1 Introduction

Funds are requested from the National Science Foundation’s Major Research Instrumentation (MRI) program to develop a computing cluster to support the research programs at the University of Richmond in astrophysics and nuclear physics. The instrument will also be available to senior users at other institutions for work on closely related projects. The research groups at Richmond support 4-8 undergraduates during the summer and the academic year (Richmond is a primarily undergraduate institution). These students routinely go on to careers in science and engineering. These research programs have external support from the US Department of Energy (DoE) (Gilfoyle in nuclear physics) and the National Science Foundation (NSF) (Bunn in astrophysics).

The astrophysics research focuses on simulation of instruments to observe the polarization of the cosmic microwave background (CMB) radiation. CMB polarimetry is one of the highest priorities in the cosmological community, due in large part to the prospect of finding direct evidence for an inflationary epoch in the early Universe. Bunn is a member of the MBI/BRAIN collaboration, which is constructing a prototype bolometric interferometer for CMB polarimetry. The proposed cluster will be used for simulation of this instrument and future extensions of it, with the goal of assessing the suitability of this technology for an eventual space-based CMB polarimeter.

The nuclear physics research is centered on unraveling the structure of the nucleon and the nature of quark confinement and is supported by DoE grant DE-FG02-96ER40980. Additional senior users in nuclear physics will work on the instrument. They are members of the CLAS Collaboration (with Gilfoyle) that is responsible for the operation of a large particle detector (CLAS) at the Thomas Jefferson National Accelerator Facility (JLab) in Newport News, VA.

We have considerable experience with high-performance computing. One of us (Gilfoyle) was co-PI on a project that developed a computing cluster at Richmond in 2001 with support from the NSF. He has been the manager of that project since then. That existing system is at the end of its useful life, but the infrastructure to support it is still sound, and the proposed system will benefit from that investment. The University is committed to software and hardware support of the proposed cluster after the lifetime of the grant, as described in Section 4.4 and the Dean’s letter of support in Appendix C.

2 Astrophysics

The proposed astrophysics research concerns analysis of observations of the cosmic microwave background (CMB) radiation. For over ten years now, CMB observations have been among the main contributors to the extraordinary advances in precision cosmology [1]. Polarimetry is the next frontier in CMB research, largely because polarization maps may provide a direct probe of an inflationary epoch in the extremely early Universe [2, 3]. Because the signal from inflation is predicted to be found on large angular scales, an all-sky polarization map will be necessary to detect it. A satellite-borne telescope is likely the only way to achieve such a data set. For this reason, such an instrument, generally dubbed CMBPol, is an extremely high priority in the cosmology community. A joint DoE/NASA/NSF Task Force recently advocated “a phased program to measure the large-scale CMB polarization signal expected from inflation” as its highest priority [4]. Similar opinions are expressed in the National Research Council’s decadal survey of astronomy and astrophysics [5] and their report Connecting Quarks with the Cosmos [6].
NASA’s *Beyond Einstein* road map for future astrophysics programs includes a satellite-borne CMB polarimeter known as the Einstein Inflation Probe (EIP) [7]. Bunn was one of the leaders, along with Peter Timbie at Wisconsin and Gregory Tucker at Brown, of a NASA-funded Mission Concept Study for the EIP, the Einstein Polarization Interferometer for Cosmology (EPIC) [8]. The members of the EPIC collaboration are currently constructing and deploying a ground-based prototype millimeter-wave bolometric interferometer (MBI) [9, 10]. The MBI collaboration recently joined forces with the BRAIN group, a European collaboration centered at the Université de Paris (Diderot), which has been working on a similar instrument [11, 12]. The MBI/BRAIN collaboration plans to build, deploy, and test a larger interferometer based on the MBI design.

Development of MBI is currently funded by NASA. Two NSF grants for further development are pending, and further funding will be sought from NASA. The BRAIN collaboration has ongoing funding from European funding agencies (CNRS and ESA), ensuring that development of MBI/BRAIN will continue in the coming years. If US funding is unavailable, the center of gravity of the experiment will shift to Europe, but in any case Bunn’s research group will continue to be heavily involved. In particular, Bunn has arranged to spend a substantial portion of his sabbatical in 2009-2010 in Paris working on MBI/BRAIN.

The work proposed herein is to develop data analysis and simulation tools in support of MBI/BRAIN. Although we will focus primarily on this specific instrument design, we expect our results to be applicable to CMB interferometers more generally. The research groups of Timbie, Tucker, and Jean-Christophe Hamilton (Paris) will be heavily involved in this research. This work is particularly timely. A variety of different designs are under consideration for next-generation CMB polarimeters, and effects of systematic errors will be of particular importance in discriminating among these possibilities as described below. Far less is currently known about propagation of systematic errors in interferometers than in single-dish imaging telescopes.

### 2.1 CMB Polarization and Inflation

Any CMB polarization map can be divided into two components, a scalar $E$ component and a pseudoscalar $B$ component [2, 3], which probe different physical phenomena. Because it is insensitive to ordinary density perturbations (to linear order), the $B$ component is predicted to be smaller than $E$ by an order of magnitude or more over all angular scales. Detection of this weak $B$ component is extremely important, because it provides a clean probe of other perturbations, especially the tensor perturbations produced by primordial gravitational waves. These tensor modes are predicted in inflationary models. If they are detected, we will have found direct evidence for inflation and will even be able to measure the energy scale and time of the inflationary epoch (perhaps as early as $\sim 10^{-33}$ s after the Big Bang). Aside from this tensor component, the dominant source of $B$ polarization is expected to be gravitational lensing of $E$ modes by large-scale structure [3], which will provide a valuable window on structure formation.

Figure 1 shows the magnitudes of the predicted $E$ and $B$ polarization signals in comparison with the temperature anisotropy. The lensing and tensor contributions to the $B$ signal are shown separately. Spectra were calculated using the CMBFAST software, with tensor-to-scalar ratio $T/S = 0.1$ and other parameters taken from the best-fit *WMAP* model [13].

All CMB polarization detections to date have been of the $E$ component; the chief goal of a future CMBPol satellite, as well as upcoming suborbital missions such as MBI/BRAIN, will be to detect the $B$ component. Although the $B$-type polarization signal is extremely weak (of order nanokelvins), the next generation of experiments will have the raw sensitivity to detect it, by combining $\sim 10^3$ low-noise detectors with long integration times. The greater challenges are removal of non-cosmological foreground signals and control of systematic errors. Bolometric interferometers raise very different systematic error issues from traditional imaging telescopes, but
the state of the art in understanding these errors is considerably less advanced for interferometers than for imagers. Bunn has developed a framework for understanding these errors [14], but simulations are necessary to characterize the errors in sufficient detail to assess the suitability of this technology for CMBPol. The primary science goal of our proposed astrophysics research is to perform such simulations of interferometric CMB polarization observations.

2.2 Interferometric CMB polarimetry

In the past, both direct imaging telescopes and interferometers have been successful in CMB observations, and both technologies are candidates for future B-mode experiments such as CMBPol. The goal of MBI/BRAIN is to combine the advantages of interferometry with bolometric detectors, the lowest-noise detector technology at millimeter wavelengths.

There are several advantages to this approach. First, interferometers have reduced sensitivity to a variety of systematic errors [8, 10, 14]. MBI/BRAIN will have simple, reflection-free optics, easily calculable, symmetric beam patterns with extremely low sidelobes, and no mechanical chopping of the telescope. Interferometers can measure the linear polarization Stokes parameters $Q$ and $U$ directly, without differencing signals from different detectors, mitigating leakage from the temperature anisotropy into the polarization channels which are 2-3 orders of magnitude weaker.

In addition, an interferometer can achieve higher resolution than is practical with a single dish. This is important for CMBPol, as imperfections in the beam shape and pointing couple the CMB temperature anisotropy into the polarization signals [15]. These effects are mitigated with small beams, as the temperature is smooth on small scales. Moreover, separating the tensor and lensing contributions requires high-resolution maps [16].

