Title: A.I. Assisted Experiment Control and Calibration

Applicant/Institution: Thomas Jefferson National Accelerator Facility (Jefferson Lab)

Street Address/City/State/Zip: 12000 Jefferson Ave / Newport News / VA / 23606

Postal Address:

CEBAF Center
12000 Jefferson Avenue
Newport News, VA USA 23606

Administrative Point of Contact name, telephone number, email:

Deborah Dowd, (757)269-7180, dowd@jlab.org

Lead PI name, telephone number, email:

David Lawrence, (757)269-5567, davidl@jlab.org

Submitting User Facility: Thomas Jefferson National Accelerator Facility

DOE National Laboratory Announcement Number: LAB 20-2261

DOE/SC Program Office: Nuclear Physics

DOE/SC Program Office Technical Contact: Dr. Manouchehr Farkhondeh

Institution	PI Name	Year 1 Budget	Year 2 Budget	Year 3 Budget	Total Budget
Jefferson Lab	David Lawrence	$366k	$369k	$376k	$1,111k

Contents

A.I. Assisted Experiment Control and Calibration

Thomas Britton[1], Naomi Jarvis[2], David Lawrence[3], and David Richards[4]

[1,3,4]Thomas Jefferson National Accelerator Facility
[2]Carnegie Mellon University

March 2020

1 Introduction

We propose to develop and deploy artificial intelligence (A.I.) systems to adjust detector controls during data acquisition in order to reduce, or eliminate, the need for calibrating the data offline. This project represents a first step towards automated facilities in which "smart" detectors communicate with each other and with A.I. embedded in the accelerator to deliver better physics results faster and more efficiently. We believe A.I. can reduce the time it takes to process data and publish a result by 3-6 months and although some related work in accelerator controls has started[6], this type of A.I. control over an experimental end station detector has never been done before. If successful, this will reduce significantly the time and effort required to calibrate large drift chamber detectors. Multiple detectors of this type are currently in use in large scale nuclear physics experiments at Jefferson Lab. Currently, the iterative, labor intensive calibration process extends the time needed to publish results. This project will require one full time physicist and one full time computer/data scientist working for 3 years under the guidance of an existing team of experienced researchers and subject matter experts. This specific project will target CLAS12 and GlueX detector systems which are currently in use and for which much of the data required for the project is already archived. This project is one component of what is planned to be a much broader A.I. program at Jefferson Lab that includes accelerator operations, experimental data quality, data analysis, and connecting experimental results directly to theory (see section 3.6).

1.1 Experimental Nuclear Physics at Jefferson Lab

The Continuous Electron Beam Accelerator Facility (CEBAF) at DOE's Jefferson Lab has four experimental halls, A, B, C and D, in which separate nuclear physics experiments are run concurrently, using the polarized electron beam from the accelerator. Halls A and C are used by many experimental groups for short-term experiments, whereas Halls B and D house permanent installations of the CLAS12 and GlueX spectrometers, respectively. These are operated by larger groups - the CLAS collaboration has around 200 members, while GlueX has around 130 - and their experiments run for long periods of time in order to acquire the high statistics necessary for high precision results. Both CLAS and GlueX use drift chambers, which are featured in this proposal.

The mission of Jefferson Lab (JLab) is to produce high quality experimental nuclear physics results. The accuracy of these results and how quickly they are produced depends heavily on the detectors and their calibration. At present, detector calibrations are a vital and extensive step between recording and analyzing the data. High-level calibrations translate measurements into physics quantities and many more low-level calibrations ensure that the measured data are consistent throughout variations in local environmental conditions. The calibrations are performed soon after the data have been collected, using software written specifically for this task. There are many iterations of calibrations, checks and corrections, first for each sub-detector (such as a drift chamber or a calorimeter), and then for the spectrometer as a whole. The actual data analysis for the experiment starts after the calibrations have been completed.

1.2 Drift Chambers for Charged Particle Tracking

Drift chambers (DC) are a common detector system in Nuclear Physics, used to track charged particles by measuring the ionization produced when the charged particles pass through a gas volume. The liberated electrons drift towards the nearest anode wire in the chamber, ionizing other gas atoms on their way to create an avalanche of electrons that forms the electrical pulse whose characteristics are recorded for that wire. The gain of the chamber determines the size of the avalanche and therefore, the height of the pulse recorded. This affects both the measured amplitude (dE/dx) used in particle ID and the measured drift time, used to determine particle momentum. Maintaining a stable gain is therefore critical to achieving stable detector performance. The chamber's gain depends primarily on the anode voltage, but is also affected by many other variables including pressure and temperature. The pre-amplifier boards that prepare the signals for digitization also contribute to the overall gain and will have a dependence on event rate and temperature. Traditionally, stability of the overall gain is achieved by attempting to keep four of these five parameters independently stable (atmospheric pressure cannot be controlled). The atmospheric pressure is the most significant of these over larger time scales of $\sim$ 1hour. The GlueX Central Drift Chamber (CDC)[4] experiences fluctuations in atmospheric pressure typically of the order of $\pm 2\%$ and the resulting gain changes in the CDC are of the order of $\pm 15\%$. Figure 1 shows the atmospheric pressure recorded throughout the data-taking period in fall 2018. Figure 2 shows the calibration gain correction factors of the GlueX CDC plotted vs. the ratio of atmospheric pressure to temperature. It should be noted that the calibration procedure used to obtain the values plotted in figure 2 was done independently of the pressure and temperature. A similar effect is seen in the CLAS12 drift chambers in Hall-B. Figure 3 shows a 3-D view of the gain factors as a function of both the atmospheric pressure and beam current. Correlations with other operational parameters are expected, but are very difficult to disentangle using traditional means.

We propose to control the high voltage (HV) on the drift chamber anode wires, dynamically, using an A.I. system that incorporates information from a variety of sources in the experimental hall during data taking. In addition, the model will also predict a set of calibration constants for the drift chambers. The input parameters will include pressure, temperature, and beam conditions

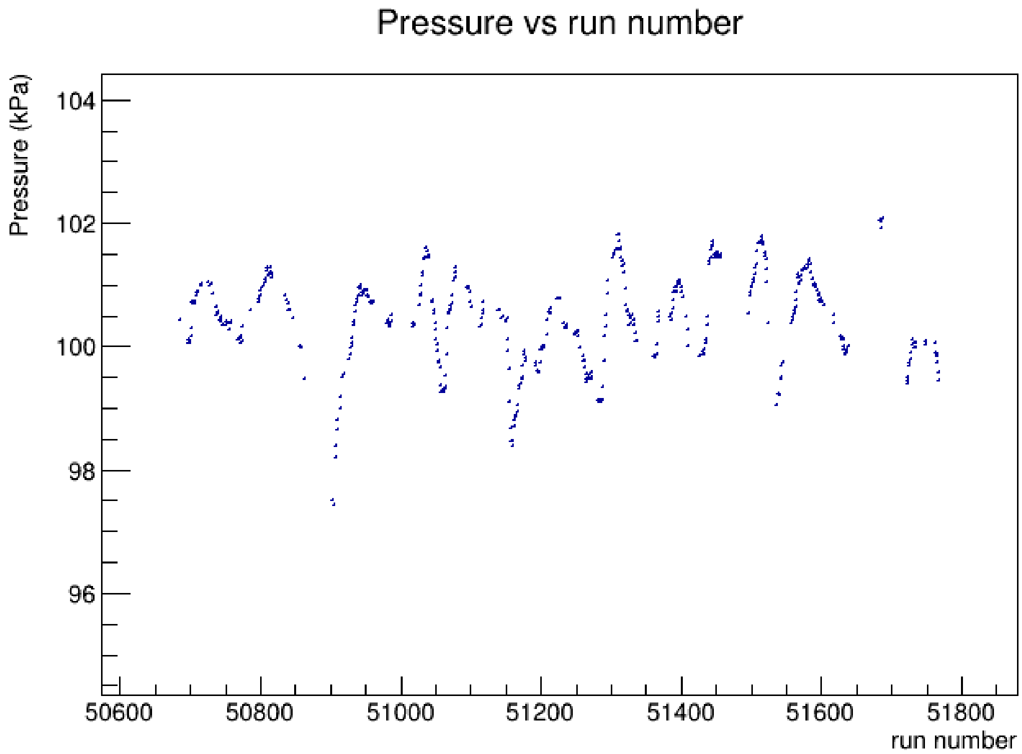

Figure 1: Data from the GlueX Central Drift Chamber showing the change in barometric pressure throughout the Fall 2018 run.

from the existing slow controls system. Inputs will also include reconstructed values from the data stream which naturally incorporates information from all detector systems. Figure 4 illustrates this concept.

