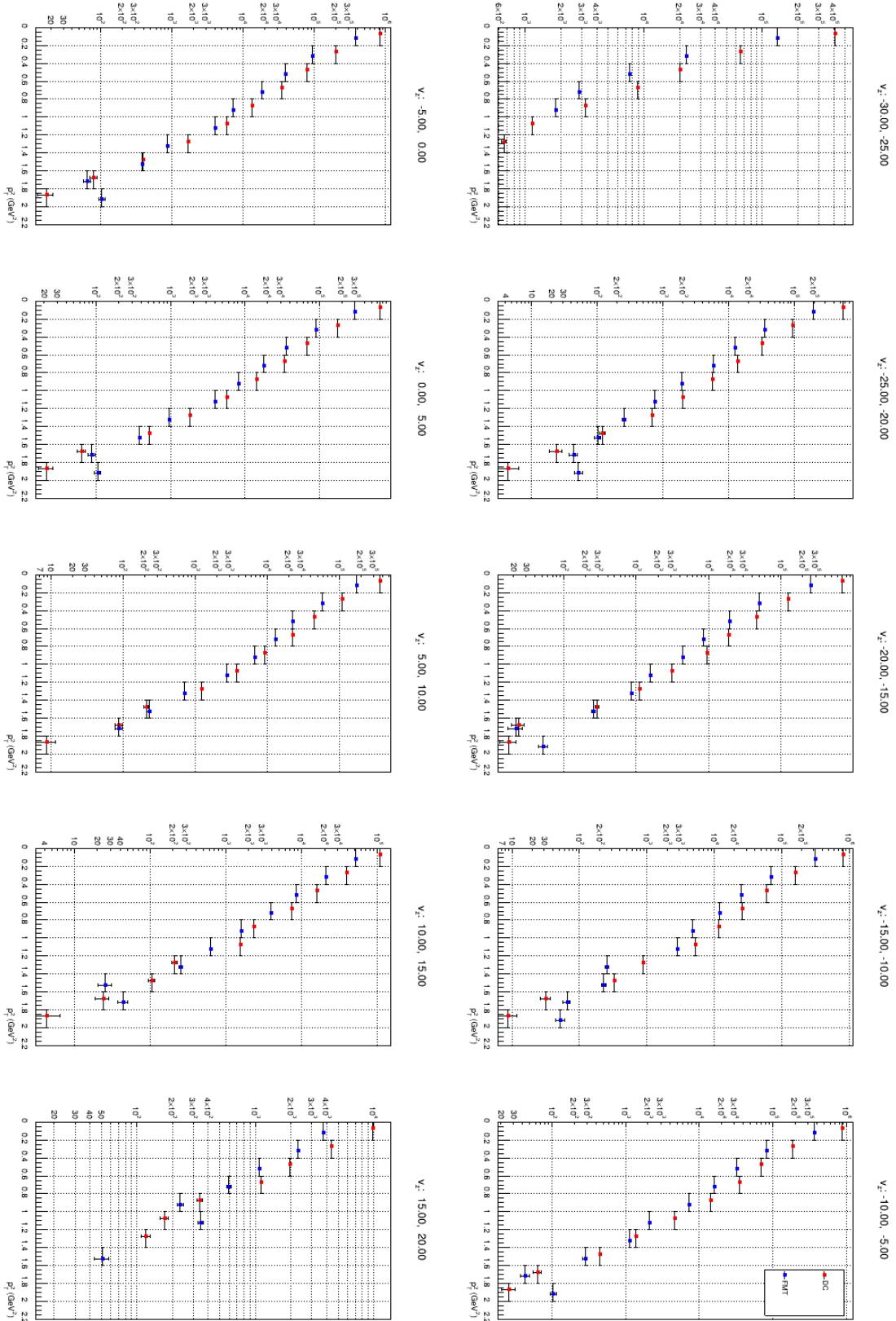


Figure 48: Acceptance-corrected p_T^2 for $e^- \pi^+$ detected by DC and FMT, separated in v_z bins. Run 12016. The bin markers are slightly shifted in x to improve legibility.

Source: Own elaboration, using the clas12-rge-analysis software.

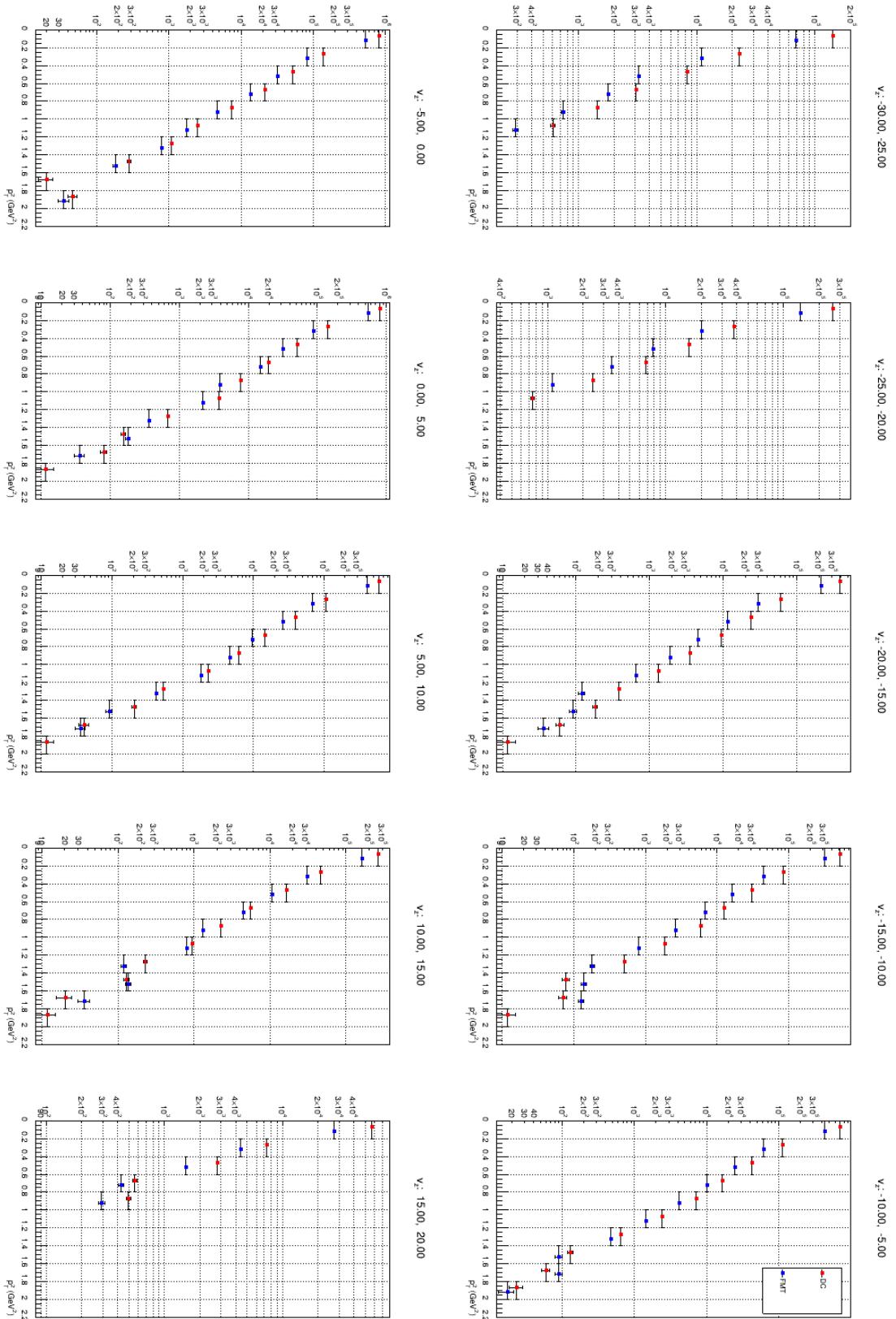


Figure 49: Acceptance-corrected p_T^2 for $e^- \pi^-$ detected by DC and FMT, separated in v_z bins. Run 12016. The bin markers are slightly shifted in x to improve legibility.

Source: Own elaboration, using the clas12-rge-analysis software.