In any incomplete sky map, there is “leakage” between the components [17–19], but interferometry mitigates this leakage [20]. To see why, note that the $E$-$B$ separation can be done trivially mode by mode in Fourier space. With incomplete sky coverage, individual Fourier modes cannot be measured. Because interferometers measure visibilities, which have narrow window functions in Fourier space, they provide cleaner separation of $E$ and $B$ modes.
2.3 Systematic error simulations

It is clearly essential to have a detailed, quantitative understanding of the effects of systematic errors and foreground contamination on data from both interferometric and imaging systems. The state of the art is far more developed for imaging systems than for interferometers. Our goal is to close that gap, so that the two technologies can be compared on an equal footing. We have established a theoretical framework for analysis of systematic errors in CMB interferometric polarimetry [14], but this work needs to be supplemented with detailed simulations. We propose to perform such simulations of CMB interferometric polarimetry in order to assess the effects of various systematic errors on MBI/BRAIN in particular and on CMB interferometers in general.

Along with the construction of the prototype interferometer, the MBI group has been developing a data analysis pipeline. The BRAIN group, working independently, has made significant progress in developing a suite of simulation software. A high priority for the newly-formed MBI/BRAIN collaboration is to merge the code from these two efforts and extend it to a single simulation pipeline. Bunn, who has led the MBI effort, will be on sabbatical for the 2009-2010 academic year and plans to spend the bulk of this time working closely with the BRAIN group on this effort. Our overall goal is to simulate the propagation of a known signal through the instrument and then analyze it in the same manner as we will with the real data. We will determine in precise detail the error properties of both the recovered Fourier-space power spectrum and the recovered image.

We now outline the key steps in the simulation of MBI/BRAIN data. We will assess the computational requirements in section 4.2.

1. Simulation of time-ordered data (TOD). Given an underlying “true” sky map, a model of the instrument as well as its attitude as a function of time, and a noise model, we need to compute the simulated output time streams from each of the detectors. This step is not computationally intensive. The BRAIN collaboration has a well-developed code for this task, although adaptations are needed to include effects of some systematic errors (e.g., mismatched beams).

2. TOD → Visibility-space “map.” The raw data from an interferometer is a set of visibilities, which are essentially samples of the Fourier transform of the map, convolved with the primary beam. Because the data are contaminated by correlated noise, the optimal recovery of a Fourier-domain visibility map from the TOD is nontrivial; however, efficient parallelized algorithms such as MADMap [22], first developed for traditional imaging systems, can be adapted for this purpose.

3. Power Spectrum Estimation. We wish to determine the maximum-likelihood power spectrum for a given visibility data vector. Once again, standard codes for imaging systems, which have been parallelized and made publicly available, can be adapted to apply to visibility data.

4. Visibility data → Image. The primary science goal of a CMB experiment is the power spectrum, which can be computed entirely in the visibility domain, without ever constructing a real-space image of the observed map. However, in order to check for errors or foreground contamination in the data, we will surely want to produce actual images from the visibility data. In addition, some CMB studies search for signals beyond merely the power spectrum and so require real images. Traditional radio astronomy techniques such as the CLEAN algorithm [23] are well-suited for data sets with sharp features, not the diffuse nearly-Gaussian structure found in CMB maps. We plan instead to use maximum-entropy reconstruction, which has been well-
developed in the CMB context [24]. Because maximum-entropy is a nonlinear method, the noise properties of the resulting maps can be computed only via simulation.

5. Component separation. The simulations described above will be our primary initial focus. Bunn will devote much of his effort during 2009-2010 to leading the development and adaptation of code described above, particularly steps 2-4. Over a longer time scale, we plan to develop code to test other aspects of the MBI data analysis and to address other problems in CMB data analysis. As MBI/BRAIN attempts to characterize $B$ polarization, the issue of component separation (i.e., removal of foregrounds) will be crucial. Both blind techniques (e.g., independent component analysis) and those based on fitting to foreground templates have been proposed for CMB component separation, but few have been adapted to the case of interferometric data. An extremely interesting question is whether these techniques are best applied in visibility space or in a real-space image produced by, e.g., maximum-entropy reconstruction. We plan to develop algorithms to address these questions. Because this work will require development of code from scratch (as opposed to adapting existing code), we anticipate seeking funding for a full-time postdoctoral researcher to work on this project.

2.4 Additional Users in Astrophysics

Bunn will lead the computational astrophysics research. The key ingredients for the simulations have been developed in Timbie, Tucker, and Hamilton’s groups. Further code development will be performed by Bunn and Hamilton’s group in Paris. (As noted above, Bunn will spend several months in Paris during his upcoming sabbatical.) The integration and testing of the code at the University of Richmond will be performed largely by Bunn and undergraduate research assistants, with support from members of the other groups. Both Bunn and Gilfoyle have extensive experience in involving undergraduates in computational physics research.

We hope eventually to hire a postdoc to work full-time on the component separation problem. If we are able to do this, the postdoc would probably be based at Richmond and work closely with undergraduates but would also spend significant time at the other collaborators’ home institutions (Brown, Wisconsin, and/or Paris).

3 Nuclear Physics

The research effort in nuclear physics is part of the program at JLab. The primary mission of JLab is to reveal the quark and gluon structure of nucleons and nuclei and to deepen our understanding of matter and quark confinement. Quantum Chromodynamics (QCD) is a highly successful description of quarks at high 4-momentum transfers or $Q^2$ [25], but at energies where the nucleons exist (the non-perturbative region), it is a daunting challenge to solve [26]. At low $Q^2 < 0.5 \text{ (GeV/c)}^2$ the “hadronic” picture of nuclei (i.e., nuclei made of protons and neutrons) has been successful [27]. However, the transition region between these extremes is poorly understood, and mapping the geography of this transition is an essential goal of nuclear physics as described in the Long-Range Plan of the Nuclear Science Advisory Committee [28].

The central instrument at JLab is a superconducting electron accelerator with a maximum energy of 4-6 GeV, a 100% duty cycle, and a maximum current of 200 μA. We work in Hall B with the CEBAF Large Acceptance Spectrometer (CLAS). This device is a large (45-ton), toroidal multi-gap magnetic spectrometer with nearly full solid angle coverage. The particle detection system consists of drift chambers [29], Cerenkov detectors [30], scintillation counters [31] for time-of-flight measurements, and electromagnetic calorimeters [32]. There are about 33,000 detecting elements capable of acquiring about 1 terabyte of data per day. The Richmond group has been part of the CLAS Collaboration (which built and operates the detector) since its inception.
The analysis of these large data sets requires significant computing resources. First pass analysis is done on the JLab computing farm, but final results require additional analysis. Demand is high for the computing resources at JLab, and it can routinely take a day for a submitted batch job to even start. We have developed our own local computing cluster so that we can analyze our data in a timely fashion (especially for our undergraduates). We also simulate the response of CLAS to separate real physics effects from artifacts of the detector. This stage requires large disk space to store the Monte Carlo events and, more importantly, considerable computing power. Our simulation generates Monte Carlo events at only about 1-3 events per second.

In 2008, the US Department of Energy approved funding for the start of construction for the 12-GeV Upgrade at JLab. The energy of the electron beam will be doubled, opening up new physics opportunities for studying QCD, nucleon structure, and the transition for hadronic to quark degrees of freedom. Part of the Upgrade includes replacing the existing CLAS detector in Hall B with a new device (called CLAS12) that will be able to take advantage of these new physics opportunities. We are part of that effort and are focused on simulation of CLAS12 and the design of the detector and planned experiments. Gilfoyle is the spokesperson and contact person for an experiment (E12-07-104) approved by the JLab Program Advisory Committee (PAC) and slated to run in the first five years after the Upgrade is complete. The JLab, 12-GeV-Upgrade is one of the highest priorities of the DoE Office of Science in the next 20 years [33]. Gilfoyle will be on sabbatical during the 2009-2010 academic year. He will be focused on the research projects described here and will make intensive use of the proposed instrument.

