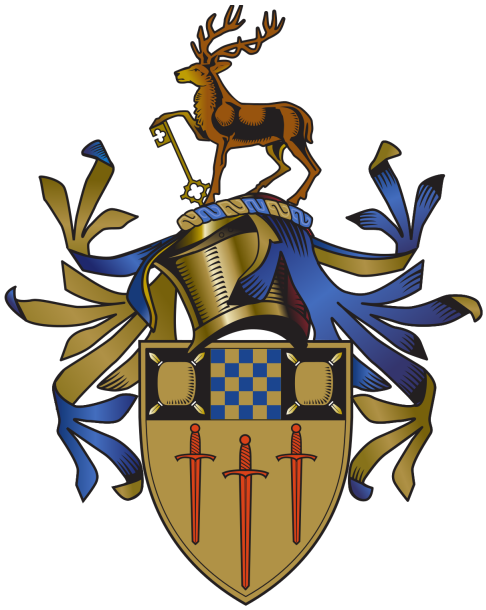
University of Surrey

CLAS12 Hough Transform Track Recognition and Drift Chamber Reconstruction Software Validation

by

Michael Armstrong



A dissertation submitted to the University of Surrey in partial fulfilment of a masters degree in physics

Department of Physics
Faculty of Engineering and Physical Sciences

January 14, 2020

Abstract

Advancements in accelerator and detector technology have for decades pushed the nuclear and particle physics envelope by enabling physicists to test their predictions at ever higher energies and with ever greater precision. Jefferson Lab's Continuous Electron Beam Facility (CEBAF) has recently been upgraded, doubling the maximum beam energy it can deliver its scientific user community to 12 GeV. The CEBAF Large Acceptance Spectrometer (CLAS12) has been upgraded to study the phenomena made observable at these higher energies in electron scattering experiments. The CLAS12 collaboration relies upon CLAS Offline Analysis Tools (COATJAVA) to reconstruct the properties of scattering products from the detector data collected in these experiments.

The subject of this dissertation is research conducted for the CLAS12 collaboration to support the continued development of this reconstruction software. This research took the form of two main projects. COATJAVA uses a software validation tool called a unit test to validate that new versions operate as intended. The unit test associated with particle reconstruction in CLAS12's Drift Chambers (DC) consistently produced false positives for software failures in the newest COATJAVA versions and so had become unusable. In the first project, the source of this failure was identified and the unit test was updated for use with the latest versions of COATJAVA. The test was also expanded by increasing the number of software failure indicators it examines.

In the second project, a new method of reconstructing the properties of charged particles from the helical patterns of detector hits they leave in CLAS12's central tracking systems was investigated. An implementation of the Hough Transform pattern recognition technique was written and its suitability for recognising the circular component of these helices from simulated data was evaluated. The Hugh Transform was combined with a version of the Taubin circle fitting algorithm to select from the circles recognised by the Hough Transform the circle most likely to best describe the circular component of the helical patterns left by

charged particles in the CD. The resolutions with which particle trajectories and momenta can be reconstructed from detector data using the Hough-Taubin combination were evaluated and compared to those achieved by the more conventional linear reconstruction method currently in use.

Declaration of Originality

This dissertation and the work to which it refers are the results of my own efforts. Any ideas, data, images or text resulting from the work of others (whether published or unpublished) are fully attributed to their originator in the text, bibliography or in footnotes. This dissertation has not been submited in whole or in part for any other academic degree or professional qualification. I agree that the University has the right to submit my work to the plagiarism detection service TurnitinUK for originality checks. I own the copyrights.

Acknowledgements

I would like to thank my supervisor Professor Jerry Gilfoyle for his tremendous support throughout my entire time living and working in the USA. My gratitude to Dr Paul Stevenson who was of great help as my visiting tutor this year and the many members of the CLAS12 Collaboration, especially Lamya Basheen who I worked closely with and Veronique Ziegler who suggested me these projects to work on.

Abbreviations

CEBAF	Continuous Electron Beam Facility
CLAS12	CEBAF Large Acceptance Spectrometer (at 12GeV)
COATJAVA	CLAS Offline Analysis Tools JAVA
JLab	Thomas Jefferson National Accelerator Facility
SLAC	Stanford Linear Accelerator
NSAC	Nuclear Science Advisory Committee
QCD	Quantum ChromoDynamics
GPD	Generalised Parton Distribution
FF	Form Factor
DVCS	Deeply Virtual Compton Scattering
FD	Forward Detector
CD	Central Detector
EVIO	Event Input Output
HIPO	High Performance Output
GEMC	Geant4 Event Monte-Carlo
CED	CLAS Event Display
DC	Drift Chamber
DOCA	Distance of Closest Approach
HBT	Hit Based Tracking
TBT	Time Based Tracking
FTOF	Forward Time-Of-Flight
SVT	Silicon Vertex Tracker
MM	MicroMega
CTOF	Central Time-Of-Flight Detectors
CND	Central Neutron Detector

Contents

Chapter 1

Background

1.1 Jefferson Lab

The Thomas Jefferson National Accelerator Facility (JLab) is a US Department of Energy nuclear physics research laboratory in coastal Virginia. JLab was originally conceived of in response to the US Nuclear Science Advisory Committee's (NSAC) 1976 "A long range plan for nuclear physics" [1]. This report highlighted the need for a "continuous beam, higher energy accelerator" which would study the interior structure of nucleons in electron scattering experiments. NSAC proposed that electrons accelerated beyond 2 GeV would have sufficiently short enough De Broglie wavelength to resolve quark-gluon matter, the existence of which was recently confirmed in electron scattering experiments carried out with the Stanford Linear Accelerator (SLAC) [2].

An incident electron may interact with its target nuclei in many different ways to produce many unique combinations of reaction products. Exclusive and semi-inclusive measurements are where all (or most) of the scattering products (rather than just the scattered electron) are captured. These measurements enable the specific interaction in which a scattering product was involved to be identified [3]. An accelerator capable of facilitating these measurements would allow for rare interactions of particular interest to be distinguished and their interac-

tion cross sections (a measure of relative probability) to be extracted from the wider data set [1].

Conventional accelerators of the time (such as SLAC) depended upon copper current carrying cavities to produce the electric fields with which they accelerated particles. These current cavities were only powered for short periods as they would otherwise melt under continuous use. The ratio of pulse time to time between pulses is known as the duty factor [4]. These accelerators operated with duty factors of $\approx 0.1\%$ [5]. NSAC argued a high duty factor ($\approx 100\%$) would be required for exclusive and semi-inclusive measurements. An electron beam current greater than $100\mu A$ would also be necessary for a high enough rare event yield that the effects of statistical fluctuations would be sufficiently minimised and precise measurements could be produced. In order to facilitate measurements of phenomena with spin-dependent cross sections, the accelerator's electron source would also need to be capable of spin polarising >85% of beam electrons [1].

This accelerator would become JLab's CEBAF facility which began construction in 1987. CEBAF was initially designed to provide a maximum continuous electron beam energy of 4 GeV, at up to $200\mu A$ with a duty factor of 100% which could be spin polarised for up to >85% of electrons [6][7]. CEBAF's beam energy went on to be upgraded multiple times, completing its latest upgrade to 12 GeV in December 2016 (see appendix 1 for overview of CEBAF's primary systems and the 12 GeV upgrade) [8].

1.2 Physics with CEBAF at 12 GeV

The original CEBAF design was accepted by the US Department of Energy in part because it could be "readily upgraded". The push for 12 GeV began as early as 1985 when it was argued at a consultation with electro-nuclear physcists at MIT to the then prospective director Hermann Grunder, that quark scale structure would be best probed at "8-12 GeV" [4]. This proposal was ultimately impeded by the limitations of the accelerator technology

of the day and the additional cost it would incur. A decade later however, NSAC would issue support for a 12 GeV upgrade citing "favourable technical developments in accelerator technology" and "sforesight in the design of the original facility" [8]. JLab's 12 GeV pre-conceptual design report describes a number of promising research opportunities that would become available at these higher energies.

1.2.1 Hybrid Mesons

Firstly, a 12 GeV beam energy would enable the verification or falsification of the presence of exotic mesons [9][10]. In Quantum ChromoDynamics (QCD) quarks and gluons interact through the strong force. These interactions are analogous to electromagnetic interactions but quarks and gluons carry three different kinds (or colours) of charge. These colour-charge-carrying particles are treated as excitations of more fundamental underlying fields [13]. Colour-carrying flux tubes are thought to form between interacting gluons. The excitations of these flux tubes give rise to short-lived exotic hybrid mesons, the properties of which are important to understanding colour confinement [11]. Colour confinement is an important prediction of QCD, describing that colour charged particles cannot be isolated (and therefore independently observed) outside of extreme conditions such as above the Hagedorn temperature at $2 \times 10^{12} K$ [13][14]. 12 GeV electron beam energies would enable the observation of these exotic hybrid mesons in 9GeV pion photoproductions [10].

1.2.2 Generalised Parton Distributions

Secondly, increased beam energies would enable extensions of previous measurements of Generalised Parton Distributions (GPDs) to higher momenta [9][10]. A parton distribution describes the probability $\mid \psi(x) \mid^2$ that hadron constituent particles (termed "partons") carry a fraction x of the momentum of a fast moving nucleon. A GPD meanwhile measures the coherence between different parton momentum states in the nucleon and so is fundamental to understanding the distribution of momentum between partons in nuclei [9][10][13].

1.2.3 Electromagnetic Form Factors

Finally, increased beam energies would also enable extensions of existing measurements of Form Factors (FF) to higher energies [9]. Electric and magnetic FFs reflect the charge and magnetisation current density distributions within nucleons [14]. 12 GeV beam energies electron scattering experiments could probe deeper into the interior structure of nucleons, revealing how charge and magnetisation current is distributed between constituent quarks [10].

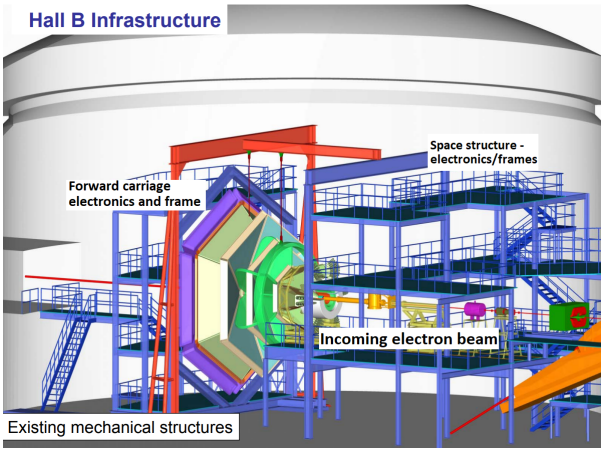
1.3 CLAS 12 GeV Upgrade

Following its upgrade CEBAF can deliver a 12 GeV electron beam to experiments in four simultaneously operating research halls (A-D) (see Appendix 1 for details of CEBAF's main systems and 12 GeV upgrade). Research Hall B's CEBAF Large Angle Spectrometer (CLAS) was also upgraded as of December 2016 to make use of the higher available beam energies [8]. CLAS12 was designed with a focus on conducting spectroscopic studies of excited baryons and of polarised and unpolarised quark spatial and momentum distributions. CLAS12 was also intended to investigate the influence of nuclear matter on propagating quarks and to extend previous measurements of GPDs and FFs in Deeply Virtual Compton Scatterings (DVCS) [15]. A DVCS event is where an electron is scattered off a proton and a virtual photon is exchanged in conservation of momentum.

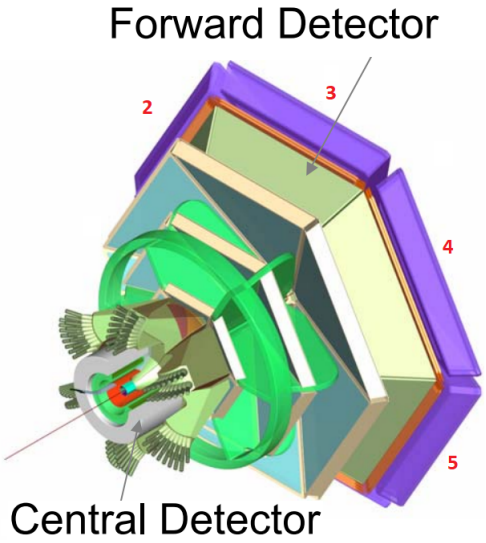
1.3.1 CLAS12 Design Overview

CLAS12 is positioned in the centre of research Hall B's large circular underground chamber and is supported by the forward carriage depicted in Figure 1.1a [15]. CEBAF's main beamline leads into CLAS12's Central Detector (CD). The CD is the first of two main sections that make up CLAS12 and contains its target. The Forward Detector (FD) is the

second of these main sections and sits behind the CD as illustrated in Figure 1.1b. Beam Electrons enter the target with nearly all their momentum directed along the beamline and so most scattering products retain this forward direction and scatter $5° - 40°$ from the beamline into the FD (See appendix 2 for CLAS12 co-ordinate system) [15].



(a) Research Hall B Infrastructure layout.



(b) CLAS12 Forward and Central Detector layout.

Figure 1.1: Illustrations of Hall B infrastructure and CLAS12 detector layouts. Adapted versions of images from [15].

Electrons that pass the target nuclei particularly closely are strongly deflected and scatter along wide angles. Those electrons and other scattering products that travel $40° - 135°$ from the beamline enter the Central Detector (CD) [15]. In order to fulfil the requirements of its first proposed experiments, the FD and CD detectors were designed to fulfil the specifications summarised in Table 1.1 [15].

	Forward Detector	Central Detector
Angular Range	5°-40°	40°-135°
Resolution	At 5 GeV/c	At 0.9 GeV/c
δp/p (%)	1	3
$\delta\theta$ (mrad)	0.5	10
$\delta\Phi$ (mrad)	0.5	6
Photon Detection(4σ)	Range	Range
Energy (MeV)	≥150	N/A
Neutron Detection	Chance	Chance
Efficiency (%)	10-60	5
Particle ID (4σ)	Range	Range
e/π	5-12 GeV/c	N/A
π/p	5-12 GeV/c	≤1.2 GeV/c
π/K	5-12 GeV/c	≤0.65 GeV/c
K/p	≤4 GeV/c	≤0.9 GeV/c

Table 1.1: Table summarising CLAS12 detector design parameters [15].

The FD is designed to reconstruct scattering products with peak efficiency at 5 GeV/c while the CD is designed for 0.9 GeV/c particle reconstructions. As shown in Table 1.1, the FD was designed to achieve a momentum resolution of $\delta p/p = 1\%$ with polar (θ) and azimuthal (ϕ) angular resolutions of 0.5 mrad and to detect photons with more than 150MeV/c momentum. The CD was designed to achieve $\delta p/p = 3\%$ with $\delta\theta = 10\ mrad$ and $\delta\phi = 6\ mrad$. CLAS12 was also designed to detect neutrons with up to 12 GeV/c momentum at a 10-60% and 5% efficiency in the case of the FD and CD respectively. The FD and CD were also designed to separate electrons, pions, kaons and protons with 99.99% (4σ) accuracy for the energy ranges listed in Table 1.1.

1.4 CLAS12 Software Structure

In order to sort, store and process data from CLAS12's 111,832 detector readout channels, the CLAS12 collaboration uses a sophisticated suite of data processing tools called the

CLAS Offline Analysis Tools (COATJAVA). COATJAVA is made up of seven main software packages, which are displayed in Figure 1.2 [17].

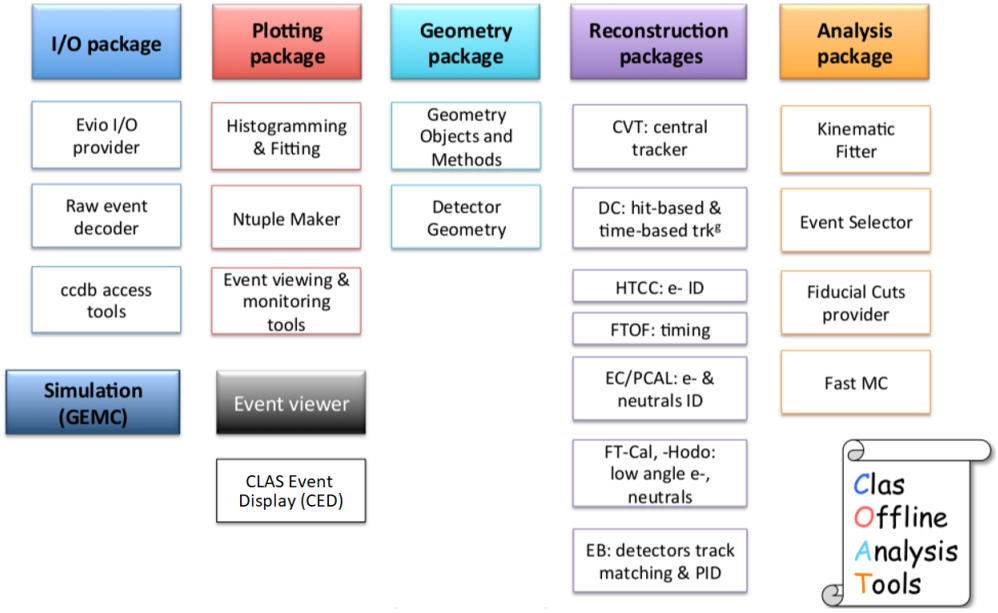


Figure 1.2: Diagram of main COATJAVA software packages and features [18].

The Collaboration uses a special file format called EVIO (Event Input Output) to store the raw signals captured by detectors in a digital format. The I/O package decodes these signals, extracting and storing the relevant information (such the detector component id or voltage peak recorded) in High Performance Output (HIPO) data banks. The HIPO file format was developed by the Collaboration to be accessibly structured and to minimise file sizes for ease of physics analysis. The reconstruction package reads signal data from these data banks and combines it with information from the detector geometry package to reconstruct the trajectories, identities and properties of the scattering products that produced the original detector signals and stores them in further HIPO data banks [17].

The analysis and plotting packages are used by the collaboration to interface with these data banks and produce particle data plots. These packages also enable the Collaboration to make fiducial cuts, where they select which particles from what part of the detector they want to include in their analyses.

The Geant4 Monte-Carlo (GEMC) package is a Geant4 particle physics based simulation package which uses information from the geometry package to build simulations of the CLAS12 detector and to produce simulated particle event detector data. GEMC is used by the Collaboration to test analyses on well understood, generated data [17].

Finally, the CLAS Event Display (CED) package displays particle event data within an intuitive visualisation of the detector to help Collaboration members to understand their results. The latest versions of COATJAVA are available on the CLAS12 collaboration github page at https://github.com/JeffersonLab/clas12-offline-software [17].

Chapter 2

Drift Chamber Reconstruction Validation Software

2.1 The Forward Detector

Particles that enter the FD travel through three regions containing Drift Chambers (DC) shown in the cross sectional view of the wider detector in Figure 2.1. The example trajectory shows how an incident beam electron might enter scatter off the target in the CD and continue through calorimeters, DC regions, time of flight scintillators and Cherenkov counters in the FD. This combination of detector subsystems enables the FD to reconstruct the trajectory, momentum and identity of incident particles (further details regarding the purpose and design of these systems can be found in Appendix 1) [10].

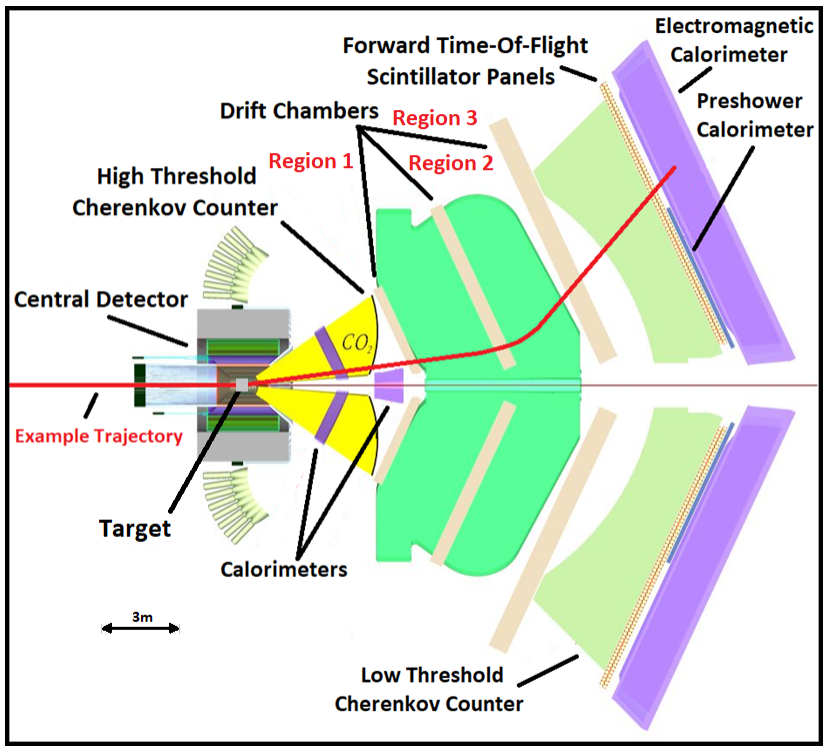


Figure 2.1: Cross section of CLAS12 detector featuring example trajectory (entering left) as it travels through the CD and FD. In this case the example particle is negatively charged and so bends away from the beamline due to the direction of the toroidal field (green). Adapted image from [4].

2.2 Reconstruction in CLAS12's Drift Chambers

The three DC regions are split into pairs of superlayers made up many hexagonal gas chambers (or cells) that are arranged into six smaller layers each 112 cells wide. These cells (illustrated in Figure 2.2) contain a 9:1 mixture of argon-CO_2 gas and a central anode sense wire [9]. Particles travelling through the DC layers leave trails of electrons as they ionize gas molecules they pass. These electrons accelerate towards the sense wires and deliver a current signal from which it can be deduced that an incident particle passed through the cell. The pattern of these sense wire hits is used to reconstruct the trajectory of incident particles.

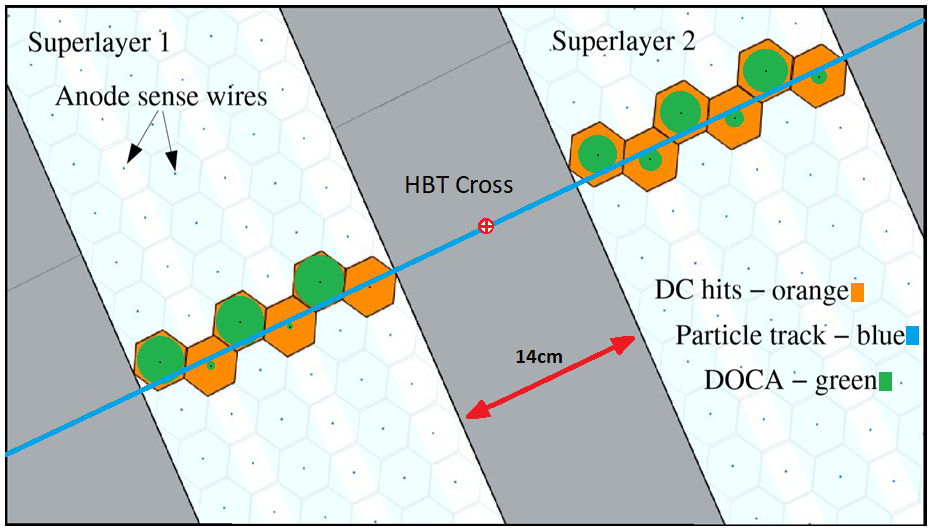


Figure 2.2: Illustration of particle reconstruction in DC region 2. Activated cells in orange, Distance of Closest Approach (DOCA) in green and reconstructed track in blue. HBT reconstruction cross positioned at midpoint between superlayers separated by 14cm.

