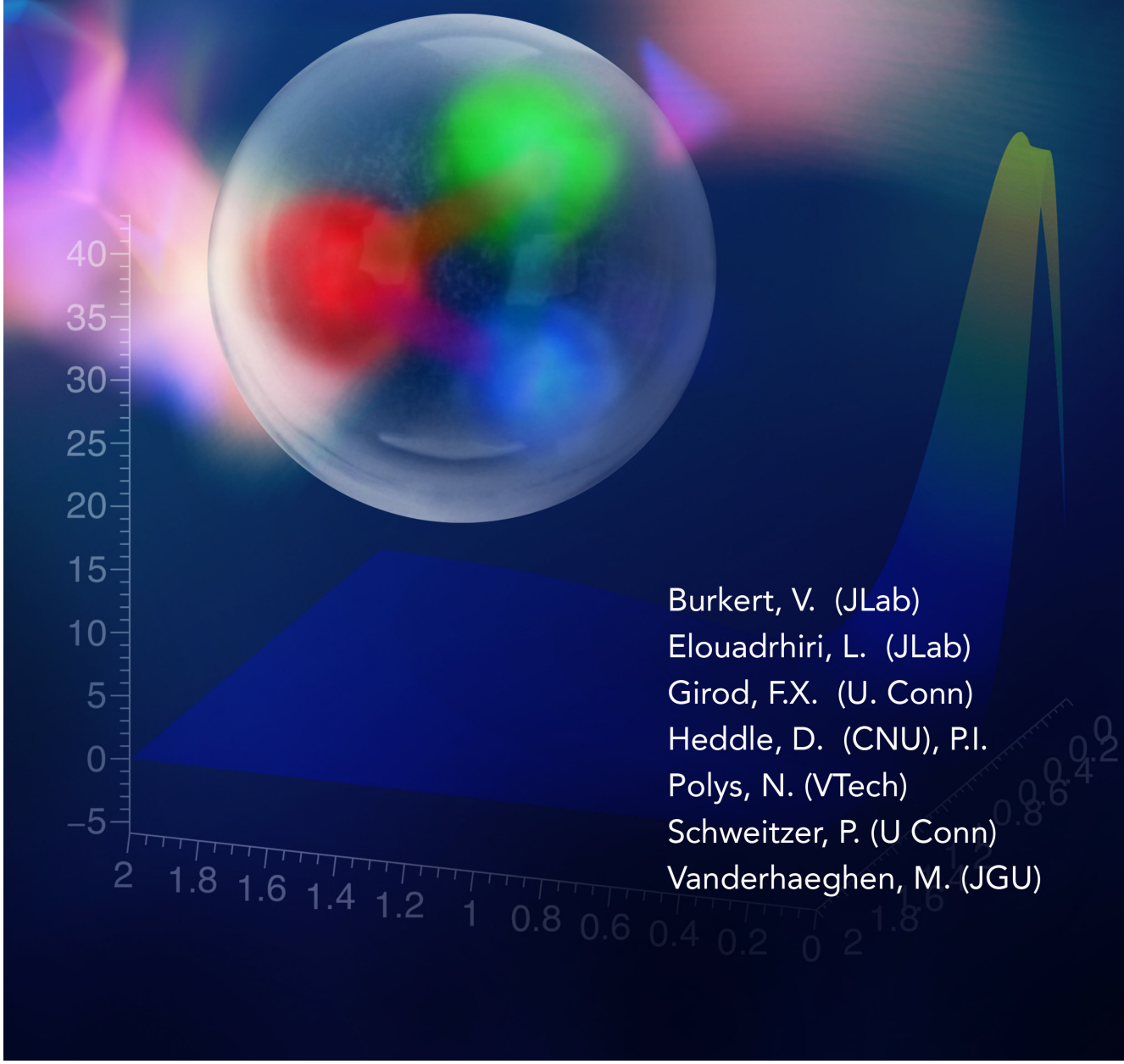
VISUALIZING FEMTO-SCALE DYNAMICS

Center for Nuclear Femtography



Burkert, V. (JLab)
Elouadrhiri, L. (JLab)
Girod, F.X. (U. Conn)
Heddle, D. (CNU), P.I.
Polys, N. (VTech)
Schweitzer, P. (U Conn)
Vanderhaeghen, M. (JGU)

Visualizing Femto-Scale Dynamics – Final Report

Center for Nuclear Femtography

Executive Summary

Overview

CNF19-09 is among the approved proposals that are the pillars of the establishment of the Center of Nuclear Femtography (CNF). Our team has been working actively on the project since the date of its approval in a coordinated interdisciplinary effort. We have met all our milestones and deliverables as outlined in the original proposal. During the first year we developed the ground work towards visualization of the Femto-Scale Dynamics, with the first 3D images emerging. We have also defined the next steps and began to elaborate the next phase of this project as an integral part of the mid-term and long-term vision of the CNF.

As the first step, we assembled a strong team of experts in the field of Nuclear Femtography, across a spectrum of disciplines, nuclear physics both theory and experiment, computing, imaging and visualization. We have bi-weekly meetings via video-conference, face to face meetings in preparation of the mid-term report and the presentations as well as focused subgroup technical meetings.

The aim of our project focuses on the study of the mechanical and gravitational sub-structure of the proton, as this is a new direction of nuclear research. Major progress in theoretical formalism of the interior structure of the nucleon over the last 15-20 years has led to breakthroughs in our understanding of the theory of quarks and gluons, through the Generalized Parton Distributions (GDPs) framework, with deeply virtual Compton scattering (DVCS) as a tool to study the structure of the nucleon from experiment. It has been demonstrated recently, for the first time, that we can access the mechanical and gravitational properties of the nucleon through DVCS data. This has given birth to a new field of research untapped so far.

This project is timely as a new generation of high precision DVCS experiments has started with the Jefferson Lab (JLab) 12 GeV upgrade and new experiment are being planned with the Electron Ion Collider. DVCS is the flagship program of the study of GPDs with the 12 GeV upgrade project at Jefferson Lab in Hall A, Hall B and Hall C. The first set of data taking has been successfully completed and data are being prepared for publication. The 12 GeV experimental program accompanied with further development of our project is poised to revolutionize our understanding of the internal structure of the nucleon. This will require further theoretical development, sophisticated and efficient computing algorithms for data analysis, physics extraction and interpretation of the results. EIC is designed to delve deeper than ever

into the fundamental building blocks of matter and the study of the GPDs via exclusive processes is the central component of its scientific program. Part of our future mission is to perform systematic studies through detailed simulations of DVCS and have timely input to the EIC machine, interaction region and detector design specifications.

Project Progress Summary

1. As part of the initial phase of the establishment of the Center for Nuclear Femtography, our group looked to the characterization and visualization of the internal mechanical and dynamical properties of the nucleon, focusing on the proton through DVCS data analysis. The data were generated from a realistic Reference Model (RM), as extracted from data already available from the 6 GeV DVCS experiments at JLab.
2. This RM has been used to simulate DVCS events based on a Fast Monte Carlo event generator that incorporated various experimental apparatus, such as CLAS12 and the JLEIC detector systems, as the basis for the Monte Carlo data convoluted with acceptance and resolution functions to obtain pseudo data as differential cross sections and beam-spin asymmetries. These "pseudo data" were then subjected to fits in order to build a new RM1 now modified from detector properties and from limited data input. This realistic RM1 was then used to obtain the mechanical and gravitational properties of the proton, in terms of normal and tangential forces on the quarks and the pressure distribution. The radial dependencies of the force were then subject to detailed visualization.
3. The workflow required conversions of data formats from the event generation and input to the visualization tools. We adapted the efficient data formats developed for CLAS12 to our project.
4. In order to obtain accurate results and a realistic visualization experience we used the simulated DVCS reaction and traced all three final state particles through the full magnetic fields and the tracking chambers of the large acceptance detector and visualized those tracks in virtual reality displays and in movies that incorporate all readout wires of the tracking chambers. This turned out to be a great educational tool for non-nuclear scientists, especially for the undergraduate students involved in the project.
5. The second part of the project involved the reconstruction of the proton's internal structure from the simulated "pseudo data" but now convoluted with the experimental resolutions and acceptances and the truncations in kinematical coverage and with incomplete sets of measurements that only employed beam-spin asymmetries and normalized cross sections. From these fits, the gravitational form factor **D(*t*)** could be determined, and used to extract the radial dependence of the ***normal*** and of the ***tangential*** forces on quarks in the proton.
6. The results have been made available through an interactive web site hosted at Virginia Tech, which will be further developed to expand the description of the science processes underlying the current results in laymen terms, and improve the visual effects.
7. The results of this analysis have been the basis for the visualization of the protons interior mechanical properties. The entire process from the initial RM through the mechanical structure of the proton and its visualization required an interdisciplinary team effort with expertise in nuclear physics, both from theory and experiment, data handling and the application of visualization tools.

8. This team brought together the broad range of expertise needed for a multidisciplinary approach to start addressing these questions in fundamental nuclear physics. The team also started working to establish collaboration with other experts in the field to plan the nature and scope of the activities that can be the basis of a future proposal to the Commonwealth of Virginia for a sustainable Center for Nuclear Femtography.

Theory and Reference Model

There are two processes contributing to the deep electroproduction of a photon off a proton ep→epɣ depending on whether the photon is emitted on the leptonic or hadronic side. In the former case, the amplitude is parameterized by the proton electromagnetic Form Factors (FFs), which is known as the Bethe-Heitler (BH) process. In the latter case is case, the process is parameterized by Generalized Parton Distributions (GPDs) and is known as Deeply Virtual Compton scattering (DVCS). In both cases the proton stays intact and is detected in the final state.

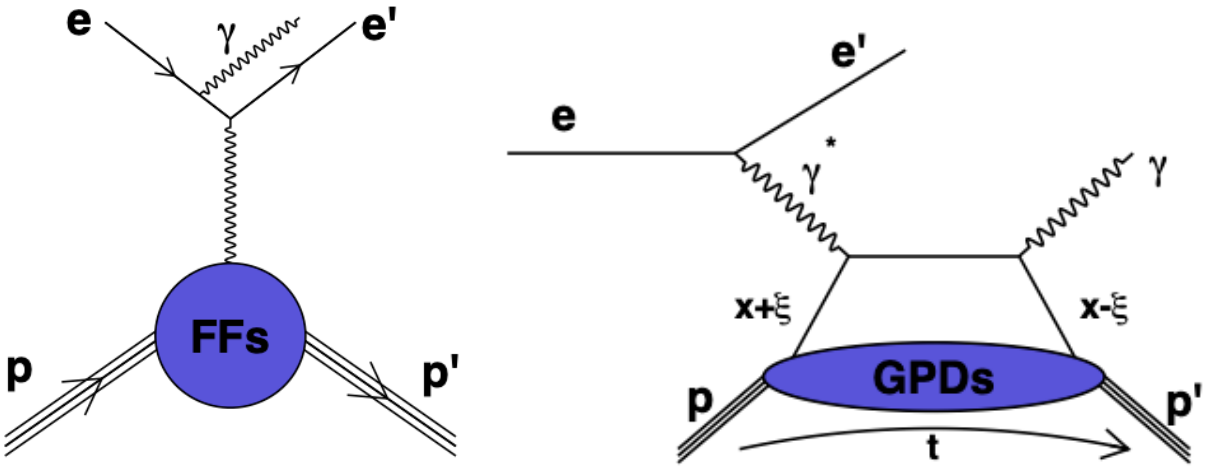


Figure 1. Two processes contributing to the deep electroproduction of a photon off a proton ep→epɣ

On the lepton side the reaction is described by the virtuality of the exchanged photon $Q^2 = -\gamma^{*2} = -(e-e')^2$ and the Bjorken variable $x_B = Q^2 / 2M\nu$ where ν is the virtual photon energy and M is the proton mass. On the hadron side, the relevant variables are the Mandelstam variable $t = (p'-p)^2$ corresponding to the momentum transferred to the nucleon, and the angle ϕ between the lepton scattering plane and the hadron scattering plane. In the physical region Q^2 is always positive and t is always negative.

In the Bjorken regime where Q^2 is large compared to M, x_B is fixed, and -t is small compared to Q^2 the DVCS amplitude factorizes as illustrated in the handbag diagram, and the nonperturbative part of the amplitude is parameterized by the GPDs, which depend on the variables x, ξ, and t. The variables x and ξ correspond to the light-cone momentum fractions of the active quark, x+ξ before and x-ξ after interaction. The variable x is integrated over in the loop and therefore not directly observable. At leading twist there are four chiral even GPDs $H, \tilde{H}, E,$ and $\tilde{E}$. The DVCS amplitude at leading order depends on the GPD H through the Compton Factor Factor (CFF)

$$\mathcal{H}(\xi,t) = \int_{-1}^{1} d\mathsf{x} H(x,\xi,t)\left(\frac{1}{\xi - x - i\epsilon} - \frac{1}{\xi + x - i\epsilon}\right)$$

The real and imaginary parts of the CFF can be related through a fixed-t dispersion relation

$$\mathrm{Re}\mathcal{H}(\xi,t) = D(t) + \frac{1}{\pi}\mathcal{P}\int_0^1 d\xi' \mathrm{Im}\mathcal{H}(\xi',t)\left(\frac{1}{\xi - \xi'} - \frac{1}{\xi + \xi'}\right)$$

The subtraction constant is related to the D-term by

$$D(t) = 2\int_{-1}^{1} dz \frac{D(z,t)}{1-z} \quad \text{with} \quad D(z,t) = (1-z^2)\sum_{n\ \mathrm{odd}} d_n(t) C_n^{3/2}(z)$$

where $C_n^{3/2}$ are Gegenbauer polynomials. The two main observables we will be working with are the unpolarized cross-sections and the beam-spin asymmetries (BSA). Their precise dependence on the CFF is somewhat lengthy to explicitly write down, suffice it to say here that the unpolarized cross-sections are mostly sensitive to the Real part of the CFF and the BSA are mostly sensitive to the Imaginary part of the CFF. We summarize the sensitivity of various observables to difference CFFs in the table below. To complete the full program of GPD extraction, all observables must be measured. We only mention the beam charge asymmetry once in the table, but note that all other polarization observables would also be accessible with a positron beam. Finally, we do not address here flavor separation or gluon GPDs, which necessitate the analysis of exclusive meson production reactions.

