Application of Machine Learning to π_0 photoproduction from CLAS/g9a

Chan Kim

The George Washington University

Igor Strakovsky and William Briscoe

CLAS Collaboration Meeting (HSWG)
November 15, 2018

Overview

1. Motivation
2. Target Selection
3. Hydrogen Contamination on Carbon Target
4. Next Steps

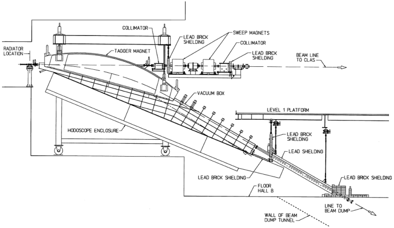
CLAS g9a/FROST Experiment

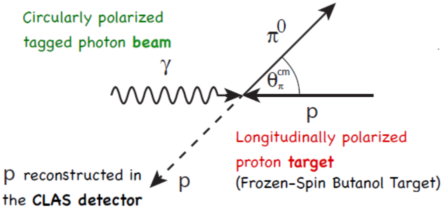
- Polarization Observables → Helicty Amplitudes → Resonances (PWA)
- Polarizable: Incoming photons, target & recoiling nucleons

	UP_T and UP_R	UP_T and P_R	P_T and UP_R	P_T and P_R
UP_B	$\frac{d\sigma}{d\Omega}$	P	T	$T_{x'}, T_{z'}, L_{x'}, L_{z'}$
LP_B	$-\Sigma$	$O_{x'}, (-T), O_{z'}$	$H, (-P), -G$	
CP_B		$-C_{x'}, -C_{z'}$	$F, -E$	

UP, P, LP, CP, B, T, R denote unpolarized, polarized, linearly polarized, circularly polarized, beam, target, and recoil, respectively.

- g9a/FROST - Circularly polarized photons with $E_\gamma \approx 0.4 - 2.4$ GeV and longitudinally polarized proton target:





Helicity Asymmetry E

- Double polarization observable E is the helicity asymmetry of the cross section:

$$E = \frac{\sigma_{3/2} - \sigma_{1/2}}{\sigma_{3/2} + \sigma_{1/2}} \qquad \text{for } \frac{3}{2} \;\&\; \frac{1}{2} \text{ are total helicty states}$$

- $\frac{d\sigma}{d\Omega}$ of polarized beam & polarized target for E (theo. & exp.):

$$\left(\frac{d\sigma}{d\Omega}\right)_{\frac{1}{2},\frac{3}{2}} = \frac{d\sigma_0}{d\Omega}(1 \mp (P_z P_\lambda)_{\frac{1}{2},\frac{3}{2}} E) \qquad \left(\frac{d\sigma}{d\Omega}\right)_{\frac{1}{2},\frac{3}{2}} = \frac{N_{\frac{1}{2},\frac{3}{2}}}{A \cdot F \cdot \rho \cdot \Delta x_i}$$

- E is measured via:

$$\boxed{E = \left[\frac{1}{D_f}\right]\left[\frac{1}{P_z P_\lambda}\right]\left[\frac{N_{\frac{3}{2}} - N_{\frac{1}{2}}}{N_{\frac{3}{2}} + N_{\frac{1}{2}}}\right]}$$

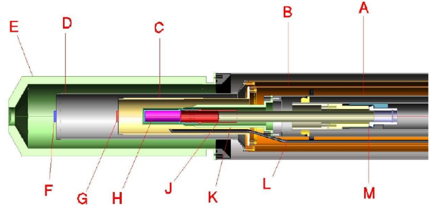
$D_f =$ dilution factor

$P_z =$ Polarization of target in $\hat{z}$

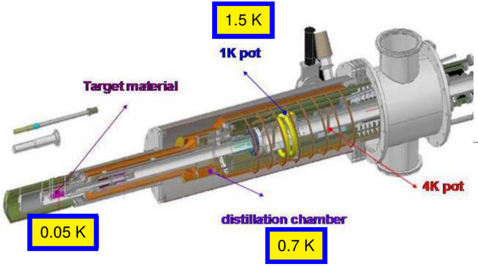
$P_\lambda =$ Polarization of beam

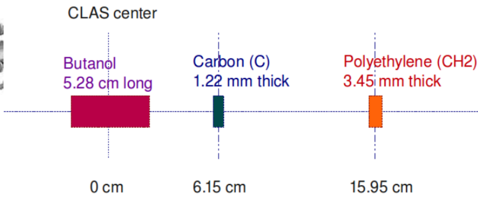
$N_{\frac{3}{2},\frac{1}{2}} = \#$ of events