2.2.1 Hit Based Tracking

COATJAVA begins this reconstruction with a process called Hit Based Tracking (HBT). In HBT the positions of activated sense wires are readout as points in 3D Cartesian space. A simplified 2D version of the process is illustrated in Figure 2.2. The process proceeds by fitting possible trajectories through the sense wire hits, (orange in Figure 2.2) each with an associated covariance matrix characterising the fitting uncertainties. The pairs of superlayers making up DC regions one, two and three are separated by 7cm, 14cm and 7cm [15]. The widths of these gaps are comparable to the width of superlayers themselves and so virtual points called reconstruction crosses are interpolated at the midpoint of where the HBT trajectory fittings cross between the superlayers. These crossing points shown in Figure 2.2 are used to better constrain possible particle trajectories between superlayers in the subsequent Time Based Tracking (TBT) process.

2.2.2 Time Based Tracking

When an incident particle ionises DC gas molecules, the ionised electrons take time to drift to the positive anode sense wire. During this drift time, the incident particle continues travelling through the detector where the timing of its passage through the Forward Time-Of-Flight scintillator (FTOF) is recorded (shown downstream of the drift chambers in Figure 2.1). This information, and the path length from the target to the FTOF (extracted from the HBT trajectory) is used to determine when the scattered particle left the target and when it later enters each DC cell along its HBT trajectory. The difference between when the scattered particle hits a DC cell t_{cell} and when ionised electrons are detected at the anode t_{anode} is taken as the time taken for the ionized electrons to drift from the ionisation site to the anode τ_{drift} (As in Equation 2.1). The Distance Of Closest Approach (DOCA) (shown in Figure 2.3) a particle may take as it passes the anode on its way through the cell is proportional to τ_{drift} with the time to distance function (f) as related in Equation 2.2. This is a measured calibration function that characterises the drift velocity from the ionisation site to the anode.

$$\tau_{drift} = t_{anode} - t_{cell} \tag{2.1}$$

$$DOCA = f \cdot \tau_{drift} \tag{2.2}$$

The DOCA for each chamber hit further constrains where within the cell the incident particle may have passed on its path through the DC. This information is used along with the HBT superlayer crossing point to more precisely reconstruct the particles trajectory.

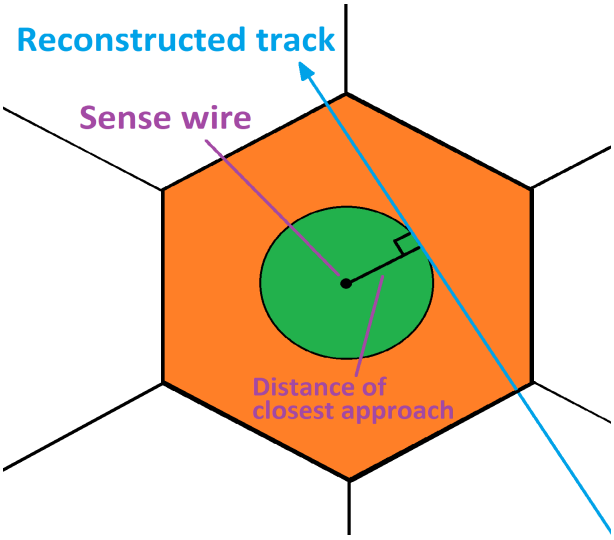


Figure 2.3: Incident particle TBT DOCA illustration, activated cell in orange, incident particle passes somewhere within green DOCA circle.

2.2.3 Momentum Reconstruction

The FD makes use of a powerful superconducting toroidal magnet to analyse the momentum of charged particles. The toroid is made up of 6 coils and takes the shape of a hexagon. The toroid encloses the second DC region in a toroidal magnetic field with a peak strength of 3.58T [19]. As shown with the example trajectory in Figure 2.1, when charged particles approach this field (shown in green) they accelerate and their trajectories bend. The curvature of a charged particles track is used to analyse its momentum.

Consider at first the simple case a point charge (Q) travelling in xy plane with momentum (P_{xy}) as it enters a region with a constant magnetic field (B_z) acting along the z axis. In the magnetic field, the charge's trajectory bends through a circle with radius R. By equating the centripetal force ($F_c = \frac{mv_{xy}^2}{R}$) to the electromagnetic force it experiences in the magnetic field ($F_B = Q \cdot v_{xy} \times B_z$), the radius of curvature can be found as a function of momentum as in Equation 2.3.

$$R = \frac{P_{xy}}{QB_z} \tag{2.3}$$

Similarly, the 3D curvature of charged particle trajectories as they travel through the toroidal field is related to their momenta. The curvature of a charged particles TBT trajectory is therefore extracted and used to determine each component of its momentum.

2.3 DC Software Unit Test

In order to make development more easily manageable, COATJAVA's 84,000 lines of executable code are split into many small modules. Each module comes with a software unit test designed to test individual units of source code and validate that each functions as intended. When a developer updates the the source code associated with the DC reconstruction module, the DC unit test takes preprogrammed detector data representing a single test particle and reconstructs its momentum from the best fitting TBT trajectory. If the momentum is sufficiently different from expectation (i.e lays outside the tolerance range), the test signals that there has been a software failure. As of March 2019 the existing unit test was found to be falsely signalling software failures every time it was run. The test needed to be updated before it could continue to be used as a tool to validate the DC reconstruction module.

2.3.1 Generation of Data for DC Software Validation

In order to understand why the unit test was producing false positives an analysis of the DC reconstruction was undertaken. Simulated detector data representing 20,000 electrons scattered from the target through the FD was generated using GEMC. These electrons were generated with the same initial starting position and momentum as the particles in the simulation from which the existing test particle was taken were generated. These initial conditions were an electron scattered from the target with a momentum of $2.5 \pm 0.5\ GeV/c$ through a polar angle $\theta = 25^\circ \pm 10^\circ$ and an azimuthal angle $\phi = 0^\circ \pm 5^\circ$ from the beamline (see appendix 2 for detector co-ordinate system definitions).

2.3.2 DC Reconstruction Analysis

This simulated data was then reconstructed using the latest version of COATJAVA (version 6.3.1 at the time) to produce histograms showing the frequency distributions of 20,000 reconstructed particle event x, y and z momenta and vertices. The vertex is a quantity used to characterise a particles trajectory from the target. By extrapolating a reconstructed trajectory back to the target region, the vertex is found as the closest point along the trajectory to the origin at the target (see appendix 2 for more quantitative definition of vertex). The histogram in Figure 2.4 shows the frequency distribution of reconstructed P_z for the 20,000 simulated particle events. The figure is superimposed with green dot representing where the existing unit test event was reconstructed ($P_z = 2.41 GeV/c$) and with green lines at the existing test event tolerance limits of $2.16 < P_z < 2.39\ GeV/c$.

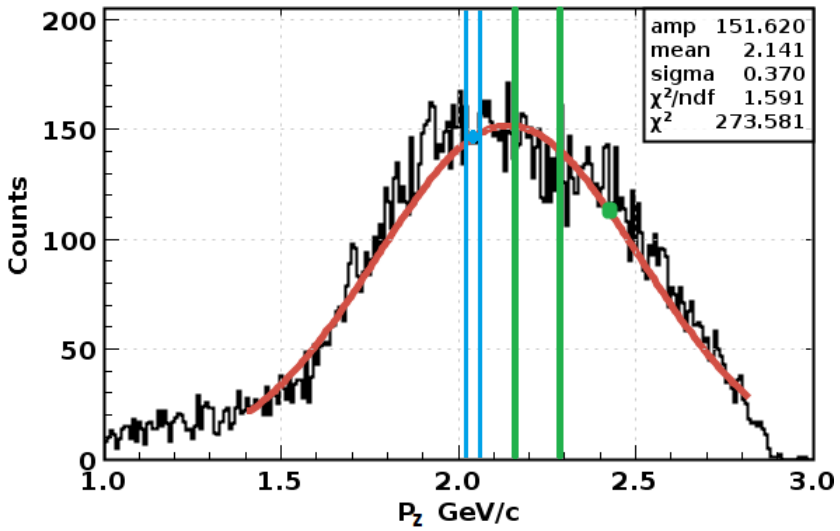


Figure 2.4: Simulated data P_z histogram, with existing unit test event reconstruction (green dot) and existing test limits (green lines). Replacement test event reconstructed at 2.04GeV/c (blue dot) with new tolerance limits $2.27 < P_z < 2.53$GeV/c (blue lines)

Figure 2.4 shows the existing unit test event P_z is reconstructed outside its tolerance limits when reconstructed with the latest version of COATJAVA. These limits were based upon a version of the DC reconstruction from 2018 which has since evolved. With alterations to the

HBT and TBT fitting algorithms and to the COATJAVA simulation of the CLAS12 detector geometry, newer versions of COATJAVA reconstruct P_z differently and so the existing limits do not match the current reconstruction. The unit test limits would therefore need to be updated to better represent more recent versions of the reconstruction.

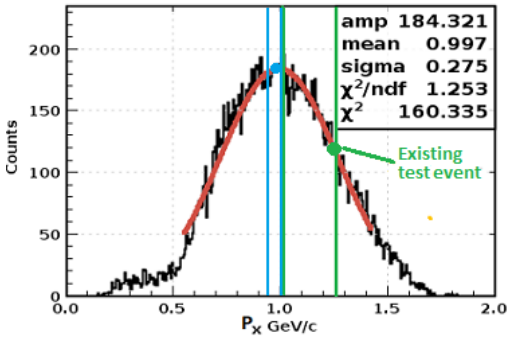
By Fitting the peak of the P_z distribution with a Gaussian function (shown in red) the mean P_z was found to be $2.14 \pm 0.39 GeV/c$. With the new reconstruction version the existing test event lay on the edge of the first standard deviation from the mean of other events (generated under the same conditions) at 2.52 GeV/c. This suggested it may be an atypical event and prompted an investigation into whether it's a suitable candidate to base the unit test on by examining its other properties.

As particles travel through CLAS12 they attenuate and lose momentum by scattering off of its systems [15]. The particles in the GEMC simulation were generated with mainly P_z and P_x momentum and so these energy losses are most apparent in the low momentum tails the P_z and P_x distributions exhibit in Figures 2.4 and 2.5. Most of the initial momentum is in the z direction and so the tail is more pronounced in the P_z distribution. These momentum losses cause a particle's trajectory to shift, this is observed as the vertex tails in the V_z distribution and less clearly in the V_x distribution. Because of these energy loss tails, only the peak of each distribution in Figures 2.4 through 2.7 could be fitted with a Gaussian function.

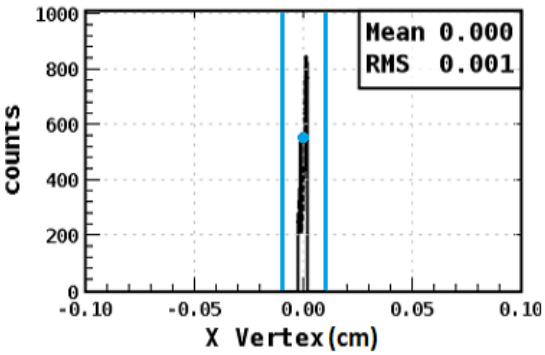
2.3.3 DC Unit Test event Analysis

Figures 2.5, 2.6 and 2.7 show the other momentum and vertex distributions reconstructed from the simulation with the position of the existing test event and limits indicated. In the case of the x component momentum and vertex, the simulated electrons were reconstructed with $P_x = 0.997 \pm 0.275\ GeV/c$ and $V_x = 0 \pm 0.001\ cm$. The electrons were generated with $\phi = 0 \pm 5^\circ$ and so V_x is expected to vary tightly around zero as with the spike in Figure 2.5b.

The existing test event was reconstructed with $P_x = 1.24\ GeV/c$ near the edge of the existing test limits at $1.02 < P_x < 1.26\ GeV/c$.

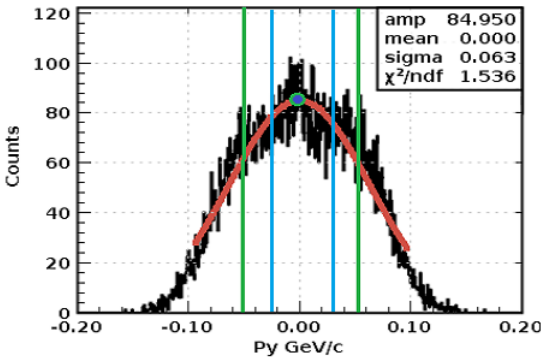


(a) Reconstructed P_x histogram with existing test event (green dot) and limits (green lines). Replacement test event at $P_x = 0.997\ GeV/c$ (blue dot) with newly implemented tolerance limits $0.993 < P_x < 1.00\ cm1$ (blue lines).

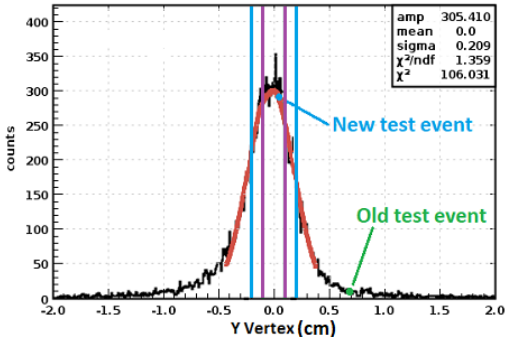


(b) Reconstructed V_x histogram spike at zero, new and existing test event both reconstructed at zero. New test limits at $0.993 < V_x < 1.001\ cm$ (blue lines)

Figure 2.5: Histograms of reconstructed momentum and vertex x components.



(a) Reconstructed P_y histogram with existing and replacement test event both at zero (overlapping green and blue dots), existing limits (green lines) and new limits at $-0.023 < P_y < 0.003\ GeV/c$ (blue lines).



(b) Reconstructed V_y histogram with existing test event with $V_y = 0.69\ cm$ (blue dot), new test event at zero and new limits at $-0.21 < V_y < 0.21\ cm$ (blue lines). New test event 0.5σ selection criteria indicated with purple vertical lines.

Figure 2.6: Histograms of reconstructed momentum and vertex y components.

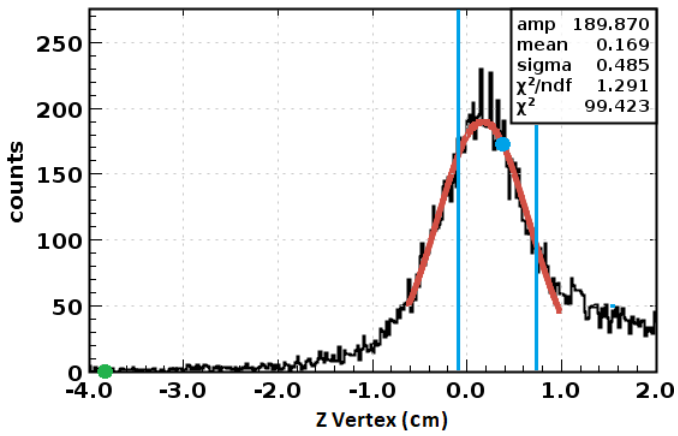


Figure 2.7: Simulated data V_z histogram with new test event at 0.39 and new limits $-0.089 < V_Z < 0.869$ cm indicated with blue dot and vertical lines, old test event more than 8 sigma from mean at -3.84 cm indicated with green dot

2.4 New DC Unit Test Event Selection

Following the recommendation of JLab staff scientist and CLAS12 software coordinator Veronique Ziegler, the new test event was selected from the GEMC simulation data with reconstructed momentum and vertex x,y and z components within 0.5 sigma from the mean of each distribution so as to be sure a typical event is selected. The vertical purple lines in Figure 2.6b show how the replacement test event was selected from the events that lay within 0.5σ of the mean of each distribution. In this case, the mean was zero with $\sigma = 0.209\ cm$ and so the selection was made from events within $0 \pm 0.1\ cm$ (as $0.5\sigma = 0.1\ cm$). The properties of the new test event selected through this criteria are summarised in table 2.1.

Property	New Test Event	New Test Limits
P_x (GeV/c)	0.997	0.997±0.004
P_y (GeV/c)	-0.01	0.01±0.013
P_z (GeV/c)	2.04	2.04±0.013
X Vertex (cm)	0.00	0.00±0.001
Y Vertex (cm)	0.00	0.00±0.210
Z Vertex (cm)	0.39	0.39±0.479

Table 2.1: New DC unit test particle properties and test particle reconstruction tolerance limits.

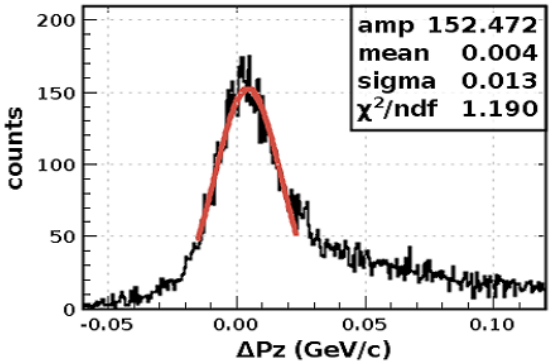
2.5 DC Reconstruction Resolution Analysis and Unit Test Tolerance Limit Selection

In order to update the DC unit test with new tolerance limits that match the new reconstruction version, the reconstruction resolution was studied. The resolution for each momentum and vertex component was determined by finding the difference between how particles are reconstructed and how they should have been based upon how they were originally generated. This difference (a measure of reconstruction error) is defined as $\Delta P = P_{gen} - P_{recon}$ and $\Delta V = V_{gen} - V_{recon}$ for momentum and vertex respectively.

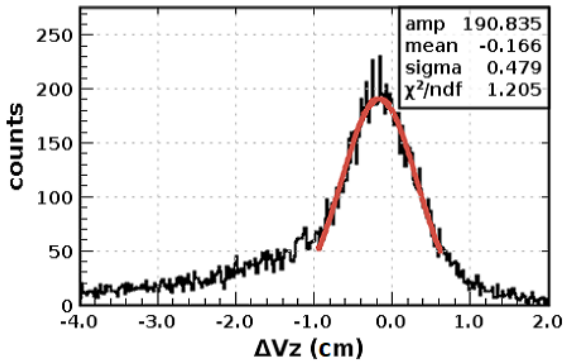
Histograms showing the frequency distribution of each ΔP and ΔV component are plotted in Figures 2.8, 2.9 and 2.10. The energy loss tails observed Figures 2.8, 2.9 and 2.10 are reversed when compared to the those in Figures 2.4 to 2.7. This is because the difference is defined as generated minus reconstructed and so a loss in reconstructed momentum becomes an increase to the difference. Figure 2.8a is a histogram of ΔP_z with a peak centred close to zero. The peak of the distribution has been fitted with a Gaussian function ($\chi^2_{reduced} = 1.19$) with $\sigma = 0.013\ GeV/c$, this means 68% of particles in that peak were reconstructed accurately to within 0.013 GeV/c of how they were known to be generated. Reconstructions of test event ΔP_z can therefore be expected to be reconstructed within a resolution of 0.013 GeV/c 68% of the time. The new unit test limits for P_z were selected at $2.04 \pm 0.013\ GeV/c$, with the one sigma resolution (0.013 GeV/c) used as the tolerance limits around where the test event is expected to be reconstructed (2.04 GeV/c).

It is possible that a test particle happens to be reconstructed outside the tolerance limits by random chance rather than a software failure and so a software failure is falsely signalled. The collaboration software group made the choice of a one sigma tolerance by balancing the concerns that too tight a tolerance would produce an unacceptably high rate of false positives and that too loose would increase the chance of software issues remaining undetected.

The same analysis of reconstruction resolution was applied to all other momentum and vertex components leading to the selection of the new unit test limits listed in table 2.1.

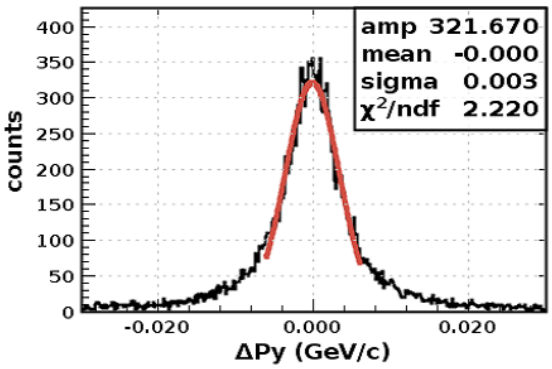


(a) Simulated data ΔP_z, distributed about mean of 0.004 GeV/c with σ = 0.013 GeV/c

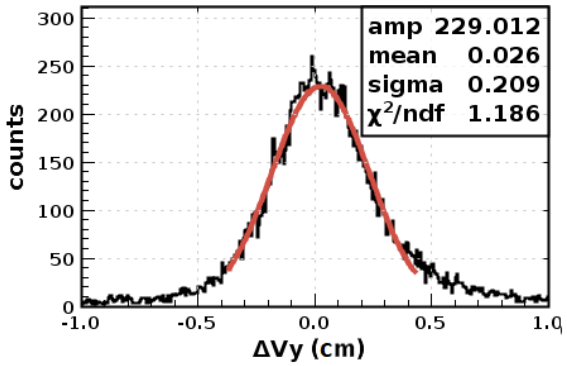


(b) Simulated data ΔV_z, distributed about of zero mean of -0.166 with σ = 0.479 cm

Figure 2.8: Reconstruction resolution analysis z component momentum and vertex histograms.

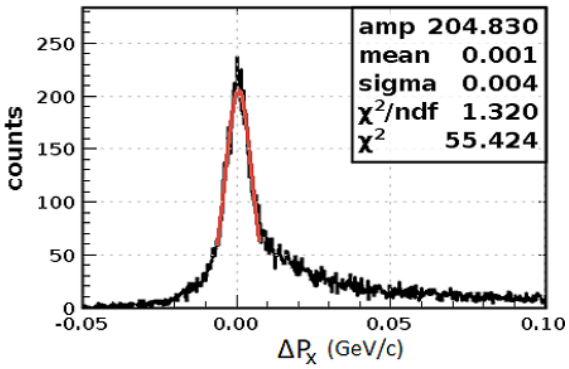


(a) Simulated data ΔP_y, distributed about mean of zero with σ = 0.003 GeV/c

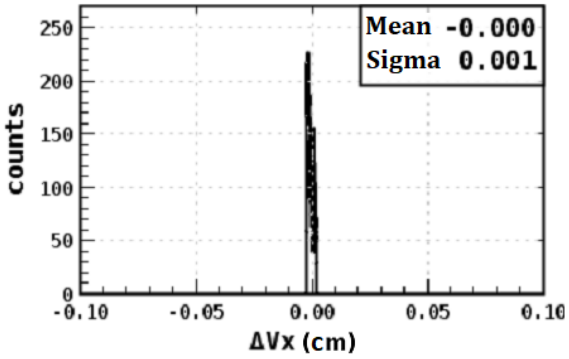


(b) Simulated data ΔV_y, distributed about mean of 0.026 cm with σ = 0.209 cm

Figure 2.9: Reconstruction resolution analysis y component momentum and vertex histograms.