Observable	**CFF leading Sensitivity**	**subleading Sensitivity**
Unpolarized cross-sections	$\mathrm{Re}\mathcal{H}$	$\mathrm{Re}\tilde{\mathcal{H}}$, $\mathrm{Re}\mathcal{E}$
Beam spin asymmetries	$\mathrm{Im}\mathcal{H}$	$\mathrm{Im}\tilde{\mathcal{H}}$, $\mathrm{Im}\mathcal{E}$
Longitudinal target single spin asymmetries	$\mathrm{Im}\tilde{\mathcal{H}}$	$\mathrm{Im}\mathcal{H}$, $\mathrm{Im}\mathcal{E}$, $\mathrm{Im}\tilde{\mathcal{E}}$
Longitudinal target double spin asymmetries	$\mathrm{Re}\tilde{\mathcal{H}}$	$\mathrm{Re}\mathcal{H}$, $\mathrm{Re}\mathcal{E}$, $\mathrm{Re}\tilde{\mathcal{E}}$
Transverse target single spin asymmetries	$\mathrm{Im}\mathcal{H}$, $\mathrm{Im}\mathcal{E}$	$\mathrm{Im}\tilde{\mathcal{H}}$, $\mathrm{Im}\tilde{\mathcal{E}}$
Transverse target double spin asymmetries	$\mathrm{Re}\mathcal{H}$, $\mathrm{Re}\mathcal{E}$	$\mathrm{Re}\tilde{\mathcal{H}}$, $\mathrm{Re}\tilde{\mathcal{E}}$
Beam charge asymmetries	$\mathrm{Re}\mathcal{H}$	$\mathrm{Re}\tilde{\mathcal{H}}$, $\mathrm{Re}\mathcal{E}$

The strategy of this analysis is to constrain the Imaginary part of the CFF using the BSA, and to extract the D-term as a subtraction constant in the dispersion relation, with the Real part constrained by the unpolarized cross-sections. The Imaginary part of the CFF in the Reference Model uses a global parameterization of the form

$$\mathrm{Im}\mathcal{H}(\xi,t) = \frac{N}{1+\xi}\left(\frac{2\xi}{1+\xi}\right)^{-\alpha(t)}\left(\frac{1-\xi}{1+\xi}\right)^{b}\left(1 - \frac{1-\xi}{1+\xi}\frac{t}{M_I^2}\right)^{-1}$$

where the parameters are a normalization constant *N*, a Regge-trajectory $\alpha(t) = \alpha_0 + \alpha_1 t$, an exponent *b* to adjust the large ξ behavior, and the mass M_I controls the *t*-dependence. The subtraction constant is parameterized similarly as a multipole

$$D(t) = D(0)\left(1 - \frac{t}{M_d^2}\right)^{-\alpha_D}$$

We make additional model assumptions to proceed with the physics extraction. First, considering that the separation of the different GPDs / CFFs necessitate additional observables (target spin asymmetries) we only treat the GPD H to start with. The study of additional polarization observable is a natural future extension of this work. Second, since the DVCS / BH processes are not directly sensitive to the quark flavors, we assume the dominance of the flavor singlet $u \approx d \approx Q/2$ noting that this can be justified for the D-term in the large N_C limit. Similarly, at the moment the only input for the gluon D-term comes from lattice QCD, until a larger Q^2 lever arm becomes available experimentally at the EIC.

In the Gegenbauer expansion of the D-term we are more particularly interested in the first term $d_1(t)$ which is expected to dominate in the asymptotic limit on infinitely large renormalization scale. This term appears in the Energy Momentum Tensor (EMT) decomposition into the Gravitational Form Factors M_2, J, and d_1: corresponding respectively to the mass, angular momentum, and force distributions:

$$\langle P' | T_{\mu\nu}^{q,g} | P \rangle = \bar{u}(P') \left[M_2^{q,g} \frac{P_\mu P_\nu}{M} + J^{q,g} \frac{i\left(P_\mu \sigma_{\nu\rho} + P_\nu \sigma_{\mu\rho}\right)\Delta^\rho}{2M} + d_1^{q,g} \frac{\Delta_\mu \Delta_\nu - g_{\mu\nu}\Delta^2}{5M} \right] u(p)$$

The pressure is given by the trace part of the spatial components of the EMT and is related to the d_1 form factor by

$$d_1(t) = 15M \int d^3\boldsymbol{r}\, \frac{j_0(r\sqrt{-t})}{2t} p(r) \quad \text{and} \quad p(r) = \frac{4}{15} \int \frac{d^3\boldsymbol{\Delta}}{2E(2\pi)^3} e^{-i\boldsymbol{\Delta r}} P_0(\cos\theta)\left[t\, d_1(t)\right]$$

where j_0 is a spherical Bessel function, and P_0 is a Legendre polynomial.

Data Exchange Framework

It was necessary that we select a technology that will accommodate transferring large, heterogeneous data sets from producers (nucleon models and simulations) to consumers (analysis software and especially visualization.) The data may be streaming or archived. In addition to the data technology selection, we defined the banks (the schema) that the models will produce, and we provided the interface between the data format's I/O package and the producers and consumers of the data, including the visualization package, Paraview.

For the model and simulation data format, we selected the HIPO4 format in use at Jefferson Lab. First deployed for CLAS12 in Hall B but now being evaluated for other experimental halls, HIPO4 is a highly

compressed format that has been thoroughly tested for speed and reliability in large (petabyte) environments.

For the data schema, we are using JavaScript Object Notation (JSON) files that describe the banks that comprise the HIPO4 data. As a first test we defined banks for the CLAS12 torus and solenoid magnetic fields, the 24 thousand CLAS12 drift chamber wires, the scintillator panels for CLAS12 calorimeter, and track parameters from simulated DVCS events. The latter were the integrated through the magnetitic fields to produce trajectories. This permitted us to perform several preliminary "end-to-end" tests with the added benefit that the resulting visualizations of detectors, magnetic fields, and particle trajectories will be part of our project's public-facing website.

We developed an application, the Data Exporter Framework (*def*) with a plugin architecture. It parses the JSON bank descriptions, reads the HIPO4 files (or streams), provides visual data inspection and 2D plotting, and most importantly it exports, under user control, the data for use by the visualization. (The user selects which banks, which columns, and the column order for export.) Users can add additional exporters, without recompiling the application, by a plugin mechanism that dynamically discovers and loads new plugins. For this work, we developed only a standard CSV format exporter.

As a quick prototype data pathway framework, HIPO4 data was exported to CSV and then passed to Python reader scripts for the row / column format. Future projects should consider a C++ importer for HIPO4 and/or additional support for VRML/X3D from GEANT tools.

We developed a four-stage "person-in-the-loop" process. The first stage is where the physics data are produced in HIPO4 format, either by theoretical modeling or simulation. That data is passed to *def*. There it can be inspected and plotted in various ways to validate the data. It is then exported, which for now is limited to CSV format. The third stage utilizes Python scripts that perform final preparation. Finally, the data is presented to the visualizing software. In future work, we should automate this process so, for example, the user can launch a simulation and observe the resulting visualizations seamlessly.

A simple schematic of the end-to-end process is shown and a *def* screenshot is shown below.

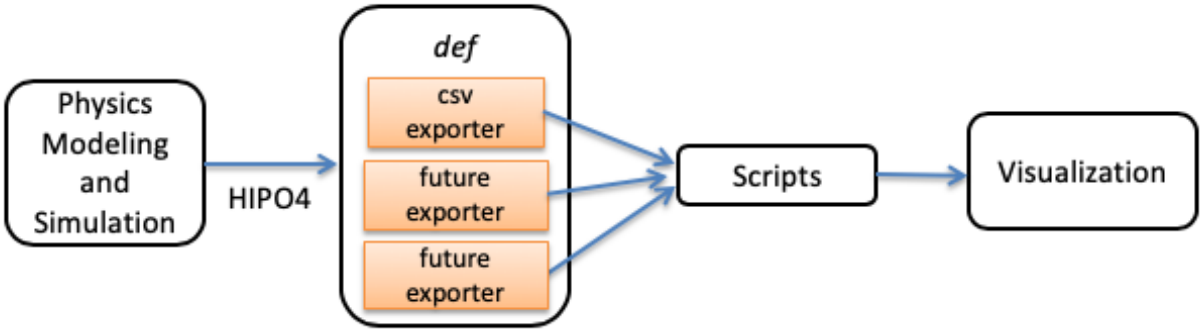


Figure 2. In a four-stage process, physics data is produced in hipo4 format, exported by def, receives final preparation through scripts, and is sent to the visualization tools.

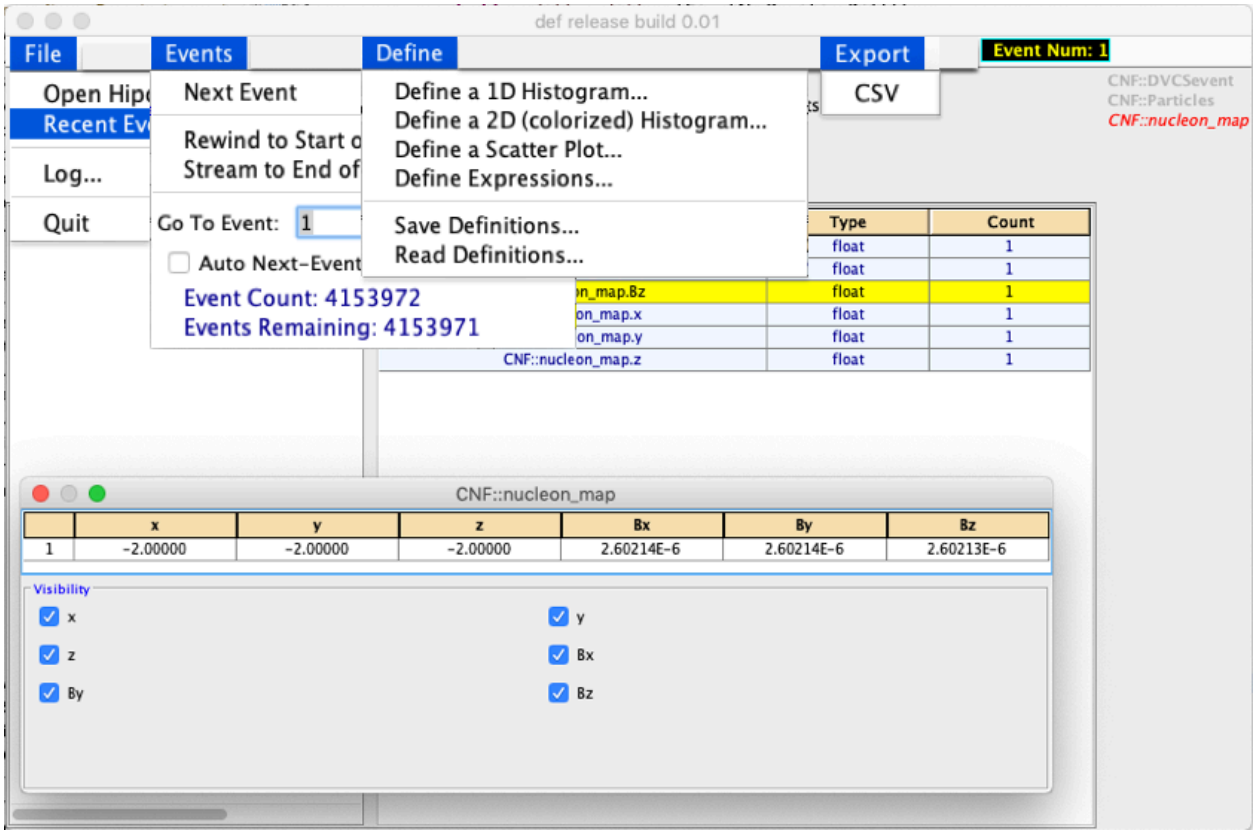


Figure 3. A screenshot of the def application, developed for this project to read physics data in hipo4 format and export it to a form more suitable for visualization.

Femto-Scale Dynamics Workflow

The entire process from the reference model through the imaging and the visualization are shown below in a Workflow diagram to illustrate how the individual parts and steps are connected.

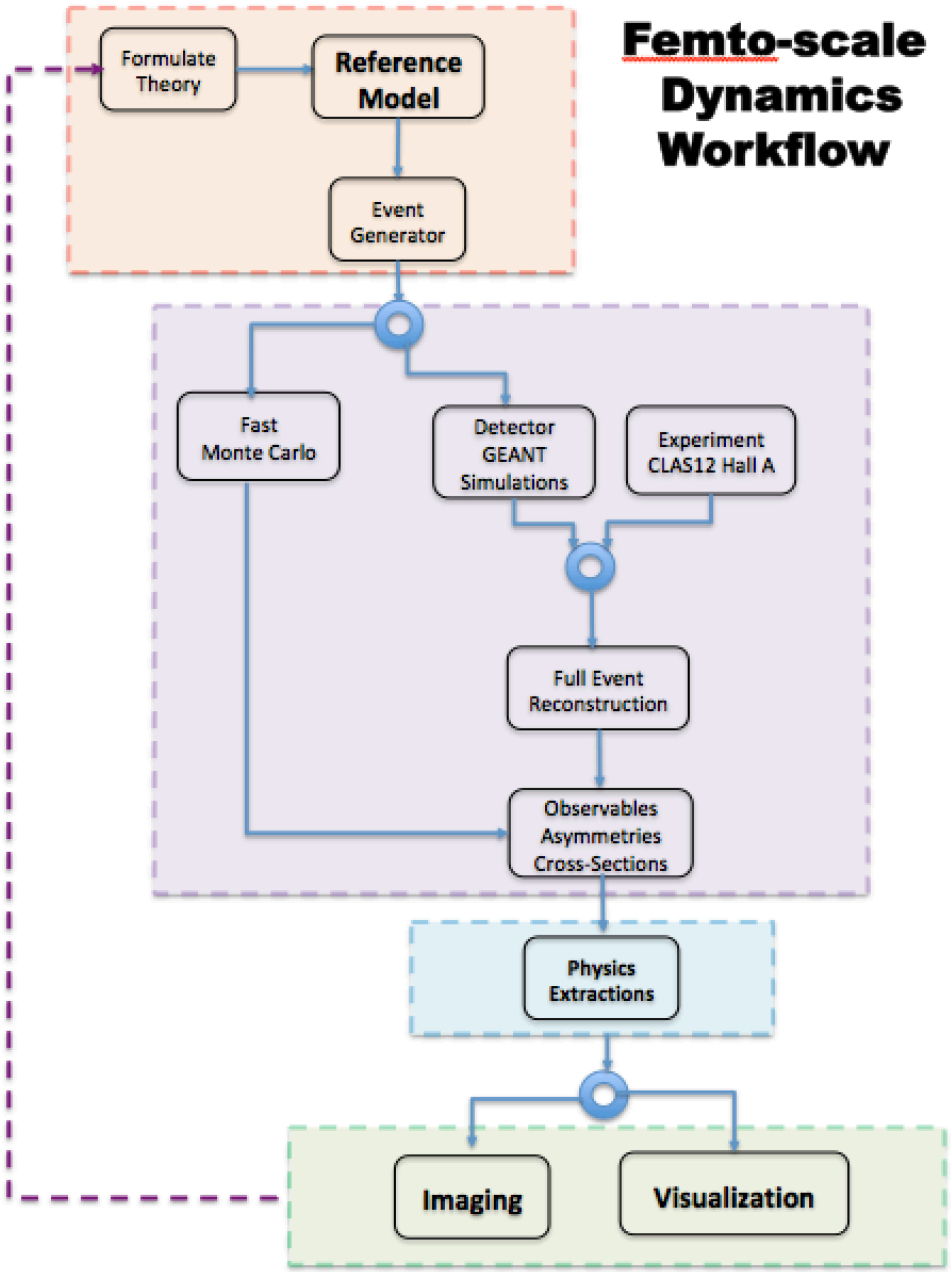


Figure 4. The workflow through which we create and direct Reference Model physics data to the visualization tools through several paths.