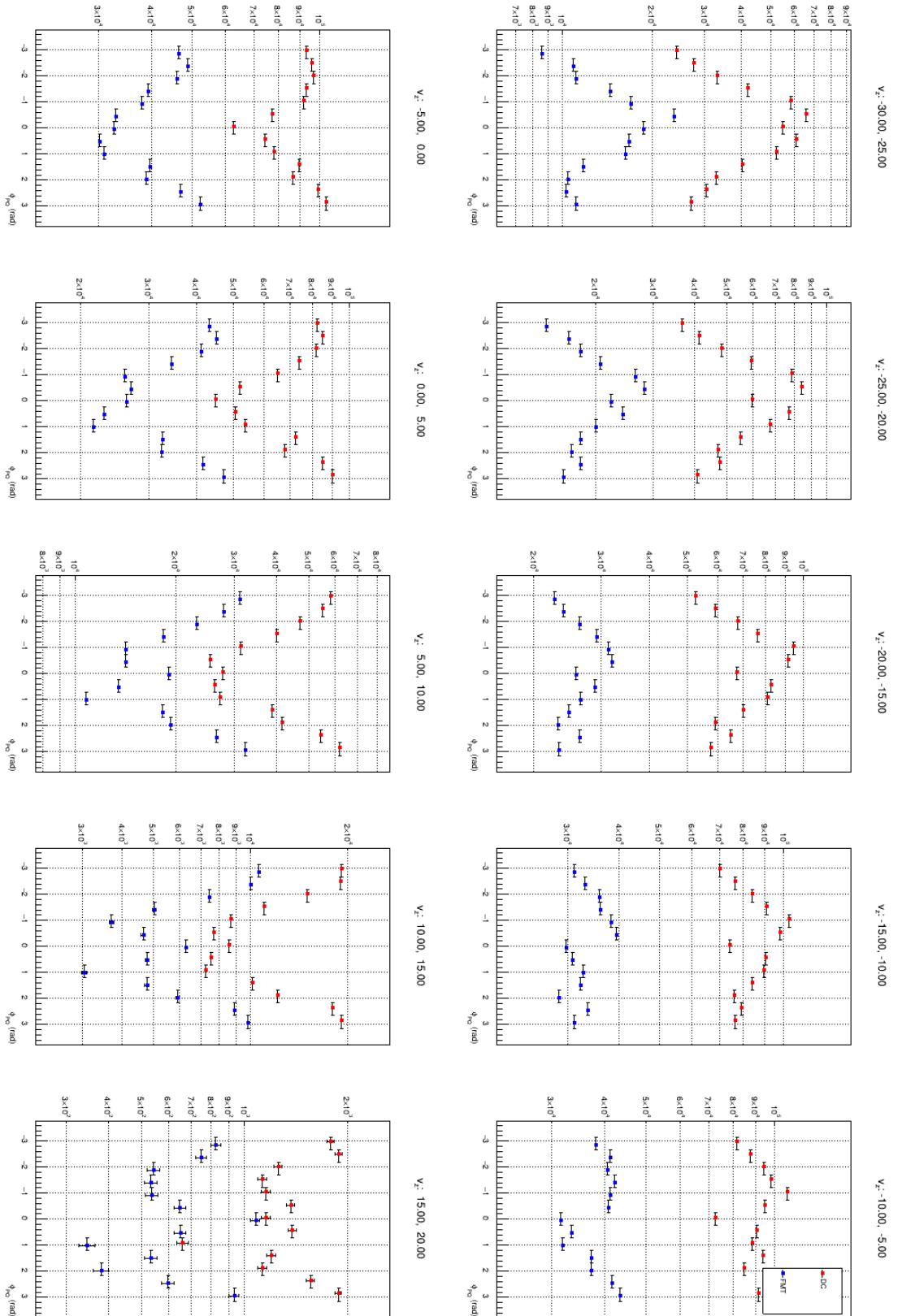


Figure 50: Acceptance-corrected ϕ_{PQ} for $e^- \pi^+$ detected by DC and FMT, separated in v_z bins. Run 12016. The bin markers are slightly shifted in x to improve legibility.

Source: Own elaboration, using the clas12-rge-analysis software.

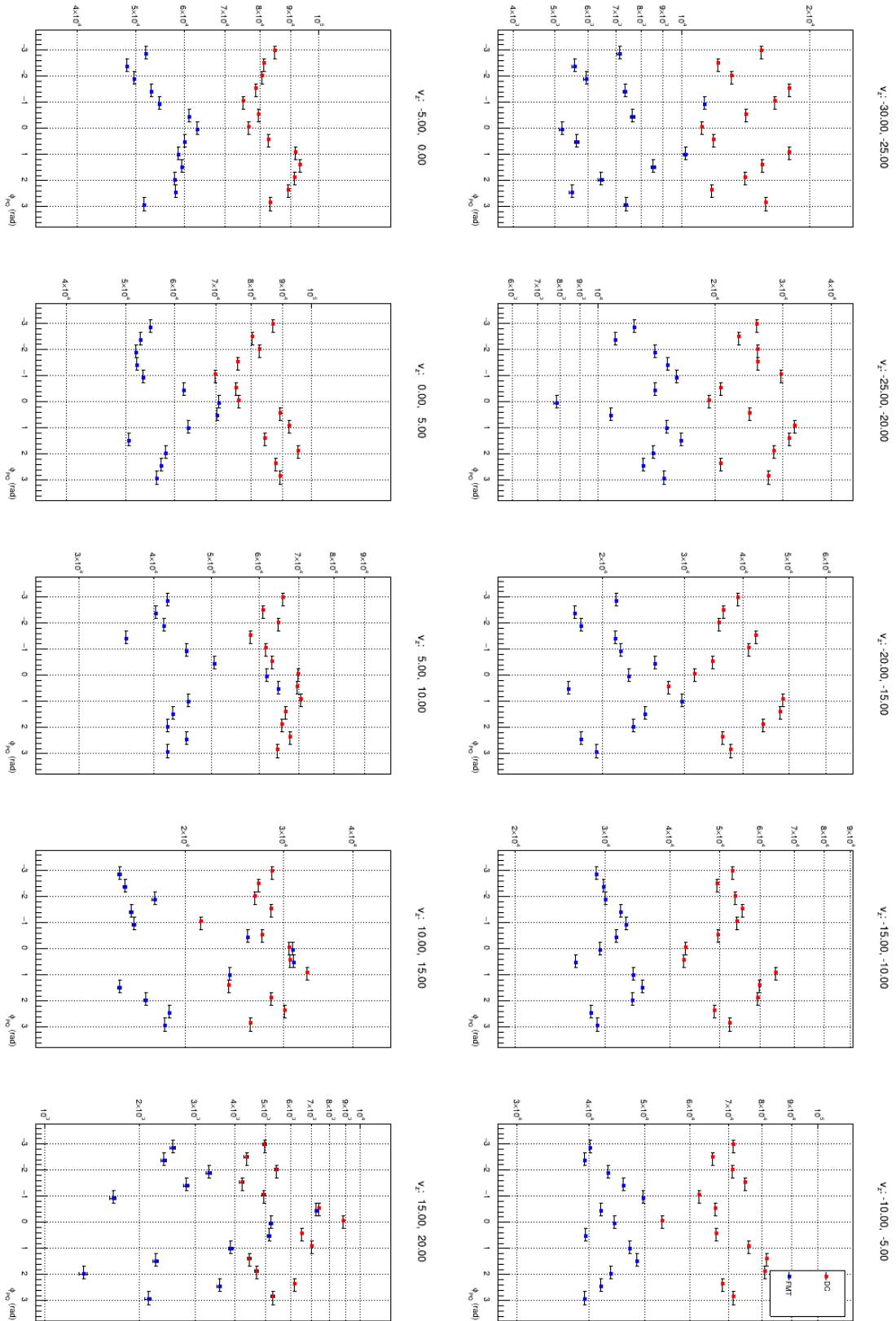
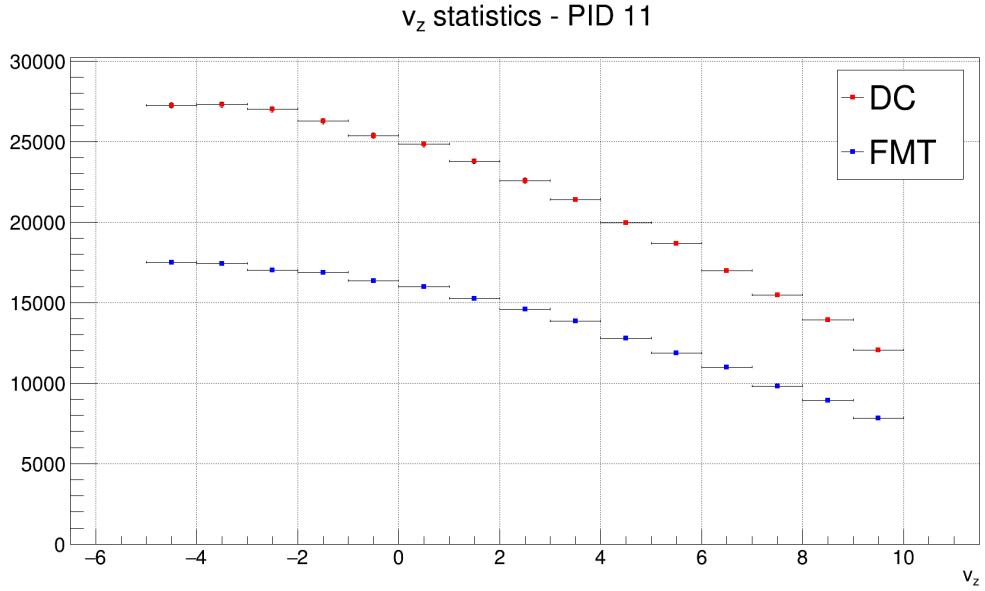
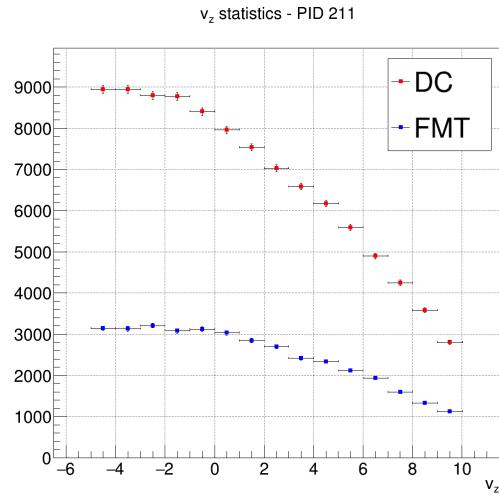


Figure 51: Acceptance-corrected ϕ_{PQ} for $e^- \pi^-$ detected by DC and FMT, separated in v_z bins. Run 12016. The bin markers are slightly shifted in x to improve legibility.

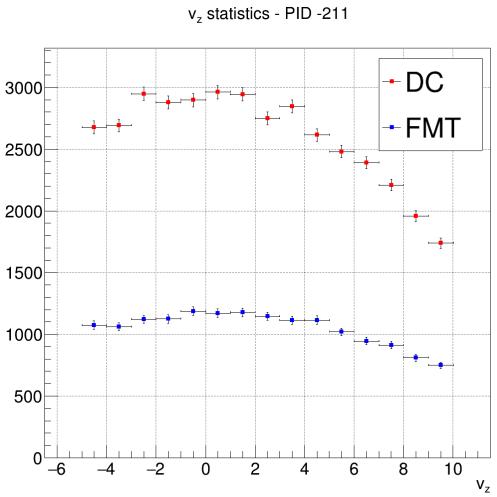
Source: Own elaboration, using the clas12-rge-analysis software.



(a) e^- statistics.



(b) $e^-\pi^+$ statistics.



(c) $e^-\pi^-$ statistics.

Figure 52: e^- , $e^-\pi^+$, and $e^-\pi^-$ statistics against v_z for run 12016. No acceptance correction was applied to these results.

Source: Own elaboration, using the clas12-rge-analysis software.