3.1 Magnetic Form Factor of the Neutron

The elastic electromagnetic form factors are the most basic observables that describe the internal structure of the proton and neutron. The differential cross section for elastic electron-nucleon scattering can then be calculated in the laboratory frame in terms of four elastic form factors (electric and magnetic ones for each nucleon) that characterize the distributions of charge and magnetization within the proton and neutron [34]. The form factors are stringent tests of non-perturbative QCD including calculations on the lattice and are connected to generalized parton distributions that enable us to perform nucleon tomography [35–38].

We are part of a broad assault on the four elastic nucleon form factors at JLab [40]. Our focus is on $G_M^n$, the magnetic form factor of the neutron. To measure $G_M^n$ we use the ratio of elastic $e - n$ to elastic $e - p$ scattering on deuterium, which is less vulnerable to uncertainties than previous methods [39, 41]. We have completed data collection and much of the analysis for a measurement of $G_M^n$ in the range $Q^2 = 1.0 - 4.8 \ (\text{GeV}/c)^2$ [39, 41–44]. An internal report describing the analysis of two out of the three sets of running conditions has been approved by the Collaboration and a paper submitted for publication [39]. Results can be seen in Figure 2 along with some of the world data for $G_M^n$. Our group at Richmond has taken on the analysis of the third data set which will make extensive use of the cluster proposed here.

We have also submitted a proposal to the JLab Program Advisory Committee (PAC) to make the same measurements at higher $Q^2$ as part of the JLab 12-GeV Upgrade. The proposal was approved by PAC32 in August 2007. We had the primary responsibility for developing this proposal. The committee report [37] summarized it in the following way:

Proposal PR12-07-104 is a measurement of the neutron magnetic form-factor $G_M^n$ in Hall B using a deuterium target. The method proposed is elegant and its physics essential to the program. The results of this experiment, if successful, will provide neutron data, which when combined with proton results determine the isovector form-factor, that is more readily computable on the lattice, having no disconnected quark contributions. This essential measurement will thus have the added benefit of provid-
This planned measurement will significantly expand the upper limit of this measurement (from $Q^2 = 4.8 \text{ (GeV/c)}^2$ to $13.5 \text{ (GeV/c)}^2$), provide important constraints on generalized parton distributions, and test the validity of lattice QCD calculations. We continue to study simulations of this experiment to support the design and construction of the new, CLAS12 detector in Hall B [45, 46].

### 3.2 Out-of-Plane Structure Functions of the Deuteron

The hadronic model of nuclear physics has been successful at low $Q^2$, but it is not well-developed in the GeV region where there are few measurements [47, 48]. We need a baseline for the hadronic model so that deviations at higher $Q^2$ can be attributed to quark-gluon effects with greater confidence [28, 49]. To this end, we are investigating the out-of-plane structure functions of the deuteron, the simplest nucleus, using the reaction $D(\vec{e}, e'p)n$ with CLAS. The cross section is

$$\frac{d^5\sigma}{d\omega d\Omega d\Omega_{pq}} = C (\rho_l f_l + \rho_t f_t + \rho_{TT} f_{TT} \cos \phi_{pq} + \rho_{LT} f_{LT} \cos 2\phi_{pq} + h\rho'_{LT} f'_{LT} \sin \phi_{pq})$$  \hspace{1cm} (1)$$

where $C$ and the $\rho_i$ are functions of the known electron parameters, $h$ is the beam helicity, and $\phi_{pq}$ is the azimuthal angle between the scattering plane (defined by the incoming and outgoing electron 3-momenta) and the reaction plane (defined by the 3-momentum transfer and the ejected proton momentum) [50, 51]. The unique, nearly-4$\pi$ solid angle of CLAS creates an opportunity to extract the $\phi_{pq}$-dependent structure functions $f'_{LT}$, $f_{LT}$, and $f_{TT}$ in a model-independent way.

These structure functions are extracted using asymmetries that reduce our sensitivity to experimental effects. For example, the asymmetry $A'_{LT} = \rho'_{LT} f'_{LT} / (\rho_L f_L + \rho_T f_T)$ can be extracted using the $\sin \phi_{pq}$-weighted moments of the angular distributions measured with CLAS over the range $Q^2 = 0.2 - 2.5 \text{ (GeV/c)}^2$ [50, 51]. We are studying the reaction in quasi-elastic kinematics first and later will investigate higher energy transfers. Our preliminary results for $A'_{LT}$ show significant structure which is reproduced by a calculation from Jeschonnek and van Orden and
disagrees with others [27, 52, 53]. The new calculation from Jeschonnek and Van Orden (JVO) is a fully relativistic calculation in the impulse approximation using the Gross equation for the deuteron ground state and the SAID parameterization of the $NN$ scattering amplitude for FSI [54]. It shows that the fifth structure function is a sensitive probe of the spin-flip scattering amplitude. The double-spin components have little effect, implying that the spin-orbit part of the interaction is the primary contributor. The extraction and analysis of the other two structure functions ($f_{LT}$ and $f_{TT}$) and investigations of different kinematic regimes are ongoing. This work is part of a CLAS Approved Analysis (spokesperson: Gilfoyle). The analysis of these large data and the complex simulations of CLAS are computationally intensive and would use the proposed cluster. We note that these data sets are the same ones used in the $G^n_M$ analysis described in Section 3.1.

3.3 Quark Propagation and Hadron Formation

The confinement of quarks inside hadrons is perhaps the most remarkable feature of QCD, and solving its mysteries is an essential goal of nuclear physics [28]. We have proposed a broad program of measurements to determine the mechanisms of confinement in forming systems. We use the nucleus as a “detector”. Measuring the ratio of hadrons produced on nuclear targets relative to the production from deuterium will enable us to extract the lifetime of a deconfined quark after it is struck by a virtual photon and is moving through the nucleus. The kinematic dependence of the transverse momentum broadening will enable us to measure the time interval required to form the hadronic color field around the struck quark. A proposal to do this experiment at high $Q^2$ as part of the 12-GeV Upgrade was approved by the JLab PAC in August, 2006 [55, 56]. Gilfoyle is a co-spokesperson on that proposal and will be responsible for analyzing the $\pi^0$, $\eta$, and $\eta'$ production. We will use the proposed instrument to simulate the physics and the upgraded CLAS detector response to prepare for the 12-GeV Upgrade.

3.4 CLAS12 Software Development

We are committed to development projects for the JLab 12-GeV Upgrade to double the beam energy of the electron accelerator and enhance the experimental equipment in Hall B [57]. We will be responsible for design, prototyping, development, and testing of software for event simulation and reconstruction. The improved CLAS detector (CLAS12) will have prodigious software requirements, and simulation is an essential aspect of the design and the eventual precision of CLAS12. For many experiments, the quality of the results will be limited by systematic uncertainties instead of statistical ones, so accurate, precise calculations of the CLAS12 acceptance and response are essential. We anticipate needing about four times as much Monte Carlo data as CLAS12 collects. The CLAS12 simulation will produce data more slowly than the detector itself by about a factor of $10^3$ ($\approx 10$ Hz for the simulation versus $\approx 10$ kHz in CLAS12).

The motivation for our group is to support our experiments that are part of the 12-GeV Upgrade in Hall B. Experiment E12-07-104 will measure the neutron magnetic form factor $G^n_M$ out to $Q^2 = 14$ (GeV/c)$^2$. The neutron measurement will be done with both the electromagnetic calorimeters and the TOF system providing an important consistency check as in our previous measurement [39]. Over most of the $Q^2$ range we will have excellent statistical precision, so that understanding the CLAS12 response to neutrons is important for extracting $G^n_M$ with the anticipated systematic uncertainty. Experiment E12-06-117 will focus on the physics of quarks moving through nuclear matter and how they evolve to fully-formed hadrons. Our responsibilities

\[ A \text{ CLAS Collaboration member can write a proposal to analyze existing data, which is reviewed by a committee of Collaboration members and defended before the Collaboration who then vote to approve it.} \]
are to study the electroproduction of $\pi^0$, $\eta$, and $\eta'$ from nuclear targets. The detection of each particle relies on resolving photons from its decay: $\pi^0 \rightarrow 2\gamma$, $\eta \rightarrow 2\gamma$, and $\eta' \rightarrow \pi^+\pi^-\eta$ where the $\eta$ in the $\eta'$ decay will also be detected via its $2\gamma$ decay. Detection of neutrons and photons will done in the existing EC (reused in CLAS12) augmented by a new pre-shower calorimeter (PCAL) located in front. The PCAL will have higher segmentation than the EC to insure adequate spatial resolution to separate the two photons from the $\pi^0$ and $\eta$ decays up to maximum momenta of 9 GeV. CLAS12 will be able to detect all charged and neutral particles emitted in the polar angular range of 5 to 40 degrees.