(a) Simulated data ΔP_x, distributed about mean of 0.01 GeV/c with σ = 0.004 GeV/c



(b) Simulated data ΔV_x, distributed about mean of zero with $\sigma = 0.001$ cm

Figure 2.10: Reconstruction resolution analysis x component momentum and vertex histograms.

2.6 DC Unit Test Reconstruction Crosses

The DC reconstruction relies upon the placement of reconstruction crosses to ensure the reconstructed trajectory is accurate as it passes between DC superlayers. The proper placement of these crosses is an important indicator that a new version of DC reconstruction software is performing as intended. The DC unit test was therefore expanded to check that reconstruction crosses are being accurately positioned.

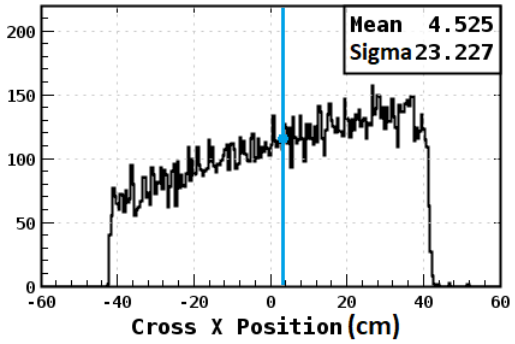
2.6.1 DC Reconstruction Cross Position Analysis

Crosses are positioned at a fixed midpoint along the z axis between the three pairs of DC superlayers and vary in the xy plane depending on the reconstructed HBT trajectory. The position of the crosses for the three DC regions was extracted from the reconstructed GEMC simulation and used to produce the x and y cross position histograms in Figures 2.11, 2.12 and 2.13.

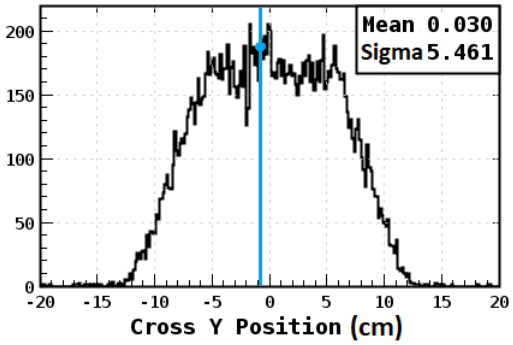
As particles travel through each DC region, they lose momentum when they ionise a DC gas molecule or scatter off the detector structure. Equation 2.3 shows charged particles

with less momentum will bend through a tighter curve with a smaller radius of curvature in a magnetic field. Electrons are negatively charged and so due to the direction of CLAS12 toroidal field their trajectories bend away from the beamline as with the example trajectory in Figure 2.1. With each successive DC region the electrons pass through, they lose more momentum and bend increasingly (in this case) towards a lower x and a higher y position. This effect manifests in the increasingly more pronounced low x cross position tails observed in Figures 2.11a, 2.12a and 2.13a.

The electrons were generated with a momentum of 2.5 ± 0.5 GeV/c directed at $\theta = 25^\circ \pm 10^\circ$ and $\phi = 0^\circ \pm 5^\circ$ and so the mean x cross position decreases from $4.53cm$ in region one to 3.12 cm and -12.312 cm in regions two and three as they travel along this trajectory and pass through different parts of each superlayer pair. The electrons were generated with little component momentum and so the y position of the crossing points stays roughly the same, going from $y = 0.03 \pm 5.46$ cm in region one to $y = 0.01 \pm 8.21$ cm and $y = 0.023 \pm 10.41$ cm in regions two and three.

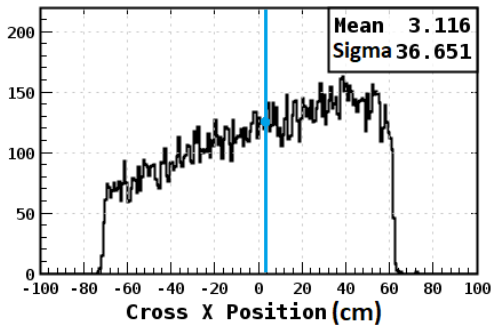


(a) Simulated data cross x position in region 1. New test event reconstruction has cross at x = 4.07 cm (blue line).

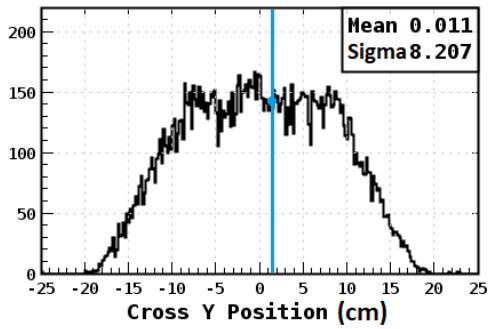


(b) Simulated data cross y position in region 1. New test event reconstruction has cross at y = -1.95 cm (blue line).

Figure 2.11: Region 1 cross position histograms from simulated data reconstructions.

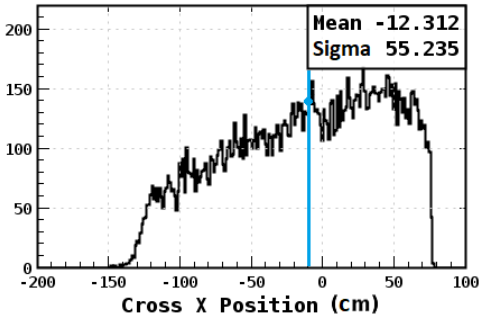


(a) Simulated data cross x position in region 2. New test event reconstruction has cross at x = 4.02 cm (blue line).

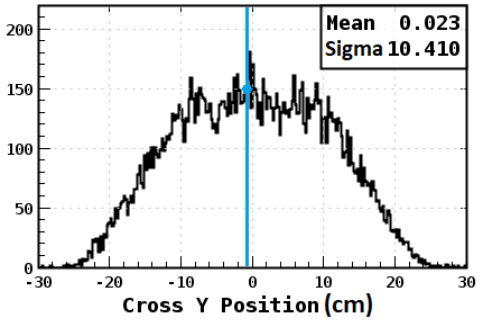


(b) Simulated data cross y position in region 2. New test event reconstruction has cross at y = -1.27 cm (blue line).

Figure 2.12: Region 2 cross position histograms from simulated data reconstructions.



(a) Simulated data cross x position in region 3. New test event reconstruction has cross at x = -11.0 cm (blue line).



(b) Simulated data cross y position in region 3. New test event reconstruction has cross at y = 2.85 cm (blue line).

Figure 2.13: Region 3 cross position histograms from simulated data reconstructions.

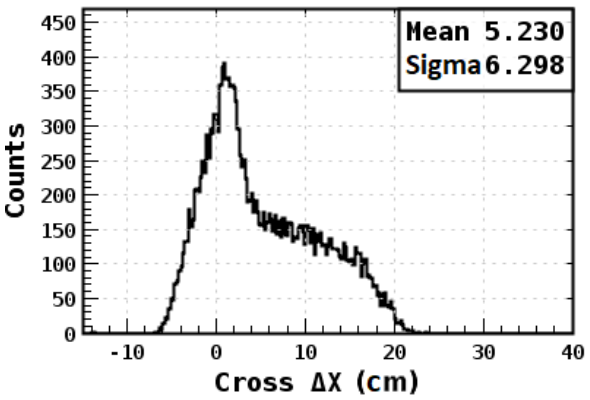
The new test event's reconstruction cross was found to be close to the mean of each distribution with x=4.07 cm and y=-1.95 cm for region 1, x=4.07 cm y=-1.27 cm for region 2 and x=-11.0 cm y=2.85 cm for region 3. In order to put tolerance limits on these cross positions the resolution of each cross position was studied.

2.6.2 DC Reconstruction Cross Positioning Resolution Analysis

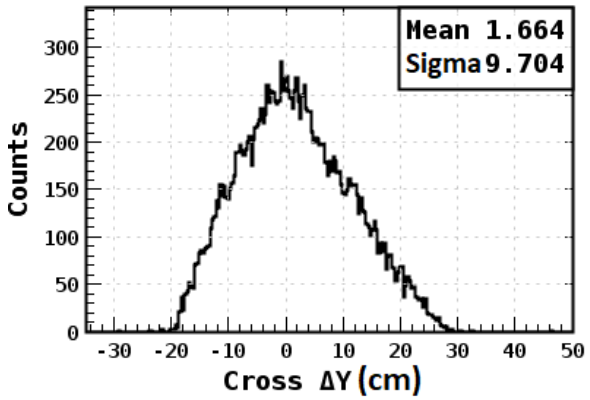
The starting point and initial momentum of each generated particle is exactly known from simulation input parameters. Using these starting points and by taking into account the geometry of the detector and toroidal field, the expected crossing points were calculated. The difference between where the crosses should be positioned and where they were are plotted in Figures 2.14,2.15 and 2.16 where $\Delta X = x_{gen} - x_{recon}$ and $\Delta Y = y_{gen} - y_{recon}$.

As with the previous reconstruction resolution analysis in Section 2.5, this difference is defined as generated minus reconstructed. The energy loss tails in Figures 2.11, 2.12 and 2.13 therefore move to the other side of the cross position histograms as a more negative cross position means an increased difference between the reconstructed and the generated.

The software working group decided that the reconstruction cross tolerance limits would be one standard deviation above and below where the new unit test particle's reconstruction crosses are currently positioned. As with the unit test limit selection in Section 2.5, this decision was reached by balancing the rate of possible false positives that could arise due to random chance with the possibility that a software failure may go undetected.

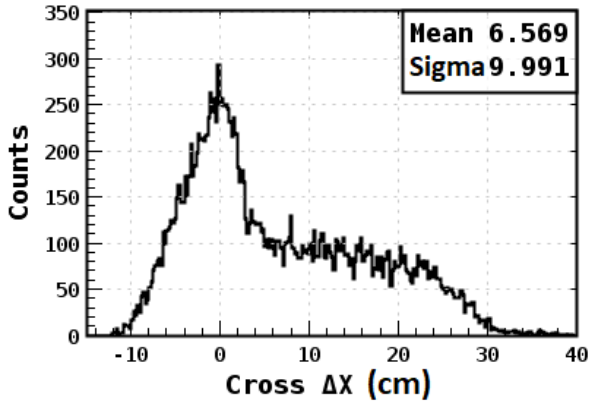


(a) Simulated data region 1 ΔX cross position. Mean of 5.23 cm with $\sigma = 6.298$ cm.

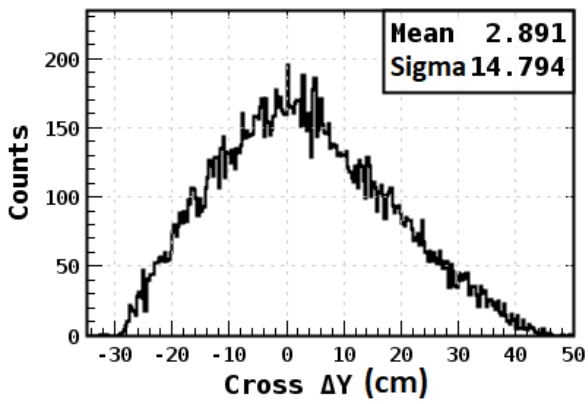


(b) Simulated data region 1 ΔY cross position. Mean of zero with $\sigma = 2.162$ cm.

Figure 2.14: Region 1 cross position resolution histograms.

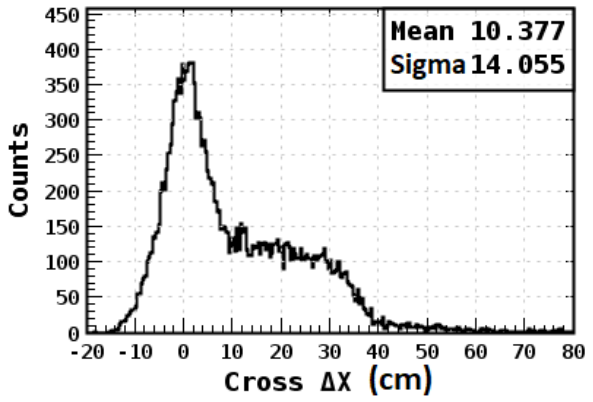


(a) Simulated data region 2 ΔX cross position. Mean of 6.569 cm with $\sigma = 9.991$ cm.

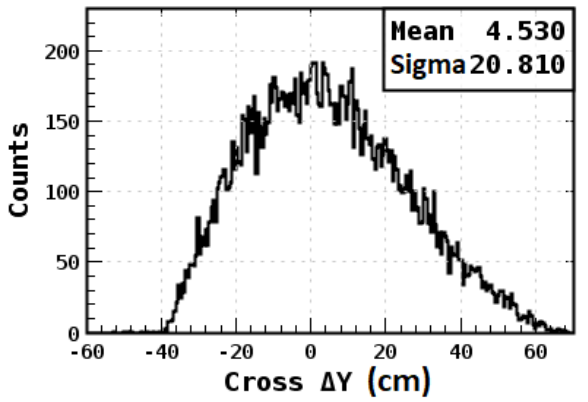


(b) Simulated data region 2 ΔY cross position. Mean of 2.891 cm with $\sigma = 14.794$ cm.

Figure 2.15: Region 2 cross position resolution histograms.



(a) Simulated data region 3 ΔX cross position. Mean of 10.377 cm with $\sigma = 14.055$ cm.



(b) Simulated data region 3 ΔY cross position. Mean of 4.53 cm with $\sigma = 20.81$ cm.

Figure 2.16: Region 3 cross position resolution histograms.

The standard deviation of the x position of reconstruction crosses for regions one, two and three were 6.3 cm, 9.99 cm and 14.01 cm. The unit test reconstruction cross x position tolerance limits were therefore put at $4.07 \pm 6.30cm$, 4.02 ± 9.99 cm and -11 ± 14.01 cm respectively. Similarly the y position standard deviations were 9.70 cm, 14.79 cm and 20.81 cm so the limits were -1.95 ± 9.70 cm, -1.27 ± 14.79 cm and -2.85 ± 20.81 cm. The new test event cross positions and tolerance limits are summarised in Table 2.2.

Property	Cross X Position	Cross Y Position	X Cross Limits	Y Cross Limits
Region 1 (cm)	4.07	-1.95	4.07±6.30	-1.95±9.70
Region 2 (cm)	4.02	-1.27	4.02±9.99	-1.27±14.79
Region 3 (cm)	-11.00	2.85	-11.0±14.055	-2.85±20.81

Table 2.2: DC unit test reconstruction cross limits

2.7 Updated DC unit test code structure

The new DC unit test works similarly to the old one, beginning by reconstructing the properties of the new test event from preprogrammed detector data. The unit test proceeds by checking that the reconstructed momentum, vertex and reconstruction crosses lay within the new tolerance limits and signalling that there has been software failure if they don't. The new unit test was written in JAVA and has been included in the latest versions of COATJAVA used by the collaboration since December 2019 (see appendix 3 for full DC unit test source code).

2.8 DC Reconstruction Software Validation Conclusion

The CLAS12 collaboration relies upon the COATJAVA software package to reconstruct the momentum, vertex and identity of particles from the raw detector data they collect in their electron scattering experiments with JLab's 12 GeV continuous electron beam. COATJAVA is split into many modules, each with a software unit test that validates new versions of source code before they are integrated into new builds. The unit test associated with the module of code that reconstructs particle momenta and trajectory from detector data collected by CLAS12's DCs was consistently falsely signalling failures in new versions and so had become unusable.

The unit test works by reconstructing a test particle's momentum from preprogrammed detector data and checking that it's reconstructed within a range of tolerated values. The test particle was found to be reconstructed outside these tolerance limits when reconstructed using the newest version of COATJAVA (6.3.1). The reconstruction software had changed sufficiently in the time since the unit test was written that the limits had become outdated and needed to be replaced.

Simulated data representing the passage of electrons through CLAS12's detector systems was generated using GEMC and reconstructed with COATJAVA. How the test particle was reconstructed was compared to how the particles in the simulation were. The existing test particle was shown to have a V_z more than eight sigma from the mean for particles generated under the same conditions meaning it was an atypical particle event. The preprogrammed data representing the existing test particle therefore needed to be replaced because an atypical particle is a bad representative to base the unit test on. A new set of preprogrammed data was selected from the simulated data representing a new test particle. This particle was selected so that it would have momentum and vertex x,y and z components within 0.5 sigma from the mean of the other particles in the simulation.

In order to select new unit test tolerance limits, the reconstruction momentum and vertex resolution was studied. These resolutions were determined by comparing how particles were reconstructed to how they should have been based on how they were generated. The new unit test limits were selected with the width of the resolution around where the new test particle is reconstructed. The unit test was expanded to examine the test particle's reconstructed vertex as well as its momentum.

An important step in the reconstruction process is the interpolation of reconstruction crosses that characterise the path of particles between DC superlayers. The placement of reconstruction crosses in the reconstruction of the simulated data was compared to how they should have been based on how the particles were generated. From this comparison the cross placement resolution was extracted. The unit test was expanded to check that reconstruc-

tion crosses are positioned accurately to within the resolution as an additional indicator of a potential software failure. The updated DC unit test has since been included in the latest development builds of COATJAVA as of December 2019.

2.9 Future Research

The COATJAVA unit test associated with particle reconstruction in CLAS12's calorimeters remains outdated, consistently returning false positives as the DC unit test used to. While the DCs are designed for spatial and momentum tracking and the calorimeters for timing and energy, they both work by examining the properties of a reconstructed test particle. It is likely the calorimeter test has begun to fail for similar reasons as the DC, with the test tolerance limits becoming outdated as newer versions of the reconstruction changed how the test particle is reconstructed. Future research into updating the calorimeter unit test could benefit by applying the same framework of analysis used to understand the DC reconstruction resolution. (Producing simulated data and comparing the generated particle properties to how they were reconstructed).

Chapter 3

Hough Transform Track Recognition

3.1 Central Detector

Essential parts of CLAS12's proposed physics program (in particular the extension of previous measurements of GPD's) require precise tracking and identification of low momentum particles [9]. Most of the momentum of incident beam electrons is directed along the beamline and so scattering products travelling through wide angles from the beamline are expected to have low momentum.

As described in Table 1.1, the CD was designed with a wide acceptance ($40° - 135°$) to detect the identity, momentum and trajectory of these low momentum particles. The CD detector takes the shape of barrel made up of concentric layers. From the CD's centre most layer outwards (displayed in Figure 3.1), these layers are the CD's Silicon Vertex Tracker (SVT), MicroMegas (MM), Central Time-Of-Flight (CTOF) scintillators, solenoid magnet and Central Neutron Detector (CND).

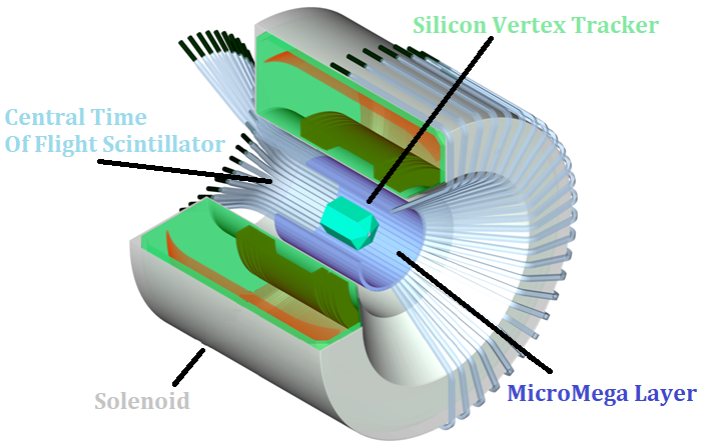


Figure 3.1: Diagram of CD showing SVT, MM, CTOF and solenoid. CND not shown, but encloses solenoid magnet. Adapted image from [20].

The CND and CTOF are designed to fulfil the CD's particle identification requirement and to provide timings for TBT reconstructions while the SVT and MM layers track particle trajectories (see appendix 1 for more detail on the design, functioning and purpose of CND and CTOF) [15] The momentum of charged particles is extracted from these trajectories by analysing how they are influenced by the solenoid's magnetic field. CLAS12's tracking systems are required to achieve an angular resolution of at least $\Delta\theta = 10\ mrad$ and $\Delta\phi = 6\ mrad$ with a momentum resolution of $\Delta P/P = 3\%$. The CD's tracking systems must also maintain a >95% detection efficiency with beam luminosities up to $10^{35} cm^{-2} s^{-1}$, all while working within the limited space available within the solenoid [15].

3.1.1 Silicon Vertex Tracker

The SVT (shown in Figure 3.2) is made up of concentric layers of double sided silicon semiconductor paddles. Passing charged particles disturb the fast responding semiconductor structure, resulting in the production of electric signals that indicate the passage of a particle. The SVT captures and sends these electric signals at a rate on the order of 42kHz. With the increased luminosity made available with the 12GeV upgrade, the maximum foreseen

physics event rate is 20MHz spread across the whole CD and so the SVT response rate is fast enough achieve the >95% detection efficiency requirement [21].

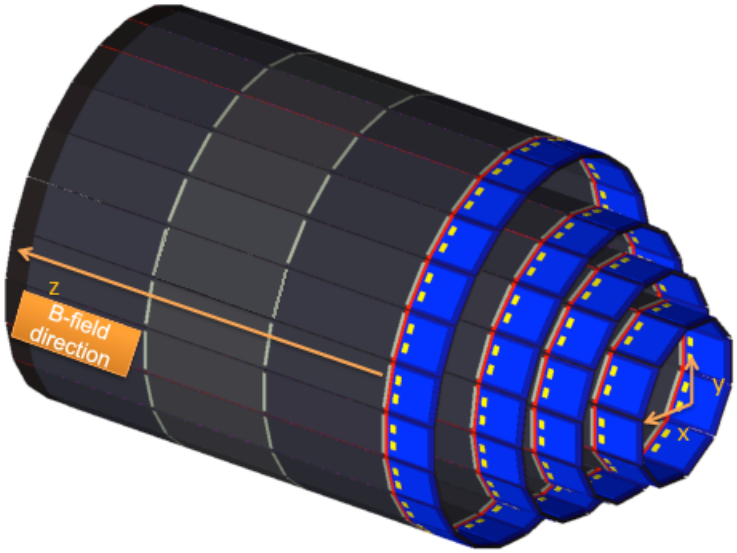


Figure 3.2: CAD image of CLAS12 SVT concentric layer barrel shape with CLAS12 coordinate system orientation indicated [21].

3.1.2 MicroMegas

A MM is a gaseous detector built of a series of parallel electrode plates that produce powerful electric fields and a set of micro-strips for signal readout. CLAS12's MM's contain ionisation cavities filled with a mixture of 19:1 Argon-iC_4H_{10} gas, which incident particles ionise as they pass. As shown in Figure 3.3, the electrons ionised from these gas molecules are accelerated by a 0.8 kV/cm electric field towards a signal amplification region. In this region the electrons are accelerated through a more powerful electric field with a strength of 40 kV/cm. These incoming ionised electrons ionise more gas molecules with the additional energy they have gained. The secondary ionized electrons go onto ionize further electrons, resulting in an avalanche of gas molecule ionisation. This now amplified electron signal travels into the sensory silicon semiconductor micro strips which convert it into an electronic signal.

The strength of these electric fields is sufficient that the MM to captures particle event data at a rate 22kHz, which is fast enough to meet the 95% detection efficiency requirement [22].