2 Project Objectives

Our ultimate goal is the development of a holistic detector control system. Such a system would decrease the time that experts are required to spend on managing their respective detector (both in control and calibration). This time is often in short supply during experimental running. The reduced load on experts would also provide a multitude of secondary benefits. For example, an environmental event may require a change in controls or calibration, or in some cases cause the detector to produce a whole host of alarms; shift crews may then call expert personnel, sometimes erroneously, to attempt to correct the problem. Replacing the humans involved with this aspect of the overall workflow could lead to increased productivity in other tasks and lower the costs associated with dealing with the alarms. The specific objectives outlined in this section form the

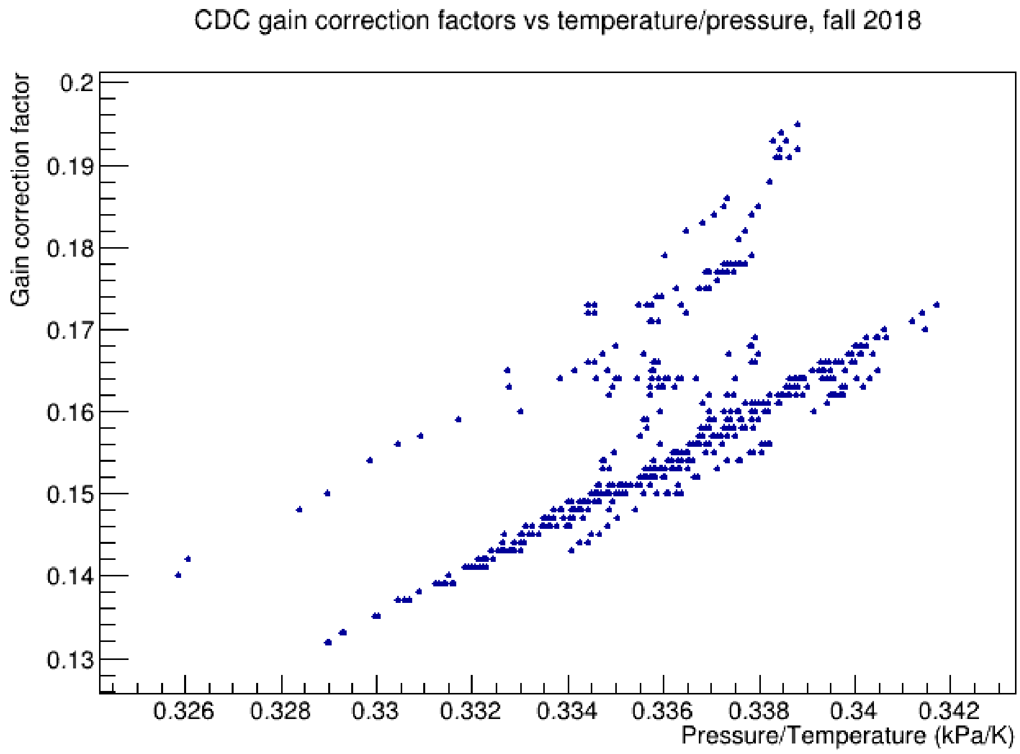

Figure 2: Gain factors for GlueX Central Drift chamber as a function of the ratio of barometric pressure to temperature during Fall 2018. A relationship between chamber gain and gas density is evident but there are several other factors at play which are not understood well. A similar effect is seen in the CLAS12 large volume drift chambers.

first steps towards that ultimate goal. The later sections 3 and 4 describe the strategy to be used and present the proposed schedule of activities.

2.1 Maintaining a Consistent CDC Response

The primary objective of this research is to maintain consistent detector response in reaction to changing environmental and experimental conditions by controlling the HV of the GlueX CDC and simultaneously generating calibration constants autonomously in near real-time.

The specific goal would be to stabilize the gain to within 5% over a 2 week period with no measurable degradation of the timing resolution. A successful application with the GlueX CDC would then be used to develop a similar system for the CLAS12 drift chambers, thus allowing the solution to be generalized.

We are already aware of a clear relationship between the chamber gain on pressure and temperature, but there are many other factors which we have not attempted to parameterize, such as

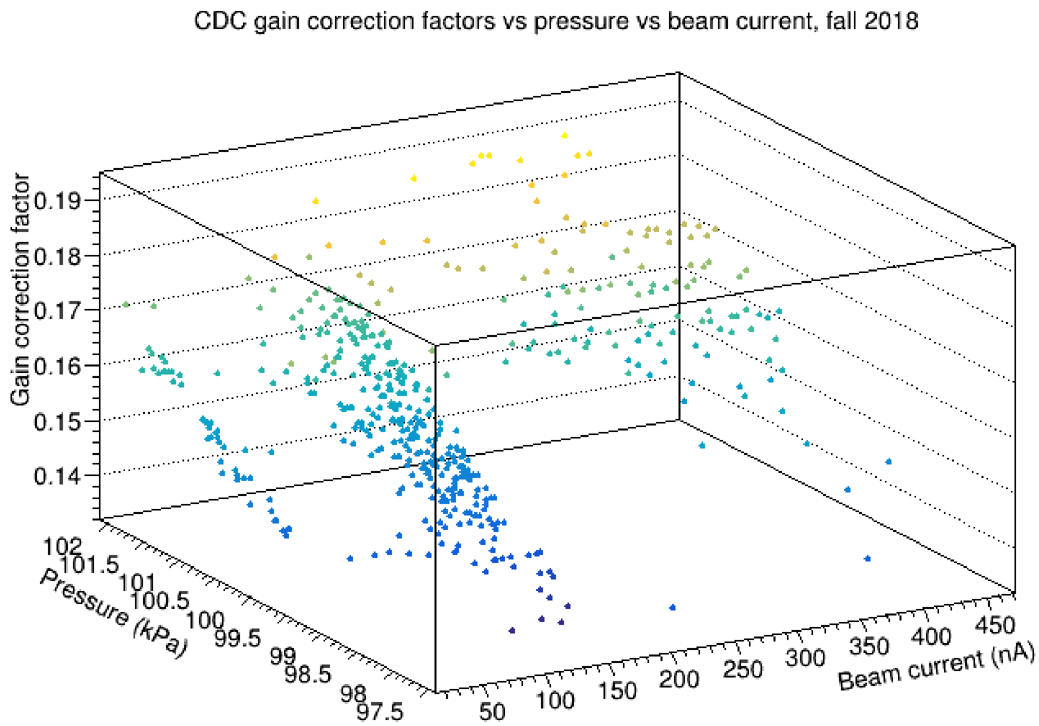

Figure 3: Gain factors for GlueX Central Drift chamber as a function of both the atmospheric pressure and the beam current during Fall 2018. This indicates the gain factors have a dependence on both of these and that the relationship is non-trivial.

event rate, beam intensity and the currents drawn by the high voltage boards. This additional unstudied dimensionality makes this an ideal project for artificial intelligence.

2.2 Reduce the Delay Between Data Acquisition and Analysis

By improving the control over the detector and generating an initial set of calibrations, this research will decrease the time interval between the end of data acquisition and the beginning of analysis.