g9a/FROST Target setup



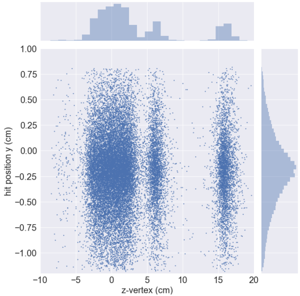
Side view of FROST target with beam entering from the right. (A) Primary head exchanger, (B) 1 K heat shield, (C) Holding coil, (D) 20 K heat shild,(E) Outer vacuum can , (F) Polyethylene target,(G) Carbon target,(H) Butanol target ,(J) Target insert, (K) Mixing chamber, (L) Microwave waveguide, and (M) Kapton cold seal.



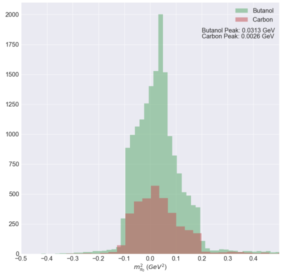


ML Objectives: Target Selection & Ice on Carbon

- Target Selection
  - Events with z-vertex $\in [2, 5]$cm, uncertain whether γ hit Butanol or Carbon

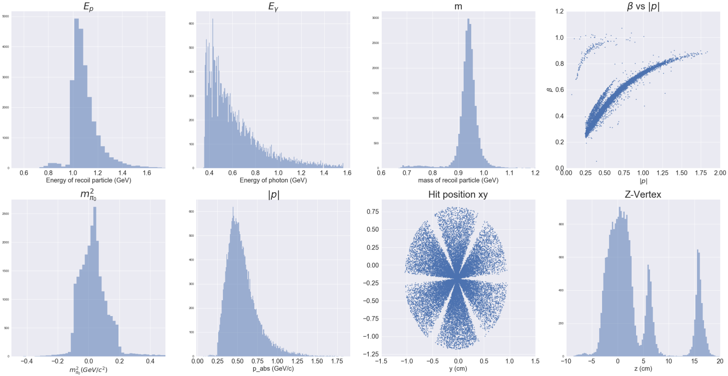


- Ice on Carbon
  - Carbon events (bound-nucleon) expected to have broader $m^2_{\pi_0}$ peak due to Fermi motion.
  - Sharp peak (free-nucleon) observed in the Carbon target region. Carbon events are scaled by $\sim$10.



Banks Used

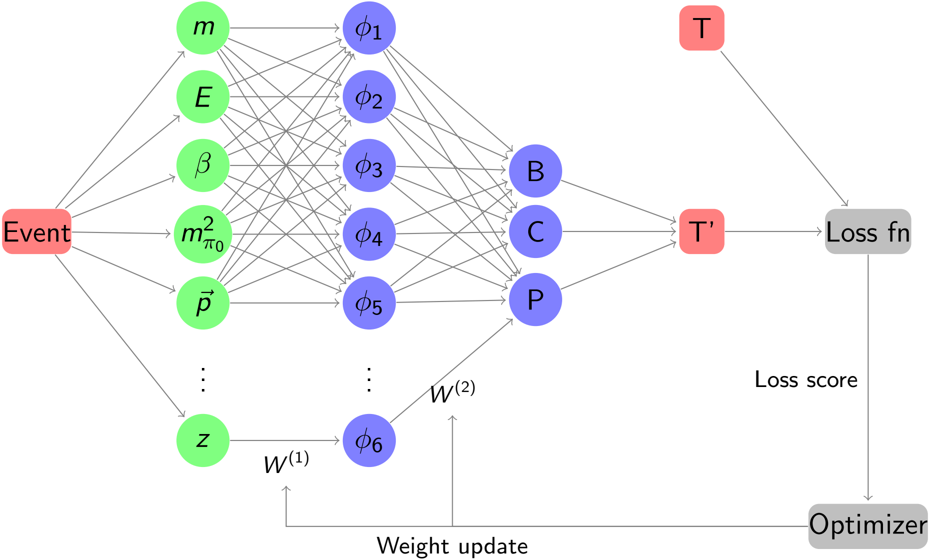
- GPID - pid, E, $\vec{p}$, β_c, β_m, m, and E_γ
- MVRT - vertex positions
- Only single tracked events with protons
- Removed events outside of target cup ($r = 7.5mm$) - He-Bath outside
- Removed events with $E_\gamma \approx 0$
- No Energy loss correction yet
- No fiducial cuts yet