Visualization Technology

We targeted open source and open standards tools to demonstrate our methods for visualizing femtoscale physics information. We leveraged the open source DOE toolkits and their C++ and Python APIs to develop mappings for interactive 3D visualization tools including Paraview, X3D [7], and VR. The Python scripts and pipeline code enable scientists to process, render, and export their data with several visualization metaphors and mappings. The results of this objective will be threefold: the interactive 3D visualization software, a set of movies highlighting the data and techniques, and a VR experience illustrating the structure and interaction of particles and forces in the nucleon. In general, Imaging and Visualization in Figure 4 can be achieved with the same tools and pipelines.

We applied these three technology pathways to three different types of CNF data:

1. the CLAS 12 instrument itself, including the magnets, solenoid, detector wires, and calorimeter
2. Event trajectories, including single particles and simulated DVCS events of Proton, Photon, and Electron
3. The mechanical forces in the hadron, including radial and tangential momentum components

Software: The Paraview visualization tool can be installed on all leading desktop and HPC platforms. Paraview enables users to compose processing and visualization pipelines and to create multi-dimensional interactive visualizations, including 3D and common 2D chart types. Custom data processing programs can be written in Python and Pipelines can be automated to run without the GUI. Paraview State files capture the entire pipeline including the data source though an absolute pathname.

Movies: Paraview supports custom animations of cameras and visualization elements in the scene and the generation of avi files at any resolution and framerate; screenshot images will also be produced. These media should be presented with some scientific explanations on the project website

Interactive 3D: is an ecosystem of tools and rendering devices. Paraview supports native VR visualization using Head Mounted Displays (HMDs) and the VRPN device bus. Paraview also allows users to export their 3D scenes to the ISO-IEC Extensible 3D (X3D) format, which can then be used in a number of tools from the WWW to CAVEs to 3D printers. For accessibility, we will publish our X3D models on the WWW so they can be viewed natively in any Web browser (web3d.org).

Results

Geant 4 Simulations

Using the reference model as input we implemented a generator based on the von Neumann acceptance-rejection method. Following the CLAS12 First Experiment parameters, we select a beam of energy of 10.6

GeV, and simulated ten million events. We divide the reconstructed kinematical coverage in the Q^2 vs x_B plane in ten irregular bins, as well as seven in the -t direction, and twenty-four in the Φ direction.

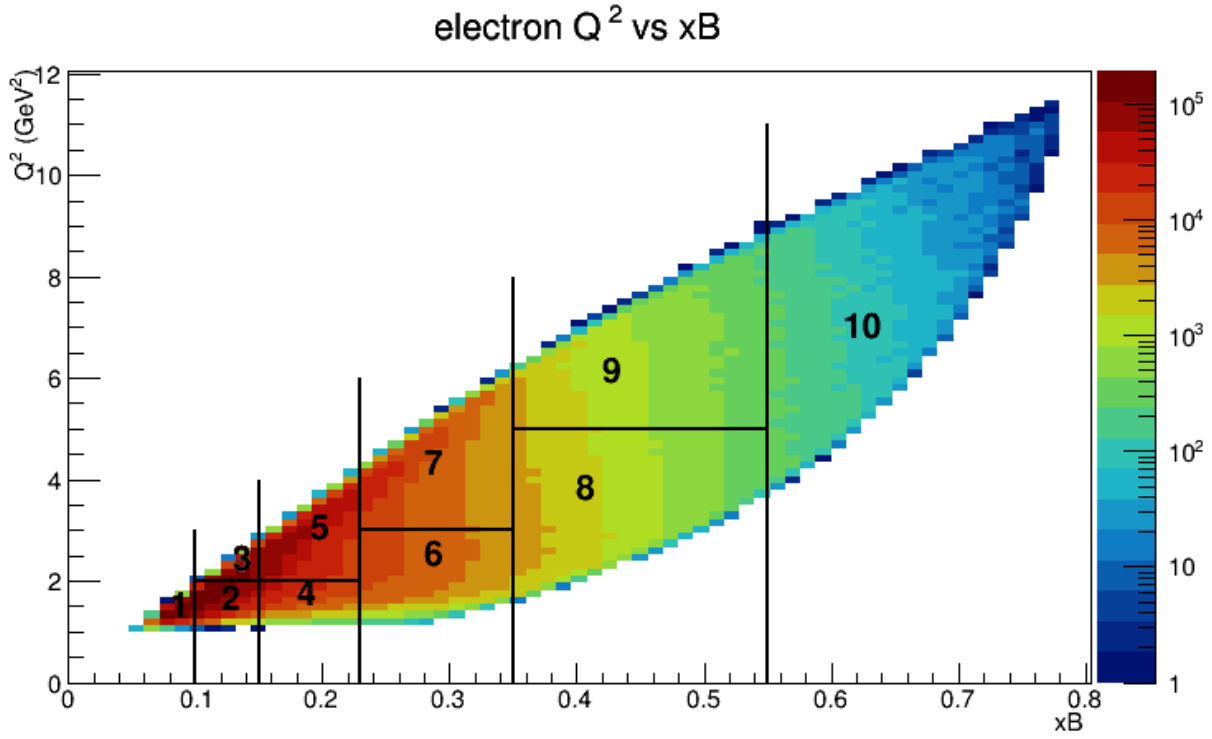


Figure 5. Geant 4 kinematic coverage and binning distribution of events in the Q^2 vs x_B plane.

Particles are identified as in the data sample. Electrons and protons have charged tracks either in the forward Drift-Chambers, or in the Central Vertex Tracker. Electrons are identified by photo-electrons in the High-Threshold Čerenkov Counter and a shower in the Electromagnetic Calorimeter, in coincidence with a negative track. Electrons are also considered trigger for the event and define the start time. Protons are identified by Time-of-Flight in Scintillator counters in coincidence with a positive track. Finally, photons are identified as isolated neutral clusters.

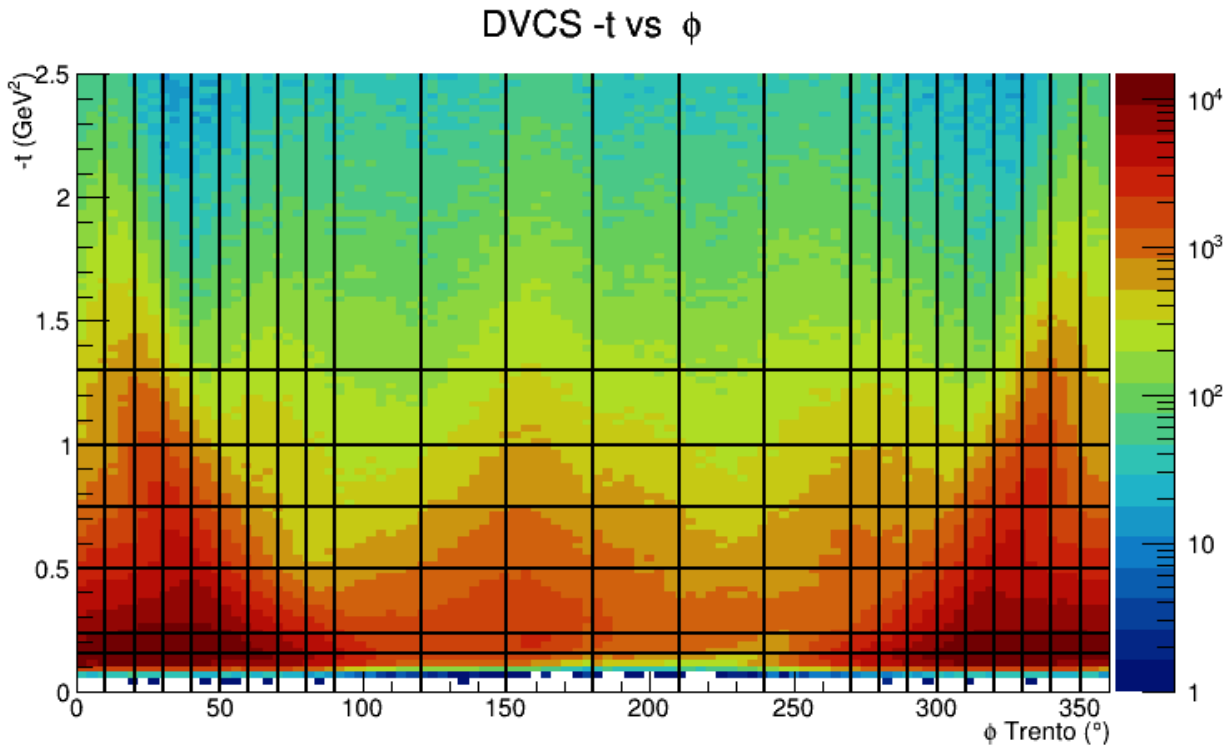


Figure 6. Geant 4 kinematic coverage and binning distribution of events in the -t vs ϕ plane.

Events are selected with all particles identified. We then form exclusivity variables whose primary purpose in the experimental data is to ensure the reaction selection. In the simulation, we do not have any other reaction to separate the sample from, however the exclusivity cuts still ensure the proper reconstruction of the kinematics, limiting for instance the radiative effects. The complete list of exclusivity cuts is as follows

- Missing energy E_X in ep→epγX : E_X < 1.25 GeV
- Missing mass squared M_X^2 : -0.03 < M_X^2 < 0.03 GeV^2
- Missing momentum transverse to the beam z-direction in ep→epγX : $\sqrt{P_x^2 + P_y^2}$ < 0.2 GeV
- Proton missing mass M_X in ep→eγX: 0.4 < M_X < 1.7 GeV
- Angle between the photon direction and the missing system in ep→epX: $\theta_{\gamma X}$ < 1.5°

We also define the Φ angle between the leptonic and hadronic scattering plane in the so-called “Trento convention” so that the angle is between 0 and 180° if the frame formed by the beam momentum, the scattered electron momentum, and the photon momentum is a right handed frame.

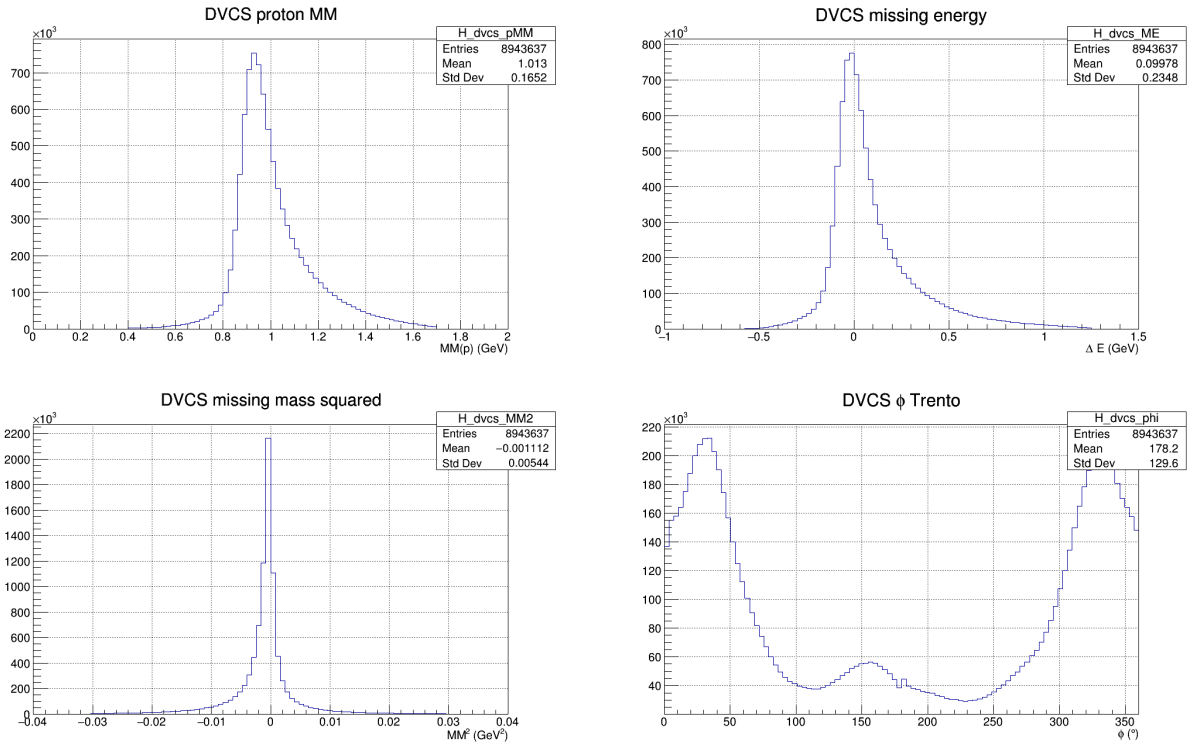


Figure 7. Geant 4 reconstruction. Top: proton missing mass and missing energy. Bottom left: missing mass squared. Bottom right: integrated ϕ distribution, convoluting acceptance and generated distributions.

A pseudo-random helicity is generated on an event by event basis, based on the Reference Model used in the generator. We then store separately left and right events reconstructed in elementary bins, and form a beam spin asymmetry at the end of the loop over all events as function of the Φ angle. The Φ modulation is fitted with the theoretical shape

$$\mathrm{BSA} = \frac{\sigma^+ - \sigma^-}{\sigma^+ + \sigma^-} = \frac{\alpha \sin\phi}{1 + \beta \cos\phi}$$

where the parameter α and β are respectively sensitive to the imaginary and real parts of the amplitude. In the Geant simulation we stopped at the extraction of the beam spin asymmetries amplitudes from the fit. The fitted result for α are shown as functions of -t for all ten bins in x_B and Q^2, and compared with the input Reference Model as red curves.

We obtain a good agreement between the input Reference Model and the extracted amplitudes. Some small residual systematic differences can be explained by finite bin corrections. The kinematics in x_B , Q^2 and -t depends on the angle Φ, and we compared the Model amplitudes at average kinematics integrated over Φ. In the data analysis, this effect is corrected for before performing the amplitude fit as functions of Φ but this requires more statistics than available from the Geant simulations.