5.4 Systematic Error Estimation

As the attentive reader will notice, no systematic errors were applied to the results of the previous section. This is only due to the broadness of the analysis, and a more detailed one would, of course, need to include them. Based on the list provided in [Osipenko et al., 2010], the sources of systematic uncertainties that we estimate to be relevant to this study are the following:

- Previous CLAS DIS measurements [Osipenko et al., 2006] on deuterium targets showed that the combined efficiency-acceptance systematic uncertainty of CLAS data averages to 4.5%. We must be cautious when extrapolating the results of this study to CLAS12 data, but the error dependence on ν/E_b allows us to estimate that the higher beam energy of CLAS12 would likely reduce this uncertainty.
- As mentioned in Section 2.2.3, particle misidentification can be a significant source of systematic error. Based on the results presented in Table 1, we assume that the e^- misidentification uncertainty would be minor, while that for π^\pm would need to be studied with closer scrutiny. To estimate this systematic error, we would need to study the particle identification scheme as described in Section 2.2.3 and understand the accuracy of each detector involved.
- No radiative corrections were applied in this document, so we can assume that a significant systematic error related to radiative effects is left unestimated in the analysis. We hope that the radiative cut described in Section 4.2.3 ($Y_b < 0.85$) addresses most of this uncertainty.
- Bin migration can arise due to experimental resolutions and uncertainties in the measurement process. To estimate the systematic error that comes from this effect, we would need to create a migration matrix that describes the probabilities of events moving from their true bin to the reconstructed bin. For this, we could use the same simulation used for acceptance correction in Section 5.2. Based on this migration matrix, we would be able to apply an event reweighting, from which we could then provide a systematic error assessment.
- Systematic uncertainties can arise in the Monte Carlo simulation described in Section 4.4. To account for these, we would need to study the precision of LEPTO in predicting the cross sections relevant to DIS. Then, we would need to study how well GEMC accounts for the CLAS12 response to these cross sections, based on publications such as [Ungaro et al., 2020b].
- There is an uncertainty associated with the process of momentum correction, such as the one done for e^- via sampling fraction, described in Section 4.3. Based on the methodology and formulae described in that same section, we could derive a function to estimate the momentum-dependent systematic error inherent in the method.

After studying all these effects in detail, we would need to sum them in quadrature. We hope to be able to perform this study in future analyses that come from this document.

5.5 Conclusions

This thesis consisted of three subjects: the alignment of the FMT detector, the development of a standard analysis software for RG-E, and the selection of a location for the RG-E target.

Regarding the FMT alignment, we developed the standard FMT alignment software, which works by minimising the track residuals. Certain modifications to the FMT reconstruction software were also developed to integrate the alignment. This work resulted in an $e^- z$ resolution of $\sigma_{\text{FMT}} = 0.387$ cm for low luminosity (50 nA) runs and $\sigma_{\text{FMT}} = 0.596$ cm for high luminosity (250 nA) runs. A detailed rundown of the subject can be read in Section 3.

Regarding the standard analysis software, the entire toolset was developed as part of the work related to this thesis. The software works as expected. As evidence of this, all the analyses done in Sections 4.3, 4.4, and 5 were made using this same code. Most of the figures presented in this thesis were obtained using this software, as shown in their sources. Additionally, other preparatory analyses for the RG-E experiment have used the software to obtain their results. The software itself is explained in Section 4.1.

Using the code just described, we selected the best location for the RG-E target based on phase space and statistics studies. For the phase space study, the range of two electron variables – Q^2 and ν – and three hadronic variables – z_h , p_T^2 , and ϕ_{PQ} – was studied for ten 5 cm bins in v_z . A range from –5 to 10 cm was obtained from this study, based on maximising the phase spaces. Then, for the statistic study, a 7 cm region was chosen in this range based on maximising the number of events. The region chosen was from –5 to 2 cm.

In order to improve the quality of the results, procedures for sampling fraction estimation and detector acceptance correction were proposed and applied. Both were done successfully, and the statistical error from the acceptance correction process was summarily propagated to the results. The application of the former is detailed in Section 4.3, and that of the latter in Sections 4.4 and 5.2.

In addition to the described analyses, a detector efficiency study was conducted on the FMT. The three main factors in this efficiency were found and studied: one is related to the detector alignment, one to its geometry, and one to its offline reconstruction. A geometry cut was derived from the second, contributing to the accuracy of the entire study presented in this document. The process and results are described in Section 5.1.

Appendices

A Reproducibility

In order to ensure the reproducibility of the research presented in this thesis, we provide access to all datasets and code used in the development of our study. We believe that transparency and accessibility are crucial for scientific integrity and to facilitate further investigations by the research community.

We encourage interested readers and fellow researchers to access and utilise these resources for the purpose of reproducibility and advancing scientific knowledge. Should there be any inquiries or issues regarding the datasets or code, please do not hesitate to contact the author at bruno.benkel@gmail.com for further assistance.

We believe that open access to data and code fosters collaboration, accelerates scientific progress, and ensures the robustness of research findings. By making these resources available, we aim to contribute to the collective effort of reproducible and transparent scientific research.

Datasets

Regrettably, there is no website or location to openly share datasets in the *Universidad Técnica Federico Santa María* (UTFSM) or the *Centro Científico Tecnológico de Valparaíso* (CCTVal). For readers with access to the JLab farm, all used datasets are available at

```
/work/clas12/users/benkel/thesis-datasets
```

For individuals who do not have access to the JLab farm, please feel free to contact the author, and we will explore alternative methods to share the relevant datasets.

Code

The sources for the code used for data processing, analysis, and generating figures are shared on Table 7. By providing the code, we aim to enable researchers to replicate our findings, perform additional analyses, or build upon our work.

Software	Link
RG-E Slow Controls	github.com/bleaktwig/rge-epics-support
CLAS12 Alignment	github.com/JeffersonLab/clas12alignment
thesis-simul	github.com/bleaktwig/thesis-simul
thesis-data	github.com/bleaktwig/thesis-data
RG-E Analysis	github.com/bleaktwig/clas12-rge-analysis
GEMC	github.com/gemc/source
Coatjava	github.com/JeffersonLab/coatjava

Table 7: Table with code locations.

B Fiducial Cuts

In detector physics, the fiducial region is defined as the region considered reliable and suitable for analysis. Fiducial cuts are constraints applied to experimental data in order to define this region. Thus, they allow us to exclude events or measurements that may be affected by experimental artefacts, detector inefficiencies, or other factors that could introduce systematic errors and biases [Leo, 1987].

Due to its 6-sector geometry, fiducial cuts are of particular importance for CLAS12 FD analysis. However, they were disregarded for this particular study. This is because of its broadness: we are only concerned with the phase space of DIS variables and the general statistics, as detailed in Section 5.3. While the cuts would likely improve the quality of the results, the data is broad enough to be considered resilient to the damage of not applying them.

To apply such cuts, we would need to follow the procedure described in [Zana, 2010]. This would involve providing ϕ vs. θ curves that cut all events at the edges of each DC sector. One curve would be needed for each p bin, for each sector. Finally, different curves would be needed for each PID being processed.

Examples of ϕ vs. θ distributions for different p bins can be seen in Figures 53, 54a, and 54b, where we show the distributions for e^- , π^+ , and π^- . As can be seen in the plots, the separation between the DC's areas and its edges are easily observed in a layer-by-layer basis.

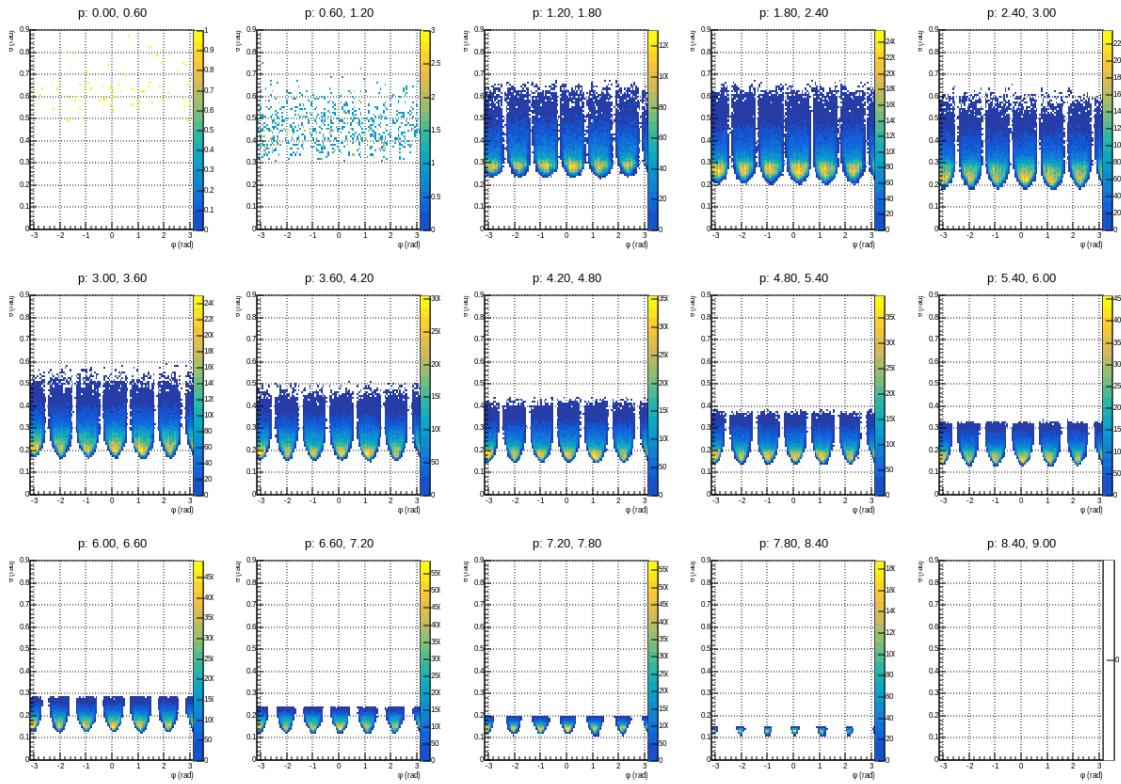


Figure 53: ϕ vs. θ of e^- detected by DC, separated in p bins. Run 12016.

Source: Own elaboration, using the clas12-rge-analysis software.