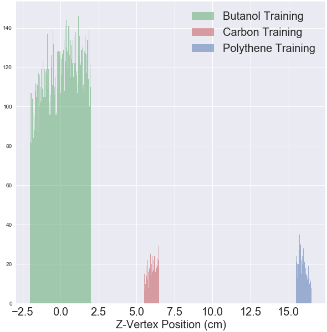
Neural Network Model Setup

- Two fully-connected (dense) neural layers
  1. Dense layer with 15 nodes - 15 parameters:
     - E, β, β_{diff}, β_m E_γ, m, $m^2_{\pi_0}$, pid,$|p|$, p_x, p_y, p_z, x, y, and z.
     - Too many parameters + insufficient train data $\rightarrow$ Too specific training $\rightarrow$ Overfitting (fail)
  2. Dense layer with 3 nodes - one for each target
     - For each event, this layer returns an array of 3 probability scores (butanol, carbon, or polythene) that sum to 1
- Optimizer used: AdamOptimizer
- Loss function used - Sparse categorical cross entropy:
  - $H_{y'}(y) = -\sum_i y'_i \log(y_i)$,where y_i is the predicted target and y'_i is the true target
- Python and Tensorflow

Neural Network Training Flowchart

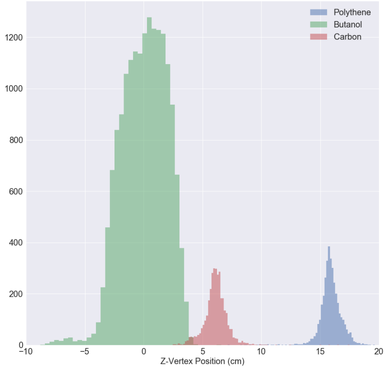


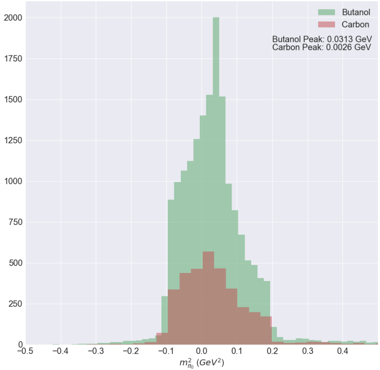
Training Data Selection



- Events with z-vertex position in close proximity of physical target region
  - Butanol $\in$ [-2, 2]cm
  - Carbon $\in$ [5.5, 6.5]cm
  - Polythene $\in$ [15.5, 16.5]cm

Result on Target Selection





- Classified Carbon events from Butanol in z-vertex $\in [2.5, 4.5]$cm
- Some Carbon events in Polythene regions & Polythene events in Butanol region.
- Tail of Butanol events in Carbon region are missing. Under review [S. Fegan].