GlueX produces a set of calibration constants for each detector for every data-taking run (this is the interval between when data acquisition is started and stopped). The duration of each run is limited to 2h or less, to keep the gain drifts due to environmental changes at or below a tolerable level. The calibrations require a considerable amount of expert attention and during the months of data acquisition the detector experts are fully occupied with the experimental program, so the calibrations start as soon as the months of data acquisition come to an end.

In the first round of calibrations, gains are calibrated for each detector independently. In subsequent rounds, information is shared between the detectors to refine the timing offsets, and

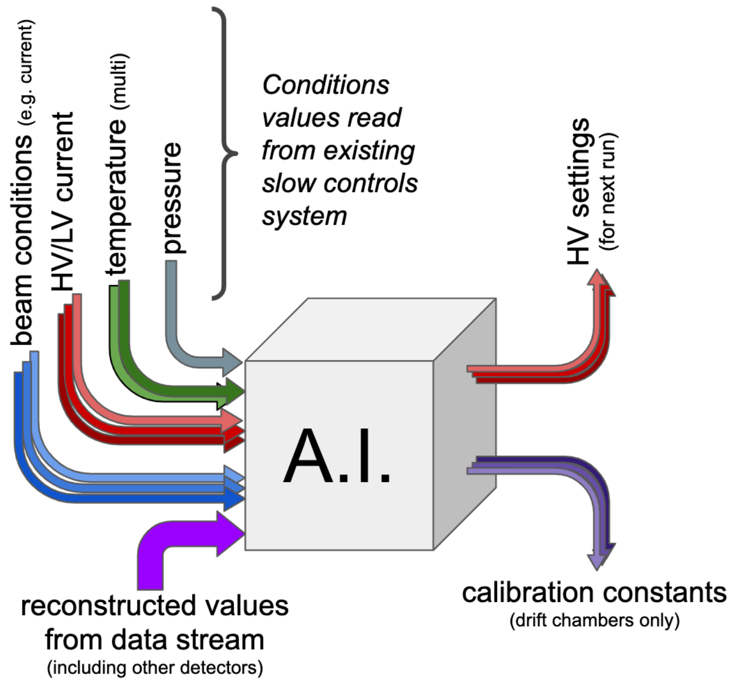

Figure 4: Illustration of the AI model which would take inputs from both slow controls (EPICS) and high speed data acquisition systems to produce a predicted set of HV settings and calibration constants that should be used together for the next run.

here the process becomes iterative as the track vertex positions from one detector are important for the path length used to determine timing offsets for others, while the overall timing is used to refine time-to-distance calibrations and hence track positions in others. The collaboration works to make the calibration process as efficient as possible, the data analysis is coordinated, one person submits all of the computer jobs for processing and a team of junior scientists monitor the results and flag anomalies for expert attention, but it still consumes a considerable amount of time from the dozen detector experts.

While speeding up the entire calibration process, an A.I. would reduce the time commitment required from the experts, who would then be free to turn their attention to data analysis and our physics program. An A.I. would also be able to take over some responsibilities from the shift crew in charge of the detectors and the data acquisition; as the A.I. would be monitoring the environmental conditions continuously it would be aware of unusual weather events, such as the rapidly falling pressure that occurs when a storm front arrives, and it would be able to alert the crew of the need to stop a run early or take the appropriate corrective steps on its own and continue running, all without requiring expert intervention.

For this project, the objective is to reduce the number of iterations required to calibrate the CDC or reduce the number of runs that must be calibrated.

2.3 Facilitate the Automation of Other Detector Systems

Other detectors, such as those of the CLAS12 collaboration, are expected to benefit from similar A.I. systems. The knowledge and experience gained here will aid the development of those systems. To this end, the A.I. system developed in this research will be tested with the CLAS12 drift chambers.

Ultimately, experiments can be viewed as a complex, highly dimensional, coupled system. In this view they become perfect targets for the application of artificial intelligence. By enabling various detectors to coordinate intelligently, experiments may be able to obtain better data, and reduce the time and effort expended to generate calibrations.

This objective will be met by producing software that is general enough to accommodate its use on other detector systems, and open sourcing any pertinent materials produced during this research.

3 Proposed Research and Methods

The research team has experience using A.I. in developing the *Hydra* system. *Hydra* is an A.I. system for data quality monitoring, currently used by GlueX. *Hydra* was developed to remove humans from the tedious task, one often left undone due to the demands of other work or human fatigue, of checking a set of histograms every minute or two. It looks at detector occupancy in finely grained time-steps and reports on the health of the detector. *Hydra* has the domain knowledge of experts, does not suffer from fatigue, and can analyze a plot every 83 ms. These are especially valuable skills for more junior shift crews. Shift crews may view a webpage which collates the results in near-realtime and take action should *Hydra* detect a problem. An example of the performance of the system, which varies by detector but is typically over 95% accurate as compared to experts, is shown in figure 5. The experience gained working on this project makes the team well positioned to construct a system for both calibration and control.

Several components are necessary for the development of an A.I. based system for detector control and calibration. These are described below.

3.1 Monitoring and Controls

An accurate and robust monitoring system is required for the A.I. system to sense its environment and take appropriate actions in given situations. It requires access to all of the quantities needed for decision-making in order to make a plan for calibration and/or control and to know when that plan has been successfully executed. We will exploit the Experimental Physics and Industrial Control System (EPICS) that is already in use at Jefferson Laboratory. This system monitors all of the conditions parameters that the A.I. would need and also provides access to the detector controls via well established APIs.

3.2 A.I. system

The A.I. system will be comprised of several models, each trained to perform a task or distill data for other A.I. processes. Use of unsupervised learning techniques would reduce the need for expensive

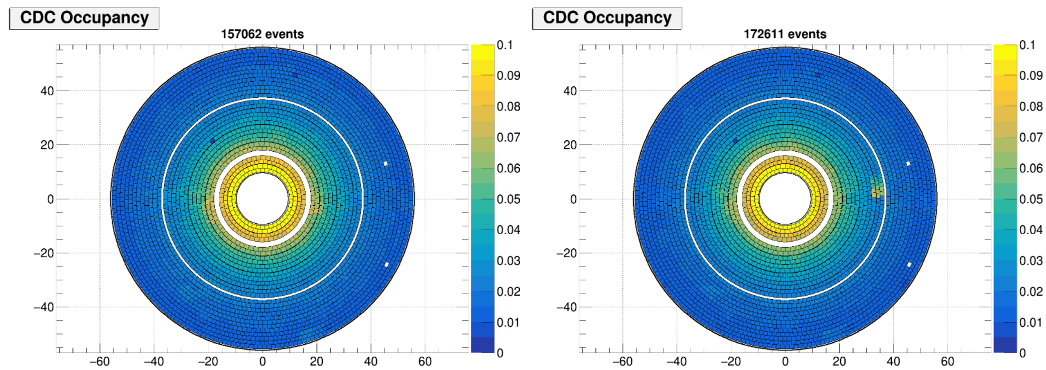

Figure 5: Examples of CDC occupancy plots, using data collected for approximately one minute. Hydra categorized these plots as good (left) and bad (right), and in the latter case Hydra was 99.99018% confident that there was a problem. The expert confirmed that Hydra was correct - note the small yellow high occupancy patch to the right of the center.

operational data acquisition. For some tasks there is ample data already available, which could be labeled by experts to employ supervised learning techniques. A mix of guided learning, in which the system is partially trained on expertly labeled data (and further in an unsupervised way), and reinforcement learning, which rewards or punishes the system based on in situ behavior, would also be appropriate. The system will likely use some form of recurrent neural network (RNN) as these are well suited for tracking trends and making predictions.

3.3 Training and Testing

Simulation of parts of the detector system would allow for a cost-effective way to generate data for both training and testing of all or part of the A.I. system. Cosmic ray data, which are easy to obtain, can be used to provide early tests of the system without relying on expensive dedicated beam time.