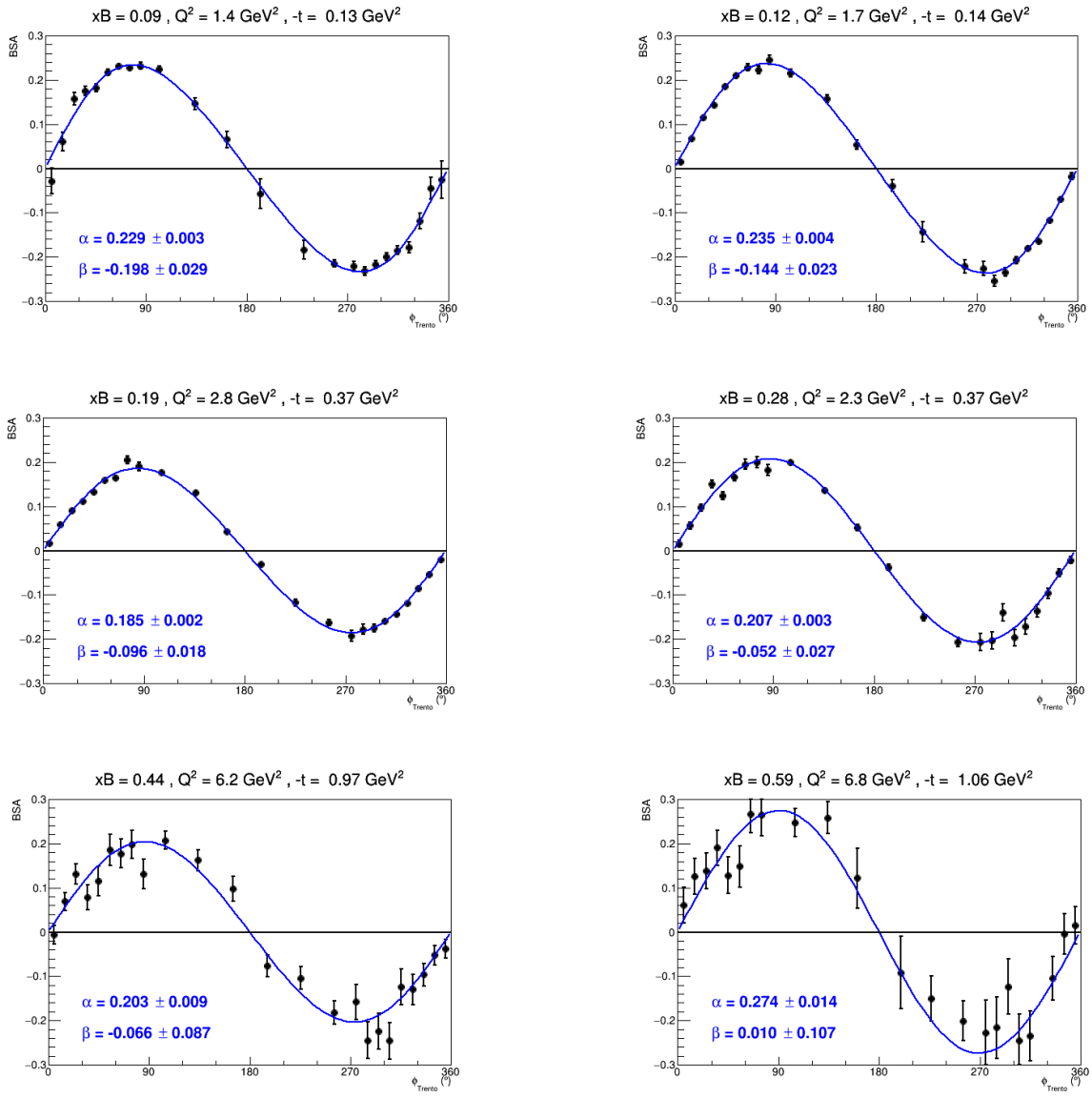


Figure 8. Geant 4 simulated beam spin asymmetries. Resulting bins at low (top), intermediate (middle) and high (bottom) x_B, Q^2, and -t.

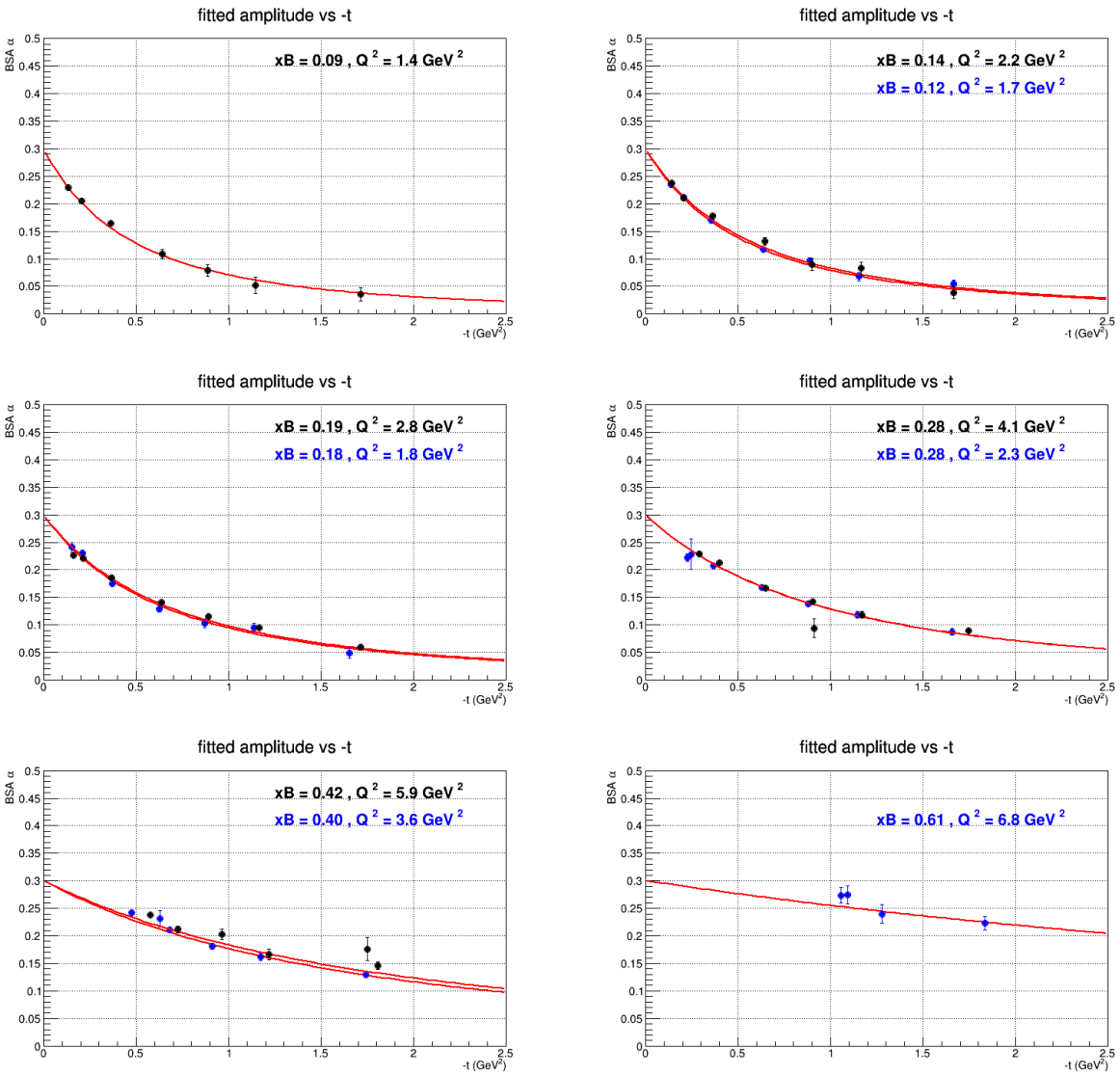


Figure 9. Geant 4 simulated extraction of amplitudes as a function of -t. The input model at average kinematics is shown in red.

Fast Monte-Carlo Simulations

The Fast Monte-Carlo simulations are based on the same Reference Model as generator. The acceptances and resolutions are directly parametrized in terms of kinematical variables at the vertex. This bypasses the Geant simulation of the physical processes in the detectors, as well as the event reconstruction from the detector signals, and considerably speeds up the computation time. We are therefore not limited in statistics in the Fast Monte-Carlo simulations, and can validate the consistency of the physics extraction strategy.

We generated events again at 10.6 GeV corresponding to hundred million events reconstructed in the spectrometer. We divided the phase space in a finer binning, of twenty five irregular cells in Q^2 vs x_B plane, eight bins in the -t direction, and thirty six bins in the Φ angle direction. The main differences in binning with the Geant simulations are the much finer grid in the Q^2 vs x_B and the regular binning in the Φ angle direction including in the central region. The central region is particularly important in the unpolarized cross-section to provide sensitivity to the real part of the amplitudes.

We illustrate the procedure with Φ angle distributions in two representative bins in the x_B , Q^2 , and -t. The distributions shown are the number of events in Φ bins, the extracted cross-sections and beam spin asymmetries. The reference model is shown with red curves, and the fit to the observables as blue curves. Additional green curves on the cross-section plots show the pure Bethe-Heitler contributions. The fit is performed simultaneously on the cross-sections and beam spin asymmetries, with the parameters being the imaginary part of the amplitude, and the subtraction constant in the real part of the amplitude computed from the dispersion relation.

The extracted real and imaginary parts of the Compton Form Factor are shown as functions of -t in six representative bins, with colored band from the propagated uncertainties from the fit. The uncertainties on the real parts are larger than those of the imaginary parts, since the real parts are calculated from the imaginary parts using the dispersion relation. Finally, the D-term extracted from the full dataset is shown in a standalone plot as function of -t. The individual points correspond to the local fits in bins. The blue band corresponds to the global fit result. The nucleon pressure distribution is obtained by a Fourier transform of the D-term. We show the pressure weighted by r^2 because the stability condition implies that this quantity should vanish when integrated, and therefore the positive and negative regions are better visualized together with this weight. We note that the pressure goes a the finite value at zero radius, given by 0.23 GeV per cubic fm which is the central pressure of the nucleon in the Reference Model.

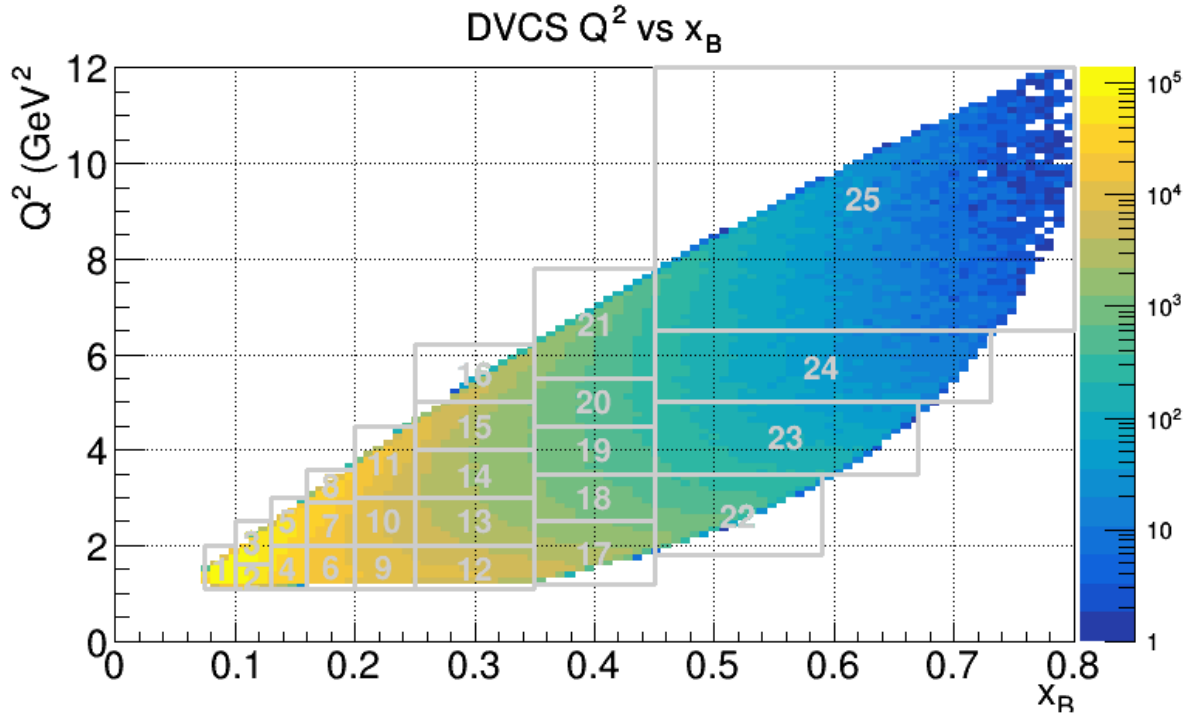


Figure 10. Fast Monte-Carlo kinematic coverage and binning distribution of events in the Q2 vs xB plane.

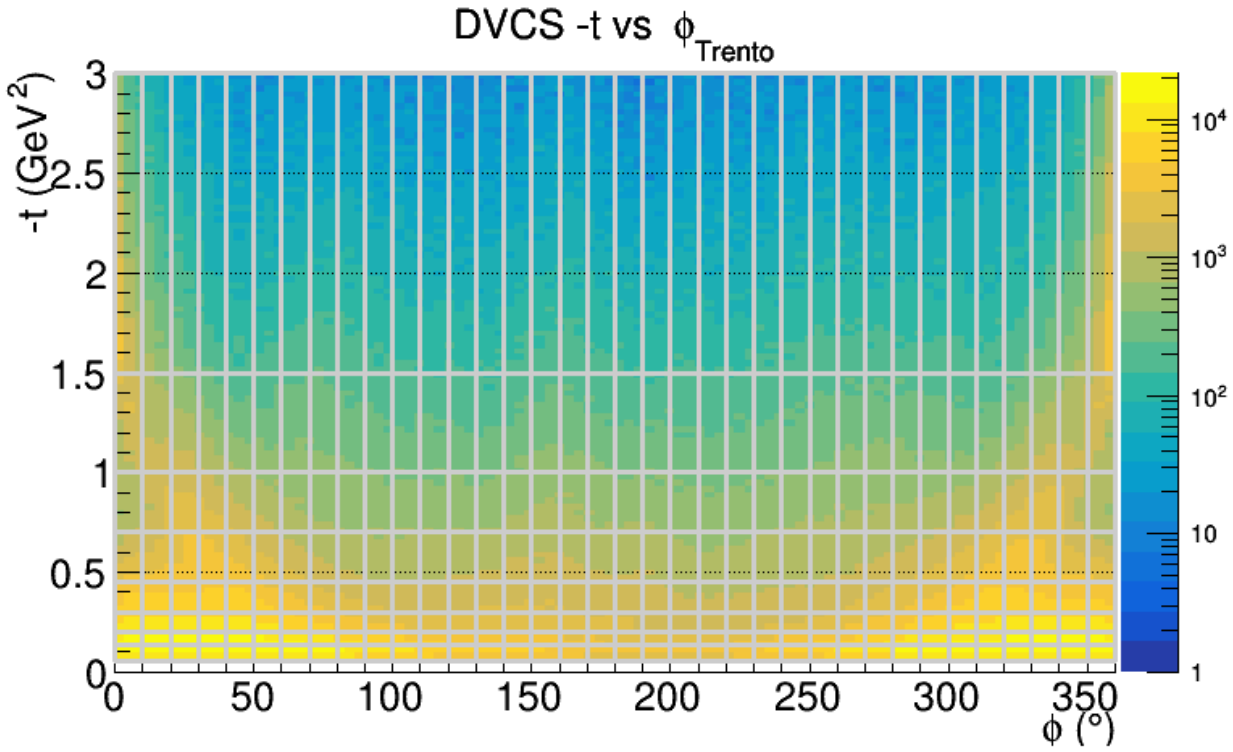


Figure 11. Fast Monte-Carlo kinematic coverage and binning distribution of events in the -t vs plane.