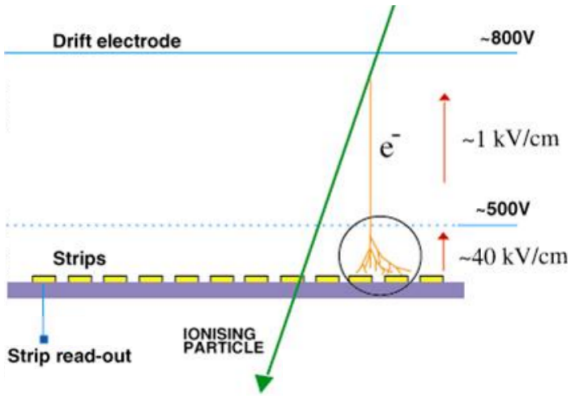


Figure 3.3: Diagram of MM operating principle. Adapted version of image from [22].

Table 3.1 contains the design resolutions intended for the SVT and MM and shows that a combination of two SVT layers and 3MM layers is sufficient to achieve the momentum and tracking resolutions required of the central tracking systems (listed in table 1.1) [22]. CLAS12's final design includes 5 MM and 3 SVT layers with the aim that required tracking precision could be exceeded [10].

	4x2 MM	**4x2 Si**	**2x2 Si + 3x2 MM**
Momentum resolution dp/p (%)	2.9	2.1	1.6
θ resolution (mrad)	1.3	15.1	1.4
φ resolution (mrad)	10.9	2.9	2.6
Spatial resolution (μm)	212	1522	267

Table 3.1: Table of resolutions achieved by different combinations of SVT and MM layers [21].

3.1.3 Solenoid

CLAS12's solenoid magnet (featured in Figure 3.1) encloses the SVT and MM in an approximately homogeneous 5.0T magnetic field [19]. The solenoid has three primary functions. Firstly, it sustains its powerful magnetic field uniformly with $\Delta B/B < 10^{-4}$ to facilitate the operation of CLAS12's dynamically polarised target for measurements of interactions with spin dependent cross-sections (see appendix 1 for more detail on target). Secondly, it provides a powerful guiding effect that keeps potentially damaging Möller electrons from sensitive tracking detectors and instead directs them down the beamline into a high-z absorption shield [23]. Finally, the solenoid maintains its powerful magnetic field to achieve momentum analysis. The field has to be sufficiently strong so that a momentum resolution of at least $\Delta p/p = 3\%$ is achieved, all while close to the beamline and within the limited space available [23].

3.2 Central Detector Circle Reconstruction

While the CD's tracking systems and their associated reconstruction software have already successfully achieved the particle recognition, trajectory and momentum reconstruction resolutions required by the original specifications, the Collaboration has begun investigating previously unconsidered reconstruction methods that may further improve these qualities.

Charged particles scattered into the CD are accelerated in a helical motion while within the solenoid field. As these particles interact with the SVT and MM layers they leave patterns of detector hits which appear circular when viewed along the direction of travel. It was proposed by the CLAS12 data reconstruction coordinator (V. Ziegler), that these circular trails that representing charged particle tracks (such as the example track in Figure 3.4) could be identified by applying a digital feature extraction technique called the Hough Transform.

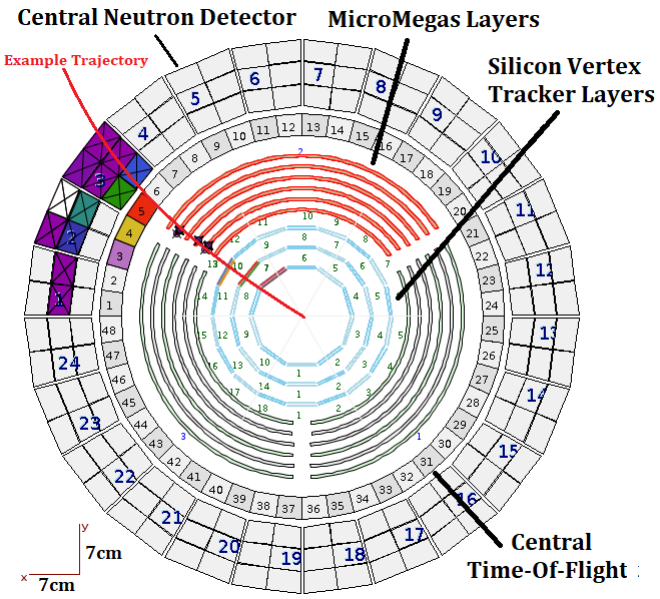


Figure 3.4: CED image of partial circle of hits left by the curving trajectory of an example high momentum charged particle. Coloured components indicate detector hits, solenoid field directed into page.

3.3 Hough Transform

The Hough transform was invented in 1962 by Paul Hough for the identification the trails left by particles in bubble chamber photographs [24]. In order to explain how the Hough transform may be applied to the identification of circles in the CD, a simpler case identifying lines is first discussed.

3.3.1 Linear Recognition

Consider a series of points $(x_i, y_i), (x_i, y_i)....(x_n, y_n)$ which lay along the line in Figure 3.5 with equation $y = Ax + B$. The variables x and y describe every point possible in Cartesian space while the specific parameters (A,B) define the specific line that constrains those possible points to only the ones that lay along it. Consider a wider, more general parameter space containing all possible $(a_i, b_i), (a_j, b_j)....(a_n, b_n)$ combinations and describing all possible lines with a different line at every (a,b) point. Going back to Figure 3.5, if this

case were to be inverted and instead all possible points in this parameter space (a, b) were constrained by specific lines defined by the (x_n, y_n) points with equations $(b = -x_i a + y_i)$, $(b = -x_j a + y_j)...(b = -x_n a + y_n)$ the results could be plotted as in Figure 3.5. The intersections of those specific (x,y) lines would be at the point (A,B) in parameter space, which defines the original line in Figure 3.5 on which all the (x,y) points lay.

To summarise, by taking the original points laying on a line y=Ax+B in Cartesian space, transforming them to represent lines in the (a,b) parameter space and identifying that they intersect at (A,B) in parameter space, the line on which all the original (x,y) points in Cartesian space lay is identified as being parametrised as (A,B).

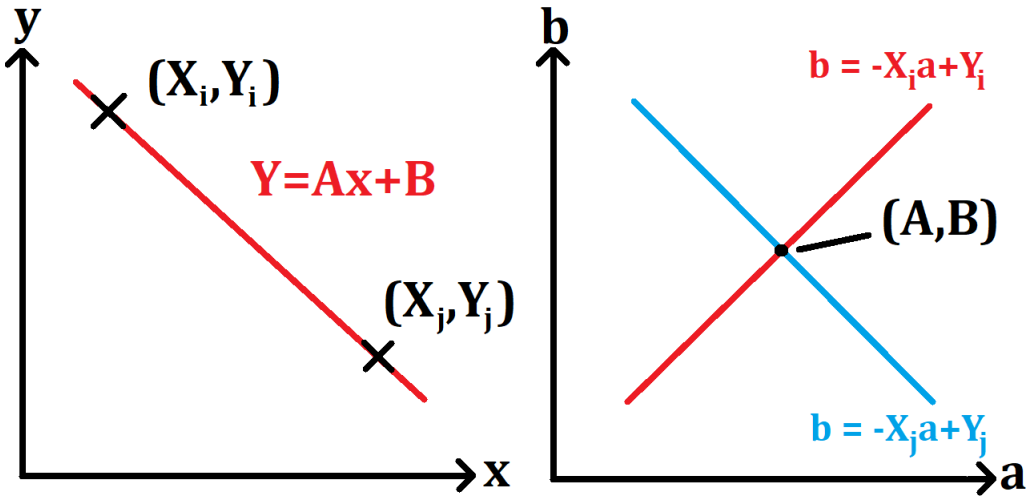


Figure 3.5: Illustration of Linear Recognition using Hough Transform, Points from Cartesian space (left) transformed to represent lines in parameter space (right), intersections of lines represents parameters that define the line on which the original points mutually lay.

3.3.2 Circle Recognition

The method applied in the linear case can be adapted more generally for the identification of curves, ellipses and circles with an appropriate choice of parametrisation. The circle of points in Figure 3.6 can be described with the classic circle equation $(x - A)^2 + (y - B)^2 = R^2$ with radius R and centre (A,B). Transforming the (x,y) points into parameter space produces a set of equations following the form $b = y_n - \sqrt{R^2 + 2ax_n - a^2 - x_n^2}$. The plot of

these equations for each point in Figure 3.6 shows that there are multiple parameter space intersection points. This arises because while there is only one possible circle on which all three lay (represented by the triple intersection point), there are other circles that pairs of points may lay upon (represented by the other intersection points). In order to identify the primary circle a hierarchy of intersection points and the circles they identify is created. The circle parametrised by the point in parameter space with most intersections (the triple intersection in Figure 3.6) is selected as the primary circle. This is because (as Figure 3.6 shows) circles made up of many points are better spatially (x,y) and radially (R) constrained by those additional points and so better represent the original circle.

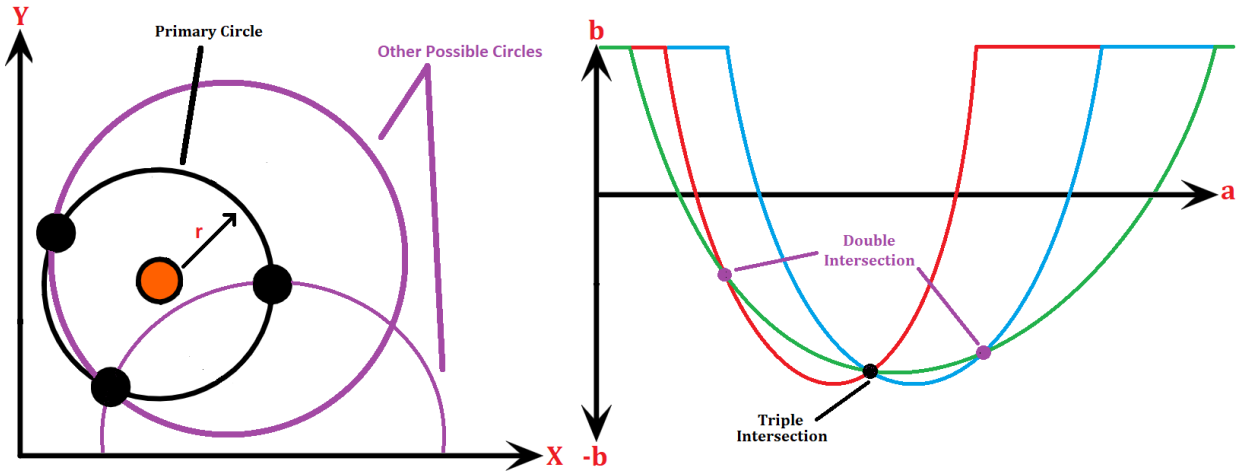


Figure 3.6: Illustration of circular Cartesian (x,y) and parameter (a,b) space diagrams. Triple intersection (right) represents to combination of parameters which describe primary circle on which all three points (left) lay. Double intersection points represent the other possible circles (left).

3.4 Circle Recognition in the Central Detector

In order to apply the Hough Transform to detector data from CLAS12's central tracking systems, the positions (and associated uncertainties) of SVT and MM detector hits are readout as points in the xy plane. Two additional complications arise when detecting circles in this data. Firstly, the radius of the circles that charged particles will follow depends upon their momentum and so is unknown. Instead of starting with a fixed radius then, the

implementation of the Hough Transfrom written by V. Ziegler iterates through a range of possible radii based upon the momenta expected for particles scattered into the CD.

Secondly, the detector hits don't lay on exact circles due to the limited resolution of the components and due to false hits that arise due to background noise. In order to accommodate these variations a tolerance width is introduced. The positions of false hits often lay far from any potential circle and so are easily distinguished and removed. The Taubin circle fitting algorithm is used to fit the sets of imperfectly circular hits recognised by the Hough Transform so that a circle can be selected to characterise the particle track being reconstructed. These circle fittings produced by the Taubin fit are ranked based on reduced chi squared (χ^2_ν) and the fitting that minimises χ^2_ν most successfully is selected as the primary circle.

3.4.1 Taubin Circle Fitting

The Taubin fitting algorithm is a non-iterative algebraic circle fit. it works by minimising the function $\mathcal{F}_T$ for n points at positions $(x_i, y_i), (x_j, y_j)....(x_n, y_n)$ with fitting parameters A,B,C and D [25].

$$\mathcal{F}_T = \sum_i^n \frac{[Az_i + Bx_i + Cy_i + D]^2}{n^{-1}\sum_i^n [4A^2(z_i) + 4ABx_i + 4ACy_i + B^2 + C^2]} \tag{3.1}$$

(where $z_i = x_i^2 + y_i^2$)

This problem is more easily approached in matrix representation, beginning by defining a parameter matrix $\mathbf{A} \equiv$ (A,B,C,D) and data matrices $\mathbf{Z}$ and $\mathbf{M} \equiv \frac{1}{n}\mathrm{Z}^T Z$ as in Equation 3.2.

$$\mathbf{Z} \equiv \begin{bmatrix} z & x_1 & y1 & 1 \\ : & : & : & 1 \\ : & : & : & 1 \\ z_n & x_n & y_n & 1 \end{bmatrix} \tag{3.2}$$

With these definitions the problem can be restated as the minimisation of the objective function $\mathcal{F}_{\mathcal{T}}$ in Equation 3.3.

$$\mathcal{F}_{\mathcal{T}} = \mathbf{A}^T\mathbf{M}\mathbf{A} \tag{3.3}$$

$$\mathbf{T} \equiv \begin{bmatrix} 4\overline{z} & 2\overline{x} & 2\overline{y} & 0 \\ 2\overline{x} & 1 & 0 & 0 \\ 2\overline{y} & 0 & 1 & 0 \\ 0 & x_n & y_n & 0 \end{bmatrix} \; where\, \overline{x} = \frac{1}{n}\sum_{i}^{n} x_i \; and \; for \; \overline{y},\, \overline{z} \tag{3.4}$$

The newly restated minimisation problem in Equation 3.3 is constrained by the Taubin fitting matrix $\mathbf{T}$ defined in Equation 3.4 [25]. In order to solve this problem, a Lagrange multiplier η is introduced so it can be reduced to the unconstrained minimisation function form in Equation 3.5 [25].

$$\mathcal{G}(A, \eta) = \mathbf{A}^T\mathbf{M}\mathbf{A} - \eta(\mathbf{A}^T\mathbf{T}\mathbf{A} - 1) \tag{3.5}$$

Differentiating with respect to $\mathbf{A}$ returns Equation 3.6 which shows that $\mathbf{A}$ is simply a generalized eigenvector for the matrix pair ($\mathbf{M}$,$\mathbf{T}$).

$$\mathbf{M}\mathbf{A} = \eta\mathbf{T}\mathbf{A} \tag{3.6}$$

The Taubin fitting algorithm operates by reading the positions of detector hits into the data matrices $\mathbf{Z}$ and $\mathbf{M}$ and solving Equation 3.6 to find the optimum combination of parameters in $\mathbf{A}$, from which can be calculated the A, B and R parameters that describes the circle of best fit for the input points.

3.5 Hough Transform Recognition Suitability Analysis

A version of the Hough transform applied to the recognition of circles was written and implemented into COATJAVA by Veronique Ziegler (see Appendix 4 for Hough Transform source code). The patterns of hits that particles leave are not perfectly circular due to the spatial resolution of detector components and background noise that can result in spurious hits that the transform needs to distinguish and exclude. A tolerance is introduced within which points that lay slightly off a potential circle are still considered to be aligned along its curve. Particles scattered into the central detector are expected to have low enough momentum that the radii of the circles of hits they leave in the CD is often around 50 cm.

By generating a circle of 100 points equally distributed around the circumference of a circle centred at zero with a radius of 50 cm, the efficiency E_{Hough} with which the Hough Transform identifies points aligned along the circle can be analysed. In order to understand the effect of sample noise on E_{Hough}, these points were randomly shifted in the xy plane to simulate normally distributed Gaussian noise with a standard deviation σ (as in Figure 3.7) and E_{Hough} was plotted as a function of σ in Figure 3.8.

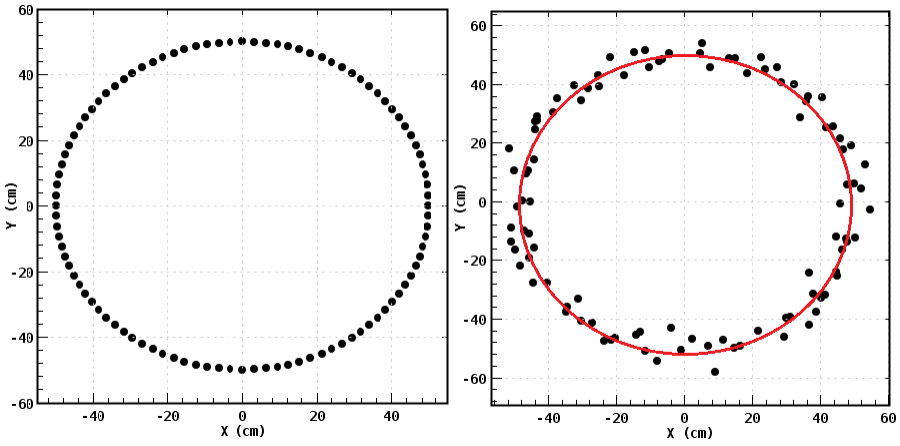


Figure 3.7: 100 point circle simulation centred around origin with radius 50cm (left) with an example of points (right) shifted to a degree randomly selected from a normal distribution (In this case with $\sigma = 3.0\ cm$).

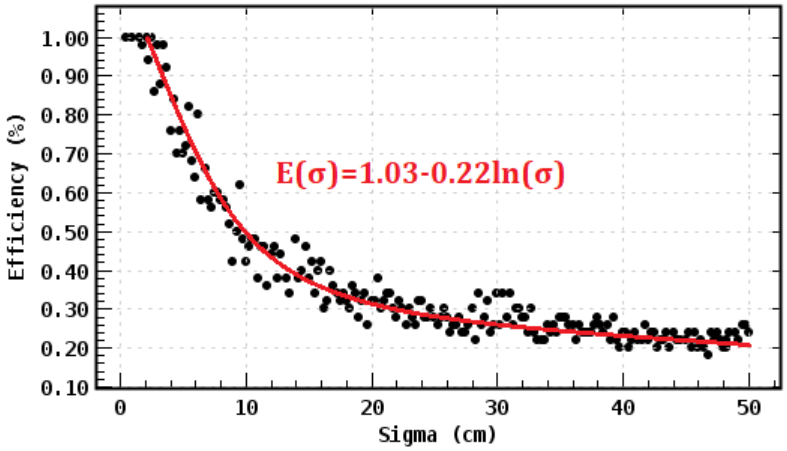


Figure 3.8: Hough Transform recognition efficiency as a function of increasing Gaussian noise σ in simulated data points. Logarithmic efficiency decay fit with function $E_{Hough}(\sigma) = 1.03 - 0.22ln(\sigma)$ with $\chi^2_\nu = 1.33$

The efficiency with which the Hough transform recognises points in noisy circle data was best fit by the function $E(\sigma) = 1.03 - 0.22ln(\sigma)$ with reduced chi squared $\chi^2_\nu = 1.33$. E_{Hough} follows a logarithmic decay, at first sharply decreasing as the sample becomes noisier until approximately $\sigma = 10cm$ where it slows and later begins to plateau. The spatial resolution of the SVT and MM from Table 3.1 are 0.15cm and 0.02cm respectively, and if for example a neighbouring SVT paddle happens to trigger and register the hit instead, the position would be shifted only much as 4.2 cm in the xy plane. Figure 3.8 shows for shifts this small the Hough transform is adequate at recognising points aligned along circles in noisy data, with E_{Hough} only dropping below 50% for values of sigma more than 10cm.

Particles usually leave $\approx$ 20 hits in the CD, though this can vary greatly depending upon the event. Low momentum particles may circle multiple times before exiting the CD and leave many more hits than a high momentum particle might. A high momentum particle may only leave a partial circle of hits as it leaves the CD such as the example trajectory shown in Figure 3.4.

The previous analysis is based upon a sample of 100 points and so it is important to understand if the chosen number of points has an effect on E_{Hough}. By fixing the noise to $\sigma = 3cm$, E_{Hough} as a function of the number of points in the sample was plotted in Figure 3.9. E_{Hough} remains spread around 85% as the number of points increases from 3 to 120 in Figure 3.9. This suggests that for data samples with noise on the order of $\sigma = 3cm$, there is no correlation between E_{Hough} and sample size. A 100 point sample for the E_{Hough} noise analysis is therefore an acceptable sample size to test the suitability of the Hough Transform for detector data particle track recognitions, despite the wide variation in the number of hits (data points) a detector data sample might have.

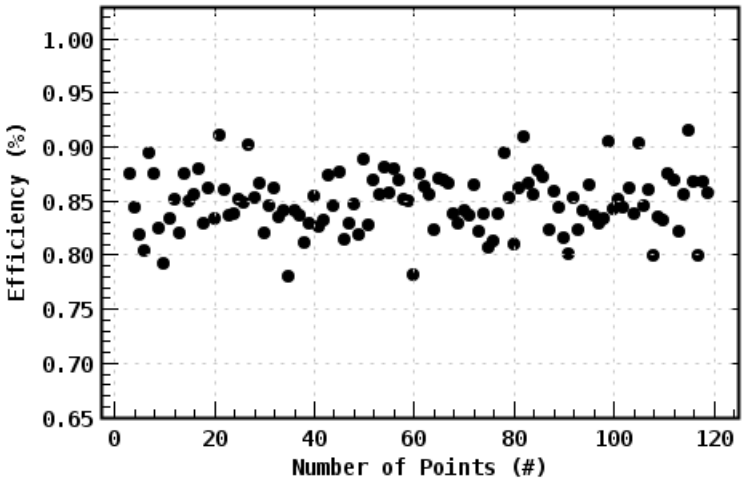


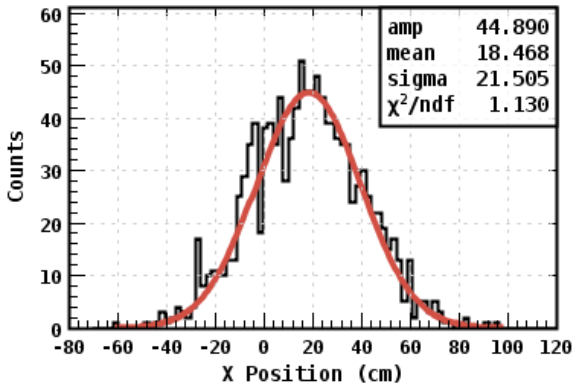
Figure 3.9: Hough Transform point recognition efficiency against sample size, extremely no correlation between E_{Hough} and the number of points in the sample.

While E_{Hough} remains high even for as little as three points, the circle which the Hough Transform has identified may not be the original circle. Positional uncertainties in the data and detector noise may have distorted the shape of the recognised circle. In order distinguish between the many sets of aligned points the Hough-Transform identifies and the circles they may represent, a version of the Taubin circle fitting algorithm was written. (see Appendix 5 for the source code of the Taubin circle fit and χ^2 analysis that was used for this purpose).