3.4 Limitation

To provide confidence in the use of the A.I. for detector control, the A.I. must be robust, fault tolerant, and forced to remain inside a predetermined operational envelope. This can be achieved by having the set of operational parameters generated by the A.I. system pass through a filter. If a parameter set suggested by the A.I. violates any constraint imposed by the filter, it would reject the parameter set and notify experts. In this way, ultimate control would remain with the experts and no harm could come to the detector. Initial implementations will only advise shift workers on a suggested change of settings, so human filtering will be required until enough confidence is gained in the system to allow for fully automated control.

3.5 Performance Feedback

The performance of the A.I. system must be monitored. The main objective is to reduce human intervention in calibration and control of the detector while optimizing stability. The challenge is in keeping the system as easy to interpret as possible. In developing a means to monitor the performance of the A.I. system we would seek to analyze the system and the models that comprise it. We would also aim to subject the models to synthesized data and understand the output produced.

3.6 Broader A.I. Program in Theoretical and Experimental Nuclear Physics

This research is the first step on the long road to "self-driving" facilities. These are experimental facilities that would be self-calibrating and managed autonomously, e.g. facilities where the beam is tuned and optimized based on the detector responses in real time.

One particular challenge that must be overcome, and which this proposed research will address, is the sociological aspect of working alongside artificial intelligence. Much of the problem space that A.I. addresses today was unfeasible just a few years ago. The detectors used for nuclear and high energy physics experiments are often very complex and very expensive. Trusting the operation of these detectors to artificial intelligence is, understandably, worrying at first glance. It may be for this reason that such research has never been performed before. It is the opinion of the proponents of this research proposal that the time is right to explore the possibilities of operating and calibrating these detectors with artificial intelligence. Deep learning and hardware capabilities have enabled A.I. to tackle ever increasingly complex challenges. As such, this research may have not been possible even 5 years ago. It will take time for researchers to grow comfortable working with A.I. in this new way. It is therefore best to introduce these systems incrementally, outside of the standard operational budget of the lab and target smaller, but no less complex, systems first. Extended safe and successful operation of an A.I. system like that proposed here will lower the sociological barriers to these types of programs and pave the way for increased adoption.

3.6.1 Other Opportunities for A.I. Driven Experiment Control

Modern detectors have numerous controls, such as high voltage, low voltage, readout thresholds, trigger thresholds, coolant temperatures, etc. The performance of any one detector will depend on some combination of these, and also on other factors that cannot be controlled easily such as environmental conditions. These conditions can be monitored and used to adjust the detector controls to obtain a more stable detector response. Here are two more examples that have been identified where such a system could be beneficial.

Digitization crate fan speed A dependence of the electronic pedestal on the temperature has been observed in some digitization electronics, for example, the 250MHz flash ADC modules used for the Barrel Calorimeter detector[1] in the GlueX experiment. Pedestal shifts can cause either increased hit rates leading to electronic noise that masks true hits, or missing small amplitude hits that are no longer above threshold. It has been shown that controlling the fan speed of the electronics crate was able to improve the stability of the pedestal. An A.I. could be used to combine temperature, hit rate, and amplitude information to control the fan speed, pedestal offsets, and thresholds of the digitization electronics ensuring continually stable performance.

PMT time sagging Photomultiplier Tubes (PMTs) often have a dependence of the gain on the rate of signals they observe. This well-known effect[5], [3], which is due to the voltage "sagging" as more current goes through the tube, is ameliorated, but not eliminated through the use of active bases. These attempt to compensate for the higher currents by adjusting the HV to stabilize the gain. What is really needed, however is to stabilize the timing. An A.I. could be trained to use the existing HV, current, measured hit rate of the tube, and the measured timing offset to adjust the HV and readout thresholds in such a way that the timing and hit efficiency are simultaneously kept stable.

3.6.2 Delivering Better Beam to Experimental Halls

An ongoing project at Jefferson Lab leverages machine learning (ML) to automate cavity trip classification. Traditional methods are effective at identifying superconducting radiofrequency (SRF) trip causes, but are labor intensive and generate results in an asynchronous fashion. Identifying and correcting faults in real-time will have numerous benefits including improving the stability of the SRF system, providing a more reliable and available accelerator, and extending the energy reach of the accelerator. All of this equates to more beam on target for experimental halls thus producing more/better physics results.

3.6.3 Connecting Theory to Experiment

The increasing volume and complexity both of the data arising from experiment, and of the computational and theoretical investigations of Quantum Chromodynamics, is demanding new methods to describe the observed particles of nature, the nucleon and nuclei, in terms of the fundamental quarks and gluons of which they are composed. The one-dimensional distribution of the quarks and gluons in terms of their light-cone momentum fraction x is encoded in the unpolarized and polarized Parton Distribution Functions (PDFs); they are important both to our fundamental understanding of the structure of hadrons, and to the interpretation of experiments at the energy frontier, such as those at the LHC. More recently, new three-dimensional measures have emerged that encapsulate both the longitudinal and transverse structure of hadrons: the *Generalized Parton Distributions* (GPDs) describing their structure in terms of x and the transverse impact-parameter space, and the *Transverse-Momentum-Dependent Distributions* (TMDs) describing their structure in terms of x and transverse momentum space. Together, they can provide descriptions of nucleon structure that elude the one-dimensional distributions, notably in revealing the role of the angular momentum of the quarks and gluons. These new measures, and the prospect of constraining and interpreting them both from experiment and through computation, have led to the new science of *Nuclear Femtography.* It is in recognition of this opportunity that the Commonwealth of Virginia has established the *Center for Nuclear Femtography* (CNF).

None of the PDFs, GPDs and TMDs is directly measurable in experiment, but they are related to experimentally measurable cross sections, and therefore to extract them requires the solution of the inverse problem. The global-fitting community, focused on extracting the PDFs from the worldwide experimental data, including in particular, those from Jefferson Lab, has historically addressed the issue of incomplete data by assuming a functional form for the PDFs, motivated by our physical understanding, thereby formulating the problem as one of optimization over a set of parameters. However, there is an increasing realization that AI proves a powerful framework both to undertake the numerically demanding effort that this requires, to tabulate the PDFs in as unbiased way as possible, and to identify the need for more data and calculation. Thus machine-learning experts at Old Dominion University and at Davidson College are collaborating with theorists at Jefferson Lab, with the support of the CNF, to apply machine learning methods to develop a mapping between experimental data and parametrizations that can rapidly see the impact of extant data, and can identify the need for new data and computation.

As we move to the era of nuclear femtography facilitated by experimental data from the 12 GeV upgrade and the future EIC, and the opportunities for *ab initio* computaton in the approach to the exascale, the challenges are still greater. The GPDs and TMDs encoding the three-dimensional structure are far more involved than the one-dimensional PDFs, the physical guidance as to their behavior far less constrained, and the data will inevitably remain incomplete. Thus AI has an ever more essential role in the experimental and theoretical campaign to reveal the structure of nucleons and nuclei.

4 Timetable of Activities

Experience with the development of *Hydra* tells us that given the complex nature of the proposed research, the inherent risk in operating a detector with artificial intelligence, and the need for interpretability (an area of A.I. research with many open questions) this program requires three years of work broken, roughly, into three major themes. Any shorter time given to the task would risk producing models which may be over fit and unable to respond appropriately to unbefore seen situations, uninterpretable and biased, producing output based on spurious correlations, and/or a system overly specialized and unable to be generalized and unable to operate in situ. The first year will be focused on data gathering and model development. Year 2 will focus on model development and refinement. This leaves year 3 to focus on the deployment and integration of the system(s) produced, culminating with a deployment of a system to calibrate and control the CDC detector or the CLAS12 drift chambers. The primary activities are given in roughly chronological order below.