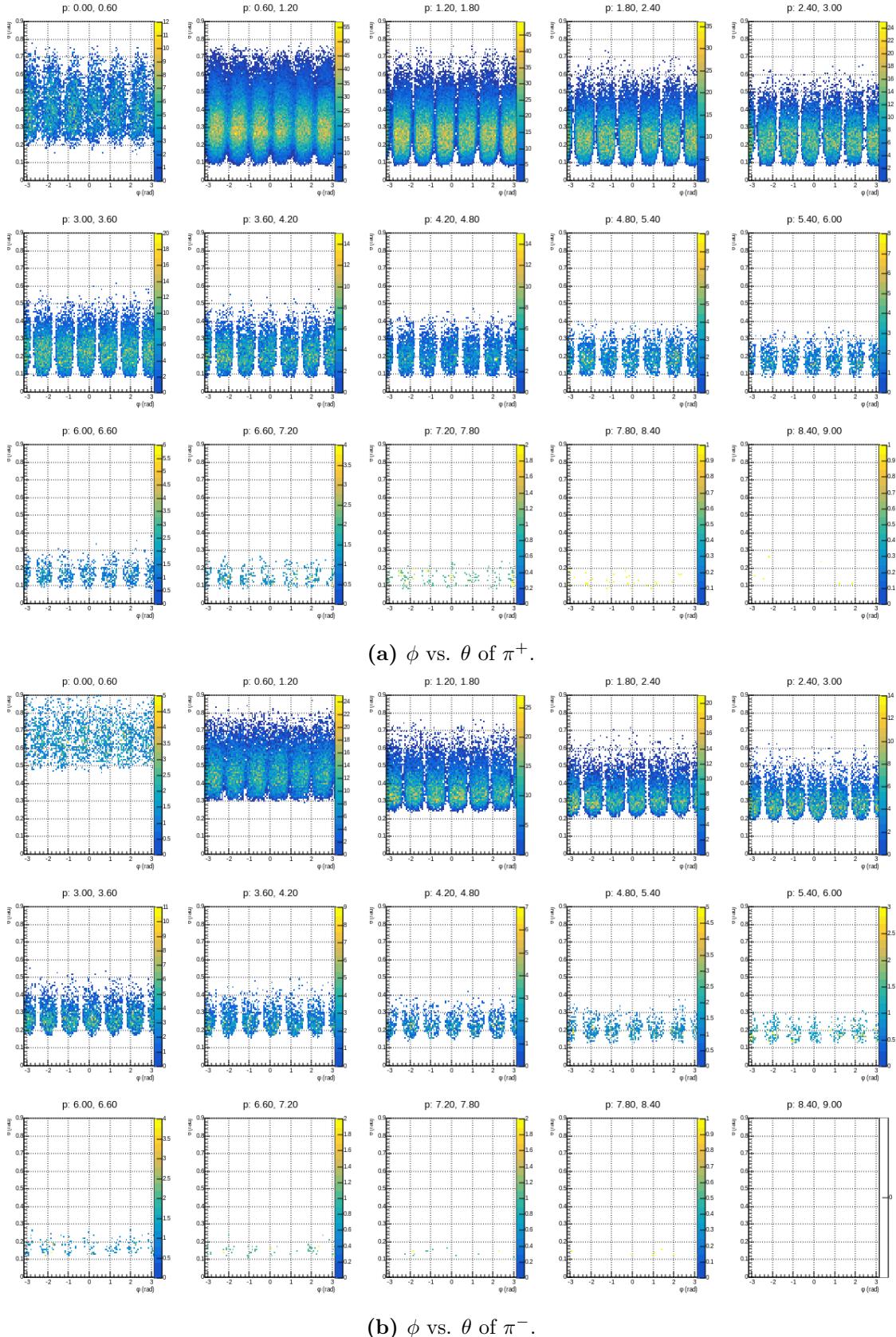


Figure 54: ϕ vs. θ of π^+ and π^- detected by DC, separated in p bins. Run 12016.

Source: Own elaboration, using the clas12-rge-analysis software.

C FMT Layer Efficiency Error Estimation

This Appendix presents the Python script described in Section 5.1.4. The script utilises the formulae outlined in that section to estimate the errors in the efficiency of the 2-layer and 3-layer tracks. It should be noted, as mentioned in the section, that the efficiency $E_{1(2)}$ (referred to as E_{12} in the code) was obtained numerically.

```

import sys
def printf(format, *args):
    sys.stdout.write(format % args)
def f_E13(E3):
    return E3**1/3
def f_E1(E12, E13):
    return (4*E12 + E13)/5
def f_dE12(E1, E12):
    return abs(E1 - E12)
def f_dE13(E1, E13):
    return abs(E1 - E13)
def f_dE2(E12, dE12):
    return (6*E12 - 9*E12**2) * dE12
def f_dE3(E13, dE13):
    return 3*E13**2 * dE13

# Input.
# Run 12933. Run 12016.
E2 = [.251,.065,.056,.375,.137,.142,.327,.111,.089,.537,.280,.295]
E3 = [.056,.003,.003,.085,.007,.007,.099,.010,.009,.164,.027,.029]
E12 = [.306,.157,.144,.402,.239,.244,.339,.206,.180,.497,.364,.377]

# Run functions.
E13 = list(map(f_E13, E3))
E1 = list(map(f_E1, E12, E13))
dE12 = list(map(f_dE12, E1, E12))
dE13 = list(map(f_dE13, E1, E13))
dE2 = list(map(f_dE2, E12, dE12))
dE3 = list(map(f_dE3, E13, dE13))

# Print.
print("dE2:")
for i in dE2:
    printf("%5.1f,", 100*i)
print("dE3:")
for i in dE3:
    printf("%5.1f,", 100*i)

```

D DIS plots in v_z bins

In Section 5.3.1, we presented the acceptance-corrected DIS variables separated into v_z bins. In this Appendix, we provide the same distributions without applying the acceptance correction. The statistics are considerably lower in v_z bins characterised by low acceptance, such as $v_z > 10$ cm. This effect is more pronounced for the hadronic variables, as one would expect. Moreover, the correction noticeably alters the shape of certain distributions, bringing them closer to the expected theoretical behaviour. A clear illustration of this can be observed in the disparity of Q^2 depicted in Figures 44 and 55, as well as in the dissimilarity of z_h demonstrated in Figures 46 and 57.

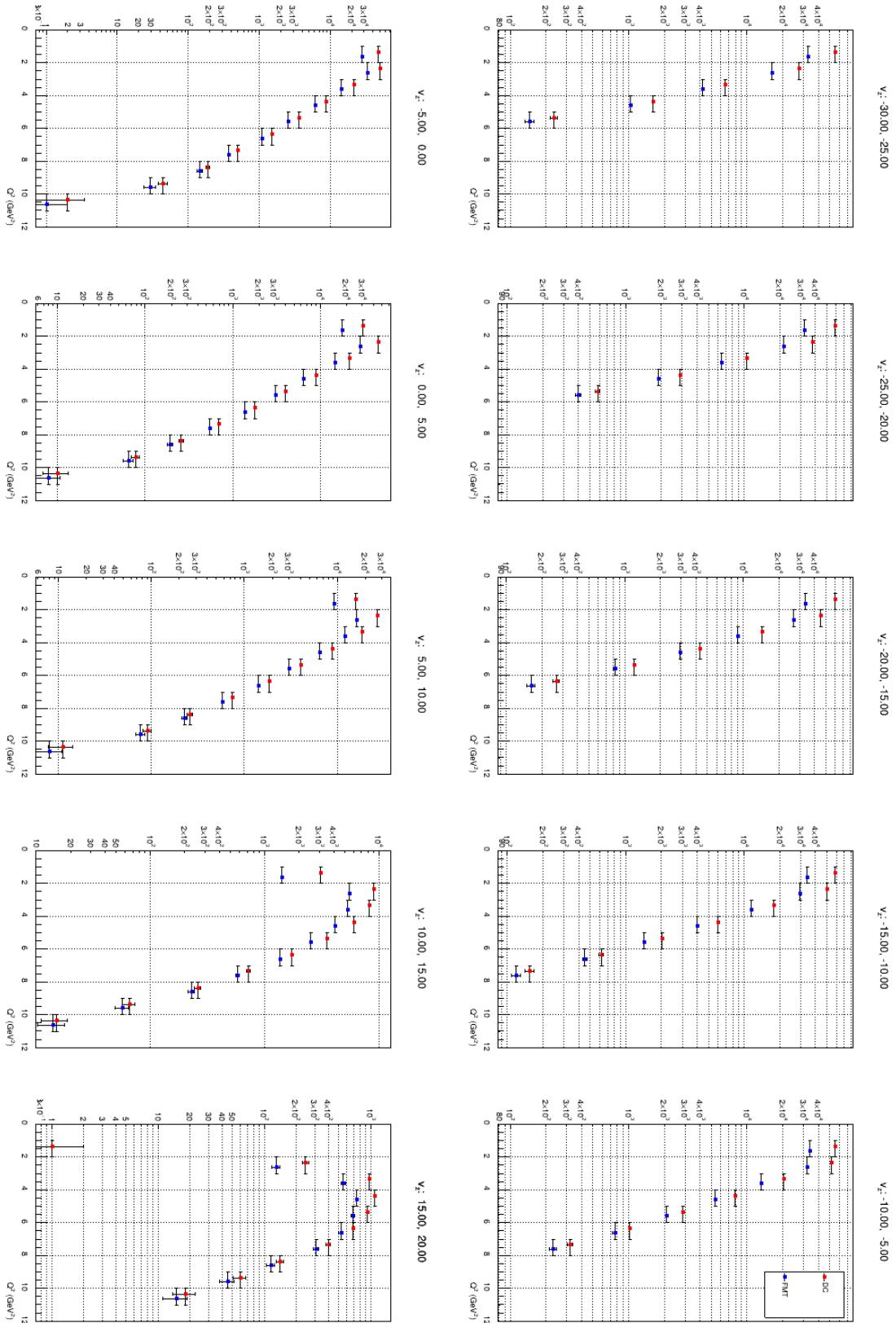


Figure 55: Q^2 detected by DC and FMT, separated in v_z bins. Run 12016. The bin markers are slightly shifted in x to improve legibility.

Source: Own elaboration, using the clas12-rge-analysis software.