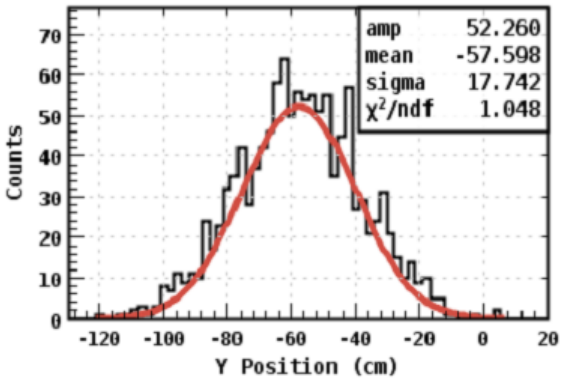
3.6 Hough-Taubin Combination Reconstruction Resolution Analysis

In order to understand how effective the Hough-Taubin combination is at particle reconstruction, the spatial and transverse momentum resolutions were analysed. The combination was applied to the reconstruction of detector data produced using GEMC. The CD is optimised to reconstruct charged particles with 900 MeV/c transverse momentum [15]. The simulated data represented 2000 electrons scattered from the target at $\theta = 85 \pm 1^{\circ}$ and $\phi = 15 \pm 1^{\circ}$ through the CD systems with a transverse momentum of 900 MeV/c. These electrons where generated so that they spiral through a helix centred at x=18.3 cm y=-57.4 cm.

The simulated detector data was read into V. Ziegler's implementation of the Hough Transform using a data handling architecture written in JAVA by M. Armstrong. The sets of points recognised by the Hough Transform as aligned along imperfect circles were fit with circles using the Taubin circle fit written in JAVA for reduced chi squared (χ^2_ν) analysis. The circle centres and transverse momenta predicted by the of the circles selected by the χ^2_ν analysis to best describe the reconstructed particle tracks were used to produce the histograms in in Figures 3.10a, 3.10b and 3.11.



(a) X position of helix centre reconstructed from simulated data.



(b) Y position of helix centre reconstructed from simulated data.

Figure 3.10: Histograms of reconstructed circle centre xy positions of circular component of simulated charged particles.

Figure 3,10a shows that the x coordinate of the centre of the circular component of the simulated particle helices was reconstructed with a mean of 18.5 cm. The uncertainty in this mean ($\frac{\sigma}{\sqrt{n}}$) was 0.48 cm while the resolution (1σ) was 21.5 cm. Similarly, Figure 3,10b shows the y position mean was -57.6 cm with an uncertainty of 0.40 cm and 17.7 cm resolution. In both cases the x and y position were on average reconstructed accurately to within the uncertainty of each mean. The diameter of the CD is only 70 cm however and so as in both cases the width of the resolution is on the same order of magnitude, the spatial resolution of these reconstructions is poor.

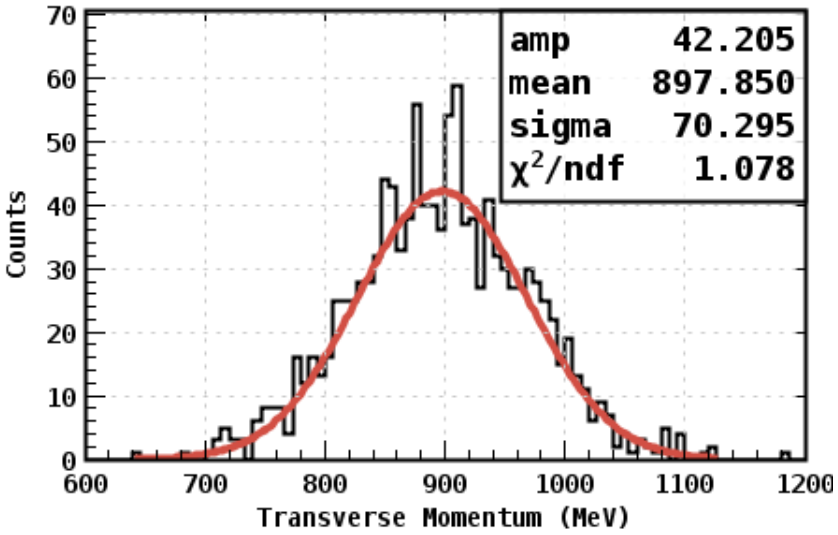


Figure 3.11: Histogram of reconstructed GEMC simulation transverse momentum.

Figure 3.11 shows the transverse momentum was reconstructed with a mean of 897.85 GeV/c. The uncertainty in this mean is 20.08 GeV/c with a resolution of 70.30 GeV/c (or $\Delta p/p = 7.8\%$). The CD was designed to achieve a resolution of $\Delta p/p = 3\%$ for charged particles at this momentum. With a difference of only 4.8%, the Hough-Taubin method shows promise in its ability to reconstruct transverse momentum at this early stage.

It is possible that the trails of hits left by the charged particles have a dependency on the longitudinal (z) component of the helical motion due to imperfections in the CD solenoid field [15]. This may cause the patterns of hits to shift in the xy plane as they travel through the detector and resulting in overlapping circles of hits. The influence of this shift on the reconstruction has not been accounted for. The Hough-Taubin implementation may be struggling to identify which hits belong to which overlapping circle and so the reconstructed circle centre may be shifted. A longitudinal dependency on xy motion may therefore be one of the sources of error in reconstructing the centres of the circular components of helices.

Furthermore, improperly grouping sets of hits from different overlapping circles may lead to stretched circles with different radii. It is from the radii that transverse momenta is extracted and so this could be a source of error in the momentum reconstruction.

3.7 CD Conclusion and reflection

CLAS12's CD is designed be able to reconstruct trajectory and momentum with a momentum resolution of $\Delta P/P = 3\%$ with $\Delta\theta = 10\ mrad$ and $\Delta\phi = 6\ mrad$ using conventional linear tracking. V. Ziegler (CLAS12 software coordinator) proposed that a new method of particle track recognition could be evaluated for its potential to improve upon these resolutions. The Hough transform was applied to the recognition of the patterns of detector hits charged particles leave as they travel through the CD. The Taubin circle fitting method was applied to select the circle most likely to best represent the particle track from the many potential circles the Hough transform may identify.

An implementation of the Hough transform was written by V. Ziegler and included in the Collaborations particle reconstruction and analysis software package COATJAVA. The ability of the Hough transform to identify points belonging to these patterns od detector hits despite a simulated sample noise was investigated. The efficiency with which points where identified was found to decay logarithmically for noisier samples. For the small degree of random noise that might shift detector hits in the CD, the Hough transform was found to be adequate at recognising circular patterns. An architecture for reading detector data into the V. Ziegler's Hough Transform and a version of the the Taubin circle fit was written for the purpose of reconstructing detector data.

The spatial and transverse momentum resolution of this reconstruction was analysed by producing and reconstructing simulated detector data using GEMC. The transverse momentum resolution was found to be $\Delta p/p = 7.8\%$ which compares well to the total momentum reconstruction resolution achieved by conventional linear tracking at $\Delta P/P = 3\%$ and shows the Hough-Taubin method has promise with transverse momentum reconstruction even at this early stage. The spatial resolution was found to be poor with the of x and y coordinate reconstruction resolution of the centre of charged particle helices ($\Delta x = 21.5\ cm$ and $\Delta y = 17.7\ cm$) on the same order of magnitude as the diameter of the CD.

3.8 Future Research

It is possible that the positions of the detector hits in the xy plane that make up the circular components of the helical patterns left by charged particles in the CD have a dependence on the longitudinal (z) component of position. This would cause the trails of hits to be shifted in the xy plane as particles traverse the CD. Future research could include investigating the extent of the influence this or other effects may have upon the spatial and momentum resolution of the Hough-Taubin combination and improving upon the current implementation by accounting for these effects. This may be achieved by introducing a bias towards circles fitted to sets of points with similar z position as these sets are less likely to include points accidentally included from other overlapping circles. The version of the Hough transform could also be rewritten fit full helices to the data rather than circles to the circular component of the helices.

Chapter 4

Bibliography

[1] National Science Advisory Committee, "A Long Range Plan for Nuclear Science", 1979, URL: clasweb.jlab.org/wiki/images/c/ce/science.energy.gov/np/nsac/.

[2] M. Riordan, "The Discovery of Quarks", 29th May 1992, Science Vol 256, Issue 5061, pp 1287-1293, DOI: 10.1126/science.256.5061.1287.

[3] R. C. Fernow "Introduction to Experimental Particle Physics", 1989, Cambridge University Pres, ISBN 978-0-521-379-403.

[4] C.Westfall, "Jefferson Lab's 1985 Switch to Superconducting Accelerator Technology", 1996, CEBAF-PR-96-29, Michigan State University

[5] J. Haimson, "Non-Cyrogenic High Duty Cycle Electron Linear Accelerators", Proceedings of 1968 Proton Linear Accelerator Conference, Upton, New York.

[6] F. Gross, "Making the case for Jefferson Lab", (2011) Journal of Physics, Conference Series, Vol 299, pp 1-20.

[7] H. Grunder, "The Continuous Electron Beam Accelerator Facility", Proceedings of 1988 Linear Accelerator Conference, Williamsburg, Virginia.

[8] C. E. Reece, "Continuous Wave Superconducting Radio Frequency Electron Linac for Nuclear Physics Research" December 2016, Phys. Rev. Accel. Beams 19 124801, DOI: 10.1103/PhysRevAccelBeams.22.129901.

[9] J. Arrington, "The Science and Experimental Equipment for the 12GeV Upgrade of CEBAF", 10th January 2005, Jefferson Lab.

[10] Jozef Dudek et al, "Physics opportunities with the 12GeV upgrade at Jefferson Lab", December 2012, European Physical Journal A, DOI:10.1140/epja/i2012-12187-1.

[11] M. Peskin, D. Schroeder, "An introduction to Quantum Field Theory", (1995), Westview Press, ISBN: 978-0-201-50397-5.

[12] W. Greinerm, S. Schramm, E. Stein "Quantum Chromodynamics", April 2007, 3rd Edition, Springer Berlin Heidelberg, ISBN 978-3-540-48525-3.

[13] M. Garcon, "An introduction to the Generalized Parton Distributions", 17th October 2002, Saclay Nuclear Research Centre, https://arxiv.org/pdf/hep-ph/0210068.pdf.

[14] C.F. Perdrisat, V. Punjabi, M. Vanderhaeghen, "Nucleon electromagnetic form factors" Progress in Particle and Nuclear Physics, October 2007, Volume 59, Issue 2, pp 694-764. DOI:10.1016/j.ppnp.2007.05.001.

[15] D. abbott et al, "Hall B 12 GeV Upgrade Preconceptual Design Report", August 2006, Preconceptual Design Report, Jefferson Lab, URL:https://www.jlab.org/div_dept/physics_division/pCDR_public/Hall_B/CDR-HallB-PAC23.pdf.

[16] V. Burkert, "Hall B 12 GeV Upgrade", Proceedings of August 2006 Jefferson Lab PAC30 Meeting, Newport News, Virginia, USA, https://www.jlab.org/exp_prog/PACpage/PAC30/talks/PAC30_HallB_12GeV.pdf

[17] V. Ziegler, "CLAS Software Overview", 23rd Feb 2016, Proceedings of CLAS12 collaboration meeting, Jefferson lab, Newport News, Virginia, USA, URL:https://www.jlab.org/indi

[18] V. Ziegler, "CLAS12 Software Document, Hall B 12 GeV Upgrade", 22nd November 2013, Version 1.4, URL:https://www.jlab.org/Hall-B/clas12-software-nov13.pdf.

[19] D. Kashy, "CLAS12 TORUS Magnet", March 2017, Jefferson Lab, URL:https://www.jlab.org/Hall-B/clas12-web/specs/torus.pdf.

[20] S. Niccolai, "Neutron Detector for the Central Part of CLAS12", 8th april 2008, Presentation for GDR Instrumentation, IPN Orsay.

[21] Y. Gotra, "Silicon Vertex Tracker for CLAS12 experiment", 2016, Proceedings of Science, 25th International Workshop on Vertex Detectors, La Biodola, Isola dâĂŹElba, Italy, URL:https://pos.sissa.it/287/008/pdf.

[22] S. Aune, D. Calvet, M. Combet, E. Delagnes, C. Lahonde-Hamdoun, S. LHenoret, O. Meunier et al, "CLAS12 Micromegas Vertex Tracker Feasibility", May 2009, CEA Saclay, URL:http://irfu.cea.fr/dphn/Csts/jun2009/Ball_CLAS12_Annexes.pdf.

[23] CLAS12 Collaboration, “CLAS12 Technical Design Report”, 15th July 2008, Version 5.1, Jefferson Lab, URL:https://www.jlab.org/Hall-B/clas12_tdr.pdf.

[24] P.V.C Hough, ”Method and means for recognising complex patterns”, 18th December 1962, US Patent 3,069,654.

[25] A. Al-Sharadqah, Nikolai Chernov, “Error Analysis for Circle Fitting Algorithms”, 2003, Electronic Journal of Statistics, Volume 3, 886-911, URL:https://projecteuclid.org/euclid.ejs/1251119958.

Chapter 5

Appendices

5.1 Appendix 1: Literature Review

Research Year Literature Review

Michael Armstrong, marmstr@jlab.org

July 30, 2019

Abstract

Developments in accelerator and detector technology have over the decades dramatically improved our understanding of nuclear and particle physics by enabling physicists to test the predictions of their models at ever higher energies and with ever greater precision. Jefferson Labs Continuous Electron Beam Facility (CEBAF) has recently been upgraded, doubling the maximum beam energy it can deliver its scientific user community to 12GeV. The CEBAF Large Acceptance Spectrometer (CLAS12) has been upgraded to make use of these higher available energies with experiments investigating the quark-gluon structure and behaviour of hadrons. This review provides an overview of the CEBAF and CLAS12 facilities at 12 GeV and explains the motivation for the upgrade. The physics behind the scintillators, photomultiplier tubes, drift chambers, Cherenkov counters and calorimeters that make up CLAS12's major detector systems is introduced, providing important background for the research I will conduct in my placement year. The CLAS12 collaboration has developed a set of software tools with which they analyse their detector data. The focus of my research is presented as it relates to the development and validation of these software tools.

I. INTRODUCTION

The Thomas Jefferson National Accelerator Facility (JLAB) was originally conceived in part as a response to the burning questions raised by the confirmation of the existence of quarks at the Standford Linear Accelerator (SLAC). [1] In its 1979 "Long Range Plan for Nuclear Physics" the US Nuclear Science Advisory Committee (NSAC) highlighted the need for a "continuous beam, high energy accelerator" that could probe the interior structure of nucleons in electron scattering experiments. Electrons accelerated beyond the 2GeV would have sufficiently short enough De Broglie wavelength to resolve this interior structure and open up quarks, gluons and their interactions to study. NSAC proposed that quark behaviour could be indirectly observed at these energies in exclusive and semi-inclusive measurements. An incident electron may interact with its target nuclei in many different ways, producing many unique combinations of reaction products. In exclusive and semi-inclusive measurements all (or most) of the scattering products, rather then just the scattered electron are captured and are used to identify the specific interaction. These measurements would mean rare interactions of particular interest could be distinguished and their cross sections (a measure of relative probability) extracted from the wider data set. [2]

NSAC argued that a high duty factor would be required to separate multi-particle states and a beam current of $\approx 100\mu A$ would be needed to yield enough particle events for the effects of statistical fluctuations to be minimised and precise measurements produced.[2] Accelerators of the time such as SLAC depended upon copper current carrying cavities to produce the electric fields with which they accelerated particles. The current cavities were only powered for short periods as they would melt under continuous use. The ratio of pulse-time to time between pulses is known as the duty factor which for SLAC was 10^{-4} at the time.[3] With a high duty factor, the proposed accelerator would provide an almost continuous electron beam with enough time between incident beam electrons for exclusive measurements of the products before secondary electrons could interfere.[2] This accelerator would later become JLab's CEBAF facility, beginning construction in 1987. CEBAF would initially provide a maximum beam continuous electron beam energy of 4GeV at up to $150\mu A$ and a duty factor of 1.5×10^{-2}.[3] CEBAF would later be upgraded with to a maximum 6 Gev beam energy and would cease operation in 2012 for its latest upgrade to 12GeV.[4]

II. CEBAF AT 12 GeV

The original CEBAF design had been accepted in part by the US Department of Energy because its design showed promise that it could be "readily upgraded". The push for 12 GeV began as early as 1985 when it was agreed that quark probing would be best achieved at "8-12 GeV" at a scientific consultation at MIT.[5] This proposal was ultimately impeded by the limitations of the accelerator technology of the day and the additional cost it would incur. A decade later however, NSAC would issue its support for the 12 GeV upgrade citing "favourable technical developments" in accelerator technology and "foresight in the design of the original facility". [3] JLab's pre-conceptual design report would identify many research opportunities offered by these higher beam energies, highlighting three of considerable interest. Firstly, a maximum 12 Gev beam energy could verify of falsify the presence of exotic mesons.[6] In Quantum Chromo-Dynamics (QCD) colour-charge-carrying particles are treated as excitations of more fundamental underlying fields. The interactions between particles are described by interaction terms in the Lagrangians that describe their corresponding fields. Colour-carrying flux tubes are thought to form between gluons describing their mutual interactions. The excitations of these flux tubes give rise to short lived exotic hydrid mesons, the properties of which hold clues to the origins of colour confinement.[7] Colour confinement is an important prediction of

QCD wherein colour charged particles such as quarks and gluons cannot be isolated (and therefore independently observed) outside of extreme conditions such as above the Hagedorn temperature of $2 \times 10^{12} K$.[8][9] 12 Gev electron beam energies would enable the observation of these exotic hybrid mesons in 9 Gev pion photoproductions. Secondly, increased beam energies would enable extensions of previous measurements of Generalised Parton Distributions (GPD) across wider kinematic ranges. A GPD describes the probability density of finding hadron constituent particles (termed "partons") with a particular quantum observable such as longitudinal or transverse momentum fraction across a four momentum squared Q^2 scale. GPDs reflect parton momentum and spatial distributions and so measuring them with higher beam energies is important for understanding the properties and interactions of hadronic constituent matter on smaller scales. Finally, increased beam energies would also enable extensions of existing measurements of form factors across wider kinematic ranges. At low energy, form factors represent the charge and magnetization current density distributions within nucleons. At 12 GeV beam energies electron scattering experiments could probe deeper into the interior structure of nucleons, revealing how this charge and magnetisation current is distributed between the constituent quarks.

CEBAF's 12GeV upgrade would be completed in October 2017, successfully producing nearly continuous electron beams with a 1.5×10^{-2} duty factor at up to 12.065 GeV and 180 μA.[4] CEBAF can be split into four major systems. The injector system shown in figure-2.1 deposits electrons into a racetrack shaped recirculation ring. Electrons are accelerated along the straights by the second system, a pair of antiparrallel Liner Accelerators (Linacs). The Linacs rely upon Superconducting Radio-Frequency (SRF) current carrying cavities to achieve this acceleration. The SRF cavities are made of niobium and so are submerged in 2.1K liquid helium cyromodules that keep them superconducting. Three primary cryogenic plants (making up the third system) cool and recirculate CEBAF's liquid helium. Central helium liquefiers one and two supply the north and south linacs respectively, while the end station refrigeration system feeds research halls A-C at the end of the beamline helium at 80K, 20K and 4.5K for their detectors, targets and various other experimental equipment. The Linacs are connected by two magnetic recirculation arcs that make up the fourth system and complete the racetrack shape. With each pass the electrons gain nearly 2180 MeV, each time recirculating through a higher energy magnetic arc. Electrons destined for halls A-C exit the southern Linac on their fifth and final circulation, before being directed into their destination hall. Electrons destined for hall D continue through the northern Linac on an additional half circulation before also being directed into their destination hall. [10]

The injector system begins with a optical-fiber laser gain switched at 499MHz that shines pulses of 750nm light on a Gallium Arsenide photo-cathode. The GaAs semiconductor materiel is capable of emitting pulses of photoelectrons spin-polarized by up to 85%. High degrees of electron beam spin polarization are required for some experiments at JLab that measure with spin dependent cross-sections. These pulses are accelerated out of the photo cathode with 130 KeV by the two electron guns shown at the start of injector system diagram figure-2.2. The oscillating nature of the electron guns electrostatic fields means not all electrons are equally accelerated and so they are fed into a pre-buncher. The pre-buncher accelerates or decelerates electrons as necessary so that they form bunches $250\mu s$ long separated by 16.7ms.[11]

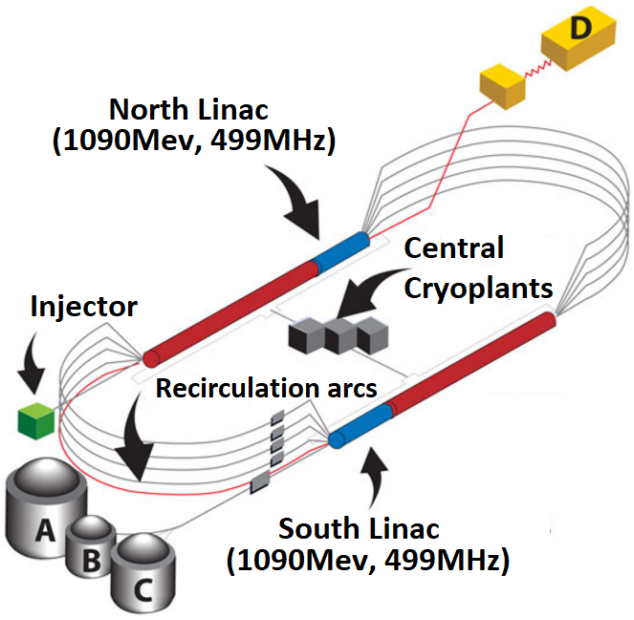


Figure-2.1 Illustration of CEBAF racetrack recirculation ring and injector system with research halls lettered A-D. Linacs accelerate electrons passing through at 499MHz by 1090MeV per pass. Blue regions indicating the five new Cyromodules added to each Linac for the 12 GeV upgrade, red indicating the older cyromodules that were upgraded.[12]

The beam is then collimated by two apertures (A1 and A2 in figure-2.1) before it is sent through a two 499MHz RF chopper cavities acting. The cavity magnetic fields act transverse to the beam line and deflect the beam bunches through slits at $0, 120°$ and $240°$ from the beamline. These slits regulate beam current as desired for experiments in halls A-D. Finally the beam is accelerated in its direction of travel by 16 SRF cavities which bring the beam energy up to 123MeV. Most of the beam is concentrated in a one-micrometer radius core surrounded by a lower current halo. The beam is deflected by four dipole magnets off its original axis of travel in a swerving chicane shape. This further concentrates the beam before final injection into the main beamline because the low current beam halo is deflected differently through the tight curve and therefore terminates against the sides of the piping that keeps the path of the beamline hermetically sealed rather than continuing with the core.