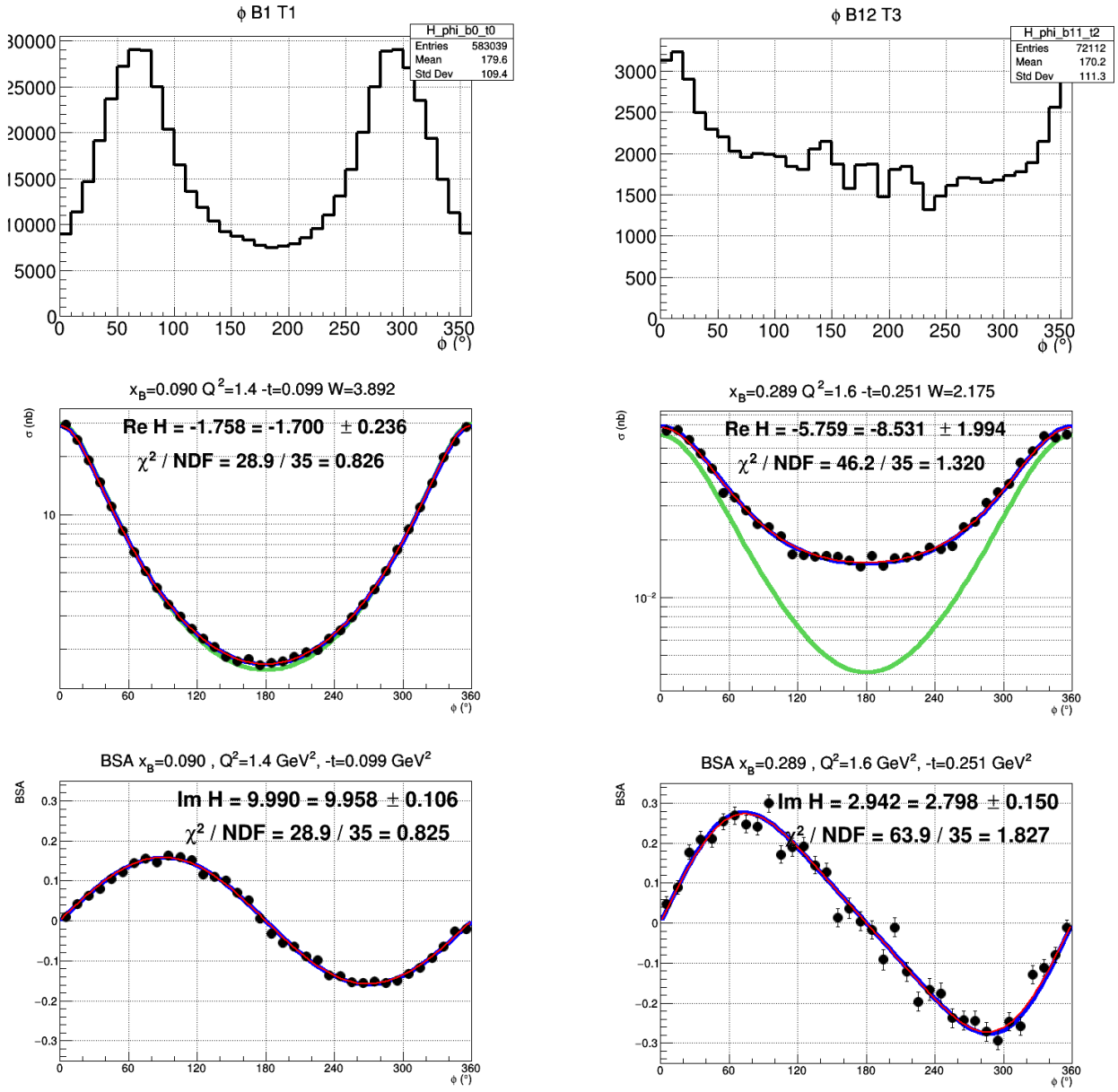


Figure 12. Fast Monte-Carlo sample of observables int two representative bins.

Left: x_B = 0.09, Q^2 = 1.4 GeV², -t = 0.10 GeV²

Right: x_B = 0.29, Q^2 = 1.6 GeV², -t = 0.25 GeV²

Top: number of events in ϕ histogram

Middle: corresponding cross-sections; the green curve shows the pure Bethe-Heitler contribution to the cross-sections

Bottom: corresponding beam spin asymmetries

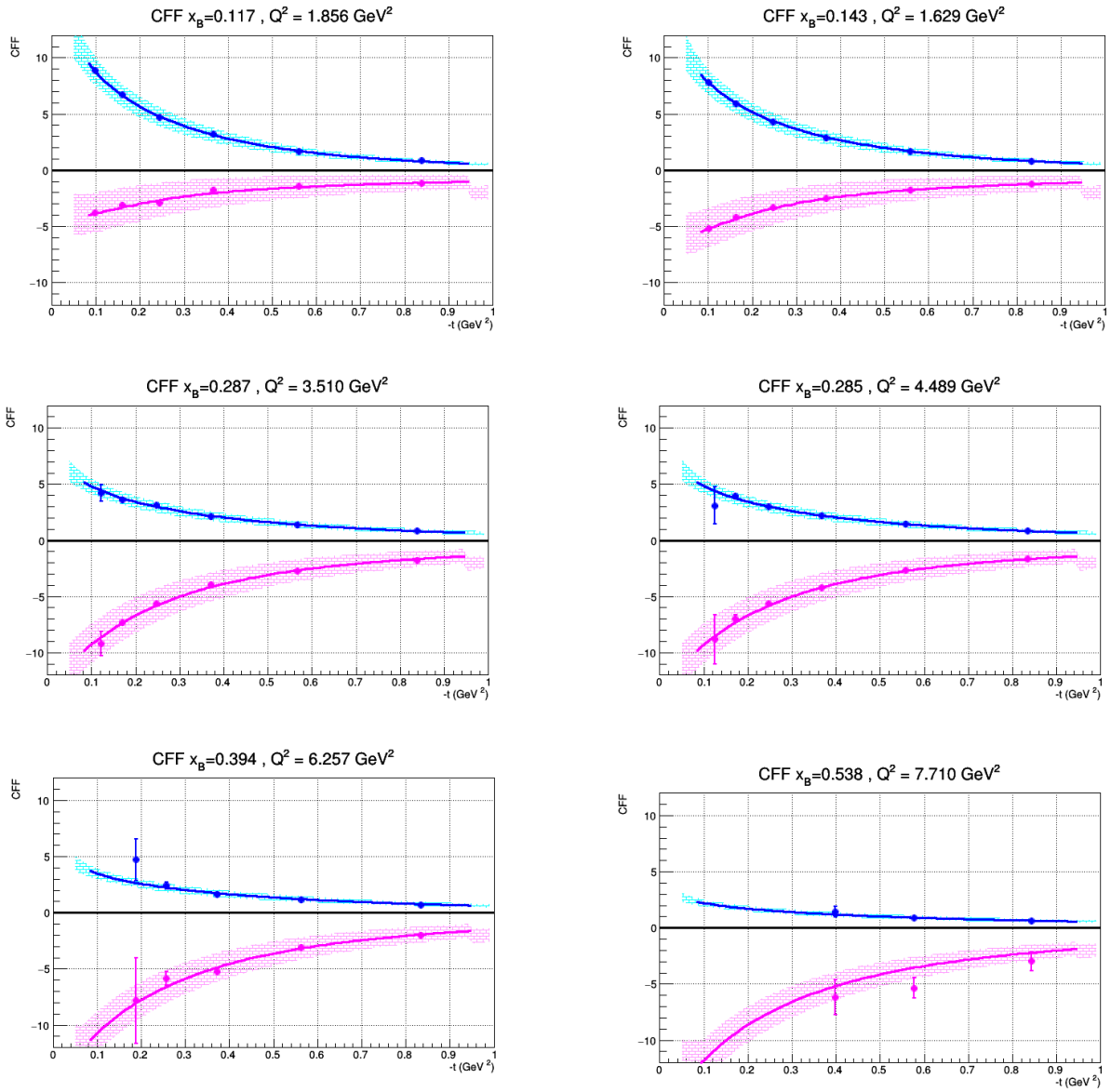


Figure 13. Fast Monte-Carlo simulations extracted real and imaginary parts of the amplitude in six representative bins as a function of -t. The color bands correspond to the propagated uncertainties from the fit.

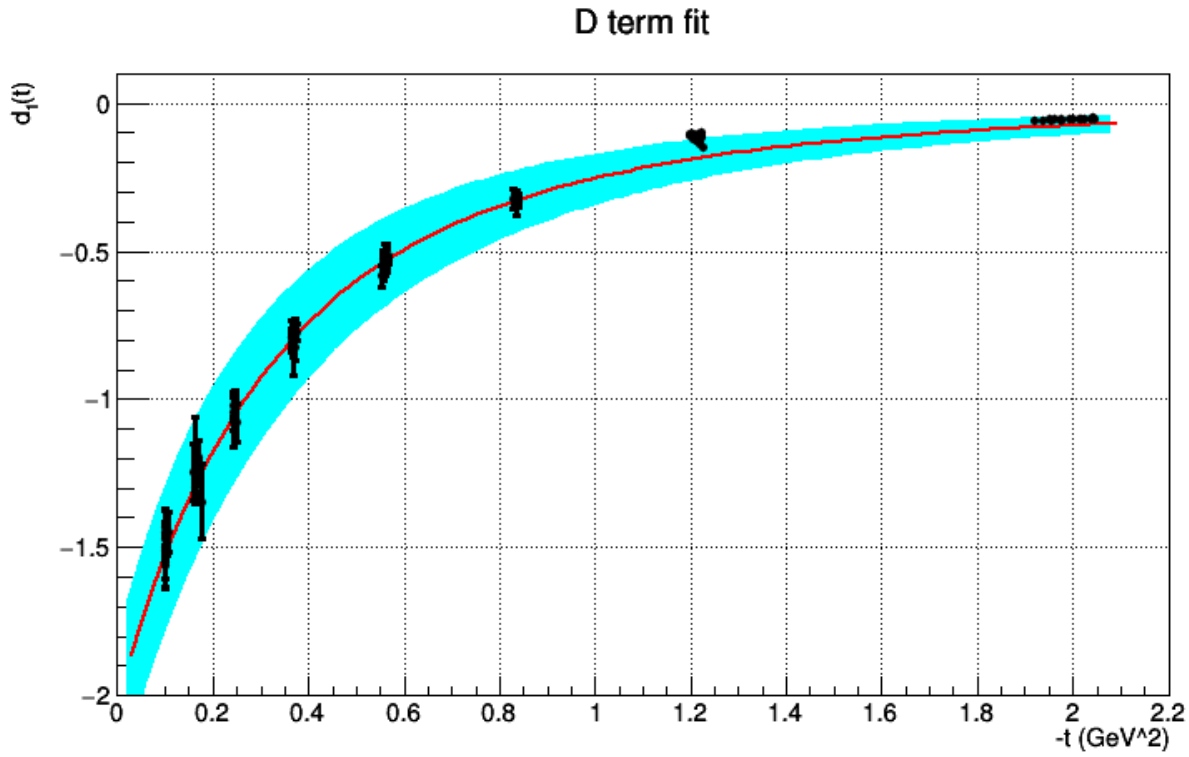


Figure 14. Fast Monte-Carlo extraction of the D-term from the dispersion analysis of the amplitudes. The blue band corresponds to the uncertainties from the global analysis.

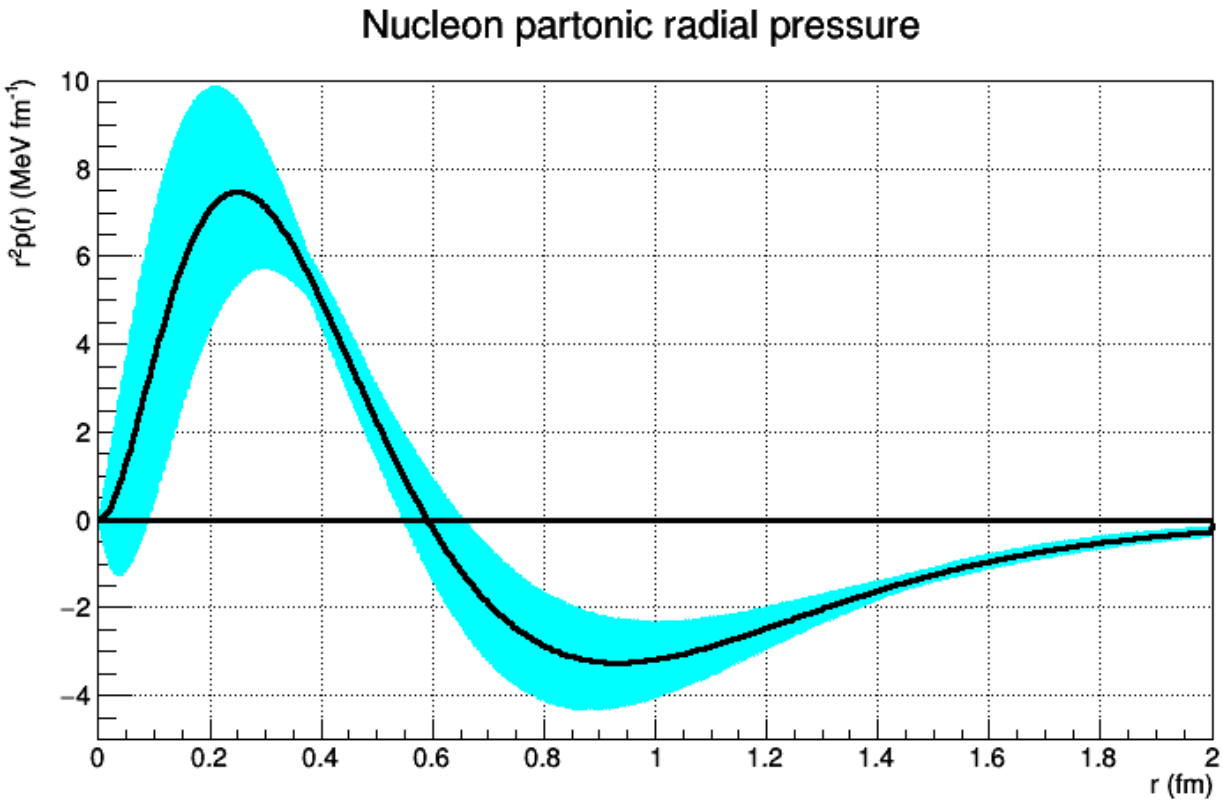


Figure 15. Corresponding radial distribution of the partonic pressure with propagated uncertainties, weighted by r^2. The stability condition imposes that this integral vanishes.

EIC

A natural extension of this work is to the future US based Electron Ion Collider. Different scenarios are currently considered for this machine, we show as an illustration the inclusive electron kinematical coverage in the (x_B,Q^2) plane for two different combinations of proton and electron energies. We also indicate on these plots the kinematical limits, the lower limit being taken at y=0.01, and the higher one to y=0.95, with an additional indicator at y=0.6. With these kinematical limits, x_B and Q^2 are strongly correlated, with the maximum Q^2 given by the center of mass energy $Q^2_{max} = s\, x_B$

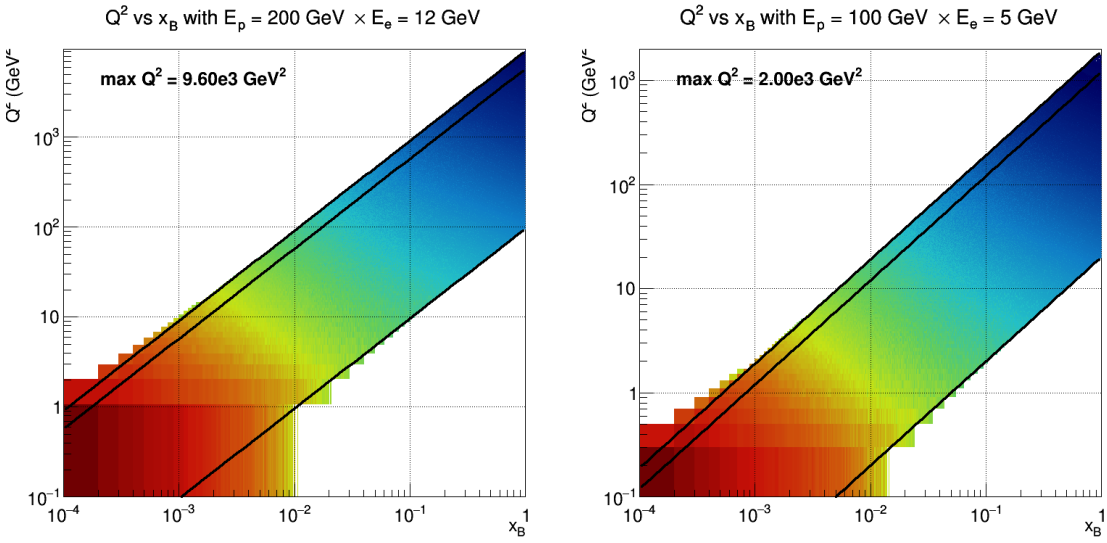


Figure 16. The inclusive electron kinematical coverage for two combinations of proton and electron energies.