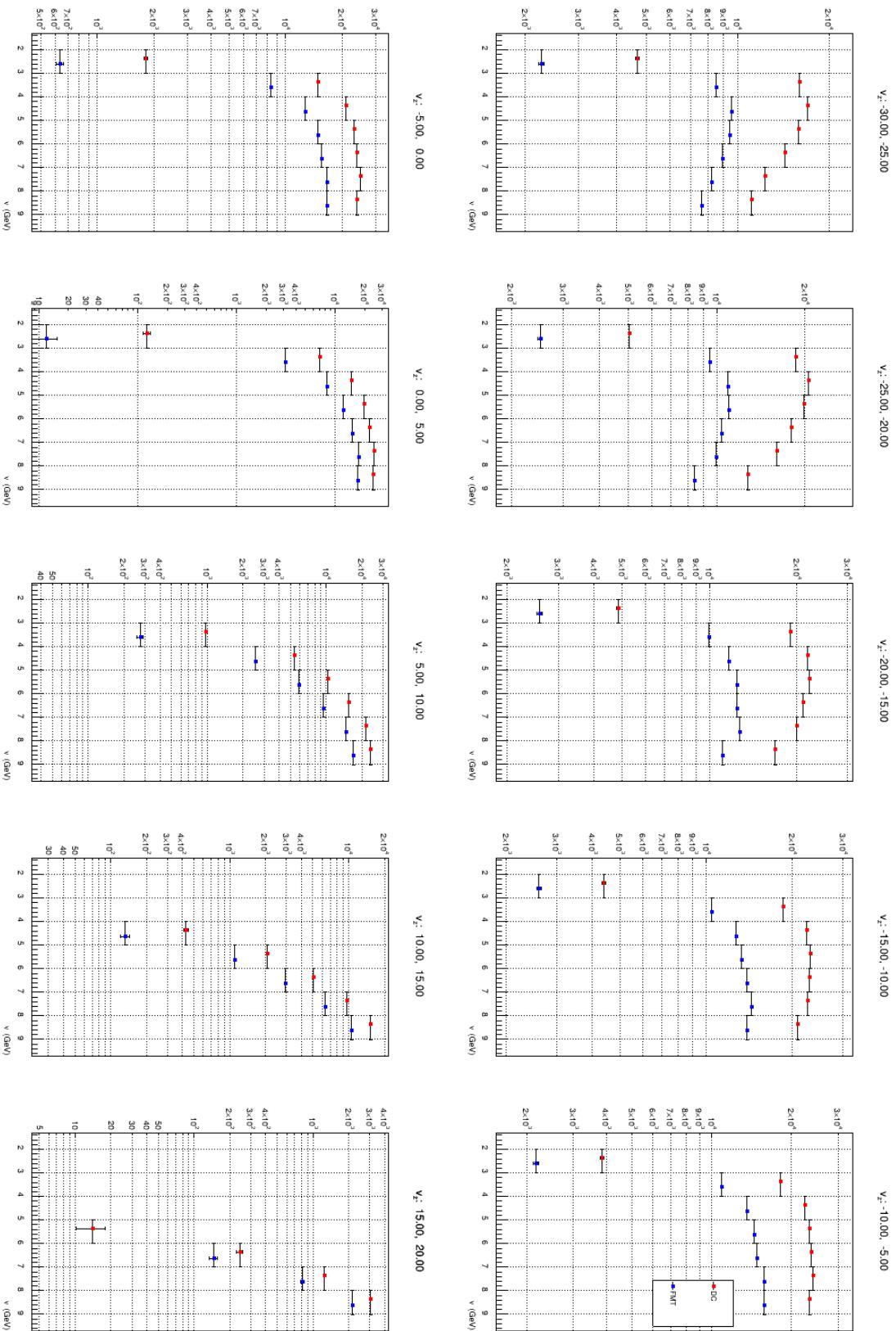


Figure 56: ν detected by DC and FMT, separated in v_z bins. Run 12016. The bin markers are slightly shifted in x to improve legibility.

Source: Own elaboration, using the clas12-rge-analysis software.

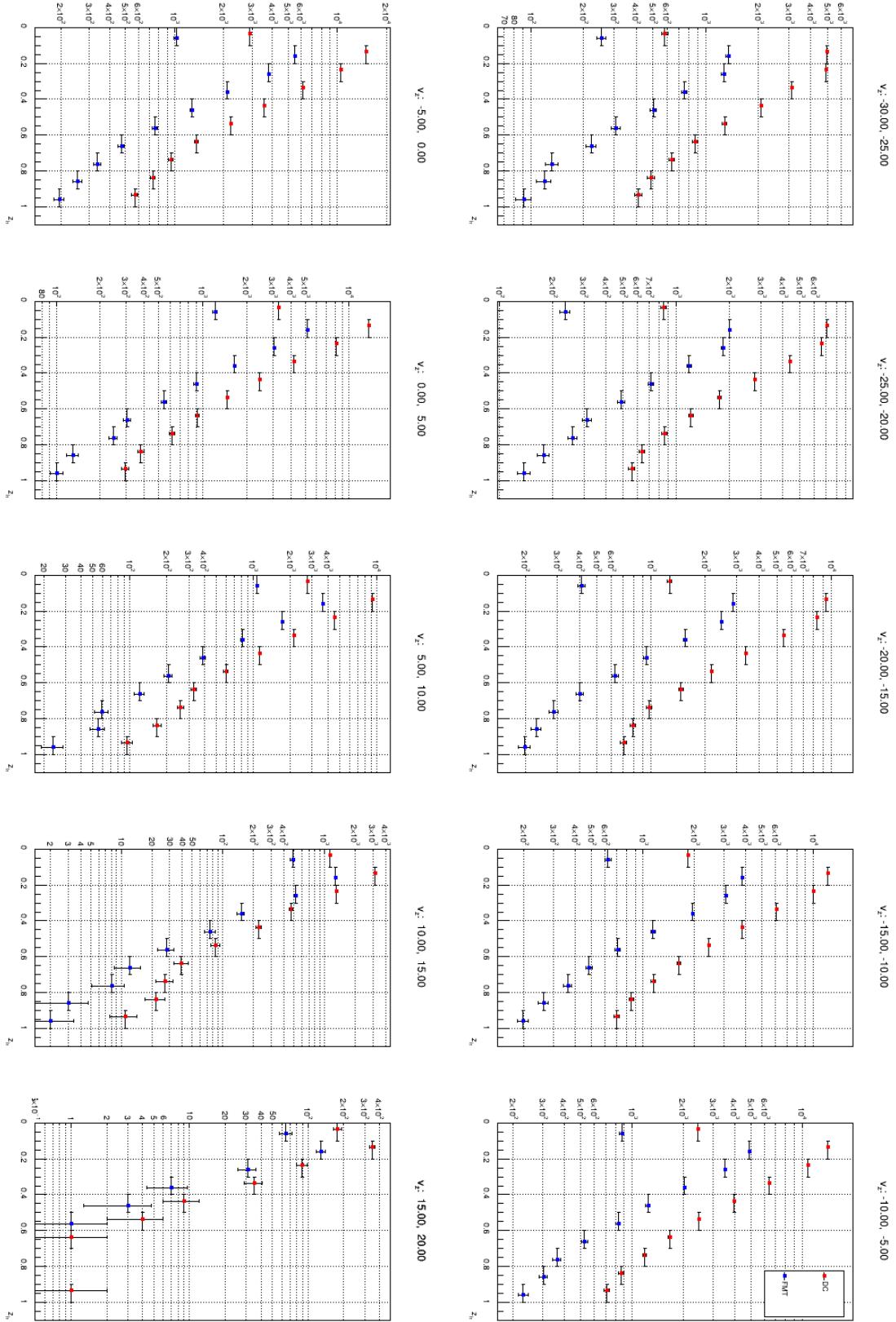


Figure 57: z_h for $e^- \pi^+$ detected by DC and FMT, separated in v_z bins. Run 12016. The bin markers are slightly shifted in x to improve legibility.

Source: Own elaboration, using the clas12-rge-analysis software.

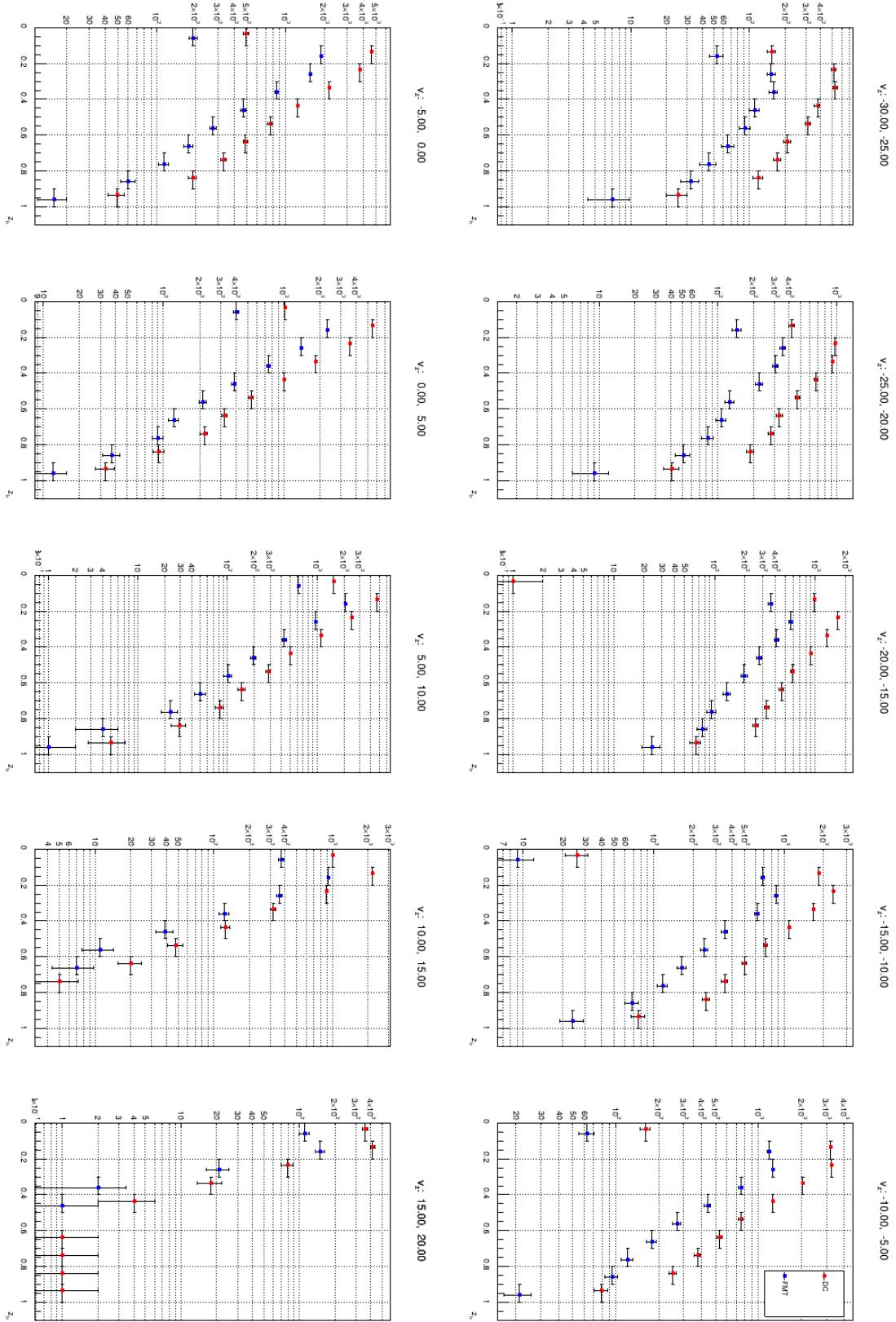


Figure 58: z_h for $e^- \pi^-$ detected by DC and FMT, separated in v_z bins. Run 12016. The bin markers are slightly shifted in x to improve legibility.