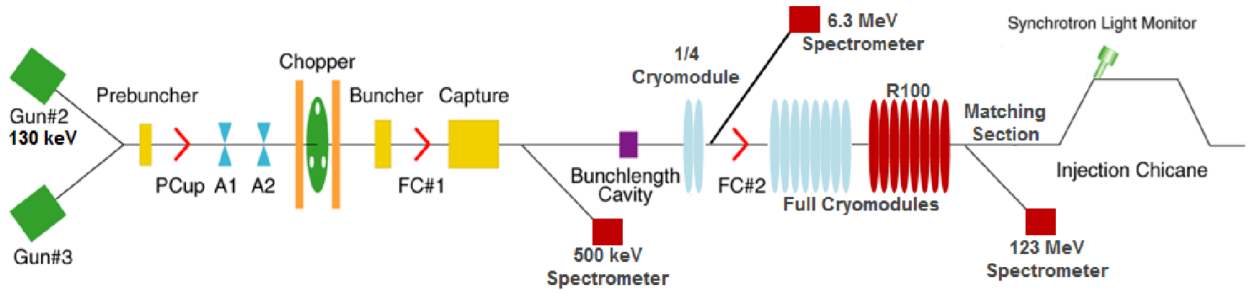


Figure-2.2 Diagram of injector system depicting main subsystems. GaAs optical fiber laser (not shown) feeds photoelectrons into the far left electron guns which accelerates them at up to 130Kev through the remaining systems.[12]

The north and south Linacs depicted in figure-2.1 are 1400m long and run 9m underground. Both Linacs house 25 cyromodules which each contain 8 SRF cavities. A cross section of one such SRF cavity in figure-2.3 shows how the cavity current produces a pattern of negative and positive charge density that push and pull the beam bunches. The current alternates at 1497MHz in sequence with the passage of bunches so the electrons are always accelerated by the cavity fields.[10]

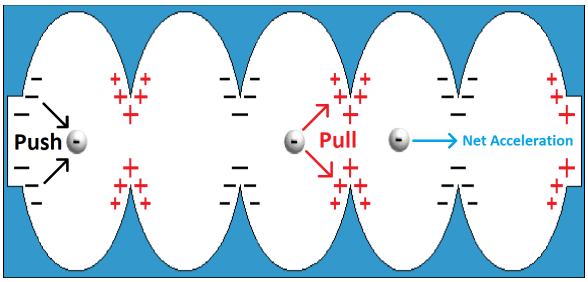


Figure-2.3 SRF cavity cross section showing how regions of charge density accelerate electron beam bunches (in this case represented by a single negative charge) through the magnetic cavities.[13]

These accelerating fields are solutions to wave equation presented in equation-2.1 and subject to the boundary conditions in equation-2.2.

$$(2.1) \qquad \left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\begin{bmatrix}\mathbf{E}\\ \mathbf{H}\end{bmatrix} = 0$$

$$(2.2) \qquad \hat{n}\times\mathbf{E} = 0\ ,\ \hat{n}\cdot\mathbf{H} = 0$$

Solving the wave equation r an idealised case of one of CEBAF's SRF cavities and selecting the fundamental frequency, we find the longitudinally accelerating electric field as a function of transverse distance from the axis of travel (r), particle distance through the cavity (z) and time (t). This solution is presented in equation-2.3 where $\omega_0 = \frac{2.405c}{R}$ and $J_0(z)$ is a Bessel function of first kind.[14]

$$(2.3) \qquad E_z(r,z,t) = E_0 J_0(z)\left(\tfrac{\omega_0 r}{c}\right)e^{-i\omega_0 t}$$

Integrating equation-2.3 across a particles path (z) through the cavity centre (i.e r=0) we find the idealised cavities accelerating voltage in equation-2.4.

$$(2.4) \qquad V_c = \left|\int_{-\infty}^{+\infty} E_z(r=0,z)e^{\frac{\omega_0 z}{\beta c}}\right|$$

The north and south Linacs contain 200 of these accelerator cavities each accelerating the beam through approximately 5.45MV while only losing 30W as heat when in continuous operation. As the beam is accelerated its bunches grow longitudinally, shifting in frequency and spreading transverse to the direction of travel. Quadrupole magnets arrays around the beamline are arranged in a FODO pattern between the cyromodules to correct for these shifts. F and D quadrupoles keep the beam vertically and horizontally focused transverse to the direction of travel, while O sections are simple drift spaces sometimes filled with deflecting dipole magnets that correct off sets from the intended axis of travel. Specialised S quadrupoles are on occasion also used to correct asymmetries due to cyromodules faults. Upon exit from the Linacs, the beam spreader sections shown in figure-2.4 use dipole magnets to deflect the beam into one of five recirculation arcs spaced in 0.5m increments. Upper arcs 1 and 2 steer the beam around a 250m, 180°curve with an array of 16 dipole magnets in series while lower arcs 3-10 guide higher energy beams with 32. The arcs have 32 quadrupole magnets along their lengths which manage three families of phenomenon. As the beam changes direction its electrons become excited, expressing this energy in the form of betatron oscillations about their equilibrium. The first set of quadrupoles tune the number of these oscillations to keep them under control and predictable. The second set keeps the achromaticity (frequency spread) of the beam within a tolerable degree, while the third regulates the momentum compaction (a measure of momentum dependence on electron recirculation path length). Diagnostic beam position monitors are placed at the entrance of each quadrupole to record the beam offset. This information is used by a second set of deflecting dipole magnets throughout the arc that keep the beam correctly aligned.[10]

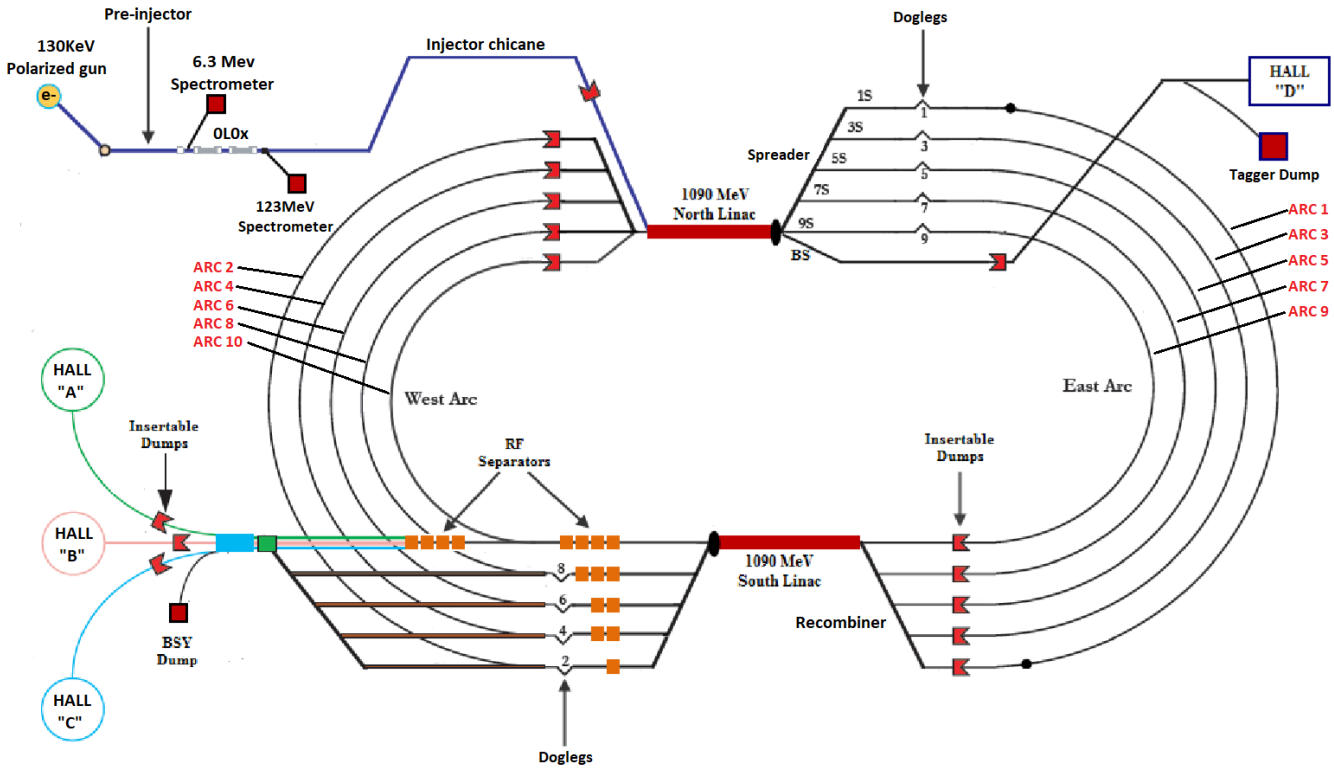


Figure-2.4 Top down view of recirculation ring showing how Injector, Recirculation Arcs, Linacs, Research Halls, Spreaders and Recombiners fit together.[10]

Beam bunches destined for halls A-C exit the south Linac on their fifth and final pass are not deflected by the spreader into another magnetic arc for recirculation and are instead enter the extraction region located bottom left of figure-2.4. In the extraction region (shown in greater detail in figure-2.5) the bunches travel through a series of horizontally deflecting RF cavities, small chicanes (doglegs) and spata (precision dipole magnets labelled YA,YB and YR in figure-2.5) for fine adjustments to the beam axis offset and to re-concentrate the beam by filtering out any beam halo that may have formed during acceleration or recirculation, just like in the injection process. Quadropoles labelled QB and QC focus the beam a final time and dipoles (labelled BM, BP, JG and JH) adjust the beam offset as required by particular experiments in the research halls. Finally, a Lambertson magnet deflects the beam bunches into three thinly separated deflecting magnetic fields that carry them to research halls A-C. Beam bunches heading to Hall D are instead deflected by the spreader through an additional recirculation arc for a final pass through the north Linac. Upon exit, the bunches are squashed by another buncher into a tighter bunch separation of only 1.33nm before travelling through a similar extraction region and delivery into hall D. The final electron beam properties CEBAF can deliver to its four research halls are summarised in table=2.1.[10]

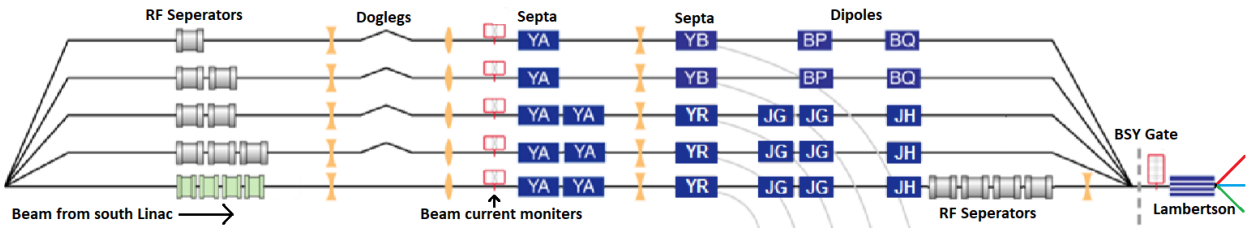


Figure-2.5 CEBAF beam extraction region for research halls A-C. Beam enters left, travelling through RF separation cavities, doglegs, spata and quadrupole magnets. Beam recombined at BSY gate after its beam axis offset and current has been selected for destined research hall. Beam bunches directed by Lambert magnet to halls A-C on far right.

	CEBAF
Max Beam Energy	12 GeV
Beam Spread (at Max)	2.16 MeV
Max Current	180 μA
Beam Bunch Frequency	499 MHz
Max Spin Polarization	85%
Beam Power	1 MW

Table-2.1 CEBAF Beam parameters at final delivery into research halls.[13][19]

III. CLAS at 12 GeV

Research Hall B's CLAS12 detector was originally designed in the 4-6 GeV era for the study of quark-gluon matter and its behaviour in exclusive and semi-inclusive electron scattering experiments, capturing particles across a wide range of momenta with its large acceptance and moderate momentum and spacial resolutions. [6] In its 12 GeV upgrade CLAS12 has been optimised for semi-inclusive and exclusive measurements with emphasis on its ability to extend its previous measurements across wider energies with increased precision. These measurements shed light on the internal dynamics of nucleons and the role of quark and gluons in the production of hadrons. The experiments in which these measurements are performed make use CLAS12's tenfold increase in operational particle luminosity $10^{35} cm^2 s^{-1}$ to collect many particle events for precise measurements across the the kinematic regions under study. Studies involving polarized solid state targets can only operate at limited luminosity and so will make use of CLAS12's large acceptance to collect particles events even at wider angles.[15]

CLAS12 is made up of a central target, a Central Detector (CD) and a Forward Detector (FD) that fit together as in figure-3.1. The target can be mounted with beads of irradiated ammonia (i.e sometimes with deuterium rather than hydrogen) that can be longitudinally spin polarized up to 90% by the powerful 5.0T solenoid field that envelopes it. A solid state hydrogen or deuterium target can be transversely spin polarised when brought to 12mK under the 5T fields. Deeply Virtual Compton Scattering (DVCS) experiments produce spin dependent cross sections which require spin polarized electrons and targets to separately measure. The combination of CEBAF's polarized beam and CLAS12's longitudinally or transversely polarized targets fulfil these requirements enabling the CLAS12 collaboration to measure these cross-sections. These quantities are significant because they test the predictions of models which describe how quark angular momentum contributes to proton spin. More conventional non-polarized targets are also available with light nuclei including hydrogen, helium and deuterium as well as heavier, solid state carbon, aluminium and lead targets. These targets are favourable for collecting vast quantities of data for GPD's across wide kinematic ranges with high beam luminosities that yield many events. The solenoid serves additional purposes beyond keeping targets polarized. It shields other detector subsystems from the electromagnetic background of the electron scatterings and enables the momentum analysis of charged particles in the CD. The CD is made up of two main detector subsystems, the Central Neutron Detector (CND) and the Central Time-of-Flight detector (CTOF). The CTOF uses an array of 38 plastic scintillations paddles arranged in a barrel shape close to the beamline (as in figure-3.3) to associate timings with the passage of particles at $35 - 125°$ from the beamline.

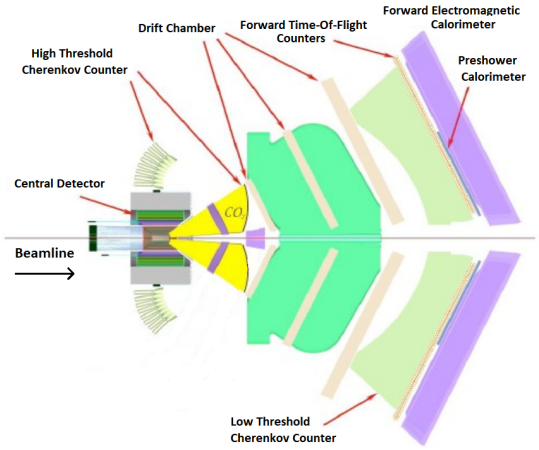


Figure-3.1 Cross-section of CLAS12 with the beamline running through the CD and FD's main subsystems including Drift Chambers (DC), Electromagnetic and Pre-shower Calorimeters (EC and PCAL), High and low threshold Cherenkov Counters (HTCC and LTCC)and the Forward Time-Of-Flight Scintillator (FTOF). The Green region shows the areas encloses by the FD's toroidal magnetic field. The target (not shown) is located in the centre of the CD.

The timing information collected by the CTOF is used in combination with information from other subsystems to help distinguish between kaons,pions and protons and to improve the spacial and momentum reconstruction resolution. [16] Scintillators work by detecting the kinetic energy converted into light that is deposited by charged particles as they interact with the scintillator volume. Incident charged particles interact electromagnetically with the scintillator volume, depositing a virtual photon in conservation of the kinetic energy transferred. The absorption of this energy by a particle in the scintillator volume can be expressed as a transition to a higher energy state. As shown in figure-3.2, particles may be excited to singlet states with a net angular momentum of zero denoted by $S_0, S_1, S_2...$ or non zero triplet states denoted by $T_0, T_1, T_2...$ (second subscript denoting vibrational energy level). These higher energy states are however not stable and the excited particles de-excite shortly after in three kinds of ways. In a florescence, a particle excited to a singlet state swiftly de-excites to its ground state on the order of picoseconds. The energy of the emitted photon is equal to the difference in energy between the states. From the Planck-Einstien relation the emitted photons energy can be found to be inversely proportional to its wavelength (as in equation-3.1).

$$E_{emmited} = \frac{hc}{\lambda_{emmited}} = E_{singlet} - E_{ground} \quad (3.1)$$

The wavelength of the emitted photon is always longer than the wavelength of the energy absorbed in the excitation. In the case the CTOF, light is usually emitted at approximately 3 eV which corresponds to a blue wavelength. [17]

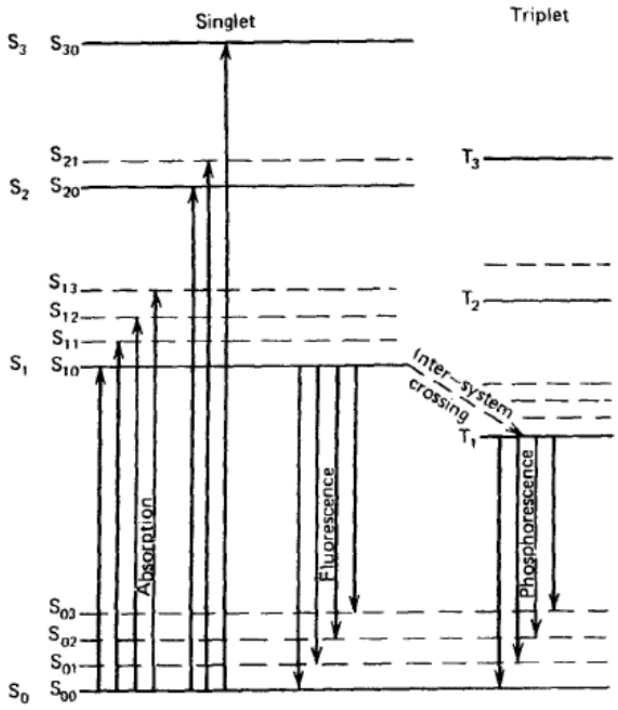


Figure-3.2 Singlet and Triplet electron states demonstrating absorption, florescence and phosphorescence mechanisms.[17]

In a phosphorescence a particle in a singlet state may transition to a triplet state in an inter-system crossing process before de-exciting back to its ground state. Triplet states have longer lifetimes than singlet states and so phosphorescences take longer than fluorescences. In a delayed florescence a particle in a triplet state absorbs nearby energy (often thermally) before undergoing a reverse crossing into a singlet state and a final de-excitation to its ground state. With the exception of S_{10} to S_{00} the magnitude of energy released in downwards transitions is less than the smallest possible upward transition. The CTOF scintillator is therefore mostly transparent to particles emitted due to these excitations. The plastic BC-408 material making up the CTOF's scintillation volume has an attenuation length of 210cm.[18] At double this length, light emitted from the longest CTOF paddles at one end, reaches the other with $\frac{1}{e^2}$ its original intensity. Upon reaching the end, light from these excitations is piped through the optical guides shown in figure-3.3 into photomultiplier tubes (PMTs) that convert it into electric currents. The incredible rapidity of this process means the timing of incident particles can be associated with current spikes measured in the CTOF with a timing resolution of 70ps.

The incredible rapidity of this process means the timing of incident particles can be associated with spikes in the CTOF's measured current to within 70ps. The characteristic photon energies produces by these transitions can be used in combination with the relationship between said energy and the size of the energy spike to identify the particular interaction the signal originates from. [16] PMT's are typically used to convert weak light signals into usable measurement electric currents while avoiding large signal background. A simple PMT is displayed in figure-3.4 illustrating its two major components, a photo-cathode and an electron multiplier structure.

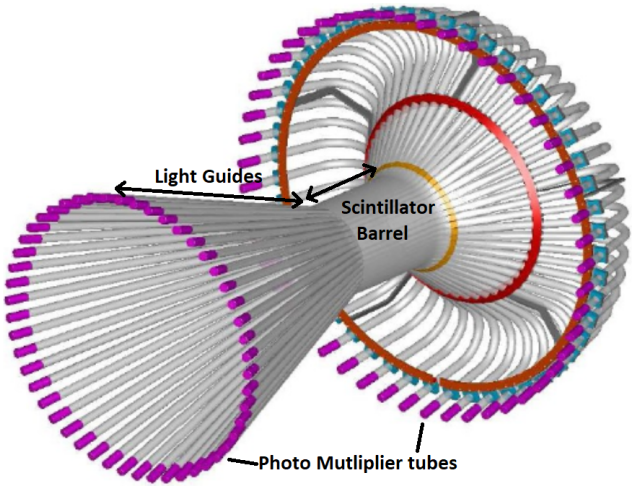


Figure-3.3 3D Computer Aided Design (CAD) model of CLAS12 CTOF. Central barrel shaped section composed of 48 scintillation paddles in white. Connected to either end of the barrel's paddles are light guides. (also in white) The light guides curve at the front and are straight at the back, leading to the purple PMT regions. Edited version of image from [16].

The photo-cathode converts photons incident on its photosensitive surface into low-energy photoelectrons through the photoelectric effect. This photoemmision process can be thought as occurring in three stages. The absorption of the incident photon and transfer of energy to electron within photoemmisive material. The migration of said electron to the photo-cathode surface and finally the emission of that photoelectron from the photo-cathode surface.

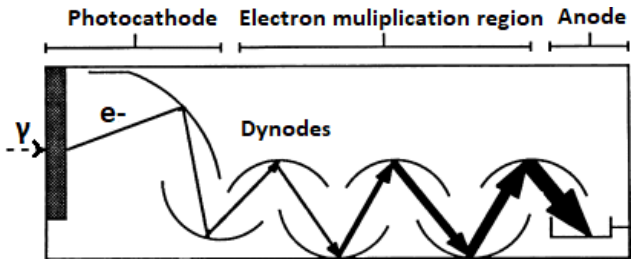


Figure-3.4 typical PMT design with photons incident on the far left photo-cathode. Photoelectron are emitted from the photo-cathode and bounce off of the dynode surfaces before registering as an electric current at the anode. Edited version of image from [18].

Light emitted from the CTOF scintillators comes in short bursts that produce similarly short photoelectron pulses with little detectable charge in the PMTs. The photoelectron signal is therefore amplified in the electron multiplier region. In this region electrons are focused through a system of high secondary emission coefficient surfaces called dynodes. In this context the secondary emission coefficient refers to the efficiency with which incident electrons induce the emission of secondary electrons from a surface. As the electrons bounce through the dynode surfaces, the signal is multiplied from a few hundred electrons to $10^7 - 10^{10}$. Following signal amplification, the electrons are collected on an anode as an electric current. PMTs have sealed outer structures that act as pressure boundaries, sustaining vacuum conditions that minimise the attenuation of travelling photoelectrons that would otherwise distort the signal. PMTs are sensitive to external electric or magnetic fields that influence the travel of photoelectrons within them, it is for this reason that the PMTs connect to the CTOF are positioned outside the main solenoid magnet. [16][17] The CND is designed to detect and identify low energy neutrons at wider, sometimes back scattering angles with $0.2 - 1GeV/c$ momentum and 40° to 120° from the beamline. The detection of neutrons is essential for many exclusive electron scattering experiments performed with CLAS12.The CND is positioned between the CTOF and the solenoid magnet and consists of three layers of 48 scintillation paddles arranged in a 70 cm long barrel shape. The end of each paddle is connected to a light guide leading to PMTs outside the solenoid field (as in the CTOF).[16] A summery of the CND and CTOF detector parameters is displayed in table-3.1.