Evidence of Hydrogen Contamination on Carbon

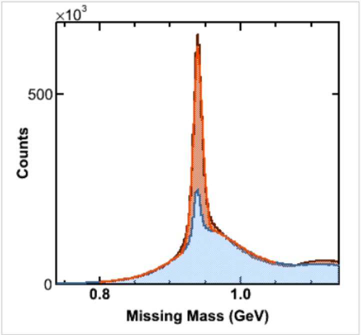


Figure 2: Missing-mass distribution for the π+n channel from FROST g9a data. W = 1.25 - 1.50 GeV, integrated over all angles. Events in the red histogram are from the butanol target and events in the blue histogram are from the 12C target with z-vertex larger 5.0 cm and smaller than 7.5 cm. The blue histogram is scaled by 5.26. The FROST distribution from the 12C target region show a **narrow peak** at the mass of then neutron.

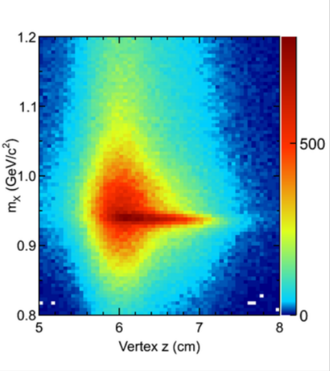
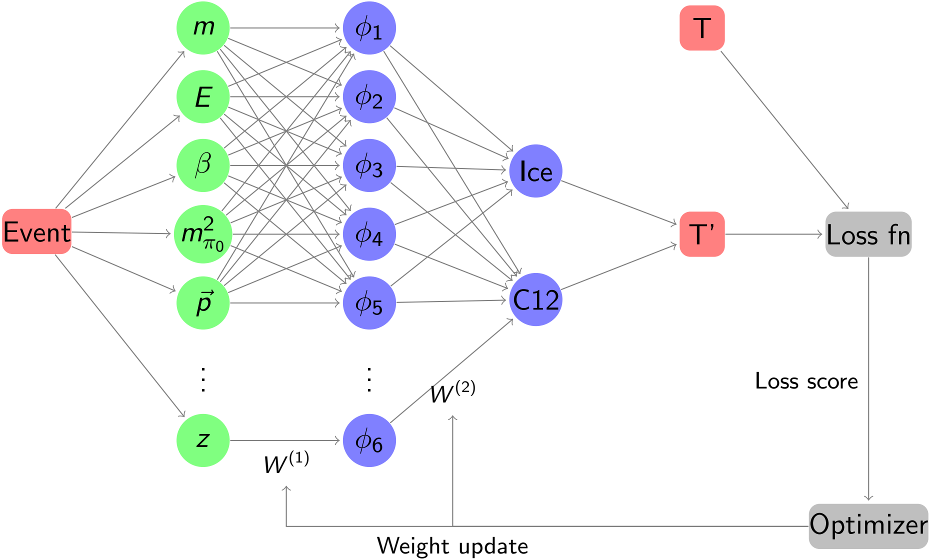


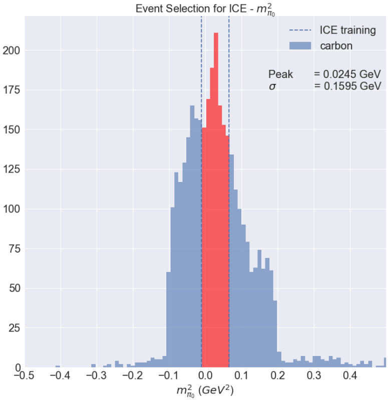
Figure 3: Missing-mass distribution as a function of the MVRT z vertex of the π+. The shape of the missing-mass distribution strongly changes with z. Event selection: $p_\pi > 0.2$ GeV/c and $\theta_\pi > 20^\circ$.