Source: Own elaboration, using the clas12-rge-analysis software.

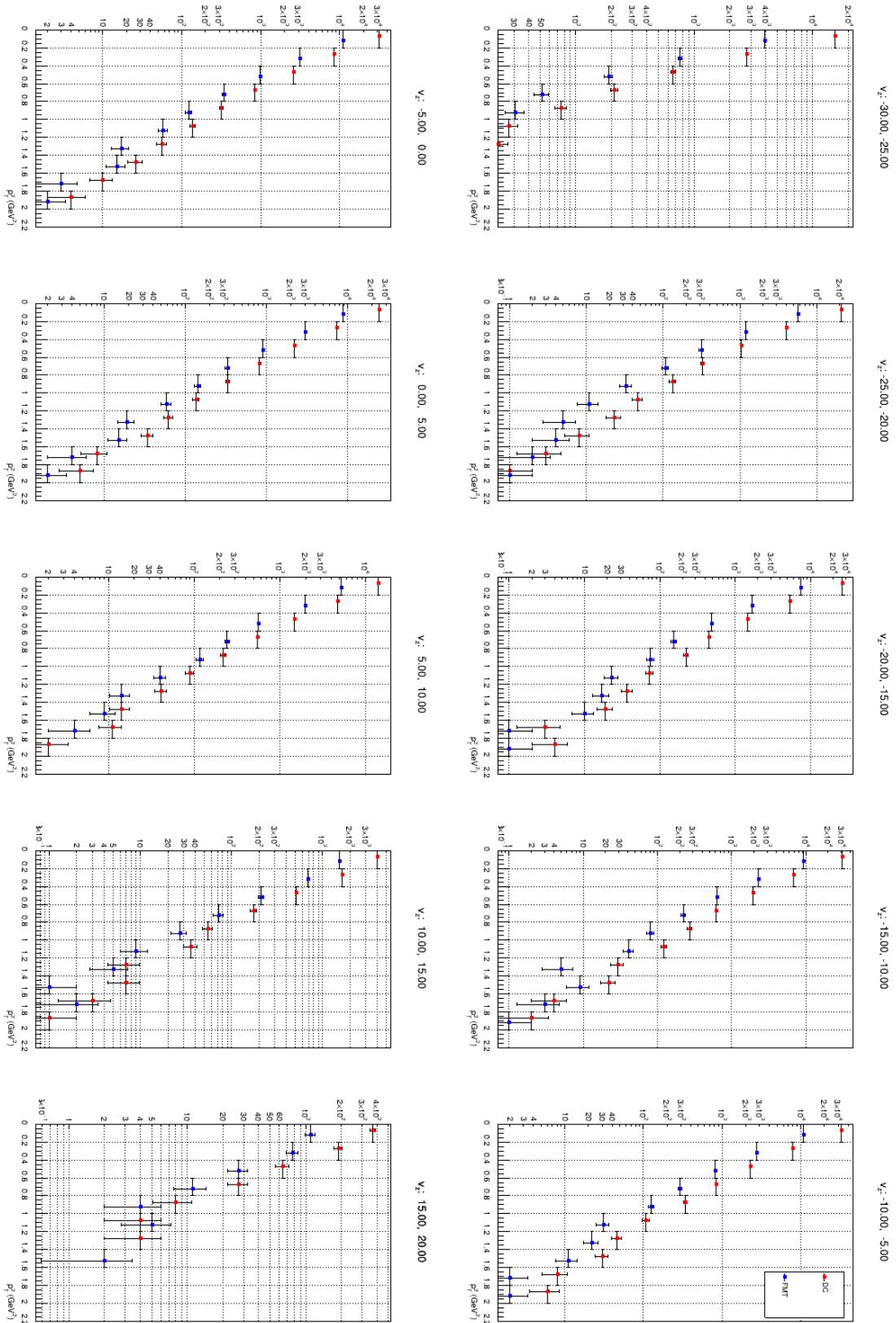


Figure 59: p_T^2 for $e^- \pi^+$ detected by DC and FMT, separated in v_z bins. Run 12016. The bin markers are slightly shifted in x to improve legibility.

Source: Own elaboration, using the clas12-rge-analysis software.

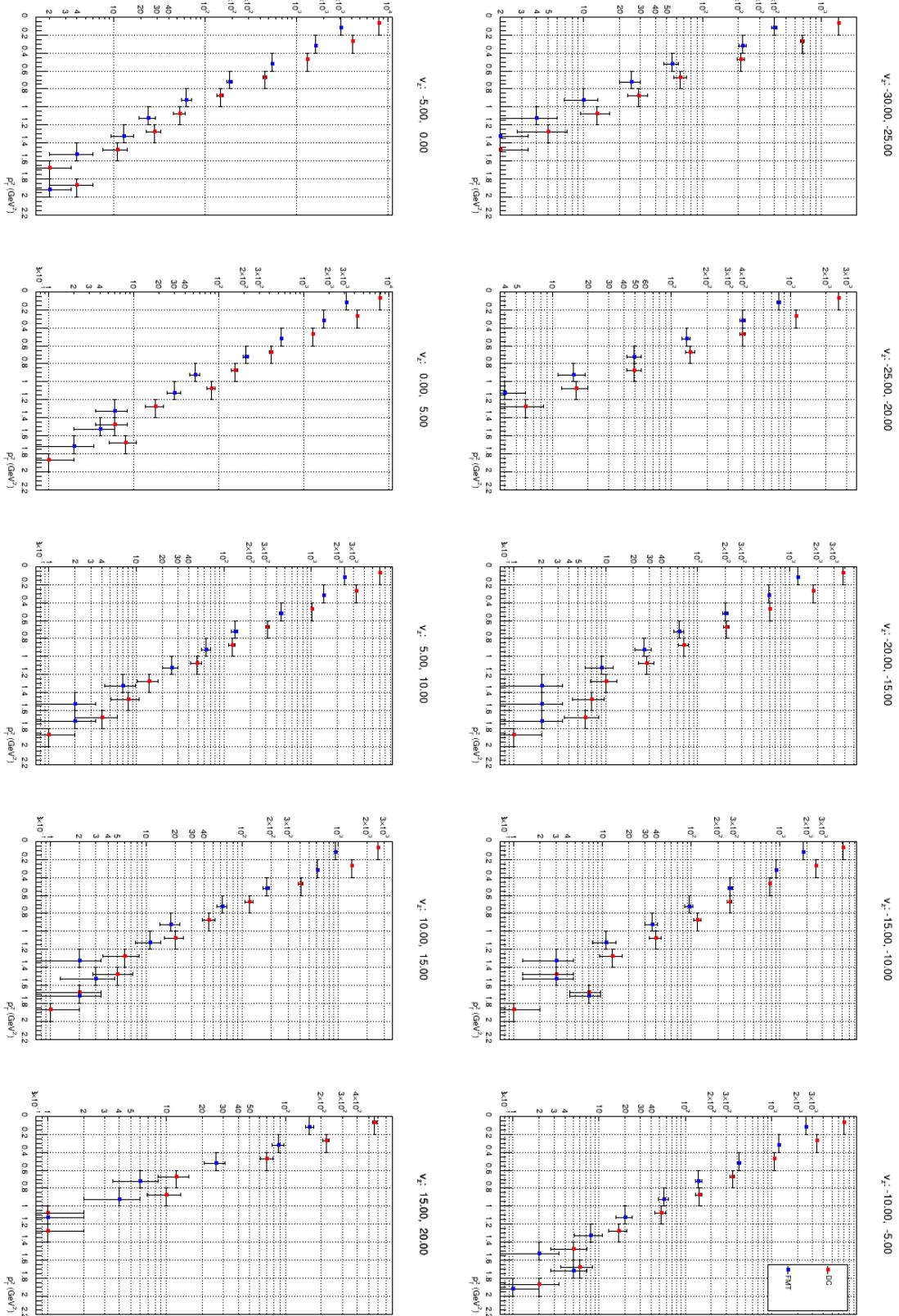


Figure 60: p_T^2 for $e^- \pi^-$ detected by DC and FMT, separated in v_z bins. Run 12016. The bin markers are slightly shifted in x to improve legibility.

Source: Own elaboration, using the clas12-rge-analysis software.