The CLAS12 simulation package called gemc (for Geant4 Monte Carlo) is a Geant4-based simulation package with the following features: C++ language, object-oriented architecture, GUI interface, mysql database used for geometry, hits, magnetic field, materials, and physics output [58–60]. The TOF system has been implemented in the code, but only limited studies of its performance have been done. The EC and PCAL code has not been written. For neutron simulation one can choose a variety of physics algorithms to describe the process, but none have been tested with the CLAS12 geometry. From our experience in CLAS we know there are differences between the neutron detection efficiency measured in CLAS [39] and the same quantity derived from the current Geant3-based CLAS simulation called GSIM [61]. We are now investigating those differences in our analysis of the low $Q^2$ $G_n^M$ data. Much work remains to be done. This part of the project will make extensive use of the proposed cluster.

### 3.5 Additional Users in Nuclear Physics

Faculty from institutions besides the University of Richmond are part of the nuclear physics portion of this project. All are members of the CLAS Collaboration with Gilfoyle and have been users of the existing cluster. Their participation here will make better use of the proposed instrument. The number of students, undergraduate and graduate, that will learn high-performance computing will increase. Dr. K.H. Hicks is a professor at Ohio University in Athens, OH. He typically has a postdoctoral fellow and 1-3 doctoral students in his group. He is co-spokesperson on three approved proposals at JLab [56, 62, 63] and two CLAS Approved Analyses [64, 65]. See Appendix B for letters of support.

### 4 Research Instrumentation Needs

We request in this proposal funds for the purchase of a cluster of 15, dual-quad-core CPU computers supported by 6 terabytes of disk storage and associated hardware and software to increase the productivity of our research efforts at the University of Richmond and to train our undergraduates in high-performance computing. We describe the current resources available to our group and our computational needs and present a detailed rationale for the proposed system.

The proposed instrument will be used by the faculty and students in the University of Richmond astrophysics and nuclear physics groups. Undergraduates will be intimately involved in using the proposed system. The University of Richmond’s primary mission is undergraduate education (there are no graduate students in Physics), and the University is committed to research involving undergraduates to broaden our students’ education beyond the scope of the classroom. In recent years we have routinely had 4-8 students each summer involved in nuclear physics and astrophysics. The proposed instrument will make their research (and the faculty’s) more productive and train them on the most modern physics instruments.
4.1 Current Computing Facilities

The current facilities in physics include a computing cluster developed in 2001 with NSF and University funds plus an array of workstations. The original system consisted of 49 remote machines (1.4 GHz) running the Linux operating system, a master node, 3 TByte of RAID storage, and 18 GByte of disk space on each node. It resides in a laboratory with a 50-ton, 60,000-BTU air conditioner, an upgraded electrical panel, and a connection to the building’s backup power. Nearby rooms provide space for workstations and our students. The system is now at the end of its useful life: only about 24 nodes and 2 out of 3 RAIDs still work. This is not surprising since the system is over seven years old. At JLab, for example, replacement of failing nodes begins after two years and all machines are replaced after four years of use. The University employs a Linux support expert who is responsible for software updates and administration.

4.2 Astrophysics Computing Needs

We now assess the computational requirements to perform the proposed astrophysics simulations. The details of the next-generation MBI/BRAIN instrument are still uncertain, but for the estimates below, we will assume a month-long flight of a balloon-borne 16-element interferometer, observing 1000 square degrees of sky with a 5° field of view in each pointing. With these parameters, the total number of samples in the time-ordered data (TOD) is \( N_{\text{tod}} \approx 3 \times 10^8 \). The number of independent visibilities, which is also the approximate number of independent pixels in the final map, is \( N_v \approx N_{\text{pix}} \approx 1 \times 10^4 \).

The memory requirements for the astrophysics simulations will easily be satisfied by the proposed cluster. The time-ordered data is always computed and read in relatively small chunks. The largest matrices are \( N_v \times N_v \), which can be stored in the 4 GByte of RAM in a single node of our proposed equipment; however, even these matrices are all sparse and almost never need to be stored in memory at a single node. We therefore focus on time requirements, not memory in this section. Memory and disk needs will be addressed in detail in the following section.

The time required for the simulations described in Section 2.3 is dominated by two steps: power spectrum estimation and maximum-entropy image reconstruction.

**Power spectrum estimation.** Naive implementations of this step would involve maximizing the likelihood over a multidimensional parameter space, with each likelihood calculation scaling as \( O(N_v^3) \). Fortunately, we can take advantage of the sparseness of the visibility-space covariance matrix [66] and use Monte Carlo Markov chains (MCMC) to replace a search of the entire likelihood parameter space [67]. Scaling the results of Ref. [66] to our parameters, we estimate that we can evaluate a single likelihood in about 1 second. A typical MCMC, requiring \( \sim 10^3 \) likelihoods, will therefore take roughly 30 min.

**Maximum Entropy Image Reconstruction.** The scaling properties of CMB maximum-entropy algorithms to the large data sets considered here are not well-known. As it happens, though, the limiting step in the algorithm (calculation of \( \chi^2 \)) involves manipulating a large sparse matrix quite similar to that required for a single likelihood evaluation of the power spectrum. We expect that the number of entropy evaluations to find the maximum is of order \( 10^4 \), meaning that a maximum-entropy map should take of order 1 hour.

The above estimates are for a full-scale production run of the simulations. In earlier testing stages, we can speed up the process in various ways (forgoing maximum-entropy reconstruction, working in a restricted likelihood parameter space, and reducing \( N_{\text{tod}} \)). With this streamlining, we can perform \( \sim 10^3 \) simulations (enough to draw reasonable statistical conclusions) in a few days. This is a reasonable turnaround time for testing purposes. For full-scale production runs, we estimate that we can perform \( \sim 10^2 \) simulations per week on the proposed cluster.
4.3 Nuclear Physics Computing Needs

The nuclear physics projects described in this proposal all have considerable computing demands. These demands involve the simulation of the CLAS detector to generate publication-quality acceptance functions and adequate disk space and CPU power to perform “second-pass” analysis of the data. To estimate the CPU demands for simulating CLAS, consider the recent experience with the analysis of deuteron structure functions described in Section 3.2. One simulation required 40 million Monte Carlo events for a single beam energy and toroidal magnetic field setting of CLAS. The typical event simulation rate in the CLAS simulation software is about 1-2 Hz on each remote node. With our remaining 24 nodes the current cluster will take about 20 cluster-days\(^3\) to complete this simulation; the calendar time can be longer because of competition from other users. The proposed cluster will reduce that time down to about one day (see Section 4.4). The JLab facilities are heavily subscribed, and our existing cluster is aging and falling short. The cluster proposed here will reduce the demand on the JLab cluster, speed the calculation of the CLAS acceptance, and complete the analysis of the CLAS data.

We have learned several lessons from our previous experience. A major bottleneck in our data analysis is the speed of the switch (about 100 Mbps). Second-pass analysis requires us to copy the data from the RAID out to the remote nodes. We have to slow the analysis rate so the switch can keep up; making it difficult for multiple users to take full advantage of the system. The switch in the proposed system will be at least 10 times faster. See Section 4.4.

The disk needs are large. We currently use 1.5 TBytes out of the remaining 2.0 TBytes available for nuclear physics. Adding the astrophysics users will only increase the demand for storage. On the remote nodes it is more efficient to temporarily store the data on those nodes if the analysis requires repeated runs through the same data set. We save the time to copy data from the RAID onto the individual nodes. This requires adequate storage on the remote nodes. We thus need more and faster remote nodes, a faster switch, and adequate storage.