Detector	CTOF	CND
Angular Acceptance	35°-125°	40°-120°
Time Resolution	70 ps	130 ps
Polar Angle Resolution	N/A	2°
Azimuthal Angle Resolution	N/A	2°
Neutron Detection Efficiency	N/A	10%
Effective Pion/Kaon Seperation Range	0-0.64 GeV	N/A
Effective Kaon/Proton Seperation Range	0-1 GeV	N/A
Effective Pion/Proton Seperation Range	0-1.25 GeV	N/A

Table-3.1 CTOF and CND detector parameters and resolutions.[16][20]

Particles that scatter from the target at a polar angle $5 - 40°$ from the beamline enter the FD. The FD contains a 6 coil, hexagonal torus magnet that encloses the three DC layers shown in figure-3.1. The green region shows the effective area of the peak 3.6T toroidal magnetic field through the DC layers, acting primarily in the direction of the beamline. The DC layers contain a mixture of 90% argon and 10% CO_2 gas. These layers are made up of 6 smaller layers containing 112 hexagonal cells, each with a central anode sense wire.[16] As particles travel through the DC layers they ionize trails of electrons from the gas molecules as they pass. These electrons are accelerated towards the sense wires and deliver a current signal from which it can be deduced that a particle passed through the cell.

As these electrons drift towards the sense wires, they randomly encounter gas molecules with much greater momentum then them and are scattered isotropically with a randomly orientated velocity c. The mean time between collisions τ can be expressed in terms of the randomly orientated velocity, the collision cross σ section and the number density of gas molecules N as in equation-3.2.[21]

$$\tau = 1/(N\sigma c) \tag{3.2}$$

The electron picks up an extra velocity u due to its acceleration along the sense wires electric field E, multiplied by the average time between the next collision. This extra velocity appears macroscopically as the drift velocity u and can be expressed as a function of electric field strength E as in equation-3.3 by substituting in for τ from equation-3.2.[21]

$$u = \frac{eE}{m}\tau = \frac{eE}{mN\sigma c} \tag{3.3}$$

It is preferable that drift velocity varies little with electric fields so that measurements depending upon drift velocity are less effected by unavoidable and unpredictable electric field gradients arising due to imperfections in the the sense wires and DC layer geometry. Argon-CO_2 mixtures exhibit this valuable characteristic and it is for this reason among safety and expense considerations that it is used. [16]

Charged particles are accelerated in a helical fashion as they move through toroidal magnetic field enclosing the drift chambers. By equating the centripetal force on an charged particle to the force it experiences as it moves in the approximately homogeneous toroidal magnetic field B acting along its direction of travel, the radius of curvature R for its path can be found as a function of its transverse momentum P_T as in equation-3.4.

$$R = \frac{P_T}{qB} \tag{3.4}$$

The angle between the longitudinal and transverse momentum (P_z and P_T) can be defined as ϕ as in equation-3.5.

$$\tan\phi = \frac{P_z}{P_T} \tag{3.5}$$

By multiplying equation-3.5 by the radius of curvature and substituting in equation-3.4 the longitudinal momentum can also be related as a function of the radius of curvature. (As in equation-3.6)

$$\tan\phi \cdot R = \frac{P_z}{P_T} \cdot R = \frac{P_z}{qB} \tag{3.6}$$

The curvature of particles path can therefore be used to determine the components of its momentum. The pattern of sense wire activations are used to reconstruct the path of particles through the DC, with which the curvature can be extracted and the momentum determined. This resolution can be further improved when the DC is used in combination with the FTOF.

The FTOF works on the same scintillation principle as the CTOF and CND, recording precise timings and aiding in the particle identification process.

The 3D image of the FTOF in figure-3.5 illustrates its main six triangular scintillation panels in orange, with an angular acceptance $5-36°$from the beamline. A second set of panels sits behind the first, covering approximately the same area, which together with the first achieves timing resolution for the passage of particles of 80 ps. A final set of scintillation panels (coloured in red) covers wider angles from 36° to 45°. Timing information from the FTOF paddles is extrapolated to when the particle passes through a specific DC cell. This is compared to the timing of detection at the sense wire (calculating for the drift time), producing a radius of closest approach that indicates the closest path particle could have taken through the cell past the sense wire. These distance further constrain path reconstruction, enabling the DC to achieve a spatial resolution of $250-250\mu m$ and a $\frac{dp}{p} < 1\%$ momentum resolution. The DC and FTOF combination produces precise enough momentum and trajectory tracking that it can be used to distinguish between negative pions, kaons and electrons at up to 2.6GeV, and between positive pions, kaons and protons up to 5.6GeV, higher than achievable with the CTOF and CND. Furthermore, the FTOF provides an independent means of identifying slower moving particles. If a particle is fully captured in the scintillator volume the quantity and shape of the energy shower deposited is characteristic of the incident particle.[16]

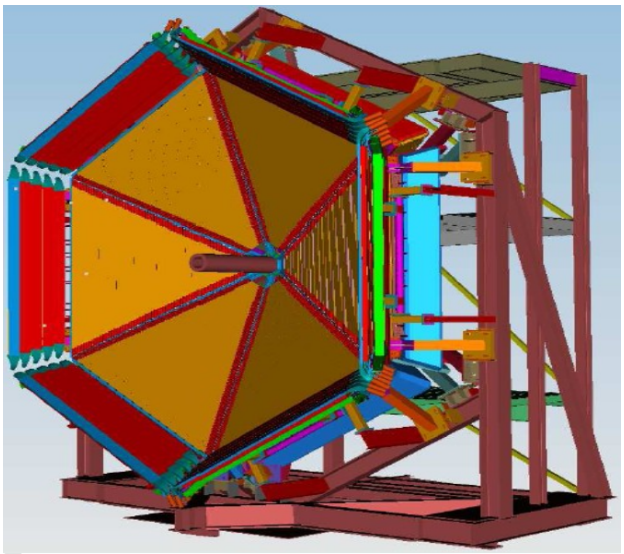


Figure-3.5 CLAS12 FTOF CAD model showing central scintillator panels in orange, symmetrically divided in 6 sectors around the beamline. Large acceptance FTOF paddles shown in red.

Located upstream of the DC, the High Threshold Cherenkov Counter (HTCC) is used to discriminate negative electrons and pions at high momentum (get number), while the Low Threshold Cherenkov Counter (LCC) separates out pions and kaons with between 3.5 to 9GeV momentum. When a charged particle travels through an dielectric medium faster than the phase velocity of light, light is emitted in a phenomenon called the cherenkov effect. The energy emitted per unit length per frequency as a charged particle moves through the material can be related with the Frank-Tamm formula in equation-3.7. (In this case for angular frequency ω independent permittivity μ and refractive index n.

$$\frac{d^2E}{dxd\omega} = \frac{q^2}{4\pi}\mu\omega\left(1-\frac{1}{\beta^2n^2}\right) \tag{3.7}$$

Integrating equation 3.7 with respect to ω the cherenkov photon intensity can be found to be proportional to $\frac{1}{\lambda^2}$ meaning most cherenkov radiation is concentrated at short wavelengths. At extremely short wavelengths the photon intensity drops off because refractive index is inversely proportional to wavelength. Cherenkov radiation is emitted in a forward cone with angle θ_c about direction of the particle travel as in equation-3.8.

$$\cos\theta_c = \frac{1}{n\beta} \tag{3.8}$$

With $\beta = \frac{v}{c}$ we have velocity as a function of angle. Comparing this velocity to momentum independently determined by the DC and CTOF enables the mass to be isolated and the particle to be identified. It is through this method that the HTCC and LTCC are able to distinguish particles with similar momenta. Cherenkov counters also have the inherent ability to discriminate particles who's mass exceeds a minimum threshold to emit cherenkov radiation in the first place. This threshold (expressed in equation-3.9) depends upon the refractive index of the chosen medium, the HTCC and LTCC combination makes great use of this phenomenon. For example, if a particle is detected in the LTCC but not in the HTCC, its mass must be large enough that it passes the threshold in the first but small enough that it doesn't pass the threshold in the second and so the range of possible masses it could have is constrained.

$$E_{th} = m_0c^2\left(\sqrt{1+\frac{1}{n^2-1}}-1\right) \tag{3.9}$$

Figure-3.6 shows a simple Cherenkov Counter, incident particles travel through the opening into the radiating medium with a carefully chosen refractive index. Particles with energy above threshold for the mediums refractive index (as related in equation 10) radiate many hundreds of Cherenkov photons. These photons are guided into a PMT where they are converted into interpretable electric currents. [16]

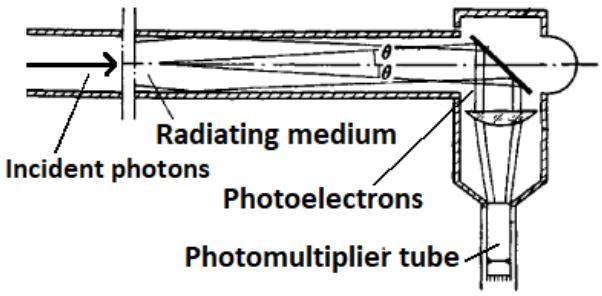


Figure-3.6 Cherenkov Counter diagram. Incident particles emit Cherenkov radiation in radiating medium, light is directed by a mirror into a lens which focuses it into a PMT. Edited from [22]

CLAS12's Electromagnetic and Pre-shower Calorimeters (EC and PCAL) are designed to fully absorb particles, collecting all of their energy and recording the location of the energy deposit from which the direction of the incoming particle can be determined. The EC and PCAL are sampling calorimeters made up of lead absorption layers that alternate with plastic scintillator detection layers. As shown in figure-3.7, incident particles excite, ionise or otherwise interact with the absorption layers to produce secondary photons, hadrons and leptons. These secondary particles go onto interact with further absorption layers resulting in electromagnetic or hadronic particle showers or are absorbed in detector layers. With each layer penetrated the secondary particles lose more and more energy until it is all absorbed.[16] By comparing the energy deposited by a particle in the calorimeters to its momentum reconstructed through the DC and FTOF, the rest mass contribution to energy can be determined and the particle identified. This enables CLAS12 to identify high energy neutral photons, neutrons and pions which is essential to inclusive measurements of interaction FF and GPD's which CLAS12 was intended to facilitate. [15][22]

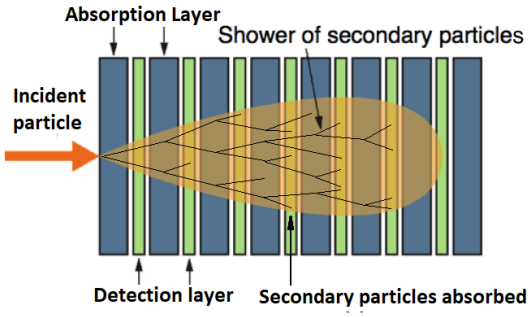


Figure-3.7 Examples of sampling calorimeter showing alternating absorption and detector layer structure and an example particle shower. Edited from [23].

The most dominant forms of energy loss for the electromagnetic component of a high energy particle shower comes in the form of Bremsstrahlung radiation and photon pair productions. A radiation length X_0 in the absorbing medium characterises the distance over which an electron (or positron) loses on average all but 1/e of its energy as Bremsstrahlung radiation. This quantity can be expressed as a function of area and mass density, satisfying equation 3.10. [23]

$$E = E_0 e^{-x/X_0}, \quad \text{Where} \quad X_0 = \frac{180A}{Z^2} \tag{3.10}$$

Differentiating with respect to distance through the absorbing medium we find that the energy loss per unit distance is proportional to the mediums density for short distances as shown in equations-3.11 and 3.12.

$$-\frac{dE}{dx} = \frac{E_0}{X_0} e^{-\frac{x}{X_0}} \tag{3.11}$$

For $x << 1$, $e^{-\frac{x}{X_0}} \rightarrow 1$

$$-\frac{dE}{dx} = \frac{E_0}{X_0} = \frac{Z^2 E_0}{180A} \tag{3.12}$$

The mean free path for a photon in the absorption medium between pair productions is approximately $\frac{9}{7}X_0$. The intensity of a photon beam in the absorption medium can therefore be related as in equation-3.13.

$$I = I_0 e^{-\frac{7x}{9X_0}} = I_0 e^{-\frac{7Z^2 x}{1620A}} \tag{3.13}$$

Evaluating equation-3.12 and equation-3.13 for lead we find that the electromagnetic components of particle showers lose 1/e of their energy to Bremsstrahlung radiation and pair productions after only 5.6mm and 7.22mm respectively. The EC and PCAL therefore use lead for their absorption layers as it makes for compact detector at low expense. When electrons and positrons slow to below 7.3MeV, energy losses due to the atoms they excite (or ionise) become greater then those due to bremsstrahlung until they and their daughter particles are finally stopped. Below this threshold, photons lose progressively more energy to Compton scatterings and in the production of photoelectrons as opposed to pair productions until they are fully absorbed.

If an incident particle produces a secondary particle shower with a hadronic component the major sources of energy loss for secondary hadrons come from elastic (i.e $N(H, H')N$) and inelastic (i.e $N(H, H')X$) collisions. Depending upon the incident particle, hadronic showers lose $\frac{1}{e}$ of their energy to these collisions after $\approx 10^1 - 10^2 cm$. [22] The penetration depth of hadronic showers is therefore considerably greater than electromagnetic showers. In sampling Calorimeters a compromise is made between compactness and efficiency of detection. The EC and PCAl have 54 layers of 2.2mm thick lead absorption layers meaning some of the hadronic products escape the calorimeter before all of their energy is collecting in a phenomenon known as leakage. [16] The resolution with which a calorimeter can detect an incident particles energy is limited by three factors. Firstly, the number of secondary particles an incident particle produces N is proportional to its energy E, the energy resolution due to these statistical fluctuations is therefore proportional to $1/\sqrt{E}$. Secondly, incident particles may strike the same calorimeter region in rapid succession and so there is an uncertainty associated with the noise of high luminosity environments proportional to 1/E as their is more noise for higher energy signals. Finally, their is a constant resolution term due to calibration errors, detector component faults and particle leakage.[22] The energy resolution can therefore be parametrised as in equation-3.14 with empirical, detector dependant constants or fit parameters c1,c2 and c3.[24]

$$\left(\frac{\Delta E}{E}\right)^2 \approx \left(\frac{c_1}{\sqrt{E}}\right)^2 + \left(\frac{c_2}{E}\right)^2 + \left(c_3\right)^2 \tag{3.14}$$

From equation-3.14 we can see calorimeter energy resolution improves at higher energies. CLAS12's calorimeters use this property to precisely capture the energies of (and therefore identify) neutral neutrons, pions and photons. The EC and PCAL detector parameters and resolutions (including energy resolution) are presented in table-3.2. [16]

Detector	EC and PCAL Combination
Energy Resolution	10%/√E Gev
Position Resolution	5 mm
Timing Resolution	500 ps
Polar Angle Acceptance	5° to 35°
Azimuthal Angle Acceptance	5° to 35°

Table-3.2 Summary table of EC and PCAL combination detector parameters and resolutions.[25]

IV. Research Focus

The CLAS12 collaboration has developed an Apache-Java based software package called CLAS12 Offline Analysis Tools (COATJAVA) with which its researchers convert raw detector signals into particle event information they can more easily perform analysis on. Within COATJAVA is the CLAS12 Reconstruction and Analysis (CLARA) framework that reads in detector signals and reconstructs the trajectory, momentum, vertex and ID of particles scattered from the target. The first reconstruction service is the event reader which reads in a digitised version of the raw detector data containing the positions and timings of detector component events. The reconstruction proceeds by plotting possible particle trajectories through the positions of detector component activations in a process called Hit Based Tracking (HBT). As previously described, timings extrapolated from the passage of particles through the FTOF are compared to their detections in individual DC cells to produce distances of closest approach from the central sense wires. In a process called Time Based Tracking (TBT), the possible trajectory fits are further constrained by these distances of closest approach for the passage of particles within the DC cells. Particle tracks that are considered likely to represent real trajectories are stored and particle ID's, momenta and vertex are determined in combination with information from the remaining detector systems. A graphical representation of Hit Based and Time based Tracking is presented in figure-4.1.

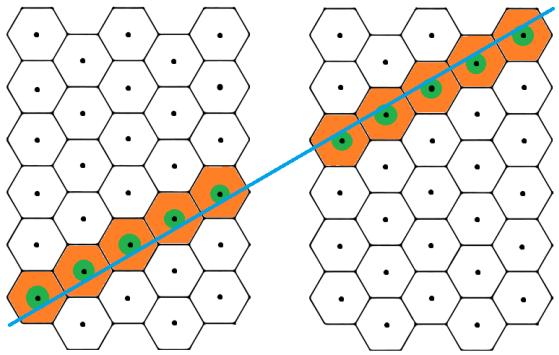


Figure-4.1 Example of linear trajectory reconstruction between two DC layers. Orange hexagons indicate DC cells that have detected the passage of particles, green circles show distances of closest approach to central sense wire. In HBT possible trajectories are fitted through the orange hexagons, in TBT these fits are limited to only those that pass through the distances of closest approach.

The focus of my research placement is to assist the CLAS12 collaboration software group in the continued development of COATJAVA. The subject of my work is intended to be twofold, to update the unit test that validates new versions of DC reconstruction code and to evaluate the effectiveness of a new method of track reconstruction in the CD. With more than 84,000 lines of executable code COAJAVA is a large software project. COATJAVA is therefore split into many smaller modules that developers can work on separately to make it more manageable. Each of these modules has a unit test that is run each time the module is updated to validate that it functions as intended. The existing DC reconstruction module's unit test falsely signals a software failure even when the reconstruction is preformed correctly. My role is to update the unit test so that is no longer consistently returns false positives and can be relied upon to validate new versions of the DC reconstruction code.

Charged particles travelling through the CD are accelerated by the solenoid magnetic field in a helical motion along their direction of travel. Setting the z axis as the longitudinal direction of travel, the circular motion of the helix about the z axis in the xy plane can be parametrised in Cartesian co-ordinates with radius r and centre (a,b) as in equation-4.1.

$$(x-a)^2+(y-b)^2=r^2 \tag{4.1}$$

The travel of one such charged particle through the CTOF of CND would leave at least partially complete circular trails of scintillator component activations that could be projected on an xy plane. The circle centres and radii would therefore parametrise charged particle tracks along their directions of travel in the CD. The radius of curvature for these tracks (as related in equation-3.4) is proportional to their transverse (xy) momentum and so could be used to reconstruct it. One possible method of recognising these circles is to apply the Hough transform to the positions of detector hits. In the context of recognising circles, the application of the Hough transform can be considered in the form of a geometrical procedure.

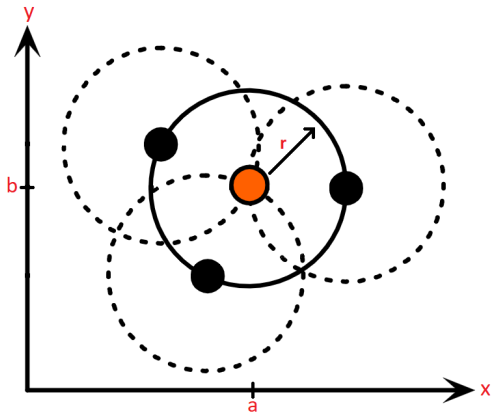


Figure-4.2 Application of geometrical procedure for Hough transform on example data. Circles of radius r plotted around each data point, most common intersection at (a,b) is the centre of the true circle.

Figure-4.2 shows three black data points in the xy plane representing detector hits produced by the travel of a single particle into the page along its longitudinal direction of travel. The travel this particle can be parametrised by a single true circle, centred at (a,b) with a known radius r. The centre of the true circle is found by plotting circles of radius r around each data point and identifying the point of most common intersection between them. [26]

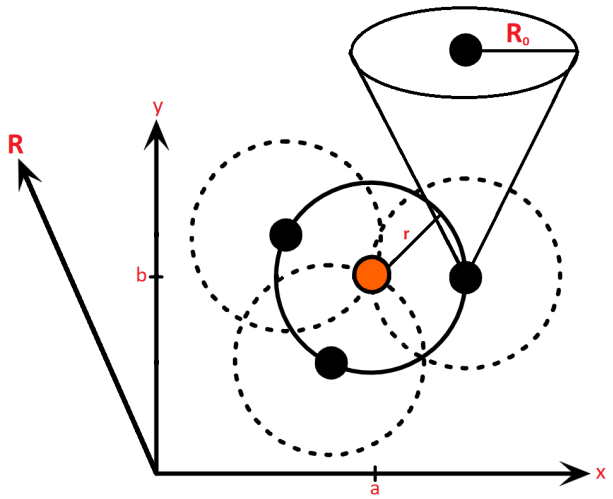


Figure-4.3 Projection of cone along third axis representing geometrical procedure for application of Hough transform for recognition of circle with unknown radius.

In the case of circular projections of the helical paths of charged particles through the CD, the radius is not known. For unknown radii, cones can be constructed at each data point along a third axis R, made up of circles of increasing radius at every level. The projection of one such cone up to a radius R_0 is shown in figure-4.3. The true centre and radius of a circle with unknown radius is found at the most common intersection of these conical surfaces. [26] It is my role to write a COATJAVA test package that uses an algorithmic version of the Hough transform to reconstruct charged particle tracks in the CD, and to evaluate the efficiency with which it does this on test data.

V. Summary

JLab's CEBAF facility has been upgraded to deliver a highly polarized continuous electron beam at a maximum beam energy and current of 12GeV and 180μA. The CLAS12 detector has been upgraded to study of hadronic matter in electron scattering experiments at these higher available beam energies with increased precision. A 12 GeV beam energy enables the CLAS12 collaboration to study phenomenon only observable at these higher energies such as the production of exotic mesons. These increased beam energies will enable CLAS12 to extend its previous measurements of GPDs and FF across wider kinematic ranges with greater precision in exclusive and semi-inclusive measurements. CLAS12 consists of Drift Chambers, Cherenkov Counters, Scintillators and Calorimeters. The CLAS12 collaboration has developed a software package that takes the raw spatial and temporal data collected by these detector subsystems and reconstructs particle trajectories, momenta, vertex and IDs. The goal of my research placement is to assist the CLAS12 collaboration software group in the continued development of this software package. My research will focus on the validation of existing code and on an investigation into the effectiveness of the application of the Hough transform to the reconstruction of charged particle tracks in the CD.

Acknowledgements

This research is funded by a US Department of Energy grant and is supported by Veronique Ziegler a JLab staff scientist and Gerard. P. Gilfoyle a professor at the University of Richmond and my supervisor for the year.

Bibliography

[1] - C. Quigg, G. Aubrecht "Elementary Particle Physics: Discoveries, Insights, and Tools, in Quarks, Quasars and Quandries", American Association of Physics Teachers. (1987)

[2] - National Science Advisory Committee, "A Long Range Plan for Nuclear Science", 1979, URL: clasweb.jlab.org/wiki/images/c/ce/s [Accessed 23/07/2019].

[3] - F. Gross, "Making the case for Jefferson Lab", Journal of Physics, Conference Series, vol 299, pp 1-20, 2011.

[4] - "History", Jlab.org, available: https://www.jlab.org/visitors/history/index.html [Accessed 24/07/2019].

[5] - C. Westfall, "Jefferson Lab's 1985 Switch to Superconducting Accelerator Technology", (1996) Michigan State Universit.

[6] - "Pre-Conceptual Design Report (pCDR) for The Science and Experimental Equipment for the 12 GeV Upgrade of CEBAF", June 11th 2004, Jefferson Lab.

[7] - M. Peskin, D. Schroeder, "An introduction to Quantum Field Theory", (1995), Westview Press, ISBN: 978-0-201-50397-5

[8] - J. Greensite, "An introduction to the confinement problem", (2011) Springer, Lecture Notes in Physics, 821, doi:10.1007/978-3-642-14382-3 ISBN: 978-3-642-14381-6

[9] - E. Cartlidge, "Quarks break free at two trillion degrees", (23rd June 2011), Physics World [10] - M. Spata, "Continuous Electron Beam Accelerator Facility Overview", 2017, JLab Summer Lecture Series, 2017.