- Sharp peak at downstream end of Carbon foil → ice built up while cooling the target
- Ice formed on the right side of Carbon target: Z-vertex ∈ [6, 7]cm
- Plots from [Steffen Strauch]'s Analysis page of FROST Wikipage

Neural Network Training Flowchart for ice vs Carbon

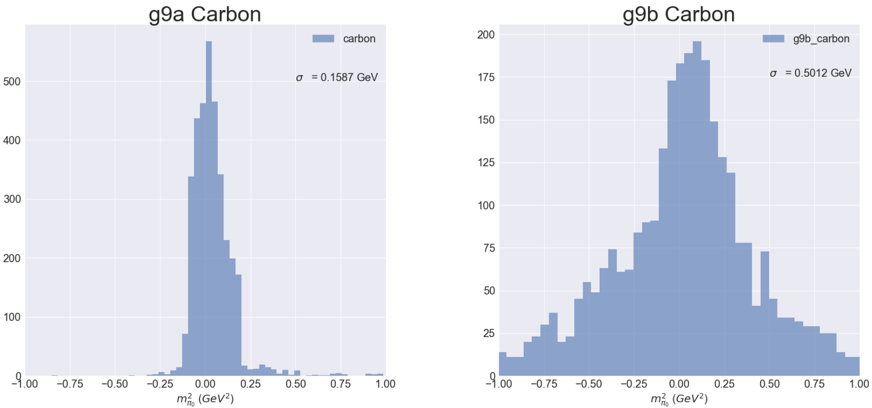


Training Data for ice



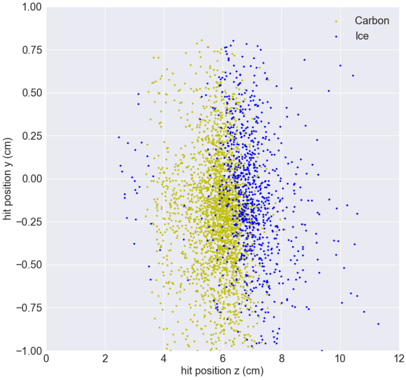
- Tight cut on the $m^2_{\pi_0}$ peak on g9a-Carbon data (or MC sim) as ice events [F. Klein].
  - Bound-nucleon (fermi p) $\rightarrow$ broader distribution $\rightarrow$ Sharper peaks from free-nucleon (ice) Broad background from bound-nucleon (carbon)

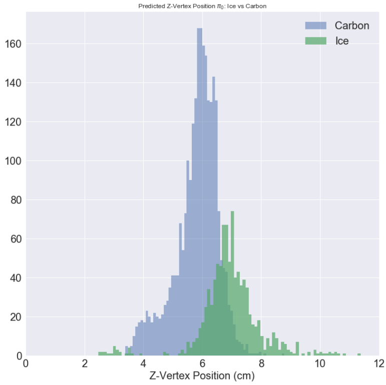
Training Data for Carbon from g9b



- g9b-carbon $m^2_{\pi_0}$ peak broader than g9a/Carbon $\rightarrow$ No ice on g9b
- During g9b, Carbon target was moved further in downstream.
- Shifted Z-vertex of g9b-Carbon events to use as training events for g9a [F. Klein].

Result on Hydrogen Contamination of Carbon Target





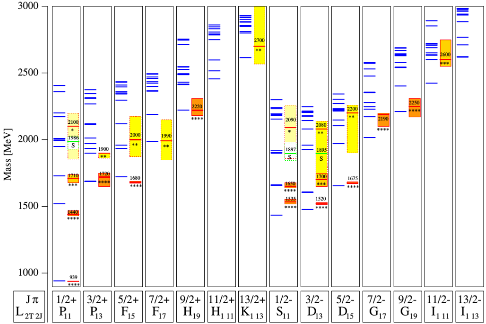
- Classified ice events from Carbon target in z-vertex $\in$ [6, 8]cm
- It is likely that ice was formed in 20 K heat shield in between Carbon and Polythene targets.

Next Steps

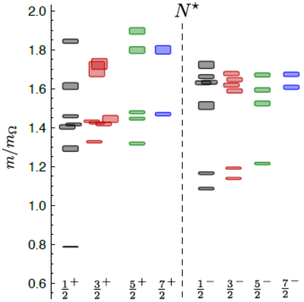
- Apparently, classification on target selection and extraction of ice events were successful. However, it is still a blackbox...
  - Need to understand underlying logic of the model (Tensorboard).
  - Find ways to implement physics (relativistic kinematics, branching ratios, forbidden angles, etc) into optimizer or loss function.
- Nuclear physics experiments data cannot have 100% confident true values for training
  - Try Unsupervised Learning techniques: Clustering
  - Find ways to measure efficiency and accuracy of predictions
- Need more data to avoid overfitting problem.
- Computational cost is too high for small fraction of the total data.
- Energy loss reconstruction via ML regression.