4.4 Proposed System

We now describe the proposed system. The components are listed in Table 1. A detailed quote for items 1-5 from the vendor LinuxLabs is in Appendix A. Below we discuss our reasoning behind the choice of the different components. The University is committed to providing $10,000 per year for four years after the lifetime of the proposed grant to support the instrument. See Appendix C for a letter of support from the Richmond Dean.

<table>
<thead>
<tr>
<th>Item</th>
<th>Number</th>
<th>Description</th>
<th>Price($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Dual, Quad-Core Xeon master node, 4 GByte RAM, 6 TByte RAID</td>
<td>11,400</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>Remote nodes, Dual, Quad-Core Xeon, 4 GByte RAM, 150 GByte storage</td>
<td>117,000</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>HP Procurve switch</td>
<td>4,460</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>cabinets</td>
<td>3,300</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>Nimbus OS license, installation, warranty, and shipping</td>
<td>24,652</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>Hardware items that cost less than $500</td>
<td>1,100</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Total Cost</strong></td>
<td><strong>161,912</strong></td>
</tr>
</tbody>
</table>

Table 1: Proposed computer cluster description and cost (see Appendix A for more details).

---

\(^3\)A cluster-day is 24 hours of time on the existing cluster with no competing calculations being performed.
The dual, quad-core Xeon processors (item 1) were chosen because of their excellent cost-to-benefit ratio. Their clock speed is about 50% faster than the speed of the current remote nodes, and architectural improvements make them overall about 10 times faster than the current ones (6.8 gigaflops versus 0.7 gigaflops). Test results from the vendor (LinuxLabs) using LinPack (a benchmarking standard for solving large, square matrices) give an equivalent speed of 1.0 teraflops for the entire instrument. This is about 13 times the speed of the original Richmond cluster and 27 times the speed of the existing, depleted cluster (see Section 4.1). The Linux operating system will be used. It is a research-quality operating system commonplace in physics. The number of machines was chosen to reduce the time to for simulating the CLAS response to a reasonable value. To generate and analyze 40 million events, we estimate about 1.5 day compared with the time required for the existing, 24-node, cluster (about 20 days) for such a calculation. Calculations of this length are routine for the analysis of the CLAS data. The astrophysics projects require hundreds of power spectrum estimations and maximum entry reconstructions which each take about a CPU-hour on the current nodes. The memory (4 GByte for the 8 cores on each node) is needed because the reconstruction and simulation packages for the nuclear physics work use large amounts of memory and the astrophysics simulations work with large matrices. A 150-GByte hard drive (item 2) will be attached to each machine to provide storage. This space is needed to store data files for analysis and the results of the astrophysical analysis. The fast ethernet switch (item 3) is needed to speed data transfer over the network (see Section 4.3). Cabinets will hold the nodes (item 4). Hardware and software installation is required (item 5). The software for managing the cluster and submitting batch jobs is Nimbus Beowulf from Linux Labs in Atlanta, GA. This is the vendor who built the current cluster, and we have had a long, fruitful relationship with them. A variety of other components each costing less than $500 (cables and tools) are included in item 6. We expect the system to have a 4-6 year lifetime. Our experience at Richmond and at JLab suggests that remote nodes will gradually fail over time and that 4-6 years is the optimum lifetime (see Section 4.1).

5 Impact of Project on Teaching and Research

This project will have a significant impact on the development of our students at Richmond and the institutions of the other senior users. We describe here the environment at Richmond, how the instrument will be used to train our students, and the impact at those other institutions.

The University of Richmond is a private, highly-selective, primarily-undergraduate, liberal arts institution in Richmond, Virginia with about 3000 undergraduates. A $36M expansion and renovation of the Gottwald Science Center was completed in spring 2006. All of the teaching and research spaces in Physics were renovated, and two new faculty positions were added in Physics (one instructor position and one tenure line). The Department of Physics consists of seven teaching faculty and graduates about 6-7 physics majors each year. The faculty are active in experimental nuclear physics, astrophysics, experimental and theoretical nuclear structure physics, surface and nano-physics, biological physics, and homeland security. All tenured and tenure-track physics faculty have external grant support from DoE (two grants), NSF (three grants), the Research Corporation, and the American Chemical Society. We all emphasize undergraduate involvement in research from early in the students’ careers. Students involved in undergraduate research are more likely to attend graduate school and be successful after college [68].

We have been successful at starting our undergraduates on their research careers. Over the last three summers about 25 of our physics majors participated in research at Richmond, JLab, Yale University, Berkeley, and the University of Notre Dame. Our students have accomplished much in their research careers. Many in physics have presented their work at national meetings and some have even obtained travel support to present their work conferences of the American Physical
Society [69–77] and American Astronomical Society [78–81]. We have also been successful in attracting under-represented groups to work in our nuclear physics group. Of the ten students who worked in our nuclear physics laboratory over the last three summers, two were women and two were African-American men.

The nuclear physics and astrophysics groups involve 4-8 undergraduates in research every summer. These students are integral parts of our research program. They receive training in sophisticated analysis methods for extracting signals from complex backgrounds, a range of programming languages (C, C++, FORTRAN, Perl, and IDL), and the Linux operating system. They learn modern high-performance computing methods by using our existing cluster. We have also recruited computer science students to help with administration of the cluster, from setting up firewalls to updating the CLAS software.

It is worth noting that for our undergraduates an on-campus cluster is essential for rapid turnaround times. The computing farm at JLab is heavily used, and it can take a day or more to get a batch job started. Such a long turnaround time is a barrier to learning and productivity, especially for these young people, that an on-campus system will overcome.

The astrophysics students currently have no access to a cluster at all. These students have developed a large code base for simulating CMB sky maps and performing a wide variety of statistical analyses on them, including a variety of tests for non-Gaussianity as well as techniques based on wavelet and radon transforms. In the past, we have been able to perform this research on individual workstations, but we have reached the point where a more powerful computing cluster is necessary for further progress. Access to a cluster will be an invaluable resource for these students in their scientific training.

The project will benefit a significant number of students beyond the University of Richmond. The group at Ohio University includes a postdoctoral fellow, 1-3 graduate students, and many undergraduates. The astrophysics research groups headed by Timbie at Wisconsin and Tucker at Brown typically have a postdoctoral fellow, several graduate students, and undergraduates working on data analysis and simulation issues.

6 Project Management Plans

The system will be managed at least initially in the same manner as the existing cluster is used now. Users will log into the master node to edit, compile, link, test, and execute their codes. They will submit jobs to the cluster from the master. All of students involved in the project, undergraduate and graduate, will have accounts on the master and be able to submit jobs.

The expertise exists in the University of Richmond nuclear and astrophysics groups to operate and maintain the proposed computer cluster. We have considerable software experience in general and with the applications described above. The University administration has adequately supported our research efforts in the past and is committed to continuing to support the University’s technology infrastructure (see Appendix C). One member of the University’s Information Services is a Linux expert who devotes half of his time to academic projects. He is responsible now for keeping the CLAS software up-to-date, updating the Linux software on the cluster and in our laboratory, and general troubleshooting. Finally, we have modeled many features of the proposed computer cluster after existing ones at JLab or within the CLAS collaboration. There is a significant amount of expertise within the collaboration that we can call on. The anticipated operating costs are for power and Linux support staff. The University has covered those costs for the existing cluster since 2001 and will continue to do so.

The laboratory that will hold the cluster is complete and in regular use now. It has adequate electrical power and cooling for the proposed instrument. It is described in more detail in Section 4.1 The usage of the current cluster runs anywhere from 5 cluster-hours/week to over 100 cluster-
hours/week if many simulations of the CLAS detector are required. The average is around 10-15 cluster-hours/week averaged over a full year with higher demand during the summer. We expect this average demand to increase with the proposed instrument. Over the last six years the downtime averages out to 3-4 days per month due to failed components, power outages, etc. The rate of failed components has, not surprisingly, increased recently as the system ages.