We generate DVCS events in the EIC kinematics, with the high energy scenario, using our reference model. The convention for the polar angle has the electron going in the negative direction, so 180° corresponds to the electron not being deflected. The photon covers a phase space similar to the electron, with the Bethe-Heitler contribution dominating at low energies.

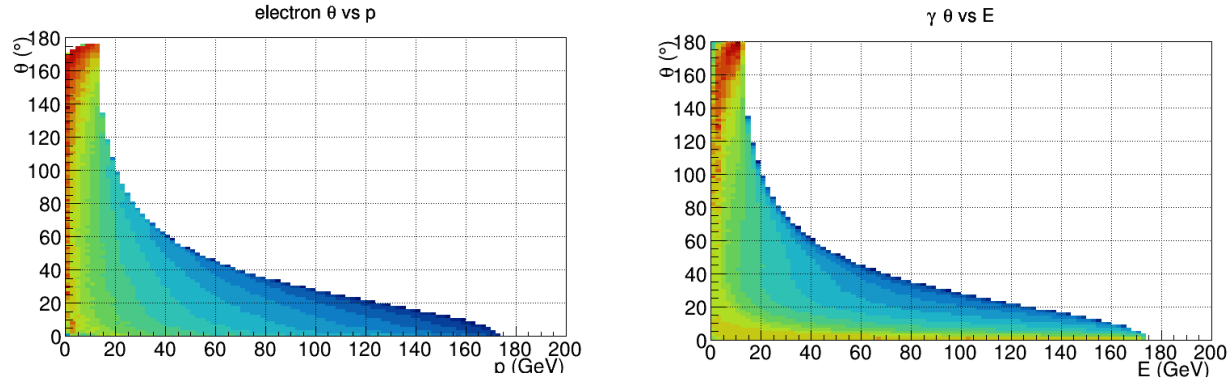


Figure 17.DVCS events in the EIC kinematics in the high energy scenario, generated using the reference model.

The proton generally goes mostly in forward at low polar angles, and high momentum. The momentum transfer t Mandelstam is restricted to the range 0 to 5 GeV^2 in this sample, and the angle between leptonic and hadronic planes ϕ Trento peaks at the endpoints when the photon is collinear with the electron.

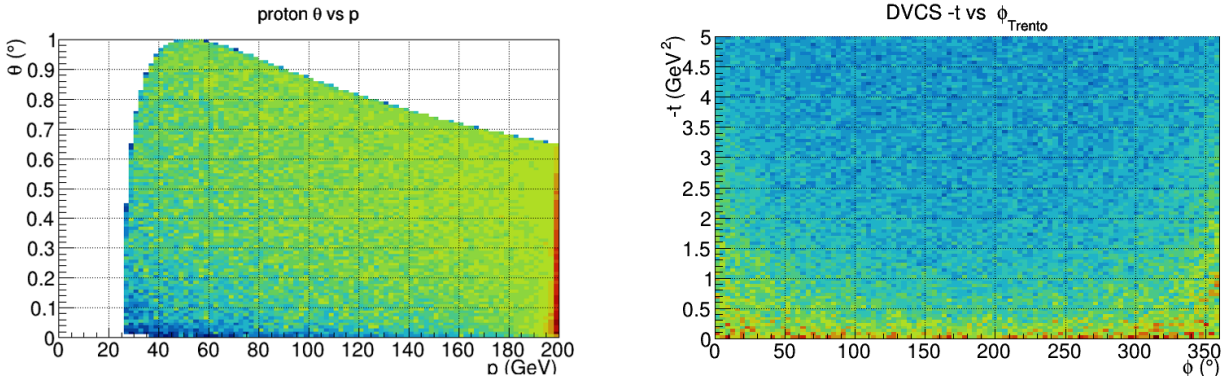


Figure 18.Left: proton polar angle vs. momentum. Right: momentum transfer t vs. the angle between the leptonic and hadronic planes.

Finally we use ad-hoc resolutions for single particles to illustrate the exclusivity variables, similarly to what was already shown for the CLAS12 analysis strategy. The knowledge of the kinematical resolutions will allow us to perform feasibility studies of this measurement with reasonable binning in t and ϕ, as well as study background separation methods.

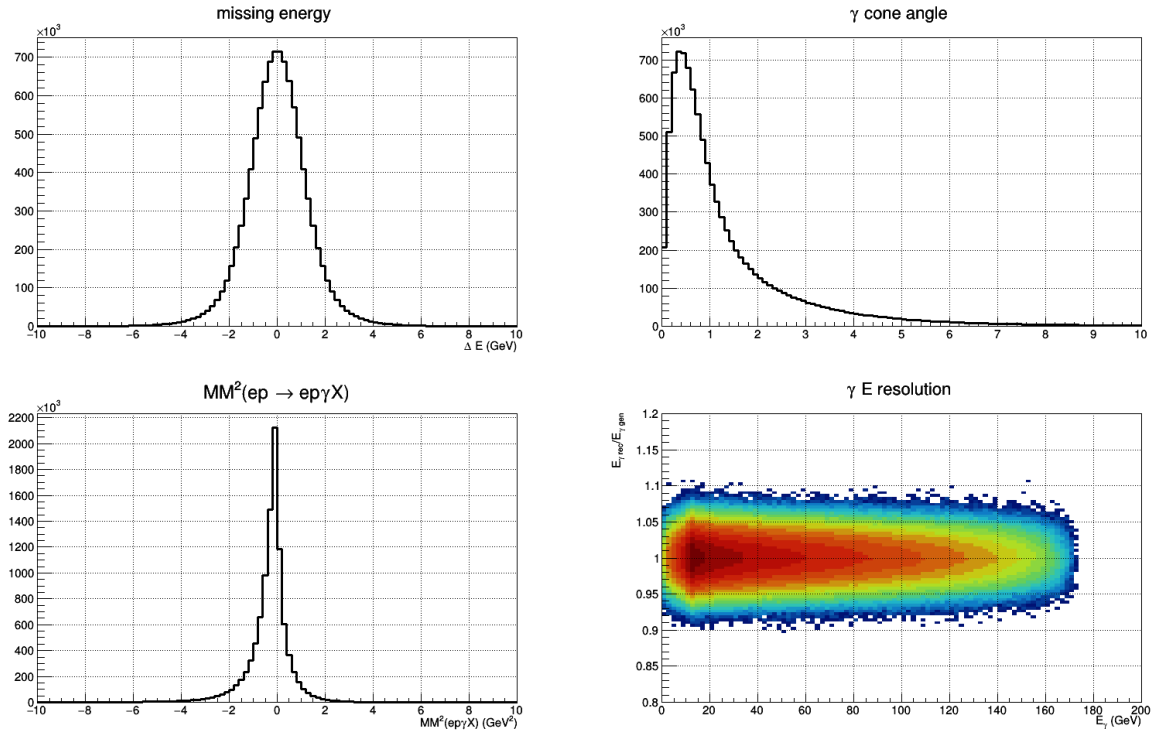


Figure 19. Exclusivity variables, similar to the CLAS12 analysis.

Visualizations

Software

Visualization can be beneficial for quickly debugging a model or algorithm. It is also the currency of communication in the information age; yet one can mis-lead with both statistics and visualization. Essential to this effort was a detailed review of the data and algorithm mappings, and vetting their assumptions and defaults with the physics experts of the team. We summarize these below:

- The graphics standards' units are assumed to be **meters** (unless otherwise declared in the X3D UNIT statement). This particle physics data is generated or measured at a 3D location; each 3D spatial location also has an associated momentum, with three (x, y, z) components. Thus, we parse the data files knowing that we have a 3D momentum vector at each point: a direction and a magnitude.
- Modern graphics programs (with the exception of geospatial coordinates, which are double precision and work in a variety of geode projections), work in **Euclidean** coordinates, with a right-handed coordinate system where +Y is 'up' and +Z is the depth toward the user's default perspective. Some data in this domain are in polar or cylindrical coordinates, and must be converted.
- In Paraview, the tabular data must be loaded and mapped to 3D points (the spatial position by the first three columns). In addition, the three separate momentum components should be combined into a single 3D vector and the magnitude of that vector is calculated. **Glyph rendering**: the momentum value can be mapped to a glyph: specifically, an arrow with a length and color scaled to its magnitude. The number and the sampling approach of glyphs are key parameters that yield different results. Specifically, the space can be uniformly sampled with a target number of glyphs, or data records can be skipped in a stride.
- In parallel in the visualization pipeline, the vector field can also be cast as a volume where a scalar magnitude value exists at every cell ; this is an interpolation process that requires a few parameters including the kernel type, its footprint and number of points considered. While the default kernel is Voronoi, we found more consistent (symmetric) results with the Linear and the Gaussian kernel with the former providing more of the expected symmetry. Paraview enables the user the specify the resolution of cells in each x y and z dimension for **Volume rendering**
- **Contours** are lines or surfaces drawn at an iso value; the number and threshold values of contours are design choices
- **Streamlines** go orthogonal to contours (default is Runge-Kutta 4); users can set the number of lines, their source (point or line), and their maximum length; streamlines can be drawn as tubes with a thickness mapped from some scalar
- **Color**s are essential for categorical information and are also valuable for portraying quantitative data. We employed several different color maps depending on the nature of the data
- **Cut-away**s provide a way to clip out portions of the visualization, which can help reveal internal structures

We created and collected the raw data, related Python code, and the Paraview State files for the CLAS 12 instrument, a set of trajectory events, and the radial and tangential nucleon forces.

As the common denominator of the knowledge economy is visual communication; we wanted to improve and expand the accessibility of these visualizations and the experience of particle physics. Thus, we adopted the ISO-IEC Standard of Extensible X3D (X3D) as a publication format for interactive 3D visualization products, which can be loaded in the 26.7 million stereo pixel CAVE at Virginia Tech's Visionarium Lab, or with any Web browser or any WebVR device (see videos). We focused on compliant X3D softwares, of which there are many. For workstation installations and in the VT Visionarium, we used InstantReality from Fraunhofer IGD for our immersive X3D runtime; for the Web and HTML5 integration, we used the open source JavaScript framework of X3DOM.org (we have also successfully tested these X3D environments with the open source JavaScript framework of X_ite).

Movies and Interactive 3D

In Paraview, all of these results can be explored with a Head-Mounted Display through the OpenVR plugin, which is distributed with the Paraview installer. Also using Paraview, we exported movies, screenshots, and X3D files. This media is available on the project webpage and our Youtube channel. A selection of results is shown below and on the project website at : vis.arc.vt.edu/projects/xxxx .

Magnet & Sensors

We processed four different 3D parts of the CLAS-12 instrument: the magnet, the solenoid, the detector wires, and the calorimeters. The detector wire geometry (14,000 rows) was converted to ISO-IEC VRML, which can be loaded into Paraview and was drawn in black. The Calorimeter geometry was also exported and converted to VRML to be loaded into Paraview and drawn in white. For the magnet and solenoid, we contoured the momentum magnitude field and colored it by magnitude to see the shapes of the magnets' strength (white to blue), which we rendered as wireframes. We then applied streamlines to the magnet data and so traced out lines orthogonal to the momentum field; in the later versions of these visualizations, we created tubes –cylinders- on the streamlines since they provide additional depth cues (more than lines) to aid pre-attentive human judgement (occlusion and shading).

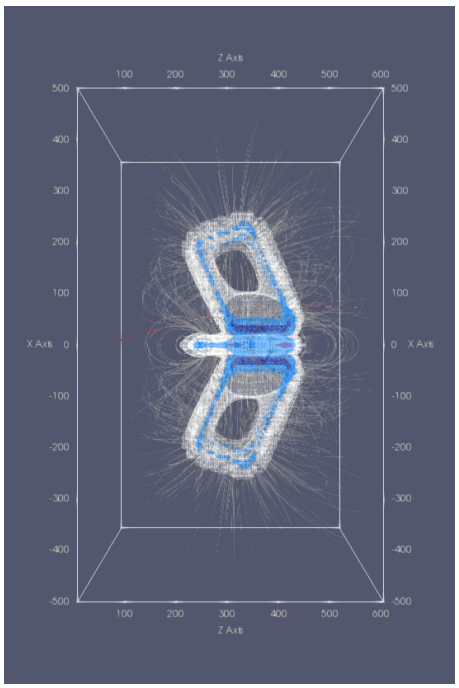


Figure 20.The momemtum field around the CLAS 12 magnet

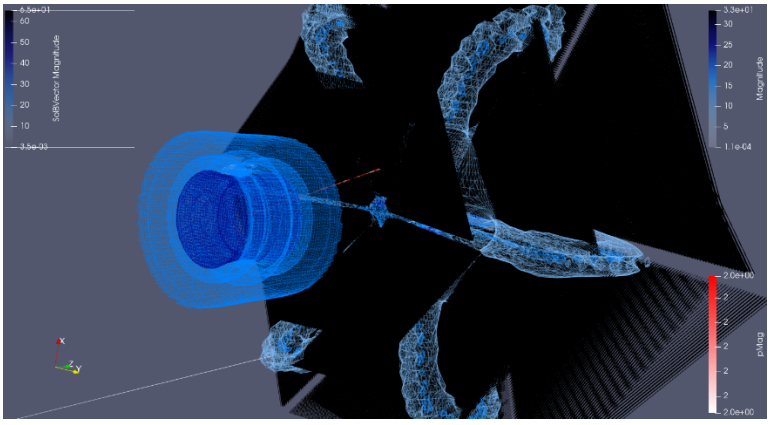


Figure 21. The Solenoid geometry with the detector wires and magnet

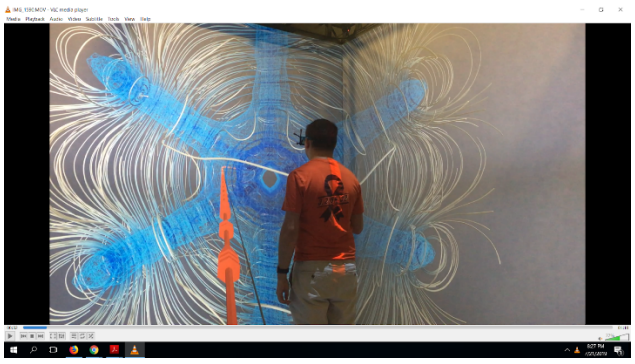


Figure 22. CLAS-12 geometry and trajectories running on the VT HyperCube projection display