Backup Slides

Backup: Constituent Quark Models & LQCD Predictions of Non-Strange Baryon Resonances



Constituent Quark Model

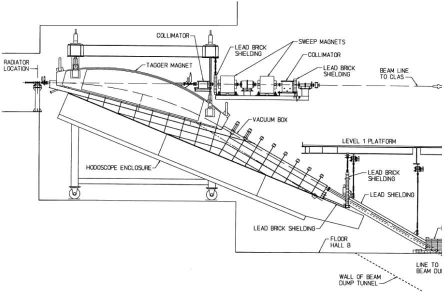


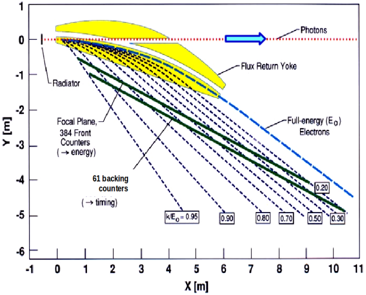
Lattice QCD

- Constituent Quark Models predicted states: 64 N^* & 22 Δ^*
- Experimentally confirmed state: 26 N^* & 22 Δ^*

Backup: Hall B Photon Tagger

- Bremsstrahlung radiation due to slowing of electrons by EM field of radiator (gold foil or thinyo diamond)
- Determine incoming photon energy of $\vec{\gamma}\vec{p} \to \pi^0 p$ by $E_\gamma = E_0 - E_e$
- g9a/FROST - circularly polarized photons with $E_\gamma \approx 0.4 \sim 2.4\ GeV$
- Tagger was built by the GWU, CUA, & ASU nuclear physics group





Backup: Circularly Polarized Photon Beam

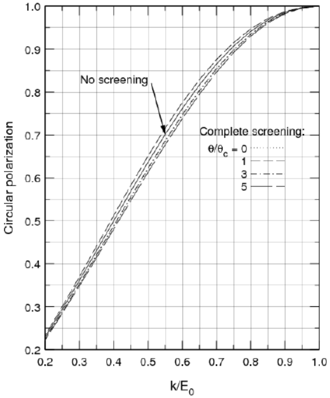
Linearly Polarized Electron Beam $\xrightarrow{\textit{Bremsstrahlung}}$ Circularly Polarized Photon Beam

- Polarization transfer:

$$P(\gamma) = P(e)\frac{4x - x^2}{4 - 4x + 3x^2}$$

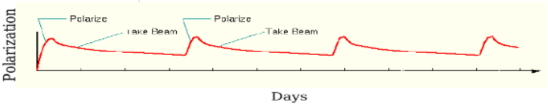
$$x = \frac{k}{E_0} = \frac{\text{photon energy}}{\text{incident electron energy}}$$

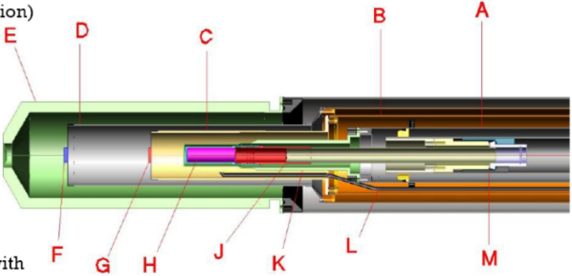
H. Olsen and L.C. Maximon, Phys. Rev. 114, 887 (1959)



Backup: Frozen Spin Target

The FroST target and its components:
A: Primary heat exchanger
B: 1 K heat shield
C: Holding coil
D: 20 K heat shield
E: Outer vacuum can (Rohacell extension)
F: CH2 target
G: Carbon target
H: Butanol target
J: Target insert
K: Mixing chamber
L: Microwave waveguide
M: Kapton coldseal