The cluster management will done with a complete set of tools from the vendor Linux Labs. The operating system will be NimbusOS which is a full, Linux-based (Fedora), operating system for High Performance Computing (HPC) clusters. The OS includes the software bproc (Beowulf Distributed Process Space) which is a set of kernel modifications, utilities and libraries which allow a user to start processes on other machines in a Beowulf-style cluster [82]. The version of bproc in NimbusOS is from Clustermatic, the HPC architecture developed at Los Alamos National Labs to which Linux Labs is a contributor [83]. There is additional software for distributing the load on the cluster (Maui batch schedule) [84], monitoring the compute nodes (modified version of Supermon) [85], and visualizing the results (Ganglia) [86]. This software has been widely used in industry and Linux Labs has extensive experience with all these tools.

We have attracted numerous other users from JLab. There are currently accounts for fifteen users from the CLAS Collaboration and other groups at JLab including the senior personnel in nuclear physics described in Section 3.5. We expect that we will have little trouble attracting new users to the proposed instrument.

See the letter from the University of Richmond dean in Appendix C committing the University to support instrument maintenance, operations, and housing.

7 Dissemination Plan

The work described above will be the subject of internal technical reports at JLab and ultimately publication in refereed journals. Our students will use their results as a springboard into their technical careers by presenting posters and talks at national and international meetings.

8 Results from Prior NSF Support

An NSF Major Research Instrumentation grant in 2001 provided funds (along with $24,000 in matching funds from the University) to purchase the existing cluster in our nuclear physics laboratory at Richmond. Gilfoyle was co-PI on that grant entitled “MRI: Development of a Computing Cluster to Support the University of Richmond Nuclear Physics Research Program at Jefferson Lab” (#6030194) for $151,758 and for the period 6/01/2001 - 5/31/2003. All of the Richmond nuclear physics work described here has made heavy use of the cluster (see Section 3) with the other JLab users. All the Richmond students used the cluster [45, 69–77, 87]. The existing cluster was crucial in completing the CLAS $G_M^n$ analysis (by Gilfoyle) which has been approved by the CLAS collaboration [88], presented at meetings [43, 44], and submitted for publication [39]. The proposed cluster will be used by him to analyze the remaining data from that experiment. The existing cluster was also essential for the CLAS12 $G_M^n$ proposal [89, 90] E12-07-104 which was approved by PAC32 [37], the analysis of the fifth structure function [51, 91, 92] and development of a new way to calculate radiative corrections for the $D(e,e'n)p$ and $D(e,e'n)p$ reactions [93]. Outside users have used the cluster for work leading to other publications [94–96] and technical reports [97–99]. Gilfoyle has been supported by DoE since 1990 (currently DoE grant DE-FG02-96ER40980). In that time he has been co-author on 81 refereed papers, 35 contributed talks and posters by him and his students, and nine JLab technical reports.

Bunn is currently supported by NSF Award AST-0507395, “Cosmic microwave background
analysis in the post-WMAP era” (July 1, 2005 - June 30, 2008, with a one-year no-cost extension). This award has funded research into a variety of cosmological topics at the interface between theory and observation:

**Modeling CMB interferometers.** Bunn has developed a detailed framework for assessing systematic errors in CMB polarimetry [14], and in collaboration with Martin White has studied curved-sky issues related to interferometric mosaicking [21]. Along with a graduate student at Wisconsin and an undergraduate at Richmond, Bunn solved the problem of finding the optimal time sequence of phase shifts for minimum-noise visibility recovery in an adding interferometer [100].

**Constraining alternative theories of gravity.** In collaboration with researchers at MIT, Bunn has examined the cosmological behavior of the class of alternatives to general relativity known as $f(R)$ theories. We find viable $f(R)$ theories in which modified gravity explains the current accelerated expansion of the Universe, but these theories are observationally hard to distinguish from a simple cosmological constant.

**Large-scale CMB anomalies.** Although the WMAP sky maps are in general remarkably consistent with the standard model, there have been several claimed detections of anomalies on large angular scales. Bunn and undergraduate Austin Bourdon have published an article [101] showing that a broad class of models, including many based on mundane issues such as foreground contamination or systematic errors as well as others based on exotic cosmology, exacerbate rather than solving some of these problems.

**Directionality in all-sky polarization maps.** Undergraduate Duncan Hanson extended previous work by Bunn and Douglas Scott [102] that devised a statistic for characterizing departures from global statistical isotropy in CMB maps. The new work [103] generalizes this technique to polarization data and applies it to the WMAP data.

**Remote quadrupole measurements.** Observations of the Sunyaev-Zel’dovich effect in a galaxy cluster provides a sample of the CMB quadrupole as it would be measured at the cluster’s location and look-back time [104]. A sample of such galaxies can probe the volume of the observable Universe on gigaparsec scales, complementing the limited information available on these scales from the CMB. Bunn has quantified the amount of extra information that a survey of such clusters can provide [105].

**Non-Gaussian statistics of Galactic dust.** Foreground contaminants in CMB maps may cause false detections in searches for CMB non-Gaussianity, even if the contamination is low enough not to affect the power spectrum. The Bunn group is attacking this problem by characterizing the non-Gaussian statistics of foreground maps such as the Schlegel-Finkbeiner-Davis dust maps [106] and assessing the degree to which low levels of dust contamination induce non-Gaussian detections in simulated CMB maps. Preliminary results on this work have been presented at numerous AAS meetings. A publication is currently in preparation.

**The meaning of the cosmological redshift.** In a pedagogical paper currently under review by the *American Journal of Physics* [107], Bunn and David Hogg argue for the rehabilitation of the much-derided interpretation of the cosmological redshift as a Doppler shift.

**Publications and presentations.** This grant has led to seven refereed publications [14, 21, 100, 101, 103, 105, 108], three of which have undergraduate coauthors. One paper [107] currently under review and at least one more is expected to be submitted in the coming months. In addition, the grant has resulted in four conference proceedings [8–10, 109] and eight contributions to AAS meetings (four presented by undergraduates) [78–81, 110–113].
References


J-M. Laget. private communication.


Biographical Sketch: Emory F. Bunn

Professional Preparation:

- Princeton University, A.B., Physics, 1989 (magna cum laude).
- University of California at Berkeley, M.A., Physics, 1993.
- University of California at Berkeley, Ph.D., Physics, 1995.
- Postdoctoral researcher, University of California, Berkeley, 1995-1996.

Appointments:

- Associate Professor of Physics, University of Richmond, 2008-present.
- Assistant Professor of Physics, University of Richmond, 2002-2008.
- Assistant Professor, Physics, Astronomy, and Engineering Science, St. Cloud State University, 1999-2002.
- Assistant Professor of Physics and Astronomy, Bates College, 1996-1999.

Selected publications related to this proposal:


Other selected publications:

**Synergistic Activities:** I am active in physics education in a variety of informal ways beyond my formal teaching duties.

- I am research mentor to 2-4 undergraduate research students at any given time.
- I am a member of a team of University of Richmond faculty developing an innovative new introductory course in Interdisciplinary Quantitative Science. The course will be team-taught by faculty in physics, biology, chemistry, mathematics, and computer science.
- I coauthored a web page on the meaning of Einstein’s equation at [http://math.ucr.edu/home/baez/einstein/](http://math.ucr.edu/home/baez/einstein/). This is a significantly extended version of an article I coauthored in the *American Journal of Physics*.
- I have recently begun maintaining a blog about topics in astrophysics, relativity, cosmology, and other areas of science at [http://blog.richmond.edu/physicsbunn/](http://blog.richmond.edu/physicsbunn/).
- I am consulted by local and national media (*Discover*, *Richmond Times-Dispatch*, *National Geographic News*) for comment on news stories related to astronomy and astrophysics.
- I appeared on Richmond area television to discuss education-related holiday gifts in 2007.
- I am one of the developers and organizers of the Richmond Physics Olympics, an annual high school outreach event for high school students.

