LDRD Director's R&D Fund Preliminary Proposal

Title: Enhancing Scientific Discovery by Integrating Simulation and Experiment in Neutron Scattering Studies at the Spallation Neutron Source

Principal Investigator: Jay Jay Billings

Lead Division: Computer Science and Mathematics (CSMD)

Co-Investigators: Garrett Granroth (NDAV), Jose Borreguero (NDAV)

Initiative: Computer Science and Math for Exascale Computing

Requested Budgets

Year	Fiscal Year	Budget (dollars)
Year-1	2016	$375,000
Year-2	2017	$375,000
Year-3		
Total Budget Request		$750,000

Project Description

Modern scientific computing can simulate materials at the nanoscale and provide key insights into the behavior of the real material. Similarly, neutron scattering, at the Spallation Neutron Source, provides a direct comparison to these simulations more quickly than ever before. Close integration between these two techniques greatly enhances scientific discovery. However, technical and temporal bottlenecks currently limit the potential knowledge to be learned from this synergy. This work will, for the first time, provide an integrated modeling and simulation tool chain that includes Density Functional Theory (DFT) to produce simulated neutron scattering spectra for comparison to data from the vibrational spectrometer, VISION, at the SNS. It will make the computational tools faster to use for the experts and more available for new users.

This proposal paves the way for ORNL to take the lead in fully predictive modeling and simulation capabilities that directly inform and direct experimental work in data management for large-scale results from simulations that are directly associated with experiments. We expect follow on funding opportunities from the U.S. Department of Energy's Office of Science and collaboration opportunities with partners at other neutron sources and even light sources as a result of this work.

Significance

This work is significant because the proposed capabilities do not exist for the user community of neutron scattering instruments or, broadly, the experimental physics and chemistry communities as a whole. Certainly tools exist that can be used to do each *individual* activity proposed in this work, but no single platform exists in the marketplace to do all of these activities holistically and to the required fidelity as this work proposes. Furthermore the existing tools have cumbersome and non-intuitive interfaces.

This is the fundamental limiter for Science production. It means we dedicate large amounts of staff time setting up each single DFT simulation to compare to data that took a few minutes to acquire. Furthermore some of the software used during this process is expensive, and/or proprietary therefore adding to ORNL costs and preventing the easy exchange for data with other tools. *The turnaround time associated with this activity prevents us from discovering and reporting critical physical results on an ideal timetable.* The effects of this are magnified when applied to the larger community of experimental physicists and synthetic chemists outside of ORNL, which has neither our resources nor our expert staff and, therefore, on the whole, may not even be able to do as thorough of an investigation of their results.

This work will change that and produce original research results for ORNL in the process. The proposed capabilities will drastically decrease the work currently performed by ORNL staff members. This work will produce original research papers that describe state-of-the-art methods for simultaneously generating, managing, correlating and analyzing experimental results from an actual instrument, (VISION), with large-scale and Leadership-class simulation results obtained from ORNL's Leadership Class Computing Facility and existing SNS compute resources. It will also produce research papers in journals and conferences that describe how the system was built and the engineering challenges involved. Finally, the physical insights generated as part of this work will be reported in relevant physics and chemistry journals and conferences.

Research Plan

The problems described above are large, general problems that cannot be completely solved in the course of a single LDRD, so our work will be limited in scope to experimental results from ORNL's vibrational spectrometer VISION and simulations with common materials simulation codes such as Quantum Espresso (DFT). Supporting more instruments and more simulation suites are good topics for follow on funding.

There are two fundamental aspects to this work: making it possible for novice users to simulate the neutron scattering spectrum of their materials and to allow them to compare the results of the simulations with results from VISION, which produces the real neutron scattering spectrum. This work separates these aspects into tasks related to modeling and simulation and tasks related to data analysis, whether that analysis be with experimental or simulated data.

Modeling and simulation in this context means performing simulations of the materials and computing a neutron scattering spectrum with the results. This can be done from, for example, DFT simulations by computing the density of states and using those results to compute the neutron scattering spectrum. While many solutions exist for computing these results, *the bottlenecks appear when constructing large, realistic inputs and post-processing the results.*

Creating inputs with larger, more realistic structures will improve the results from the simulation. This process currently requires either manually creating the structure of the material - atom by atom - in a text file or using cumbersome, proprietary solutions that do not scale well to large structure sizes and make it impossible (or at least difficult) to export the structure to other software products. Post-processing is not limited to generating the neutron scattering spectrum from the the DFT results. It may also include additional simulations with other codes or applying sets of mathematical operations (fourier transforms, scaling, etc.) to the output. In the case of applying sets of mathematical operations, this is similar to performing regular data analysis.

This work will address the following areas, based on the discussion above: creating models of large structures that reflect real materials for simulations, performing DFT simulations on these materials with codes such as Quantum Espresso and directly comparing the computed neutron scattering spectra to measured spectra. This work will deploy the software developed as part of ORNL's Eclipse Integrated Computational Environment, which already supports fifteen different high-performance computing efforts.

Year one of this work will focus on developing and deploying data analysis capabilities for experimental results from VISION and from the simulations. Year two of this work will focus on developing a new molecular model builder that is based on requirements that are gathered in year one in tandem with the other activities. This molecular model builder will be optimized for generating large structures and surfaces that can be used in high-performance computing simulations using DFT. This work will also develop input-output translation tools so that results generated experimentally and virtually can be stored in open a single, cross-platform file that includes relevant metadata, experimental infrared, neutron and backscattering data as well as down-sampled simulation results.

The 10-month deliverables will include 1.) plugins to ICE to enable data analysis of VISION results and DFT simulations and 2.) the initial input-output tools for data translation.

The 24-month deliverables will include 1.) the molecular model builder for constructing 3D molecules for the DFT simulations, 2.) a peer-reviewed journal or conference publication detailing the new capabilities and 3.) deployment of the new capabilities to SNS users.

One or more students will be asked to work on this project through ORNL's GO! program.

Future Funding

There are several opportunities for future funding for this project and the expected return on investment is five years. The U.S. Department of Energy's Office of Science is the obvious target for future funding because it currently funds software infrastructure development for the SNS and we expect at least two opportunities from them in the future. First, we expect that this activity will be highly synergistic with CAMM and help it compete in upcoming future funding opportunities. We also expect that the Office of Advanced Scientific Computing Research will be issuing a call in the Fiscal Year 2017 time frame that looks at the issues associated with knowledge discovery in modeling and simulation and this work would allow ORNL to compete directly in that call.

This technology will could also be easily extended to light sources instead of neutron sources and this work would enable ORNL to join with partners at light sources on funding opportunities. It will also be possible to extend this work to include electron microscopy with relative ease, which opens up additional future funding opportunities.