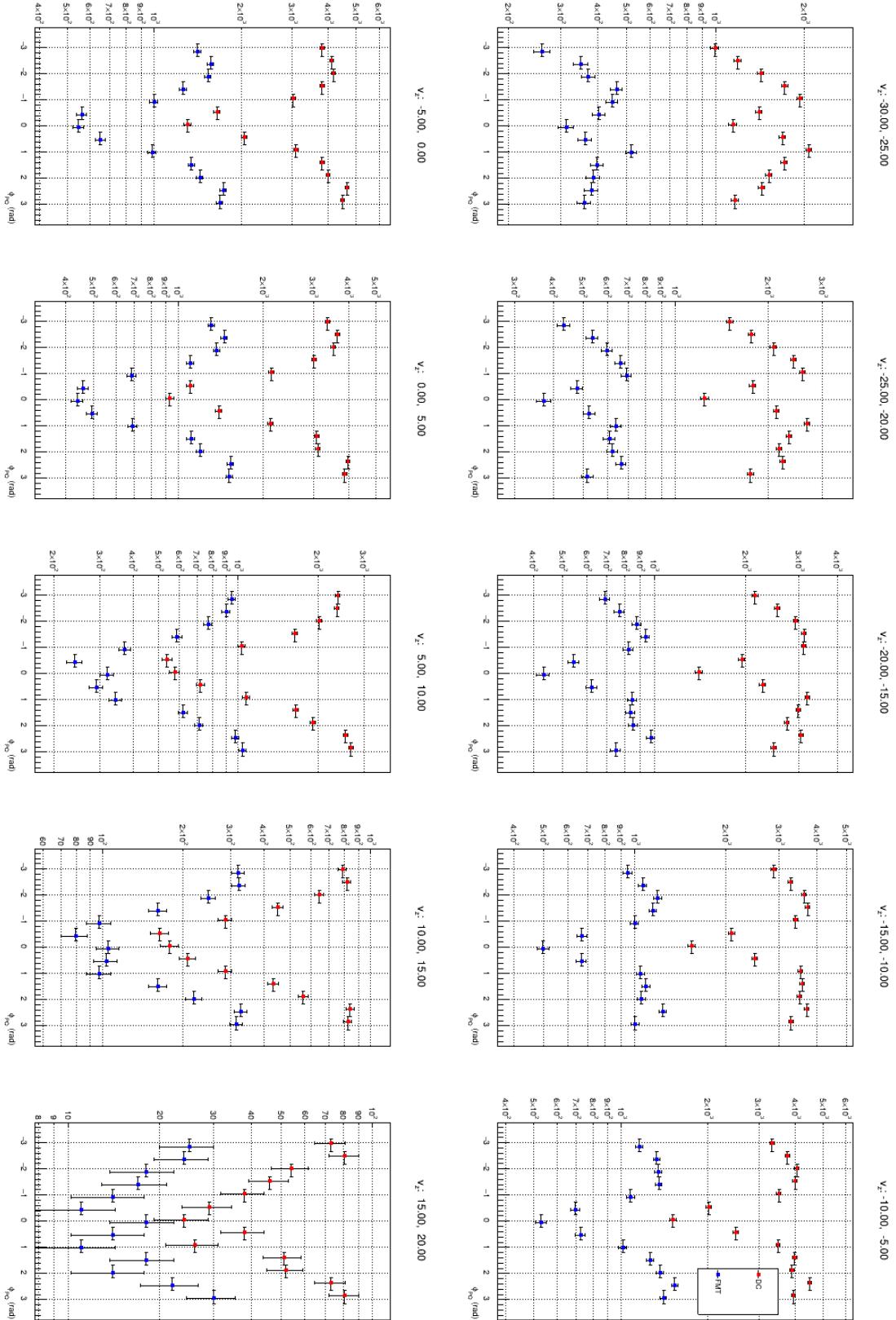


Figure 61: ϕ_{PQ} for $e^- \pi^+$ detected by DC and FMT, separated in v_z bins. Run 12016. The bin markers are slightly shifted in x to improve legibility.

Source: Own elaboration, using the clas12-rge-analysis software.

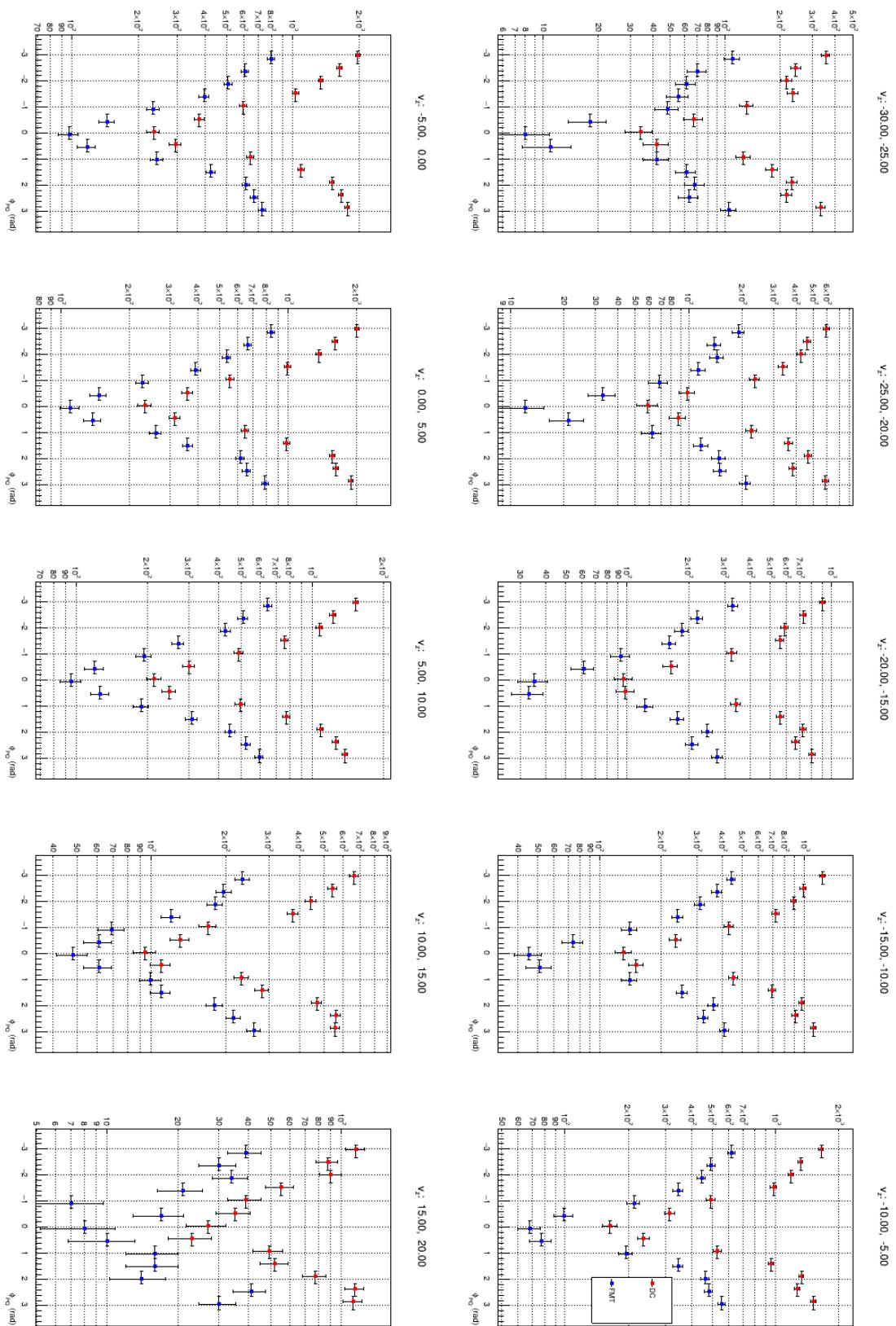
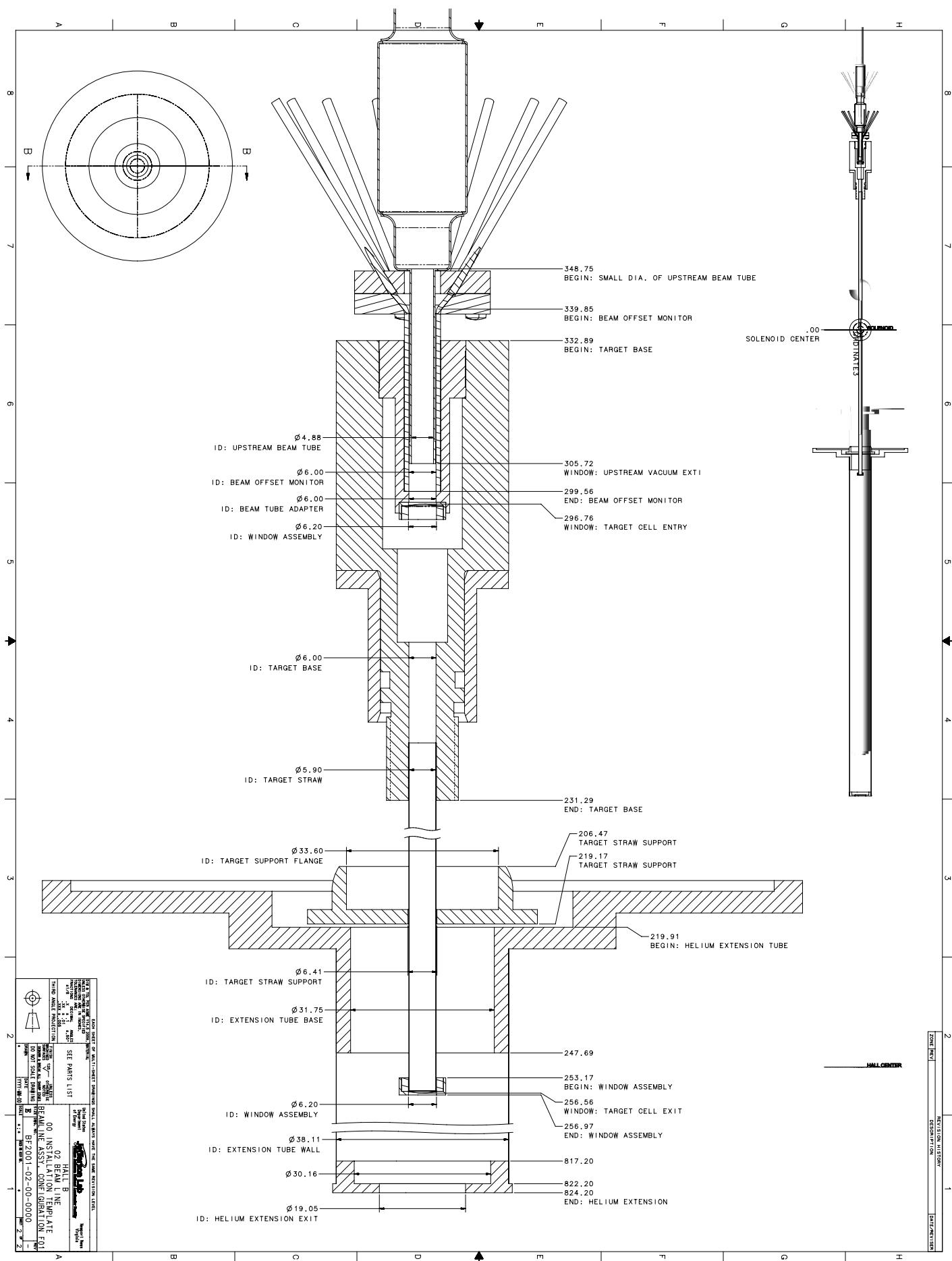


Figure 62: ϕ_{PQ} for $e^- \pi^-$ detected by DC and FMT, separated in v_z bins. Run 12016. The bin markers are slightly shifted in x to improve legibility.