**Collaborators (past 48 months):** P.A.R. Ade (Cardiff), J.C. Baez (U.C. Riverside), E. Bierman (U.C.S.D.), A. Bourdon (Richmond), C. Calderon, T. Faulkner (M.I.T.), B. Follin (Richmond), A.C. Gault (Wisconsin), D. Hanson (Cambridge), D.W. Hogg (N.Y.U.), P. Hyland (McGill), B.G. Keating (U.C.S.D.), J. Kim (Niels Bohr Inst.), A. Korotkov (Brown), S. Malu (IUCAA), Y. Mao (M.I.T.), P. Mauskopf (Cardiff), J.A. Murphy (Natl. Univ. of Ireland, Maynooth), C. O'Sullivan (Natl. Univ. of Ireland, Maynooth), L. Piccirillo (Manchester), D. Scott (U.B.C.), M. Tegmark (M.I.T.), P.T. Timbie (Wisconsin), G.S. Tucker (Brown), B.D. Wandelt (Illinois), M. White (Berkeley).

**Graduate and postdoctoral advisor:** J. Silk (Oxford).
Biographical Sketch: Dr. Gerard P. Gilfoyle

Professional Preparation:
SUNY, Stony Brook, Postdoctoral Fellow in Experimental Heavy-Ion Physics, 1985-1987.

Appointments:
2008-present - Clarence E. Denoon Professor of Science.
2006-present - Chair, Nuclear Physics Working Group of the CLAS Collaboration.
2004-present - Professor of Physics, University of Richmond.
2000-2006 - Chair, Department of Physics, University of Richmond.
1993-2004 - Associate Professor of Physics, University of Richmond.
Summer, 1988 - Visiting Research Professor, University of Pennsylvania.
1987-1993 - Assistant Professor, University of Richmond.

Awards and Honors:
1990-present - US Department of Energy ($1,421,000).
2004 - Who’s Who Among America’s Teachers.
2003 - University of Richmond Distinguished Educator Award.
2002-2003 - SURA Sabbatical Support ($10,000).
2001-2002 - National Science Foundation Major Research Instrumentation Program ($175,000).
1995-1997 - National Science Foundation, Instrumentation and Laboratory Improvement Program ($14,986).
1992-1995 - National Science Foundation, Instrumentation and Laboratory Improvement Program ($49,813).

Selected Publications Related to the Proposed Research:
See Reference [48] in ‘References Cited’ for a list of the members of the CLAS Collaborations.
Selected Other Publications:
See Reference [48] in ‘References Cited’ for a list of the members of the CLAS Collaborations.


Synergistic Activities:
We have made broader impacts. Gilfoyle is chair of the Nuclear Physics Working Group of the CLAS Collaboration and manages a portion of the Collaboration’s physics program. He served in government (1999-2000) as a scientific consultant on weapons of mass destruction for the US Department of Defense applying his physics skills to a range of policy issues. Our teaching has been illuminated by this work. We have added computational methods to our upper-level physics curriculum and computer-based data acquisition and analysis to our introductory physics sequence (with teaching grants from NSF). Finally, we have recruited women and African-American students to our group in nuclear physics. A former female students is now a staff scientist at the Jet Propulsion Lab in California and over the last two years two women and two African-American men have worked in our laboratory. One female graduate (Greenholt) is combining nuclear physics and public policy.

List of Recent Collaborators:
See Reference [48] in ‘References Cited’ for a list of the members of the CLAS Collaborations. Below we list any current Collaboration members not on Reference [48] and additional collaborators.

A. Afanasev Hampton University J. Arrington Argonne National Lab
E. Bunn University of Richmond L. El Fassi Argonne National Lab
A. Freyberger Jefferson Lab M. Fetea University of Richmond
D. F. Geesaman Argonne National Lab K. Hafidi Argonne National Lab
R. J. Holt Argonne National Lab S. Jeschonnek Ohio State University
P. Kroll Universität Wuppertal B. Mustapha Argonne National Lab
H. Nebel University of Richmond D. H. Potterveld Argonne National Lab
P. E. Reimer Argonne National Lab P. Rubin George Mason University
P. Solvignon Argonne National Lab J.W. Van Orden Old Dominion University
H. Arenhoevel Institut für Kernphysik, Mainz

Graduate and Postdoctoral Advisors
Graduate Advisor - Dr. H.T. Fortune, University of Pennsylvania.
Postdoctoral Advisor - Dr. R.W. McGrath, SUNY, Stony Brook.
Thesis Advisor and Post-Graduate Advisor

None. The University of Richmond is a primarily undergraduate institution.
Budget justification

We now describe the proposed system that will satisfy our computational needs. The components are listed in Table 1. A detailed quote for items 1-5 is in the supplementary documents from Linux Labs. Below we discuss our reasoning behind the choice of the different components.

<table>
<thead>
<tr>
<th>Item</th>
<th>Number</th>
<th>Description</th>
<th>Price($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Dual quad-core Xeon master node, 4 GByte RAM, 6 TByte RAID</td>
<td>11,400</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>Remote nodes, 4 GByte RAM, 150 GByte storage</td>
<td>117,000</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>HP Procurve switch</td>
<td>4,460</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>Cabinets</td>
<td>3,300</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>Nimbus OS license, installation, warranty, and shipping</td>
<td>24,652</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>Hardware items that cost less than $500</td>
<td>1,100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total Cost</td>
<td>161,912</td>
</tr>
</tbody>
</table>

Table 2: Proposed computer cluster description and cost (see quote in supplementary documents for more details).

The dual, quad-core Xeon processors (item 1) were chosen because of their excellent cost-to-benefit ratio. Their clock speed is about 50% faster than the speed of the current remote nodes, and architectural improvements make them overall about 10 times faster than the current ones (6.8 gigaflops versus 0.7 gigaflops). Test results from the vendor (LinuxLabs) using LinPack (a benchmarking standard for solving large, square matrices) give an equivalent speed of 1.5 teraflops for the entire instrument. This is about 13 times the speed of the original Richmond cluster and 27 times the speed of the existing, depleted cluster (see Section 4.1). The Linux operating system will be used. It is a research-quality operating system commonplace in physics and computer science. The number of machines was chosen to reduce the time to for simulating the CLAS response to a reasonable value. To generate and analyze 40 million events, we estimate about 1.5 days compared with the time required for the existing, 24-node, cluster (about 20 days) for such a calculation. Calculations of this length are routine for the analysis of the CLAS data. The astrophysics projects require hundreds of power spectrum estimations and maximum entry reconstructions which each take about a CPU-hour on the current nodes. The memory (4 GByte for the 8 cores on each node) is needed because the reconstruction and simulation packages for the nuclear physics work use large amounts of memory and the astrophysics simulations work with large matrices. A 150-GByte hard drive (item 2) will be attached to each machine to provide storage. This space is needed to store data files for analysis, and the results of the astrophysical analysis. The fast ethernet switch (item 3) is needed to speed data transfer over the network (see Section 4.2). Cabinets will hold the nodes (item 4). Hardware and software installation is required (item 5). The software for managing the cluster and submitting batch jobs is Nimbus Beowulf from Linux Labs in Atlanta, GA. This is the vendor who built the current cluster, and we have had a long, fruitful relationship with them. A variety of other components each costing less than $500 (cables and tools) are included in item 6. We expect the system to have a 4-6 year lifetime. Our experience at Richmond and at JLab suggests that remote nodes will gradually fail over time and that 4-6 years is the optimum lifetime (see Section 4.1). See the Facilities document and Section 4 of the Project Description for further details. The University of Richmond is committed to software and hardware support for the instrument of $10,000 per year for four years after the grant expires.
Facilities, equipment, and other resources

The current facilities in physics include a computing cluster developed in 2001 with NSF and University funds plus an array of workstations. The original system consisted of 49 remote machines (1.4 GHz) running the Linux operating system, a master node, 3 TByte of RAID storage, and 18 GByte of disk space on each node. It resides in a laboratory with a 50-ton, 60,000-BTU air conditioner, an upgraded electrical panel, and a connection to the building's backup power. Nearby rooms provide space for workstations and our students. The system is now at the end of its useful life: only about 24 nodes and 2 out of 3 RAIDs still work. This is not surprising since the system is almost eight years old. At JLab, for example, replacement of failing nodes begins after two years and all machines are replaced after four years of use.