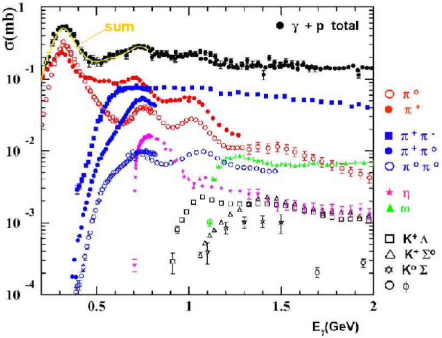




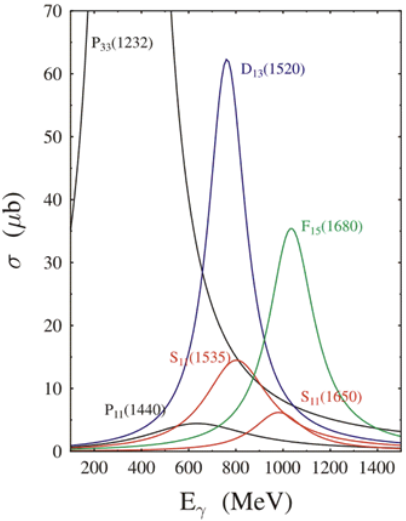
Performance Specs:
Base Temp: 28 mK w/o beam, 30 mK with
Cooling Power: 800 μW @ 50 mK, 10 mW @ 100 mK, and 60 mW @ 300 mK
Polarization: +82%, -90%
1/e Relaxation Time: 2800 hours (+Pol), 1600 hours (-Pol)
Roughly 1% polarization loss per day.

C. Keith et al. Nucl Instrum Meth A 684, 27 (2012)

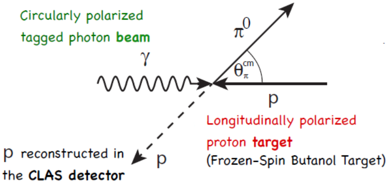
Backup: CLAS g9a/FROST Data



- Select only $\vec{\gamma}\vec{p} \rightarrow \pi^0 p$ events
- $\vec{\gamma}\vec{p} \rightarrow \pi^0 p$ resonance channels
- Appropriate enegy bins - include all resonances ($\lesssim$ 1500 MeV)



π^0 photoproduction



- From *T Matrix* to *Helicity Amplitudes* of $\vec{\gamma}\vec{p} \to \pi^0 p$:

$$\langle \mathbf{q}\ m_{s'} |\, T\, | \mathbf{k}\ m_s\ \lambda \rangle = \boxed{\langle m_{s'} |\, \mathbf{J}\, | m_s \rangle} \cdot \epsilon_\lambda(\mathbf{k}) \quad \Rightarrow \quad H_i(\theta) \equiv \langle \lambda_2 |\, \mathbf{J}\, | \lambda_1 \rangle$$

- 4 Complex Helicity Amplitudes:

$$H_1(\theta) = \left\langle +\frac{3}{2} \middle| \mathbf{J} \middle| +\frac{1}{2} \right\rangle \qquad H_2(\theta) = \left\langle +\frac{1}{2} \middle| \mathbf{J} \middle| +\frac{1}{2} \right\rangle$$

$$H_3(\theta) = \left\langle +\frac{3}{2} \middle| \mathbf{J} \middle| -\frac{1}{2} \right\rangle \qquad H_4(\theta) = \left\langle +\frac{1}{2} \middle| \mathbf{J} \middle| -\frac{1}{2} \right\rangle$$

Backup: Complete Experiment - 8 Polarization Observables

- Polarizable: incoming photons, target & recoiling nucleons
- 8 well chosen observables at fixed E_γ & angle $\rightarrow$ 4 helicity amplitudes

	UP_T and UP_R	UP_T and P_R	P_T and UP_R	P_T and P_R
UP_B	$\frac{d\sigma}{d\Omega}$	P	T	$T_{x'}, T_{z'}, L_{x'}, L_{z'}$
LP_B	$-\Sigma$	$O_{x'}, (-T), O_{z'}$	$H, (-P), -G$	
CP_B		$-C_{x'}, -C_{z'}$	$F, -E$	

UP, P, LP, CP, B, T, R denote unpolarized, polarized, linearly polarized, circularly polarized, beam, target, and recoil, respectively.