4.1 Orientation and Model Review ~ 1 *month*

This will take place first. It will be the time in which Naomi Jarvis will spend one week at Jefferson Laboratory directly working with and orienting the new hires on the GlueX CDC as well as the specifics of control and calibrations. Similar information will be gathered for the CLAS12 DC. It will also be a time for the identification of any models which may have potential for transfer learning in part or whole.

4.2 Analysis of Historical Data ~ 1 *to* 2 *months*

This must take place early in the research project. Much of the data will be found in the Experimental Physics and Industrial Control System (EPICS) archive, which stores a large number of the variables in the experimental run. Other data will be mined from each of the CLAS12 and GlueX Calibration and Control Databases. During this time data sources will be accessed and simple data analyses will be performed. This will give the entire research team specific knowledge about the detector systems and will give ways to effectively judge produced models.

4.3 Candidate Algorithm Selection ~ 1 *to* 2 *months*

This must take place before work on the system can begin in earnest. The specific topology and algorithm(s) used will need to be determined in the course of this research. It is unlikely that a 'one size fits all' approach will yield desired results and thus by the end of this period candidate algorithms and techniques should be found. Any additional data used by the chosen technique(s) should also be identified. During the preceding Analysis of Historical Data period a basic understanding of the data which, given the overall challenge, will evoke possible algorithms. Given the coupled nature of this and the preceding section, 1 to 2 months should be ample.

4.4 Infrastructure Development ~ 1 *month*

This period includes the production of any necessary data sets to be employed in training the models. If expert labels need to be produced, they will be obtained during this time. The quantities for judging the efficacy of the produced models will be identified. Any computing infrastructure, including tool-kits such as Tensorflow and Keras, will be obtained. The choice of A.I. framework will be based on the needs of the research team and industry developments in the intervening period between the submission of this proposal and the start of work. The time allocated for this is comparable to that which was required for *Hydra*.

4.5 Candidate Model Development ~ 6 *months*

The remaining time during the first year will be wholly dedicated to the development of the candidate model(s). For any unsupervised learning, historical data will be used and presented to the model as if it were live data. Supervised learning will take place in the usual fashion. These candidate models will be subjected to tests involving cosmic ray data to ensure that problems can be diagnosed and addressed early. Hydra needed about 6 months of iteration and testing to build up its system of candidate models. The early testing with cosmic data will be invaluable. At the end of the first year, candidate models with potentially non-optimized hyperparameters will be produced.

4.6 Model Refinement ~ 6 *months*

The candidate models produced should be refined and iterated on during this time. This may involve selecting the best candidate model developed for a given task (or subtask). It may involve some fundamental changes to the data that the model uses to improve efficacy. The models from this period will continue on to further hyper parameter optimization (HPO).

4.7 Hyper Parameter Optimization and Interpretability ~ 6 *months*

This work includes optimization of the models to reduce inference times (the time it takes to pass input data through the model), improvement of training methods to produce better generalized models, and the fashioning of the entire system to be deployed. For each model produced, care must be taken to ensure that, to the highest degree possible, the models are able to be interpreted, i.e. in order to ensure the safe operation of the system, the research team must be able to test the system's operation in a bevy different situations and understand why the system behaves as it does. This is an important and non-trivial part of A.I. development [2] and an active area of research in the A.I. community. This activity is likely to blend with Candidate Model Development. In the early stages of *Hydra* development, simple proof of concept models were developed utilizing chosen algorithms. It took as much time proving out a candidate model and improving the training of the model to make it suitable for deployment in the hall as it did to develop the candidate model.

4.8 Integration and Deployment ~ 12 *months*

A system for maintenance and monitoring of the A.I. system will be developed and deployed to ensure that the system adapts, or can be adapted, to drastic changes in experimental conditions or detectors. It is often easy to train a model and run inference ad hoc. Ensuring that the system can operate within the necessary time-frame (set by the problem being solved) given additional operational overheads (e.g. communicating with the controls system) brings a whole host of new challenges. For example, during the development of *Hydra* it was discovered that the first inference of a model took considerably longer than subsequent inferences. This was due to the dynamic loading of models and their necessary libraries, and it was corrected by developing a system to load them in advance and force a single inference. The integration and deployment period will culminate in the active deployment and use of a system to calibrate and control drift chambers in either the CLAS12 or GlueX detectors, whichever appears most promising at the time.

5 Project Management Plan

This project will be carried out as a joint effort between Jefferson Lab and Carnegie Mellon University with Jefferson Laboratory serving as the primary base of operations. The project participants will meet bi-weekly via video conference (or more often as needed) to review progress. Jarvis will visit JLab at least once, the first visit being close to the start of the project.

The duties of the individuals will be:

David Lawrence (PI) Lawrence will spend 5% of his time on this project. The primary responsibilities will be to oversee the project, ensure milestones are met, report on project progress, and provide guidance where needed. The PI was a full time staff member in Hall-D for more than a decade and is familiar with the software and data systems there and will provide support and expertise as needed for accessing the data.

Thomas Britton (Co-PI) Britton will spend 10% of his time on this project. He spent a 3 year post-doc in Hall-D with the GlueX experiment. During this time he acted as online monitoring coordinator. After these 3 years he became full staff in IT where he led development of *Hydra*, an A.I. system for data quality monitoring. He will work directly with the hired post-doc and computer/data scientist providing guidance on developing the A.I. system and integrating it with the GlueX data acquisition workflow.

Naomi Jarvis (Co-PI) Jarvis will spend up to 10% of her time on this project. She spent the last decade working on the construction, readout, monitoring and calibration of the CDC for GlueX. Her role will be to develop limits for the A.I., provide technical expertise on the drift chamber and feedback on the performance of the A.I.

Post-doc (to be hired for project) The post-doc will be an employee of Jefferson Lab paid with funding from this project. The post-doc, in conjunction with the Computer/Data-scientist, will perform much of the day-to-day work related to the project. These include:

- Gathering the relevant data and formatting it for use in a training system
- Coordinating with the computer/data scientist on the development of the A.I. system.
- Ensuring that the A.I. system respects physics constraints and, with the help of Naomi Jarvis, testing the efficacy of any model produced to render valid physics results.
- Evaluation and communication of the results

Computer/Data-scientist (to be hired for project) The individual will be an employee of Jefferson Laboratory funded through this project. They will work closely with the hired postdoctoral fellow and the PI/Co-PIs of this proposal on much of the day-to-day work. These responsibilities include:

- Developing appropriate models for the A.I. algorithm and training them
- Ensuring that models are well generalized and not over fit.
- Ensuring that models are interpretable
- Evaluation and communication of the results

Appendix 1: Biographical Sketch(es)

David Lawrence Ph.D. (PI)

David Lawrence Ph.D.
Staff Scientist III, EPSCI Group Lead
Thomas Jefferson National Accelerator Facility, Newport News VA
e-mail: davidl@jlab.org **phone: (757)269-5567**

Education and Training:

Ph.D. in Physics, Arizona State University, 1998
Subfield: Experimental Subatomic Physics
Dissertation: Initial Tests of the pb Decay Detector
Advisor: Barry Ritchie

M.S. in Physics, Arizona State University, 1995

B.S. in Physics, University of Oklahoma, 1992

Research and Professional Experience

2020-present: Staff Scientist(SSIII 2020-present)
Experimental Physics Software and Computing Infrastructure Group Lead
Jefferson Lab, Newport News,Virginia

2005-present: Staff Scientist(SSI 2005-2009; SSII 2009-2020)
Jefferson Lab, Newport News,Virginia

2004-2005: Data Acquisition Physicist(SSI)
Jefferson Lab, Newport News, Virginia

2005-2014: Adjunct Professor of Physics
Christopher Newport University

2004-2007: Adjunct Assistant Professor, Experimental Nuclear Physics
University Massachusetts

2003-2004: Research Assistant Professor.
Experimental Nuclear Physics University Massachusetts
(stationed at Jefferson Lab Accelerator in Newport News, VA)