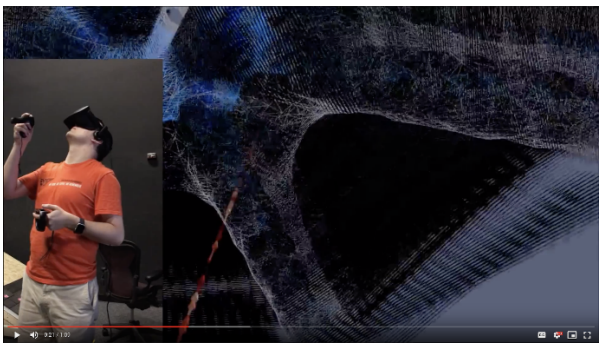


Figure 23. CLAS-12 geometry and trajectories running on an Oculus Rift VR setup

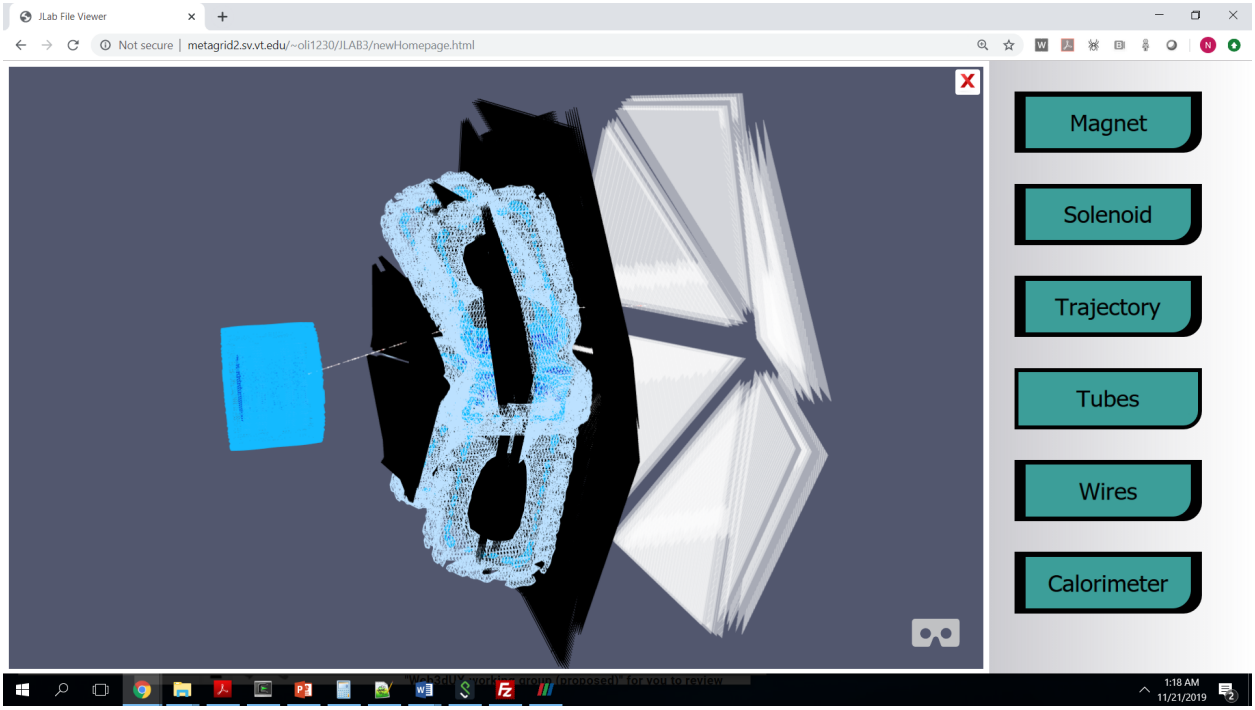


Figure 24. CLAS-12 geometry including detector wires and calorimeter

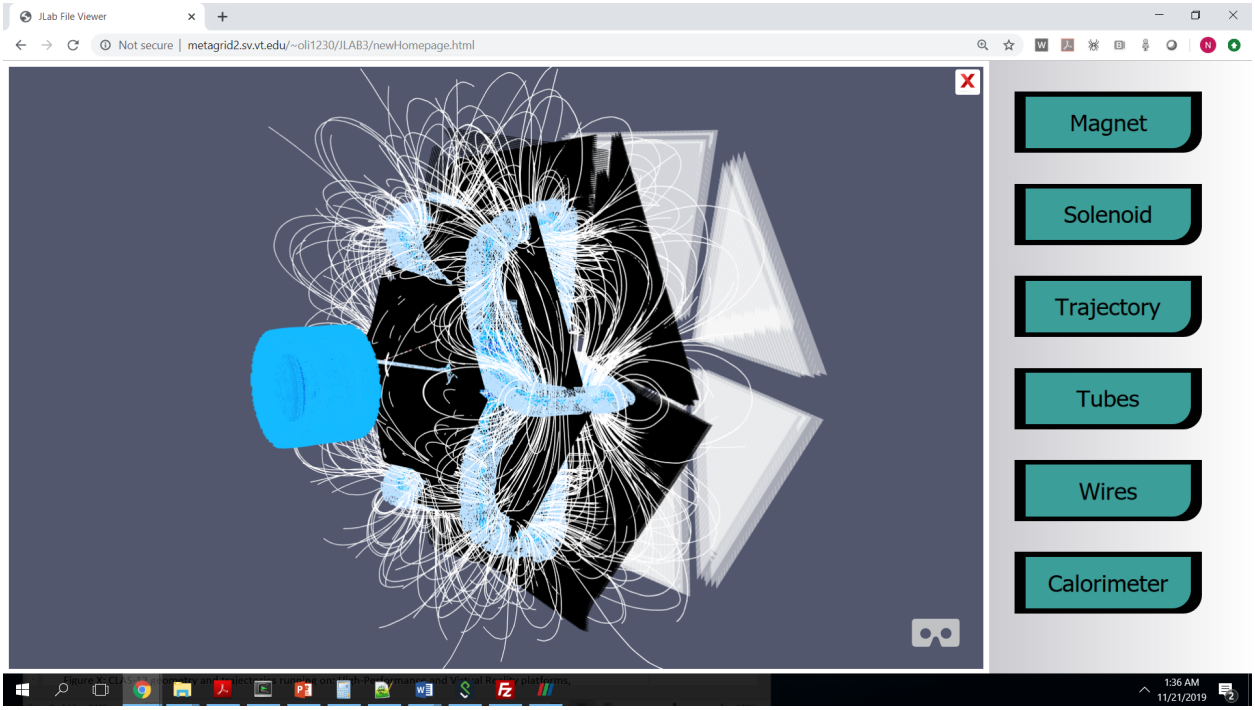


Figure 25.CLAS-12 geometry and trajectories running on: High-Performance and Virtual Reality platforms, the interactive Web/WebVR interface

Trajectories

The initial reference data trajectory was a simple, incomplete track of one particle through the magnet and detectors. We used a Glyph arrow, scaling size and red saturation by the vector magnitude. The second set of trajectories we examined were complete events of photon, proton, and election interactions. We created a Python script to automate Paraview to analyze a set of trajectories and output X3D models. These models were added to the website with a toggle interaction. We used the convention of the Particle Physics field, which is the photon is white, the proton is red and the election is a blue/turquoise.

Nucleon Forces

This project also examined visualization approaches for the distribution of forces in the nucleon. We applied Glyphs to both Radial and Tangential forces, scaling their attributes by magnitude. Again, we wrestled with Glyph crowding and testing/portraying the symmetry with contours. After our SURA presentation and feedback from the the CNF community, we resolved to create simpler visualizations that would show fewer things at a time and not require training to read and interpret. Thus, our latest portrayals focus on either the Radial or the Tangential forces and show how they work one-at-a-time. Because segmenting data is hard, this was also helpful to process the opposite directions as separate files loaded into the visualization. This made it easier to show a categorical difference between Radial vs Tangential forces in the nucleon and also the directionality of the Radial forces: Radial in vs Radial out. However, we needed to make sure that color scale were applied consistently and shared among the different loaded data sources in Paraview.

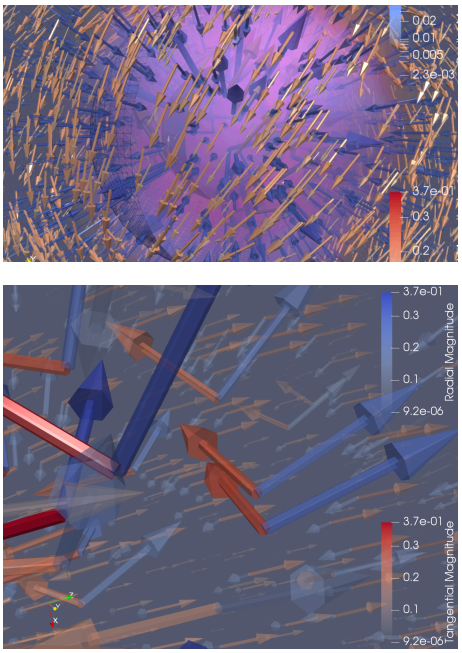


Figure 22: An early, zoomed-in view of a spatially-uniform sampling for Glyph seeding; note the orthogonal Radial and Tangential components at each sampled point, each colored with its (divergent) color scale and fully transparent at zero.

The figure below combines the techniques to illustrate how Radial forces are distributed in the nucleon.

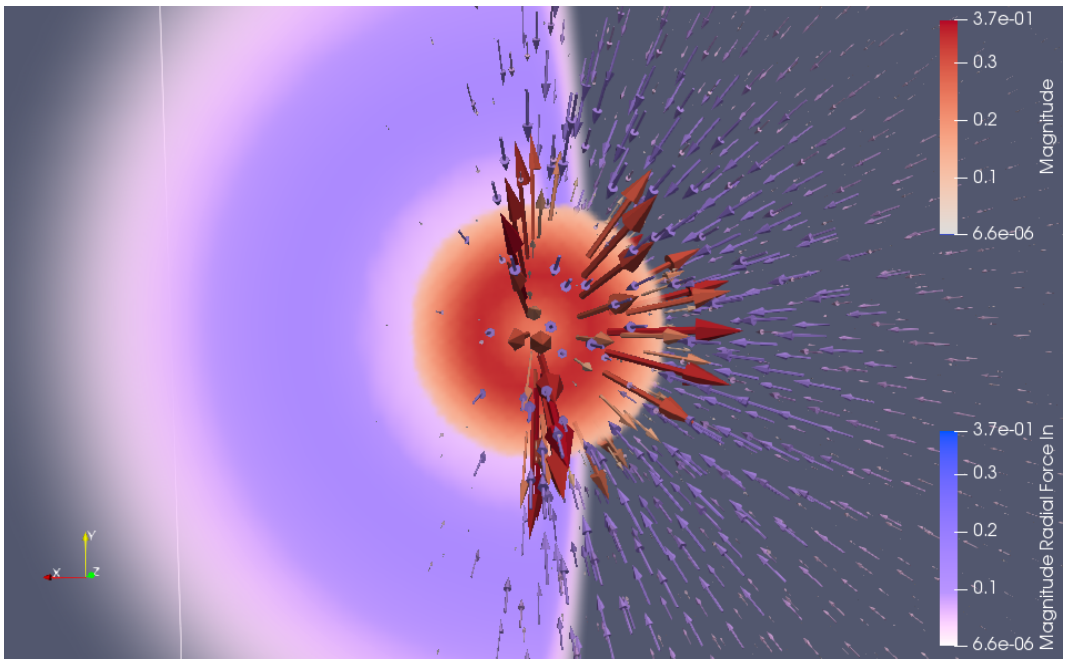


Figure 23: A split cut-away of the Radial forces (based on R-squared; i.e. Figure 13) where the red scale is the radial forces pushing out and the blue scale is the radial forces pushing in. On left is a volumetric rendering of the magnitude scalar, on right

is a Glyph-based rendering of the same (symmetric) vector field shown for reference. Arrow Glyphs and volume magnitude are scaled by a factor of .8 for portrayal.

Reflections

Many visualization algorithms bring some notion of re-sampling; this is especially true for working between points and volumes. Volume rendering techniques are now common with 256 cubic voxel spaces, but even the most modern workstations are challenged at the order of 512 or 1024 cubed at interactive rates. Viewing full-detail or even densely-sampled models (rendering every point) at future data scales is only possible with HPC hardware and remote visualization. Thankfully, Paraview can be used remotely on DOE and VT machines with good performance and effect [Abidi et al, 2018]. For this project, the team examined the spatial extent, distributions, and details of the data using contours on the volume data. Upon finding several contour asymmetries in the 3D nucleon data, we investigated with a variety of *data extents and resolutions and contour threshold levels*. This helped us resolve our Paraview pipeline assumptions and have confidence in our visualization results.

Further, we resolved that while such interrogation by contouring *should be used* for data quality control and debugging, they should be used with restraint in production, and only when they represent an actual surface, a physically-relevant threshold value, or serve as a reference frame for measurement. Once confident of the expected dynamics and visual effect, we resolved to use a uniform spatial sampling of the data for glyph seeding wherever possible and to also mitigate Glyph crowding with scale factors such as size and transparency, and also to use data cut-aways to mitigate occlusion. Striding Glyph seeds has the downside that it depends on the lexical order of the sampled points, but it can be useful for illustrative purposes rather than scientific analysis.

There was much debate, discussion, and exploration in the group about the physical quantities, their digital representation, and their visual portrayal. We followed perceptual and usability guidelines in our color and volume rendering transfer functions [Harrower & Brewer, 2003 ; Moreland, 2009]. The diverging color scale was useful to jointly portray and distinguish: Radial vs Tangential forces in the nucleon and also the directionality of the Radial forces: Radial in vs Radial out.

Outreach

Website

We have created a persistent VT website for the project at ***url*** the currently published videos reside at on Dr. Polys's VT playlist: *youtube* .

Students

VT Undergrads Oliver Stein and Nicolas Gutkowski worked on this project during Summer and Fall 2019 at Virginia Tech Blacksburg

Future Work

1) The team effort has resulted in the first-ever realistic visualization of force distribution on the quarks inside the proton in 3D views and animation. To further enhance the scientific value of this study, the difference of the original RM and of the finally extracted RM1, which is obtained after the limitations in acceptances, resolution effects, and lack of information of the complete set of Compton Form Factors are quantitatively incorporated, analyzed and visualized. As new measurements have already been conducted, the analysis of these measurements in terms of the mechanical properties should significantly reduce uncertainties, both in statistics and systematic. Here the need arises to visualize the size of the uncertainties as more experimental information becomes available.

2) The previous effort focused on the proton, which is generally described by four complex Compton form factors, in the future we plan to include light nuclei, especially $4He$, which is a J=0 nucleus, and is described by just a single Compton form factor H(ξ,t). This will make the analysis of the internal quark structure in terms of mechanical properties much easier. The beam spin asymmetry has already been measured, but the cross section remains to be analyzed. This would be the first nucleus where mechanical properties on the pressure distribution can be accessed.