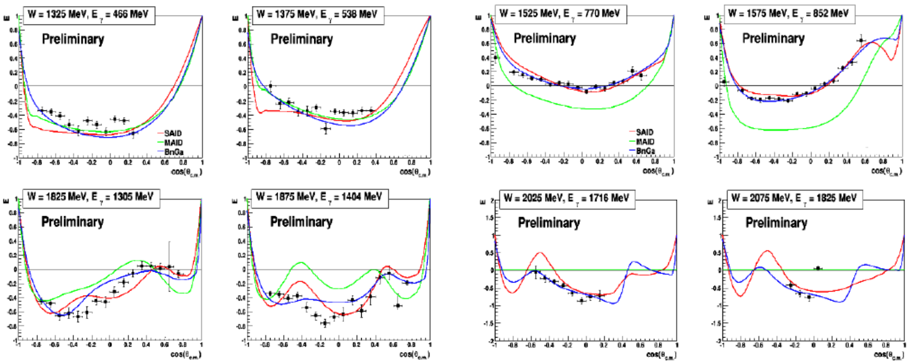
- Helicity asymmetry E related to other observables via Fierz identities:

$$E^2 + F^2 + G^2 + H^2 = 1 + P^2 - \Sigma^2 - T^2$$

$$FG - EH = P - \Sigma T$$

$$\vdots$$

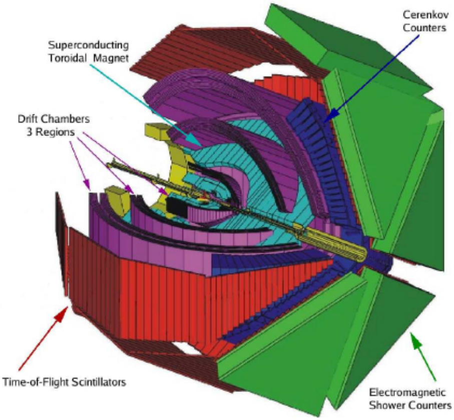
Comparison with PWA (SAID, MAID, BnGa) predictions

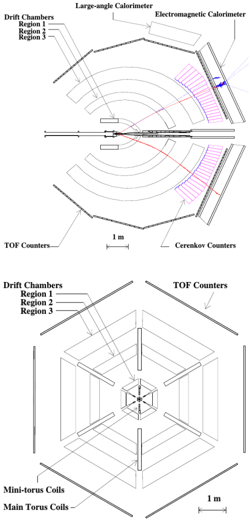


Sample of Helicity Asymmetry E versus CM angle, θ_{cm} in ranges of $E_\gamma = 466 - 1825\ MeV$ and $W = 1325 - 2075\ MeV$. The red, green, and black lines correspond to results from **SAID**, **MAID**, & **BnGa** respectively.

- Measured asymmetry E will be compared to PWA predictions
- Measured asymmetry E into SAID data base → new pole positions
- **SAID**, **MAID** & **BnGa** models agree at lower energies & deviate at higher energies

CEBAF Large Acceptance Spectrometer





Overtraining Limits

- Overtraining:

Excess training with only specific training data

$\downarrow$

Classification succeeds on training data, but fails on actual data

- Must determine adequate classifying variables & size of training data
- Rule of thumb for Decision Tree algorithm:

$$L_D(h) \quad \leq \quad L_S(h) + \sqrt{\frac{(n+1)\log_2(d+3) + \log(2/\delta)}{2m}}$$

$L_D(h)$ = Error of classification on actual data set
h = Error of classification on a training data set
δ = Confidence level of randomly selected training data points
n = Number of nodes

$L_S(h)$ = Error of classification on a training data set
d = Number of variables
m = Size of training data sets
• n & d inversely proportional to L_s