1998-2003: Senior Postdoctoral Research Associate
University of Massachusetts
(stationed at Jefferson Lab Accelerator in Newport News, VA)

Publications: (https://orcid.org/0000-0003-0502-0847)

JANA2 Framework for Event Based and Triggerless Data Processing
David Lawrence, Amber Boehnlein, and Nathan Brei
2020 EPJ conference series (epjconf200650 in process)

Measurements of Meson Polarizabilies
D. Lawrence 2020 Proceedings of Science (vol 317 Feb)
doi:10.22323/1.317.0032
https://pos.sissa.it/317/032/

Measuring the charged pion polarizability in the $\gamma\gamma \to \pi^+\pi^-$ ***reaction***
D. Lawrence, R. Miskimen, E.S. Smith, A. Muskarenkov 2013 Proceedings of Science (vol 172 Jul)
doi:10.22323/1.172.0040
https://pos.sissa.it/172/040

The JANA Calibrations and Conditions Database API
D. Lawrence 2010 J. Phys.: Conf. Ser. 219 042011 (6pp)
doi:10.1088/1742-6596/219/4/042011
https://iopscience.iop.org/article/10.1088/1742-6596/219/4/042011

Multi-threaded event processing with JANA
D. Lawrence 2008 Proceedings of Science (Vol 070 Oct)
DOI: https://doi.org/10.22323/1.070.0062
https://pos.sissa.it/070/062

Multi-threaded event reconstruction with JANA
D. Lawrence 2008 J. Phys.: Conf. Ser. 119 042018 (6pp)
doi: 10.1088/1742-6596/119/4/042018
https://iopscience.iop.org/article/10.1088/1742-6596/119/4/042018

C++ Introspection and Object Persistency Through JIL
D. Lawrence, D. Abbott, V. Gyurjyan, E Jastrzembski, C. Timmer, E. Wolin, JLab Proceedings of the Conference on Computing in High Energy and Nuclear Physics, CHEP06 (ISBN 0230630160) (2006)

Synergistic Activities:

- SBIR Reviewer 2017-2019
- NERSC User's Group Executive Committee 2019-present
- Co-organizer of JLab quarterly Machine Learning challenges 2019-present
- EIC Software Consortium/EIC Software Group 2017-present
- GlueX Collaboration Board member 2005-2006 and 2014-2017
- Organizer *Parallelsim in Experimental Nuclear Physics* workshop at Christopher Newport University Jan. 2011

Identification of Potential Conflicts of Interest or Bias in Selection of Reviewers

- Abbott, David - Jefferson Laboratory, CODA/DAQ
- Austregesilo, Alexander - Jefferson Laboratory, GlueX
- Britton, Thomas - Jefferson Laboratory, GlueX
- Brei, Nathan - Jefferson Laboratory, EPSCI
- Chudakov, Eugene - Jefferson Laboratory, GlueX
- Dalton, Mark - Jefferson Laboratory, GlueX
- Diefenthaler, Markus - Jefferson Laboratory, EIC
- Dobbs, Sean - Florida State University, GlueX
- Deur, Alexandre - Jefferson Laboratory, GlueX
- Furletov, Sergey - Jefferson Laboratory, GlueX
- Furletova, Yulia - Jefferson Laboratory, EIC
- Gasparian, Ashot - NC A&T, PrimEx
- Gavalian, Gagik - Jefferson Laboratory, CLAS12
- Gyurjyan, Vardan - Jefferson Laboratory, CODA/DAQ
- Ito, Mark - Jefferson Laboratory, GlueX
- Jarvis, Naomi - Carnegie Mellon University, GlueX
- Jones, Richard - University of Connecticut, GlueX
- Larin, Iliya - University of Massachusetts, Amherst, GlueX/CPP
- Lersch, Daniel - Florida State University, GlueX
- Meyer, Curtis - Carnegie Mellon University, GlueX
- Mestayer, Mac - Jefferson Laboratory, CLAS
- Miskimen, Rory - University of Massachusetts, Amherst, GlueX/CPP
- Pentchev, Lubomir - Jefferson Laboratory, GlueX
- Phelps, William - Christopher Newport University, GlueX/CLAS12
- Romanov, Dmitry - Jefferson Laboratory, GlueX/EIC
- Shepherd, Matthew - Indiana University, GlueX
- Smith, Elton - Jefferson Laboratory, GlueX
- Somov, Alexander - Jefferson Laboratory, GlueX
- Stevens, Justin - Jefferson Laboratory, GlueX
- Tennant, Christopher - Jefferson Laboratory, Accelerator
- Timmer, Carl - Jefferson Laboratory, CODA/DAQ
- Taylor, Simon - Jefferson Laboratory, GlueX
- Zihlmann, Benedikt - Jefferson Laboratory, GlueX

Thomas Britton, Ph.D. (CO-PI)

Education and Training:

- **Ph.D.**, Syracuse University, Experimental high energy physics (LHCb) *Amplitude Analysis of* $B \rightarrow J/\psi\phi K$, 2016 - Advisor: Tomasz Skwarnicki

- **B.A.**, Coe College, *major*: Physics, Mathematics *minor*: Computer Science, 2009

Research and Professional Experience:

- **Staff Scientist I** - Currently a Staff Scientist with the Experimental Physics Software and Computing Infrastructure group in Scientific Computing at Jefferson Laboratory. Responsibilities include the development of Hydra, the A.I data quality monitoring system, and the support of MCwrapper, the automated platform for Monte Carlo production for GlueX.

- **Postdoctoral Fellow** - Between 2016 and 2019 was a postdoctoral in Hall-D at Jefferson Laboratory. Responsibilities included online monitoring lead, which involved data quality assurance when acquiring data and developing web pages to collate monitoring histograms. During this time he was responsible for developing MCwrapper as the standard of GlueX Monte Carlo production. He was also responsible for measuring the cross section of ϕ.

- **Graduate Research Assistant** - Between 2009 and 2016 was a graduate assistant with the LHCb group at Syracuse University. He was responsible for carrying out an Amplitude Analysis of $B \rightarrow J/\psi\phi K$. He was responsible for developing a laser test stand for testing silicon detectors. Additionally, was responsible for wire bonding detectors for testing.

Publications:

- **Hydra** (https://github.com/JeffersonLab/Hydra) - Hydra is an A.I. system currently deployed in Hall-D for the GlueX collaboration. Hydra augments shift crews by providing a fine grained look at the quality of data by looking at detector occupancy, which shows most of the underlying data quality problems. It integrates with the EPICS system, writing the state of the detector for archiving.

- **MCwrapper** (https://github.com/JeffersonLab/gluex_MCwrapper)- MCwrapper is the definitive frame work for Monte Carlo production at GlueX. It is usable independently and also as part of an automated system which take requests for simulation via a web app and automatically tests, submits, and monitors jobs across different platforms. It was designed and built, frontend and backend, by Thomas Britton and involves code written in python, cshell, bash, HTML, CSS, php, JavaScript and utilizes a MySQL database.

Synergistic Activities

- Co-founder of the A.I. Lunch series at Jefferson Lab.

- Organizer of the quarterly machine learning challenges, which challenge participants to tackle a problem with machine learning.

- Organizer of the 2020 A.I hackathon which preceded the 2020 A.I. Workshop at Jefferson Lab
- Developing Hydra, an A.I. data quality monitoring system deployed in Hall-D.
- Facilitating the use of machine learning techniques at Jefferson Lab by bench marking systems with Hydra and developing tools for wider machine learning adoption at the lab.