3) The current effort has focused on the forces and pressure on the quarks. The contributions of gluons are unknown and require measurements with final states others than DVCS photons. Vector mesons such as exclusive production of $\phi(1020)$ are sensitive to the gluon distribution in the nucleons. An interesting problem is how pressure coming from quarks and from gluons can be visualized together, or separately.

4) So far, the focus has been of forces and pressure distribution. Other mechanical properties are the angular momentum distribution in the nucleon. This requires employing data from transverse target polarization asymmetries and further developments of phenomenological analysis techniques, which have not been developed yet, and require new efforts. In this regard, the EIC, especially at the lower center-of-mass energies will allow for measurements of the double polarization asymmetries with polarized electrons (P_e=85%) and transversally polarized protons (P_p= 80-85%) at high luminosities. Accessing the transverse distribution of spin on the proton will enable us to finally solve the famous proton spin puzzle, as it accesses the orbital angular momentum contributions of quarks to the proton spin. Simulations with realistic fast Monte Carlo and detector implementations will provide important projections for this Flagship Program at the EIC.

5) The team also took a first look into how the proposed Electron Ion Collider (EIC) will allow scientists to explore these questions using precisely defined and controlled collisions of very-high energy ions and electrons including polarized protons and deuterons at high luminosity. Using reactions such as ep -> epJ/ψ we will be able to explore the gluon contributions to the mechanical properties of the nucleons in a very different kinematics range from the currently studied valance quark domain. To address these new challenges of vector meson production in a strictly scientific way to access the gluon contributions to the

protons mechanical properties requires attracting more theorists with so far untapped expertise in this new area of nuclear physics into the team.

6) This collaboration between Physics and Computer Science has enabled new explorations and innovations in the visualization of Particle Physics data. Through open-source and open-standard software, the graphics pipelines we have demonstrated are repeatable and can scale to larger data sizes, sampling fidelity, and rendering size using VT and DOE HPC systems. Lastly, we identified several software components that would improve compatibility in this ecosystem: native Paraview importers for HIPO and X3D, improved Paraview and GEANT X3D exporters, and upgrading all VRML software to support X3D.

7) A overall framework should be developed replacing the "person-in-the-loop" system used for this work, wherein data was transferred by files over email or cloud storage. Existing comprehensive frameworks, such as the CLARA framework developed at Jefferson Lab, should be evaluated. However, some time should also be spent investigating whether a more appropriate solution is to develop a service based architecture based on current industry web services platforms, such at the Java based RESTful service architecture. In this model, the different nodes in our current workflow would be available over the internet as web services.

Talks and Publications

Compilation movie featured in the Virginia Tech exhibit booth at SuperComputing 2019 in Denver CO

References

Faiz Abidi, Nicholas Polys**,** Srijith Rajamohan, Lance Arsenault, and Ayat Mohammed. (2018). Remote high-performance visualization of big data for immersive science. In *Proceedings of the High-Performance Computing Symposium* (HPC '18). Society for Computer Simulation International, San Diego, CA, USA, Article 5, 12 pages.

Harrower M, Brewer CA. ColorBrewer. org: an online tool for selecting colour schemes for maps. The Cartographic Journal. 2003 Jun 1;40(1):27-37.

Moreland K. Diverging color maps for scientific visualization. In International Symposium on Visual Computing 2009 Nov 30 (pp. 92-103). Springer, Berlin, Heidelberg.

Polys, Nicholas F. "Information Visualization in Virtual Environments: Tradeoffs and Guidelines". In: *Handbook of Virtual Environments, Second Edition* (eds.) Kelly Hale and Kay Stanney. CRC Press, 2014.

Polys, Nicholas F. “**Publishing Paradigms with X3D”.** In: *Information Visualization with SVG and X3D*, (eds.) Chanomei Chen and Vladimir Geroimenko, Springer-Verlag, 2005.

Addendum: List of Co-PI Participants

NAME: Burkert, Volker
INSTITUTION: Jefferson Lab
ADDRESS: 12000 Jefferson Avenue, Newport News, VA 23606
EMAIL: burkert@jlab.org
PHONE: (757) 269-7540

NAME: Elouadrhiri, Latifa
INSTITUTION: Jefferson Lab
ADDRESS: 12000 Jefferson Avenue, Newport News, VA 23606
EMAIL: latifa@jlab.org
PHONE: (757) 269-7303

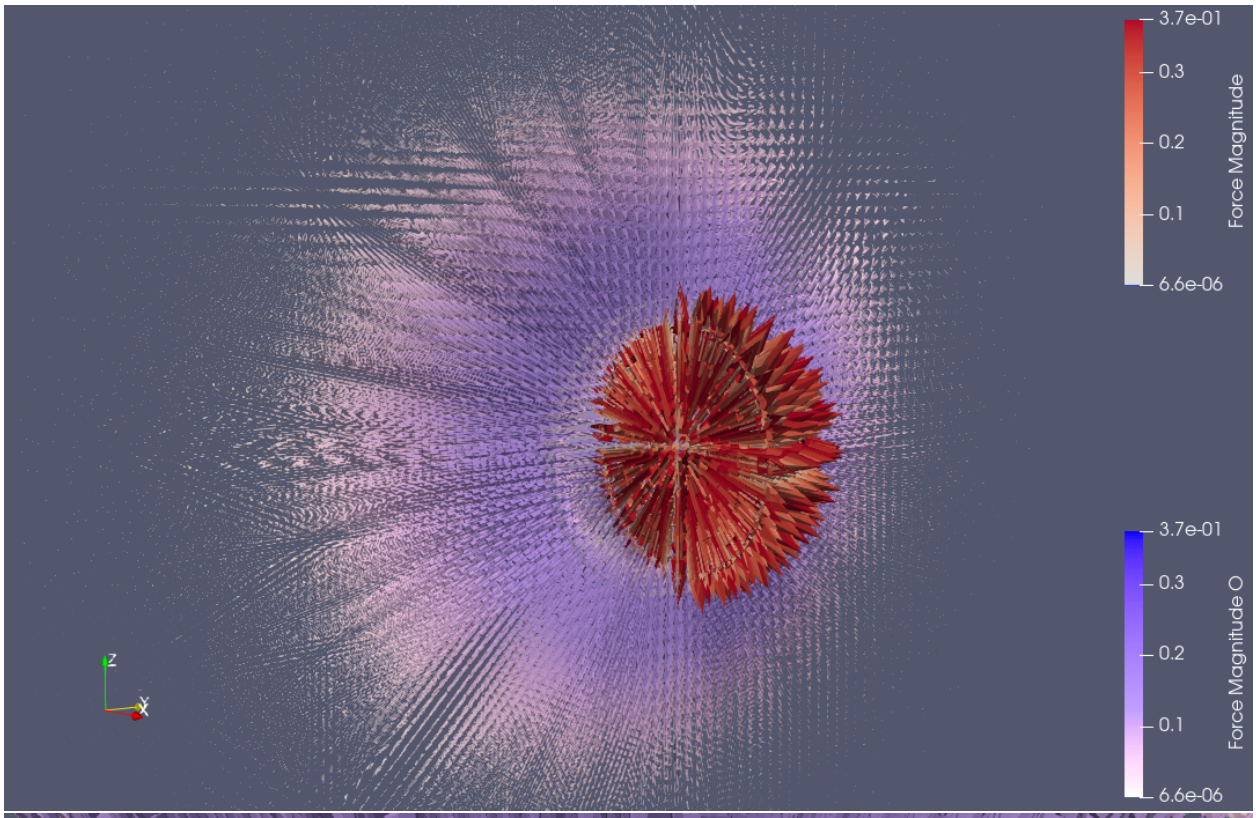
NAME: Girod, F. X.
INSTITUTION: University of Connecticut
ADDRESS: Dept. of Physics, Univ. of Connecticut, Storrs, CT 06269
EMAIL: fxgirod@jlab.org
PHONE: (757) 927 8150

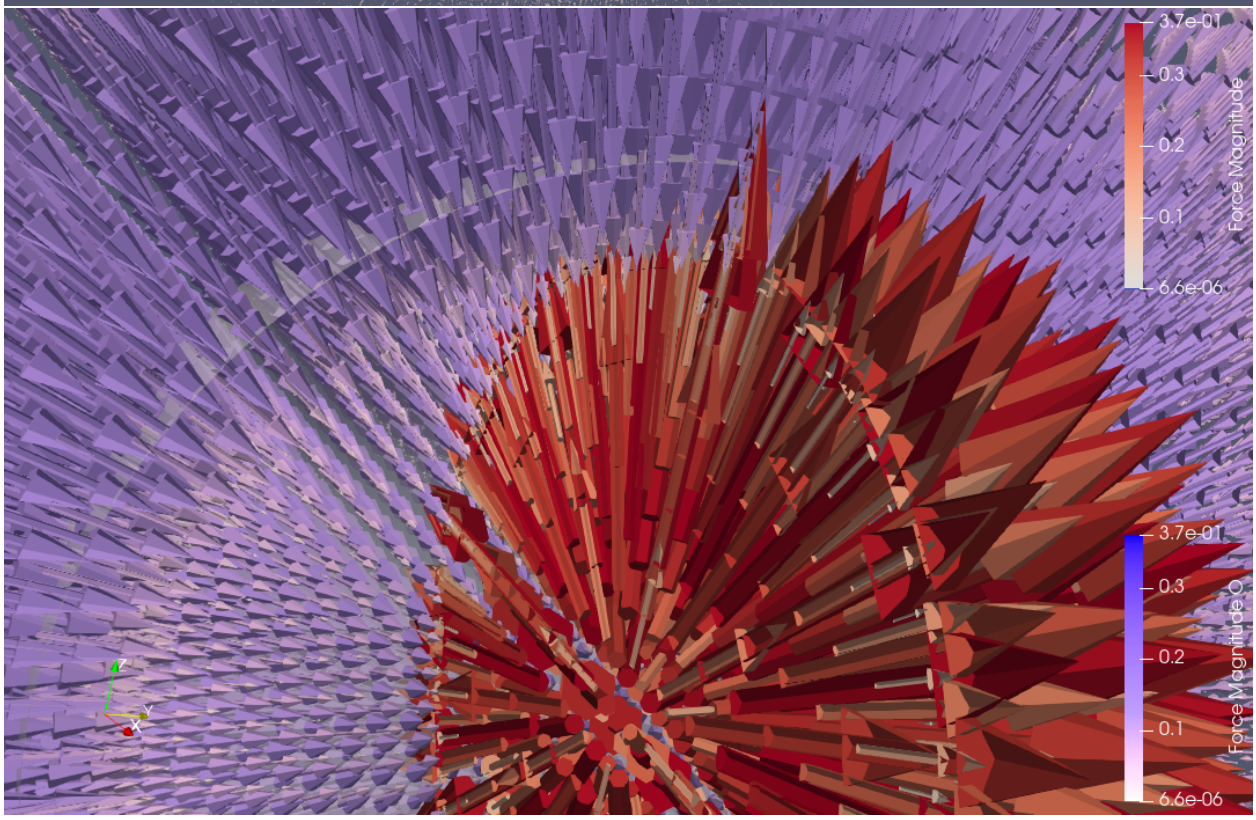
NAME: Polys, Nicholas
INSTITUTION: Virginia Tech
ADDRESS: Virginia Tech Visionarium Lab, (MC0718) Blacksburg, VA 24060
EMAIL: npolys@vt.edu
PHONE: (540) 231-0968

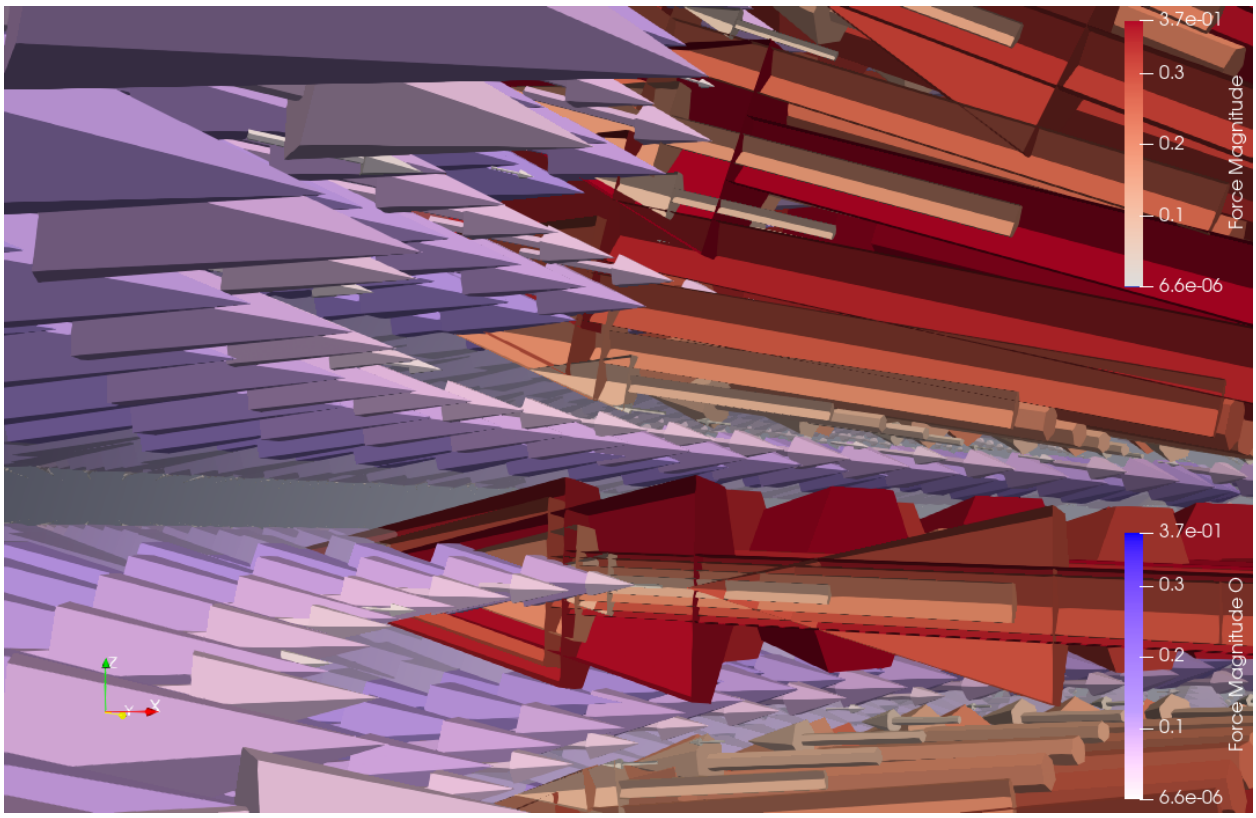
NAME: Schweitzer, Peter
INSTITUTION: University of Connecticut
ADDRESS: Dept. of Physics, Univ. of Connecticut, Storrs, CT 06269
EMAIL: peter.schweitzer@uconn.edu
PHONE: (860) 486-0443

NAME: Vanderhaeghen, Marc
INSTITUTION: Johannes Gutenberg University
ADDRESS: D-55099 Mainz, Germany
EMAIL: vandma00@uni-mainz.de
PHONE: +49 6131 3923695

ADDITIONAL Figures of Radial Force Glyphs






3.7e-01
0.3
0.2
0.1
6.6e-06
Force Magnitude
3.7e-01
0.3
0.2
0.1
6.6e-06
Force Magnitude O