[11] - M. Baylac, "Calibration of the CEBAF injector high voltage power supply versus the 100 keV polarized source Faraday cup", 2004 Jeerson Lab Technical Notes, Newport News, VA.

[12] - "ARRA", Jlab.org, URL: www.jlab.org/fm/arra.html [Accessed 25/07/2019].

[13] - Y. Wang , "CEBAF overview", 2010, Hampton University Graduate Studies Program, URL: www.jlab.org/hugs/archive/Schedul [Accessed 25/07/2019]
[14] - H. Padamsee, J. Knobloch, T. Hays "RF Superconductivity for Accelerators", John Wiley & Sons Inc, ISBN: 0-47115432-6.

[15] - D. Abbott et al. "The Hall B 12GeV Upgrade: pre-conceptual design report", (2003), Jefferson Lab, Newport News, VA.

[16] - "CLAS12 Technical Design Report", 2008, Jlab.org, URL: jlab.org/Hall-B/clas12_tdr.pdf, [Accessed: 25/07/2019].

[17] - G.F. Knoll, "Radiation Detection and Measurement 3rd edition", 2000, Wiley. ISBN:978-0-471-07338-3.

[18] - "BC-400,BC-404,BC-408,BC-412,BC-416Premium Plastic Scintillators", Febuary 2018, Saint-Gobain Crystals, URL: www.crystals.saint-gobain.com/sites/imdf.crystals.com/files/documents/bc400-404-408-412-416-data-sheet.pdf, [Accessed:26/07/201

[19] - J. Benesch et al, "Beam Physics For the 12 GeV CEBAF Upgrade Project", 2005, Proceedings of 2005 Particle Accelerator Conference, URL: accelconf.web.cern.ch/accelconf/p05/PAPERS/TPPP016.PDF [Accessed: 26/07/2019].

[20] - S.Niccolai,L. Elouadrhiri, G. Young. V.D. Brukert, "CLAS12 CND", 24th March 2017,, Jefferson Lab, Newport News, VA, URL:www.jlab.org/Hall-B/clas12-web/specs/cnd.pdf, Accessed: [27/07/2019]

[21] - W. Blum, W. Riegler, L. Rolandi, "Particle Detection with Drift Chambers", 2008, Spinger, ISBN: 978-3-540-76683-4

[22] - "Cherenkov Counter", Encyclopedia2, URL: encyclopedia2.thefreedictionary.com/Cherenkov+Counter, [Accessed 29/07/2019].

[23] - M. Krammer, "Calorimeters", Detectors for Particle Physics. Lecture 6, URL: www.hephy.at/fileadmin/user_upload/VO-6-Calorimeters.pdf, [Accessed: 27/07/2019]

[24] - D. Cockerill, "Introduction to Calorimeters" 4th may 2016, University of Southampton Lecture, URL: indico.cern.ch/event/51847 lecture-to-Southampton-students-4May2016-Cockerill-compressed.pdf, [Accessed: 27/07/2019]

[25] - C.Smith, S. Stepanyan, L. Elouadrhiri, G. Young. V.D. Brukert, "CLAS12 ECAL-PCAL (EC)", 16th March 2017, Jefferson Lab, Newport News, VA, URL: www.jlab.org/Hall-B/clas12-web/specs/ec.pdf, [Accessed 27/07/2019].

[26] - "Hough Circle Transform", 11th October 2005, Harvey Rhody, Rochester Institute of Technology. URL:www.cis.rit.edu/class/sim [Accessed: 29/07/2019].

5.2 Appendix 2: CLAS12 Coordinate System and Vertex

The origin of the CLAS12 coordinate system is at the centre of its target system. As shown in Figure A the z axis of this coordinate system extends along the direction of the incoming beam.

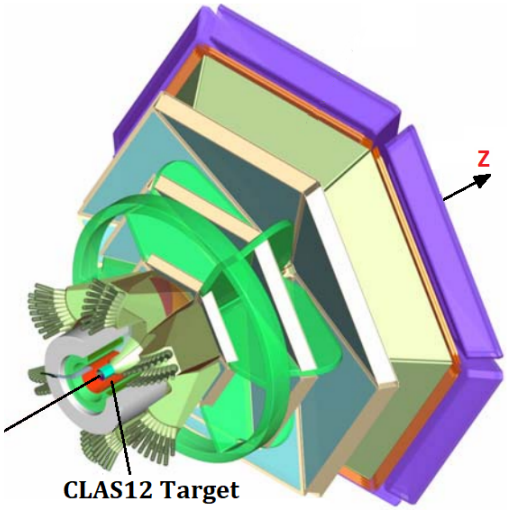


Figure A: Illustration of CLAS12 coordinate system showing placement and direction of z axis. Adapted image from [16]

CLAS12 is split into 6 triangular sectors around this z axis with the numbering convention shown in Figure B. In Figure B the z axis is directed into the page perpendicular to the xy plane.

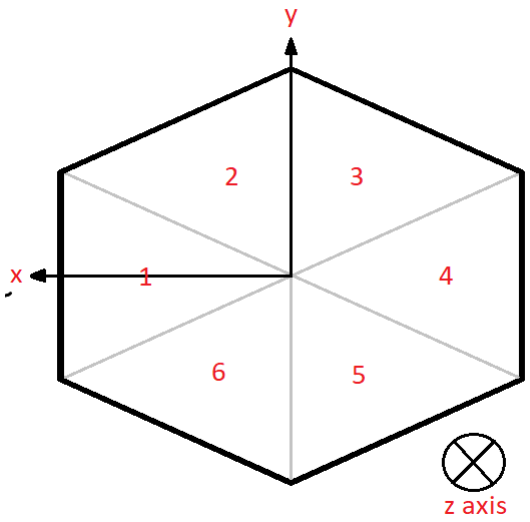


Figure B: Representation of the CLAS12 coordinate system's xy plane with view facing down the beamline.

This system can be extended into spherical coordinates by defining a polar angle θ, azimuthal angel ϕ and distance r such that $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$ and $z = r\cos\phi$ as shown in Figure C.

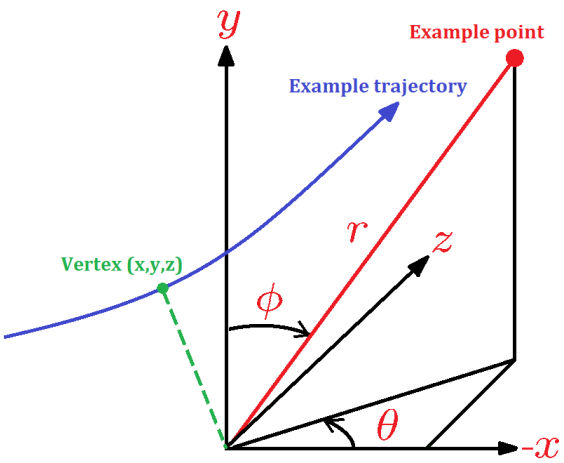


Figure C: Diagram of spherical CLAS12 coordinate system showing vertex of an example trajectory.

The trajectories of particles through CLAS12's detector systems reconstructed by COATJAVA can be extrapolated back to the target region. The closest point along a trajectory to the centre of the target is defined as its vertex.

5.3 Appendix 3: New DC Reconstruction Unit Test Source Code

```
package org.jlab.service.dc;

import cnuphys.magfield.MagneticFields;
import org.junit.Test;
import static org.junit.Assert.*;

import org.jlab.io.base.DataEvent;
import org.jlab.service.dc.DCHBEngine;
import org.jlab.service.dc.DCTBEngine;

import org.jlab.analysis.physics.TestEvent;
import org.jlab.analysis.math.ClasMath;

import org.jlab.clas.swimtools.MagFieldsEngine;

import org.jlab.clas.swimtools.ProbeCollection;
import org.jlab.clas.swimtools.Swim;
import org.jlab.clas.swimtools.Swimmer;


import org.jlab.utils.CLASResources;
/**
 *
 * @author nharrison, updated by marmstr june 2019
 */
public class DCReconstructionTest {
  @Test

  public void testDCReconstruction() {

    //Setup Magnetic Fields
    String mapDir = CLASResources.getResourcePath("etc")+"/data/magfield";
        try {
            MagneticFields.getInstance().initializeMagneticFields(mapDir,
                    "Symm_torus_r2501_phi16_z251_24Apr2018.dat","Symm_solenoid_r601_phi1_z1201_13June2018.dat");
        }
        catch (Exception e) {
            e.printStackTrace();
        }

    String dir = ClasUtilsFile.getResourceDir("CLAS12DIR", "etc/bankdefs/hipo4");
    SchemaFactory schemaFactory = new SchemaFactory();
    schemaFactory.initFromDirectory(dir);


    //Read in test event
    DataEvent testEvent = TestEvent.getDCSector1ElectronEvent(schemaFactory);
    MagFieldsEngine enf = new MagFieldsEngine();

    //HB reconstruction
    enf.init();
    enf.processDataEvent(testEvent);
    DCHBEngine engineHB = new DCHBEngine();
    engineHB.init();
    engineHB.processDataEvent(testEvent);

    //Compare HB momentum to expectation
    assertEquals(testEvent.hasBank("HitBasedTrkg::HBTracks"), true);
    assertEquals(testEvent.getBank("HitBasedTrkg::HBTracks").rows(), 1);
    assertEquals(testEvent.getBank("HitBasedTrkg::HBTracks").getByte("q", 0), -1);
    assertEquals(ClasMath.isWithinXPercent(16.0, testEvent.getBank("HitBasedTrkg::HBTracks").getFloat("p0_x", 0), 1.057), true);
    assertEquals(testEvent.getBank("HitBasedTrkg::HBTracks").getFloat("p0_y", 0) > -0.1, true);
    assertEquals(testEvent.getBank("HitBasedTrkg::HBTracks").getFloat("p0_y", 0) < 0.1, true);
    assertEquals(ClasMath.isWithinXPercent(16.0, testEvent.getBank("HitBasedTrkg::HBTracks").getFloat("p0_z", 0), 2.266), true);

    //TB reconstruction
    DCTBEngine engineTB = new DCTBEngine();
    engineTB.init();
    engineTB.processDataEvent(testEvent);

    //Compare reconstructed test event TB momentum and vertex to expectation
    assertEquals(testEvent.hasBank("TimeBasedTrkg::TBTracks"), true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBTracks").rows(), 1);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBTracks").getByte("q", 0), -1);
    assertEquals(ClasMath.isWithinXPercent(27.9, testEvent.getBank("TimeBasedTrkg::TBTracks").getFloat("p0_x", 0), 0.997), true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBTracks").getFloat("p0_y", 0) > -0.0702, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBTracks").getFloat("p0_y", 0) < 0.0438, true);
    assertEquals(ClasMath.isWithinXPercent(17.5, testEvent.getBank("TimeBasedTrkg::TBTracks").getFloat("p0_z", 0), 2.04), true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBTracks").getFloat("Vtx0_x", 0) < 0.2, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBTracks").getFloat("Vtx0_x", 0) > -0.2, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBTracks").getFloat("Vtx0_y", 0) < 0.18, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBTracks").getFloat("Vtx0_y", 0) > -0.228, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBTracks").getFloat("Vtx0_z", 0) < 0.885, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBTracks").getFloat("Vtx0_z", 0) > -0.0753, true);

    //Compare TB reconstruction cross position and direction to expectation

    //Region 1
    assertEquals(ClasMath.isWithinXPercent(72.6, testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("x", 0), 4.02), true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("y", 0 ) < 9.25, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("y", 0 ) > -11.78, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("ux", 0 ) < 0.214, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("ux", 0 ) > -0.186, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("uy", 0 ) < 0.054, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("uy", 0 ) > -0.054, true);
    assertEquals(ClasMath.isWithinXPercent(36.5, testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("uz", 0), 1 ),  true);

    //Region 2
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("x", 1 ) < 6.785, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("x", 1 ) > 1.357, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("y", 1 ) < 13.9, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("y", 1 ) > -17.8, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("ux", 1 ) < 0.14, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("ux", 1 ) > -1.41, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("uy", 1 ) < 0.077, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("uy", 1 ) > -0.077, true);
    assertEquals(ClasMath.isWithinXPercent(33.7, testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("uz", 1), 1) , true);

    //Region 3
    assertEquals(ClasMath.isWithinXPercent(44.3, testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("x", 2), -11.0), true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("y", 2 ) < 19.3, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("y", 2 ) > -25.0, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("ux", 2 ) < 0.0997, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("ux", 2 ) > -0.342, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("uy", 2 ) < 0.106, true);
    assertEquals(testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("uy", 2 ) > -0.118, true);
    assertEquals(ClasMath.isWithinXPercent(34.0, testEvent.getBank("TimeBasedTrkg::TBCrosses").getFloat("uz", 2), 1) , true);

  }
}
```

5.4 Appendix 4: Hough Transform Source Code

```
 * @author ziegler
 */
public class CircleHoughTrans {

    private double _RMin        = 10; // minimum helix radius
    private double _RWidth      = 2200;// maximum helix radius
    private double _RBinWidth   = 5;  // R bin width

    private int width  = 1400; // x range for helix (p~<3.25)
    private int height = 4100; // y range

    double R_Array[];
    int[] acc;
    int[] results;

    public CircleHoughTrans() {
        int nR = (int) (_RWidth/_RBinWidth);
        R_Array = new double[nR];
        for(int i = 0; i < nR; i++) {
            R_Array[i] = _RMin + (0.5 + i)*_RBinWidth;
        }
        acc     = new int[width * height];
        results = new int[width * height*3];
    }
    private Map<Integer, Integer> peaks = new HashMap<Integer, Integer>();
    public int minAccVal = 8;
public void fillAccumulator(List<Double>X, List<Double>Y) {
    double t;
    double x;
    double y;
    double x0, y0;
    for(int j = 0; j < X.size(); j++) {
        x = X.get(j);
        y = Y.get(j);

        for(int i = 0; i < R_Array.length; i++) {
            for (int theta=0; theta<360; theta++) {
                t = (theta * 3.14159265) / 180;
                x0 = (x - R_Array[i] * Math.cos(t)) ;
                y0 = (y - R_Array[i] * Math.sin(t)) ;
                if(x0 < width/2 && x0 > -width/2 && y0 < height/2 && y0 > -height/2) {
                    acc[(int)x0+ (int) (width/2) + (((int)y0 + (int) (height/2)) * width)] += 1;
                    if(acc[(int)x0+ (int) (width/2) + (((int)y0 + (int) (height/2)) * width)]>=minAccVal) {
                        peaks.put((int)x0+ (int) (width/2) + (((int)y0 + (int) (height/2)) * width),
                                acc[(int)x0+ (int) (width/2) + (((int)y0 + (int) (height/2)) * width)]);

                    }
                }
            }
        }
    }
}
List<Point2D> xy; // xy of seed points
public void findCircles(List<Double>X, List<Double>Y, List<Integer>L) {
    for(int j = 0; j < set.size(); j++) {
        set.clear();
    }
    for(int j = 0; j < CHTSeeds.size(); j++) {
        CHTSeeds.clear();
    }
    fillAccumulator(X, Y);
```

```
    // find the peaks of the accumulator array
    Map<Integer, Integer> sorted = peaks.entrySet()
                                        .stream()
                                        .sorted(Collections.reverseOrder(Map.Entry.comparingByValue()))
                                        .collect(
                                            toMap(Map.Entry::getKey, Map.Entry::getValue, (e1, e2) -> e2,
                                                LinkedHashMap::new));

    for(Map.Entry entry:sorted.entrySet()){

        xy = new ArrayList<Point2D>((int)entry.getValue());

        double t;
        int l;
        double x;
        double y;
        double x0, y0;
        for(int j = 0; j < X.size(); j++) {
            l = L.get(j);
            x = X.get(j);
            y = Y.get(j);

            for(int i = 0; i < R_Array.length; i++) {
                for (int theta=0; theta<360; theta++) {
                    t = (theta * 3.14159265) / 180;
                    x0 = (x - R_Array[i] * Math.cos(t)) ;
                    y0 = (y - R_Array[i] * Math.sin(t)) ;

                    if((int)x0+ (int) (width/2) + (((int)y0 + (int) (height/2)) * width)==(int)entry.getKey()) {
                        xy.add(new Point2D(x,y,l,R_Array[i],t));
            }
        }
    }
}

if(this.contains(xy,set)==false){
    set.add((ArrayList<Point2D>) xy);
    List<ArrayList<Point2D>> newset = this.split();
    for(int i = 0; i < newset.size(); i++) {
        if(newset.get(i).size()>=this.minAccVal &&
                this.contains(newset.get(i),CHTSeeds)==false) {
            CHTSeeds.add(newset.get(i));
            System.out.println("=================");
            for(int j = 0; j < newset.get(i).size(); j++) {
                Point2D p =newset.get(i).get(j);
                System.out.println(p._layer+" "+p._x+" "+p._y+" rad "+p._radius+" pt "+(p._radius*(5*Constants.LIGHTV
            }
        }
    }
}
private List<ArrayList<Point2D>> split() {
    hxy.clear();
    List<ArrayList<Point2D>> newset = new ArrayList<ArrayList<Point2D>>();
    //get list with single hit per layer
    List<Point2D> base = new ArrayList<Point2D>(); // xy of seed points
    base.add(xy.get(0));
    for(int i = 1; i < xy.size(); i++) {
        if(xy.get(i-1).getLayer() != xy.get(i).getLayer()) {
            base.add(xy.get(i));
        }
    }
    List<Point2D> offbase = new ArrayList<Point2D>();
    offbase.addAll(xy);
    offbase.removeAll(base);

    //get same layer hits
    for(Point2D p : offbase) {
        hxy.put(p.getLayer(),new ArrayList<Point2D>());
    }
    for(Point2D p : offbase) {
        hxy.get(p.getLayer()).add(p);
    }
    newset.add((ArrayList<Point2D>) base);
```

```
    for (Map.Entry<Integer, List<Point2D>> entry : hxy.entrySet()) {
        for(Point2D p : entry.getValue()) {
            List<Point2D> base2 = new ArrayList<Point2D>();
            base2.addAll(base);
            for(int k = 0; k < base2.size(); k++) {
                Point2D bp = base2.get(k);
                if(p.getLayer()==bp.getLayer()) {
                    base2.remove(k);
                    base2.add(k,p);
                }
            }
            newset.add((ArrayList<Point2D>) base2);
        }
    }
    return newset;
}
private boolean contains(List<Point2D> x0y0, List<ArrayList<Point2D>> set) {
    boolean isc = false;

    x0y0.sort(Comparator.comparing(Point2D::getLayer).thenComparing(Point2D::getR));

    for(int i = 0; i < set.size(); i++) {
        List<Point2D> xy = set.get(i);

        xy.sort(Comparator.comparing(Point2D::getLayer).thenComparing(Point2D::getR));

        if(xy.size() == x0y0.size()) {
            int n = 0;
            for(int k =0; k < xy.size(); k++) {
                if(xy.get(k).equals(x0y0.get(k)))
                    n++;
            }
            if(n == xy.size())
                isc = true;
        }
    }

    return isc;
}
```

5.5 Appendix 5: Taubin Circle fit and χ^2_ν Analysis Source Code

```
public class CircleFit {

    public static double[] taubinFit( double[][] points) {
        //Performs taubin algebraic circle fit

        int nPoints = points.length;
        if (nPoints < 3)
        throw new IllegalArgumentException("Too few points");

        //Calculate elements of data matrix M
        double[] centroid = getCentroidND(points);
        double Mxx = 0, Myy = 0, Mxy = 0, Mxz = 0, Myz = 0, Mzz = 0;
        for (int i = 0; i < nPoints; i++) {
            double Xi = points[i][0] - centroid[0];
            double Yi = points[i][1] - centroid[1];
            double Zi = Xi*Xi + Yi*Yi;
            Mxy += Xi*Yi;
            Mxx += Xi*Xi;
            Myy += Yi*Yi;
            Mxz += Xi*Zi;
            Myz += Yi*Zi;
            Mzz += Zi*Zi;
        }
        Mxx /= nPoints;
        Myy /= nPoints;
        Mxy /= nPoints;
        Mxz /= nPoints;
        Myz /= nPoints;
        Mzz /= nPoints;

        //Define parameter matrix A
        double Mz = Mxx+Myy;
        double Cov_xy = Mxx*Myy - Mxy*Mxy;
        double A0 = Mxz*Mxz*Myy + Myz*Myz*Mxx - Mzz*Cov_xy - 2*Mxz*Myz*Mxy + Mz*Mz*Cov_xy;
        double A1 = Mzz*Mz + 4*Cov_xy*Mz - Mxz*Mxz - Myz*Myz - Mz*Mz*Mz;
        double A2 = -3*Mz*Mz - Mzz;
        double A3 = 4*Mz;

        //Iterate through center x,y combinations
        double xnew = 0;
        double ynew = 1e+20;
        double epsilon = 1e-12;
        double iterMax = 20;

        for (int iter = 0; iter < iterMax; iter++) {
            double yold = ynew;
            ynew = A0 + xnew*(A1 + xnew*(A2+xnew*A3));
            if (Math.abs(ynew) > Math.abs(yold)) {
                //Moves in wrong direction
                xnew = 0;
                break;
            }
            double Dy = A1 + xnew*(2*A2+xnew*3*A3);
            double xold = xnew;
            xnew = xold - ynew/Dy;
            if (Math.abs((xnew - xold)/xnew) < epsilon) {
                break;
            }
            if (iter >= iterMax) {
                //Fit doesn't converge
                xnew = 0;
            }
            if (xnew < 0) {
            //Negitive root error"
                xnew = 0;
            }
        }
        //Return circle center and radius
        double[] centreRadius = new double[3];
        double det = xnew*xnew-xnew*Mz+Cov_xy;
        double x = (Mxz*(Myy-xnew) - Myz*Mxy)/(det*2);
        double y = (Myz*(Mxx-xnew) - Mxz*Mxy)/(det*2);
        centreRadius[0] = x+centroid[0];
        centreRadius[1] = y+centroid[1];
        centreRadius[2] = Math.sqrt(x*x + y*y + Mz);
        return centreRadius;
    }
```

```
    public static double[] getCentroidND(double[][] points) {
    //Finds centroid for set of points with arbitrary dimension
        int nPoints = points.length;
        int nDimensions = points[0].length;
        double[] centroid = new double[nDimensions];
        double[] sums = new double[nDimensions];

        for (int n = 0; n < nPoints; n++) {
            if (points[n].length != nDimensions)
                throw new IllegalArgumentException("All points must have equal dimention");
        for (int i = 0; i < nDimensions; i++) {
            sums[i] += points[n][i];
            }
        }
        for (int i = 0; i < nDimensions; i++) {
            centroid[i] = sums[i] / nPoints;
        }
        return centroid;
    }
}
public static double circChiSquared(double[][] points, double[] centerRadius) {
    //Calculates reduced chisquared for a circle fit
    double chi2 = 0;

    //Caculate residuals
    for (int i=0; i<points.length; i++) {
        chi2 += Math.pow(centerRadius[2]-Math.sqrt(Math.pow(points[i][0]-centerRadius[0],2)
                -Math.pow(points[i][1]-centerRadius[1],2)),2)/(Math.pow(points[i][0]-centerRadius[0],2)
                -Math.pow(points[i][1]-centerRadius[1],2));
    }
    //Account for degrees of freedom
    return chi2/(points.length-4);
}

}
```