Potential Conflicts of Interest

- Artuso, Marina - Syracuse University, LHCb
- Austregesilo, Alexander - Jefferson Laboratory, GlueX
- Blusk, Steven - Syracuse University, LHCb
- Chudakov, Eugene - Jefferson Laboratory, GlueX
- Dalton, Mark - Jefferson Laboratory, GlueX
- Dobbs, Sean - Florida State University, GlueX
- Deur, Alexandre - Jefferson Laboratory, GlueX
- Furletov, Sergey - Jefferson Laboratory, GlueX
- Gui, Bin - The Ohio State University
- Ito, Mark - Jefferson Laboratory, GlueX
- Jarvis, Naomi - Carnegie Mellon University, GlueX
- Jurik, Nathan - University of Oxford, LHCb
- Lawrence, David - Jefferson Laboratory, GlueX
- Lersch, Daniel - Florida State University, GlueX
- Meyer, Curtis - Carnegie Mellon University, GlueX
- Mountain, Raymond - Syracuse University, LHCb
- Pentchev, Lubomir - Jefferson Laboratory, GlueX
- Phelps, William - Christopher NEwport University, GlueX/CLAS12
- Romanov, Dmitry - Jefferson Laboratory, GlueX/EIC
- Skwarnicki, Tomasz - Syracuse University, LHCb
- Smith, Elton - Jefferson Laboratory, GlueX
- Somov, Alexander - Jefferson Laboratory, GlueX
- Stevens, Justin - Jefferson Laboratory, GlueX
- Stone, Sheldon - Syracuse University, LHCb
- Taylor, Simon - Jefferson Laboratory, GlueX
- Wang, Jianchun - Institute of High Energy Physics
- Zihlmann, Benedikt - Jefferson Laboratory, GlueX

Naomi Jarvis, DPhil (CO-PI)

Naomi Jarvis DPhil CPhys MInstP
Research Scientist
Department of Physics, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213
e-mail: nsj@cmu.edu

Education and Training

- D.Phil. in Experimental Nuclear Structure Physics, *Breakup studies with 23Na*, University of York, UK, 1991
- B.Sc.(Hons) in Physics, University of Birmingham, UK, 1987

Research and Professional Experience

- **2020** Research Scientist
  Carnegie Mellon University, Pittsburgh, PA
- **2013-2020** Research Associate
  Carnegie Mellon University, Pittsburgh, PA
- **2009-2014** Postdoctoral Research Associate
  Carnegie Mellon University, Pittsburgh, PA
- **2001-2008** Contractor to Health Protection Agency, UK
- **2003-2005** Contractor to ACJ & Associates, Richland, WA
- **1992-2001** Senior Scientific Officer, Biokinetic Modelling Group, National Radiological Protection Board, UK

Research Activities:

- **2010-2020** Project Scientist for GlueX CDC construction, commissioning and operations
  Developed algorithms for readout firmware. Responsible for CDC calibrations, data corrections and monitoring guidance.
- **1992-2008** Internal dosimetry research, software development and guidance
  Developed, published and supported biokinetic modelling and internal dosimetry software. Participated in international intercomparison exercises in internal dosimetry.

Significant Publications

- Jarvis, N.S., Meyer, C.A., Zihlmann, B., Staib, M., Austregesilo, A., Barbosa, F., Dickover, C., Razmyslovich, V., Taylor, S., and Van Haarlem, Y., Visser, G. and Whitlatch, T. *The Central Drift Chamber for GlueX* NIM A **962** 163727 (2020)

- Lawrence, D., Dickover, C., Jarvis, N., Pentchev, L., Zihlmann, B. *FADC125 Operation and Data Format Requirement.* Technical Report GlueX-doc-2274, Jefferson Lab, June 2014.

- Jarvis, N.S., Watson, D.L., Gyapong, G.J., Jones, C.D., Bennett, S.J., Freer, M., Fulton, B.R., Karban, O., Murgatroyd, J.T., Tungate, G., Rae, W.D.M., and Smith, A.E. *Breakup studies with 23Na* Phys. Rev. C **51** 2606 (1995).

Synergistic Activities

- GlueX Collaboration, 2010-present

- CDC calibrations, monitoring software development and guidance.

Identification of Potential Conflicts of Interest or Bias in Selection of Reviewers

- Austregesilo, Alexander - Jefferson Laboratory, GlueX
- Britton, Thomas - Jefferson Laboratory, GlueX
- Cornejo, Juan Carlos - Carnegie Mellon University
- Dalton, Mark - Jefferson Laboratory, GlueX
- Dickover, Cody - Jefferson Laboratory, GlueX
- Dobbs, Sean - Florida State University, GlueX
- Lawrence, David - Jefferson Laboratory, GlueX
- Meyer, Curtis - Carnegie Mellon University, GlueX
- Pentchev, Lubomir - Jefferson Laboratory, GlueX
- Quinn, Brian - Carnegie Mellon University
- Schumacher, Reinhard - Carnegie Mellon University, GlueX
- Visser, Gerard - Indiana University
- Taylor, Simon - Jefferson Laboratory, GlueX
- Zihlmann, Benedikt - Jefferson Laboratory, GlueX

Appendix 2: Current and Pending Support

David Lawrence, Ph.D. (PI)

David Lawrence is a Staff Scientist in Jefferson Laboratory's Experimental Physics Software and Computing Infrastructure Group, under Scientific Computing, and is supported by DOE contract DE-AC05-06OR23177, under which Jefferson Science Associates, LLC, operates the Thomas Jefferson National Accelerator Facility. He is obligated to spend 20% of his time (2.4 person-months) on Jefferson Laboratory's LDRD 2013 project *"Development of Next Generation Parallel Event Processing Framework"* until it completes at the end of FY20. The award for this project in FY20 was $180k for work to be performed over a 1 year period starting Oct. 1st, 2019. The project involves writing JANA, a multi-threaded software framework for experimental particle physics data processing using C++. There is no direct overlap with the research for the current proposal. The JANA framework will be used for the reconstruction of experimental data for the GlueX experiment which will be fed into the AI models. The data reconstruction will be done independent of the current proposal and neither activity depends on the other.

Thomas Britton, Ph.D. (CO-PI)

Thomas Britton is a Staff Scientist in Jefferson Laboratory's Experimental Physics Software and Computing Infrastructure Group, under Scientific Computing, and is 100% supported by DOE contract DE-AC05-06OR23177, under which Jefferson Science Associates, LLC, operates the Thomas Jefferson National Accelerator Facility.

Naomi Jarvis, DPhil (CO-PI)

Naomi Jarvis is a Research Scientist in Carnegie Mellon University's Nuclear Physics group, which is supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics, DOE Grant No. DE-FG02-87ER40315.

Appendix 3: Bibliography and References Cited

References

[1] T.D. Beattie et al. "Construction and performance of the barrel electromagnetic calorimeter for the GlueX experiment". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 896 (2018), pp. 24–42. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2018.04.006. URL: http://www.sciencedirect.com/science/article/pii/S0168900218304807.

[2] Leilani H. Gilpin et al. *Explaining Explanations: An Overview of Interpretability of Machine Learning*. 2018. arXiv: 1806.00069 [cs.AI].

[3] Hamamatsu. *Photomultiplier Tubes: Basics and Applications*. 2007. URL: https://www.hamamatsu.com/resources/pdf/etd/PMT_handbook_v3aE.pdf.

[4] N.S. Jarvis et al. "The Central Drift Chamber for GlueX". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 962 (2020), p. 163727. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2020.163727. URL: http://www.sciencedirect.com/science/article/pii/S0168900220302771.

[5] V. Popov and H. Mkrtchyan. "New photomultiplier active base for Hall C jefferson lab lead tungstate calorimeter". In: *2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC)*. 2012, pp. 1177–1179.

[6] Alexander Scheinker et al. *Advanced Control Methods for Particle Accelerators (ACM4PA) 2019 Workshop Report*. 2020. arXiv: 2001.05461 [physics.acc-ph].

[7] Chip Watson. *Jefferson Lab's Data Management Plan*. 2014. URL: https://scicomp.jlab.org/DataManagementPlan.pdf.