The remainder of the nuclear physics and astrophysics laboratories consists of nine Linux machines for student and faculty use. Most of the software used in the both the nuclear physics and astrophysics research is non-proprietary. The University and other grants maintain licenses for the needed proprietary software (IDL, Mathematica).

One member of the University’s Information Services is a Linux expert who devotes half of his time to academic projects. He is responsible now for keeping the CLAS software up-to-date, updating the Linux software on the cluster and in our laboratory, and general troubleshooting.

The anticipated operating costs are for power and Linux support staff. The University has covered those costs for the existing cluster since 2001. See the letter from the University of Richmond dean supporting this project in Appendix C.

The University supplies and maintains office space for all faculty and laboratory space for faculty and students, as well as necessary services such as secretarial support, postage, Internet, telephone, printing, and photocopying. University funds are available for both student and faculty travel to conferences. Summer stipends for research students are available on a competitive basis; students in the astrophysics and nuclear physics groups have a very high success rate when competing for these stipends.
# Cluster Price Quote

Linux Labs International Inc.
3276 Buford Road
#104-335
Buford, Ga 30519

Name/Address
University of Richmond
Gerard P. Gilfoyle
Science Center room N104
University of Richmond, VA 23173

### Estimate

<table>
<thead>
<tr>
<th>Date</th>
<th>Estimate No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/18/2008</td>
<td>1546-2012R</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Description</th>
<th>Qty</th>
<th>Rate</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Master node - 4 U Intel 5345 Dual Quad Core Xeon with 4 Gb DDR2 ECC RAM, DVD R/W, RAID 6 with (12) 500 GB Caviar SE16 SATA HD's. Dual power supplies and SATA controller cards, Monitor, keyboard, mouse.</td>
<td>1</td>
<td>11,400.00</td>
<td>11,400.00</td>
</tr>
<tr>
<td>Compute nodes - 2 u Intel 5345 Dual Quad Core Xeon with 4 Gb DDR2 ECC RAM,(2) 150 GB Hard Drive. HP ProCurve 48 port Gig E switch with cables</td>
<td>15</td>
<td>7,800.00</td>
<td>117,000.00</td>
</tr>
<tr>
<td>Equipment Cabinet - Black with Glass front panel</td>
<td>2</td>
<td>1,650.00</td>
<td>3,300.00</td>
</tr>
<tr>
<td>Onsite Installation and training. Includes travel expenses for one technician.</td>
<td>1</td>
<td>2,600.00</td>
<td>2,600.00</td>
</tr>
<tr>
<td>Licenses for NimbusOS with new Iron Penguin features</td>
<td>16</td>
<td>472.00</td>
<td>7,552.00</td>
</tr>
<tr>
<td>2 year extended service agreement for 24 x 7 phone and email support with depot repair service.</td>
<td>1</td>
<td>12,800.00</td>
<td>12,800.00</td>
</tr>
<tr>
<td>Shipping via ground insured</td>
<td>1</td>
<td>1,700.00</td>
<td>1,700.00</td>
</tr>
</tbody>
</table>

This system is (28) dual quad core nodes which is 224 processing cores.

Total $160,812.00

We thank you for the opportunity to provide this proposal. Feel free to contact us at 866 824 9737
Paris, January 17th 2009

To Whom it may concern

**Letter of support for the NSF Major Research Instrumentation proposal MRI**

My team in Paris is heavily involved in the data analysis and simulations for the future BRAIN-MBI instrument. We have started a fruitful collaboration with Prof. Bunn's team and are willing to continue working with him. The NSF Major Research Instrumentation proposal "MRI Acquisition of a computing cluster for astrophysics and nuclear physics research at the University of Richmond" proposes among other activities to perform and develop simulations of the BRAIN-MBI interferometer. We would be happy to collaborate and participate to the proposed simulations providing support, manpower and technical help as well as sharing the codes we have already developped. The simulations proposed in this applications will be huge help in finalizing the design of the instrument and in developping the data analysis pipeline.

Sincerely,

Jean-Christophe Hamilton
14 January 2009

Prof. Emory F. Bunn
University of Richmond
Richmond, VA 23229
(804) 287-6486

Dear Ted;

I understand that among the proposed investigations in the NSF Major Research Instrumentation proposal “MRI: Acquisition of a computing cluster for astrophysics and nuclear physics research at the University of Richmond” are simulations of the BRAIN/MBI interferometer. These simulations will be of great value in the development of this instrument. My research group will participate in the proposed simulations by providing technical guidance and code as available. We look forward to working with you on these simulations.

Sincerely,

Peter Timbie
Professor of Physics
pttimbie@wisc.edu
608-890-2002
16 January 2009

Prof. Emory Bunn
Department of Physics
University of Richmond
Richmond, VA 23173

Dear Ted,

Among the proposed investigations in the NSF Major Research Instrumentation proposal “MRI: Acquisition of a computing cluster for astrophysics and nuclear physics research at the University of Richmond” are simulations of the BRAIN/MBI interferometer. These simulations will be of great value in the development of this instrument. My research group will participate in the proposed simulations by providing technical guidance and code as available.

Sincerely,

Gregory Tucker
Associate Professor of Physics
January 20, 2009

Dr. Gerard P. Gilfoyle  
Physics Department  
University of Richmond  
Richmond, VA 23173

To Whom It May Concern:

I acknowledge that I am very interested in being a user for the NSF Major Research Instrumentation proposal ‘MRI: Acquisition of a computing cluster for astrophysics and nuclear physics at the University of Richmond.’ The proposed investigations will be valuable in the ongoing collaboration between my research group and co-PI Gilfoyle. My research group and I will participate in the proposed analysis and simulations and use the proposed instrument.

Sincerely,

[Signature]

Kenneth Hicks

Professor of Physics  
Project Director, University Research Priority Project
January 16, 2009

Dr. Randy Phelps, Staff Associate
Office of Integrative Activities
National Science Foundation
4201 Wilson Blvd.
Arlington, VA 22230

Dear Dr. Phelps:

I am writing to provide the University of Richmond’s commitment to the National Science Foundation-Major Research Instrumentation proposal, “MRI: Acquisition of a Computing Cluster for Nuclear Physics and Astrophysics Research at the University of Richmond,” being submitted by Drs. Emory F. Bunn and Gerard P. Gilfoyle in the Department of Physics.

Acquisition of this instrument is critical to the research activities of our faculty and to the research training of our students. This computing cluster will be housed in our Gotwalt Center for the Sciences, a newly renovated and enhanced building with extensive new facilities for learning and research. The university is committed to the operation and maintenance of the computing cluster beyond the warranty with ongoing software support and replacement of two nodes per year for four years after the grant expires.

For more than nine years, I have worked very closely with the Physics Department and other science faculty, developing and implementing our Strategic Plan’s Initiative for Scientific Discovery, the flagship initiative for our institution’s renewed focus on student learning and research. Faculty members have diligently revised the curriculum in each program with an emphasis on experiential learning where students are challenged with hands-on and real-world applications. In this curriculum, students are encouraged to join research partnerships under the guidance of our faculty, allowing active engagement and leading to peer-reviewed presentation and publication. Up to date instrumentation is a vital component of this plan for both faculty and students.

As we realize our aspiration of moving from the ranks of the very good science undergraduate programs to one of the very best, each enhancement in instrumentation affecting research takes us closer to our goals. Our facility and programmatic improvements through the Initiative for Scientific Discovery will enhance our students’ learning process while on campus and will encourage our students to continue their scientific study beyond their four years at Richmond. At the same time we need partners, highly respected external agencies and foundations that are also committed to improving scientific research, study and education.

I thank you for the opportunity to be considered under the NSF-MRI program.

Sincerely,

[Signature]

Andrew F. Newcomb, Ph.D.
Dean

Office of the Dean
University of Richmond, VA 23173
(800) 269-8165 Fax (804) 289-8418
www.urichmond.edu