Source: Own elaboration, using the clas12-rge-analysis software.

E: RG-F Target Layout



References

- [Aad et al., 2012] Aad, G., Abajyan, T., Abbott, B., Abdallah, J., Khalek, S., Abdelalim, A., Abdinov, O., Aben, R., Abi, B., Abolins, M., et al. (2012). Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*.
- [Acker et al., 2020a] Acker, A., Attié, D., Aune, S., Ball, J., Baron, P., Bashkanov, M., Battaglieri, M., Behary, R., Benmokhtar, F., Bersani, A., et al. (2020a). The clas12 forward tagger. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*.
- [Acker et al., 2020b] Acker, A., Attié, D., Aune, S., Ball, J., Baron, P., Bertrand, Q., Besin, D., Bey, T., Bossù, F., Boudouin, R., et al. (2020b). The clas12 micromegas vertex tracker. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*.
- [Antonioli et al., 2020] Antonioli, M., Baltzell, N., Boyarinov, S., Bonneau, P., Christo, S., Cuevas, C., Defurne, M., Derylo, G., Elouadrhiri, L., Eng, B., et al. (2020). The clas12 silicon vertex tracker. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*.
- [Artru and Mennessier, 1974] Artru, X. and Mennessier, G. (1974). String model and multiproduction. *Nuclear Physics*.
- [Asryan et al., 2020] Asryan, G., Chandavar, S., Chetry, T., Compton, N., Daniel, A., Dashyan, N., Gevorgyan, N., Ghandilyan, Y., Giovanetti, K., Griffioen, K., et al. (2020). The clas12 forward electromagnetic calorimeter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*.
- [Aune et al., 2009] Aune, S., Ball, J., Combet, M., El Yakoubi, M., Konczykowski, P., Lahonde-Hamdoun, C., Meunier, O., Procureur, S., and Sabatié, F. (2009). Micromegas tracker project for clas12. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*.
- [Boyarinov et al., 2020] Boyarinov, S., Raydo, B., Cuevas, C., Dickover, C., Dong, H., Heyes, G., Abbott, D., Gu, W., Gyurjyan, V., Jastrzembski, E., et al. (2020). The clas12 data acquisition system. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*.
- [Burkert et al., 2020] Burkert, V., Elouadrhiri, L., Adhikari, K., Adhikari, S., Amaryan, M., Anderson, D., Angelini, G., Antonioli, M., Atac, H., Aune, S., et al. (2020). The clas12 spectrometer at jefferson laboratory. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*.
- [Carman et al., 2020a] Carman, D., Asryan, G., Baturin, V., Clark, L., De Vita, R., Kim, W., Miller, B., and Wiggins, C. (2020a). The clas12 central time-of-flight system. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*.
- [Carman et al., 2020b] Carman, D., Clark, L., De Vita, R., Fedotov, G., Gothe, R., Hollis, G., Miller, B., Phelps, E., Tian, Y., Trivedi, A., et al. (2020b). The clas12 forward time-of-flight system. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*.

- [Chatagnon et al., 2020] Chatagnon, P., Bettane, J., Hoballah, M., Hull, G., Imre, M., Marchand, D., Mathon, B., Murdoch, G., Naidoo, P., Niccolai, S., et al. (2020). The clas12 central neutron detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*.
- [Chekanov et al., 2021] Chekanov, S., Gavalian, G., and Graf, N. (2021). Jas4pp—a data-analysis framework for physics and detector studies. *Computer Physics Communications*.
- [Contalbrigo et al., 2020] Contalbrigo, M., Kubarovskiy, V., Mirazita, M., Rossi, P., Angelini, G., Avakian, H., Bailey, K., Balossino, I., Barion, L., Benmokhtar, F., et al. (2020). The clas12 ring imaging cherenkov detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*.
- [Dalesio et al., 1991] Dalesio, L., Kozubal, A., and Kraimer, M. (1991). Epics architecture. Technical report, Los Alamos National Lab., NM (United States).
- [Dokshitzer et al., 1991] Dokshitzer, Y., Khoze, V., Mueller, A., and Troyan, S. (1991). Basics of perturbative qcd. *Gif-sur Yvette*.
- [Farnsworth, 2009] Farnsworth, R. (2009). Motion controls using epics and galil controllers.
- [Giomataris et al., 1996] Giomataris, Y., Rebourgeard, P., Robert, J., and Charpak, G. (1996). Micromegas: a high-granularity position-sensitive gaseous detector for high particle-flux environments. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*.
- [Group et al., 1998] Group, P. D. et al. (1998). Review of particle physics. *The European Physical Journal C-Particles and Fields*.
- [Group et al., 2020] Group, P. D., Zyla, P., Barnett, R., Beringer, J., Dahl, O., Dwyer, D., Groom, D., Lin, C., Lugovsky, K., Pianori, E., et al. (2020). Review of particle physics. *Progress of Theoretical and Experimental Physics*.
- [Gyurgyan et al., 2016] Gyurgyan, V., Mancilla, S., and Oyarzun, R. (2016). Clara: Clas12 reconstruction and analysis framework. In *Journal of Physics: Conference Series*.
- [Hakobyan et al., 2008] Hakobyan, H. et al. (2008). A double-target system for precision measurements of nuclear medium effects. *Nucl. Instrum. Meth. A*.
- [Halzen and Martin, 1991] Halzen, F. and Martin, A. (1991). *Quarks and leptons: an introductory course in modern particle physics*.
- [Hattawy et al., 2019] Hattawy, M., Amaryan, M., Bültmann, S., Dodge, G., Dzbenski, N., Hyde, C., Kuhn, S., Payette, D., Poudel, J., Weinstein, L., et al. (2019). Neutron dvcs measurements with bonus12 in clas12. *arXiv preprint arXiv:1908.00949*.
- [Ingelman et al., 1997] Ingelman, G., Edin, A., and Rathsman, J. (1997). Lepto 6.5—a monte carlo generator for deep inelastic lepton-nucleon scattering. *Computer Physics Communications*.

- [Kasemir et al., 2007] Kasemir, K. et al. (2007). Control system studio applications. In *Proc. 11th Int. Conf. on Accelerator and Large Experimental Physics Control Systems (ICALEPCS'07)*.
- [Konczykowski et al., 2010] Konczykowski, P., Aune, S., Ball, J., El Yakoubi, M., Lahonde-Hamdoun, C., Meunier, O., Procureur, S., Sabatié, F., Cazaux, S., Delagnes, E., et al. (2010). Measurements of the lorentz angle with a micromegas detector in high transverse magnetic fields. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*.
- [Kopeliovich et al., 2004] Kopeliovich, B., Nemchik, J., Hayashigaki, A., and Predazzi, E. (2004). Nuclear hadronization: within or without? *Nuclear Physics A*.
- [Leemann et al., 2001] Leemann, C., Douglas, D., and Krafft, G. (2001). The continuous electron beam accelerator facility: Cebaf at the jefferson laboratory. *Annual Review of Nuclear and Particle Science*.
- [Leo, 1987] Leo, W. (1987). *Techniques for nuclear and particle physics experiments: a how to approach*.
- [Mestayer et al., 2020] Mestayer, M., Adhikari, K., Bennett, R., Bueltmann, S., Chetry, T., Christo, S., Cook, M., Cuevas, R., Dodge, G., El Fassi, L., et al. (2020). The clas12 drift chamber system. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*.
- [Morán et al., 2022] Morán, S., Dupre, R., Hakobyan, H., Arratia, M., Brooks, W., Bórquez, A., El Alaoui, A., El Fassi, L., Hafidi, K., Mendez, R., Mineeva, T., Paul, S., et al. (2022). Measurement of charged-pion production in deep-inelastic scattering off nuclei with the clas detector. *Phys. Rev. C*.
- [Osipenko et al., 2006] Osipenko, M., Ricco, G., Simula, S., Battaglieri, M., Ripani, M., et al. (2006). Measurement of the deuteron structure function F_2 in the resonance region and evaluation of its moments. *Phys. Rev. C*.
- [Osipenko et al., 2010] Osipenko, M., Ricco, G., Simula, S., Ripani, M., Taiuti, M., Adhikari, K., Amaryan, M., Anghinolfi, M., Avakian, H., Baghdasaryan, H., et al. (2010). Measurement of the nucleon structure function f_2 in the nuclear medium and evaluation of its moments. *Nuclear Physics A*.
- [Perkins, 2000] Perkins, D. (2000). *Introduction to high energy physics*.
- [Rode and Team, 2010] Rode, C. and Team, J. . G. P. (2010). Jefferson lab 12 gev cebaf upgrade. In *AIP Conference Proceedings*.
- [Segarra et al., 2020] Segarra, E., Hauenstein, F., Schmidt, A., Price, K., Vega, I., Barlow, J., Barriga, E., Beck, A., Brooks, W., Cohen, E., et al. (2020). The clas12 backward angle neutron detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*.
- [Stanley et al., 1998] Stanley, P., Anderson, J., and Kraimer, M. (1998). Epics record reference manual. Technical report, The University of California, The University of Chicago.

[Ungaro et al., 2020a] Ungaro, M., Anderson, D., Asryan, G., Antonioli, M., Bonneau, P., Burkert, V., Christo, S., Cook, M., Duran, B., Eng, B., et al. (2020a). The clas12 low threshold cherenkov detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*.

[Ungaro et al., 2020b] Ungaro, M., Angelini, G., Battaglieri, M., Burkert, V., Carman, D., Chatagnon, P., Contalbrigo, M., Defurne, M., De Vita, R., Duran, B., et al. (2020b). The clas12 geant4 simulation. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*.

[Wigmans, 2000] Wigmans, R. (2000). *Calorimetry: energy measurement in particle physics*.

[Yost et al., 1988] Yost, G., Barnett, R., Hinchliffe, I., Lynch, G., Rittenberg, A., Ross, R., Suzuki, M., Trippe, T., Wohl, C., Armstrong, B., et al. (1988). Review of particle properties. *Physics Letters B*.

[Zana, 2010] Zana, L. (2010). *Search for the onset of color transparency through neutral rho meson electroproduction on nuclei*.

[Ziegler et al., 2020] Ziegler, V., Baltzell, N., Bossù, F., Carman, D., Chatagnon, P., Contalbrigo, M., De Vita, R., Defurne, M., Gavalian, G., Gilfoyle, G., et al. (2020). The clas12 software framework and event reconstruction. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*.