Appendix 4: Facilities and Other Resources

The proposed research will take place primarily at Jefferson Laboratory with staff stationed at the laboratory and solely utilize the facilities of Jefferson Laboratory. The CO-PI Jarvis will work primarily at her home institution, Carnegie Mellon University with the exception of brief periods of travel to Jefferson Lab to collaborate on this project. All members of the research team currently have or will have access to a portion of the computational resources of the IT division of the laboratory. These resources include:

- 3 dedicated machine learning nodes each with 4 Titan RTX GPUs.
- The equivalent of 7000 cores of AMD Rome.
- Nearly 10PB of online disk storage.
- A tape library capable of over 100PB of storage.

The fraction of the listed resources needed for this project are modest (<1%) and will not disrupt current operations at the lab. For the needs of the proposed research the research team will have ready access to the dedicated machine learning nodes, ample CPU compute power, and sufficient storage (several TeraBytes at most).

Additionally, the research team will have access to the existing calibration and conditions databases of both Hall-B and Hall-D for the purposes of this project. Access to Hall-B and Hall-D where the detectors and control systems are housed will be available if needed. The systems can be controlled remotely so physical access will not be strictly required.

Appendix 5: Equipment

The equipment needs for this research are modest given that the primary focus is in the production of software. Office equipment and desktop/laptop computers are provided by each team members' home institution. The exception being personal computers for the new hires which will be purchased using funds from this proposal.

Appendix 6: Data Management Plan

The proposed research will take place and thus any data management plan will conform to that of Jefferson Lab [7].

Data Types and Sources

The data used by this research will come directly from, or be derived from, quantities stored in the EPICS archive system at Jefferson Lab and the respective calibrations and conditions databases of the CLAS12 and GlueX experiments. These include values read from sensors in and around the CLAS12 DC and GlueX CDC detectors and their respective electronics. These also include derived values representing calibration constants used in the processing of raw data. The trained models will nominally be stored in HDF5 (Hierarchical Data Format), along with any meta data required for loading or interacting with the model (e.g. the version numbers of any used packages). These records, the models and their metadata, will be permanently archived on magnetic tape at Jefferson Lab and a copy made accessible to the public upon request.

Content and Format

This research has no need for any personally identifiable information (PII), nor does it have the need for any sensitive information related to national security interests or any other interests outside of the advancement of the scientific objectives of the Department of Energy (DOE). Any data gathered or produced in the course of this research will thus be stored in an unencrypted, human readable (where appropriate), format suitable for the training and/or testing of the models produced in the course of this research.

Any codes developed during this research will be version controlled by and stored on github.com. Upon completion of the research the code will be made open source and released, with suitable documentation to the public. In this way future projects may benefit from the models and code developed in pursuit of the proposed research.

Data Sharing and Preservation

Believing that the sharing of data fosters innovation and advancement all data used in the course of this research will be made public. Produced codes will be preserved in an open github repository. Any models produced along with their metadata will be permanently archived on magnetic tape at Jefferson Lab. Items stored as "production" in the archive are automatically copied to a second backup tape that is stored in a protected vault ensuring preservation in the event of catastrophe.

Data in Publications

Datasets used in publications will have clearly identifiable tags that can be used to retrieve them from the JLab permanent tape archive. The archive is not directly publicly accessible, but is available to all JLab users. Datasets whose size is not too large that they can be reasonably posted to the web will be. Otherwise, the datasets will be made available upon request.

Data Storage Requirements

Permanent data storage will use the existing JLab tape archive. Much of the initial input data needed for this project will be mined from existing tape archives, MySQL databases, and EPICS archive servers. All of these have redundant backup systems already in place at Jefferson Lab. The mined data which will be modest in size compared to the large experimental datasets, will also be stored in the tape archive. In addition, trained models will be stored along side their meta data permanently on the Jefferson Lab tape archive.

5.1 Intellectual Property

Allocation of Intellectual Property rights will be in accordance with Jefferson Science Associates' (JSA's) Prime Contract with the U.S. Department of Energy or under a separate Cooperative Research and Development Agreement (CRADA) between the Participants.

Appendix 7: Letters of Commitment from co-PIs and Collaborators

Mac Mestayer

Senior Staff Physicist

Physics Division, Jlab

Funding Opportunity Committee

Dear Committee Members;

I am writing to support the proposal, "A.I. Assisted Experiment Control and Calibration", submitted to you by David Lawrence, Thomas Britton, Naomi Jarvis and David Richards. This is a well-founded and exciting proposal with the potential to improve the quality and consistency of our experimental data and speed up the publication of physics results.

I have designed, built, operated, monitored and calibrated the drift chambers used in Hall B and know that the calibration step is complicated and time-consuming. The process of turning raw data (TDC signals) into a reliable space point on a particle trajectory is complicated because it involves so many variables. We rely on the experience of subject matter experts to identify which variables are most important to monitor and include in our models. Still, there are unidentified factors which influence our calibrations in unknown ways. AI can help to find these hidden correlations. The result would be more reliable calibrations done in a shorter time.

Of course, any particular correlation (for example, chamber gain's dependence on atmospheric pressure) can be studied "by hand", but I can say from experience that there is always that one variable which was not identified as being important. I can also say that my colleagues who want to do high-level physics analysis with our data are always impatient about any delays in calibration.

Just as important as the ultimate quality of the calibrations is their consistency. A self-consistent calibration is essential for a reliable simulation of the data. Here especially, AI can assist us in finding the underlying causes of calibration changes. Once identified, we can deal with them. It's a truism in physics that the most interesting findings were unexpected, and I expect the same from the results of this pilot project. It is certainly worthwhile to support it.

Sincerely,

Mac Mestayer

Curtis A. Meyer
The Otto Stern Professor of Physics
Associate Dean for Research
Carnegie Mellon University
Pittsburgh, PA 15213-3890
(412) 268-2745
cmeyer@cmu.edu

April 22 , 2020

Dr David Lawrence
Staff Scientist III, EPSCI Group Lead
Thomas Jefferson National Accelerator Facility

Subject: Organizational Letter of Commitment, Funding Opportunity Number (FOA) LAB 20-2261

Dear Dr Lawrence,

This letter serves as an organizational letter of commitment for Carnegie Mellon University as a collaborator in the proposal to the U.S. Department of Energy (DOE), LAB 20-2261, entitled "A.I. Assisted Experiment Control and Calibration" being led by Thomas Jefferson National Accelerator Facility (Jefferson Lab.).

If this proposal is awarded, Carnegie Mellon University intends to collaborate as detailed in the proposal and will fully comply with the terms of the FOA and applicable DOE requirements. Key personnel participating from Carnegie Mellon University is Naomi Jarvis, as co-P.I. The participation of Carnegie Mellon University is free of organizational conflicts of interest.

Sincerely,

Curtis A. Meyer

Curtis A. Meyer
Associate Dean for Research

David Lawrence
Jefferson Lab
12000 Jefferson Avenue
Newport News, VA 23606

SUBJECT: Letter of Individual Commitment in Support of Jefferson Lab's Application to the FOA LAB 20-2261, 'A.I. Assisted Experiment Control and Calibration'

Dear Dr. Lawrence,

This letter confirms my commitment to fully support and participate in the execution of the Proposal 'A.I. Assisted Experiment Control and Calibration', FOA LAB 20-2261 if awarded funding by the Department of Energy.

I will be the Co-PI for the Carnegie Mellon contributions. My role will be to develop limits for the A.I., provide technical expertise in the use of the GlueX Central Drift Chamber and feedback on the performance of the A.I. My commitment will involve 10 percent of my time and effort. As one of the detector experts for this drift chamber, I am very motivated to improve its performance wherever possible. Calibrations are a time-consuming part of my work and the success of this project would make those substantially less onerous.

Sincerely,

Dr Naomi S Jarvis
Research Scientist
Department of Physics
Carnegie Mellon University
412 268 6949
nsj@cmu.edu