Annual Report for Blue Waters Allocation

Project Information

- Title: Quantum Simulations
- PI: David Ceperley (BW professor), Department of Physics, University of Illinois Urbana-Champaign
- Co-PI Norm Tubman (postdoc), University of Illinois Urbana-Champaign
- Corresponding author: David Ceperley, ceperley@illinois.edu

Executive summary (150 words)

The majority of our research on Blue Waters is related to the "Materials Genome Initiative," the federally supported cross-agency program to develop computational tools to design materials. We employ Quantum Monte Carlo calculations that provide nearly exact information on quantum many-body systems and are also able to use Blue Waters effectively. This is the most accurate general method capable of treating electron correlation, thus it needs to be in the kernel of any materials design initiative. Ceperley's group has a number of funded and proposed projects to use Blue Waters as listed below. In the past year, we have been running benchmark calculations for dense hydrogen in order to develop analytic representations of the electronic energies and, in so doing, develop accurate models needed to simulate systems for longer length and time scales.

Description of research activities and results

We have been awarded, submitted, or will submit in the near future, several projects that involve Blue Waters calculation:

- "Warm dense matter" DE-NA0001789. Computation of properties of hydrogen and helium under extreme conditions of temperature and pressure. Ceperley is sole PI. This is the funding that has supported computations in 2014. A proposal for renewal is pending at DOE.
- SCIDAC, Predictive Computing for Condensed Matter DE-SC0008692. So Hirata is the PI, involves 5 other UIUC faculty. Renewal is pending at DOE.
- Network for Ab Initio Many-body Methods: for development of the QMC software, QMCPACK. Kent at Oak Ridge is the PI, involves scientists at 4 national labs as well as UIUC. The renewal pending at DOE.
- EFRC: Center for Emergent Superconductivity. A very ambitious project to discover new superconducting materials. This grant was funded in July 2014, the PI is Johnson at Brookhaven National Lab.
- Materials Innovation Platforms. To be submitted to NSF in March 2015, Abbamonte will be the PI. Will involve many PIs from UIUC and NCSA.

<u>Access to Blue Waters is crucial for success of these projects.</u>

Our work on Blue Waters during the past year has focused on a "multiscale" project: a way to use highly accurate energies from quantum Monte Carlo to treat larger systems at longer length scales. Atomistic simulations are widely used across many different scientific disciplines, including physics, material science, chemistry, and biology. Large-scale simulations and highly accurate simulations cannot

generally be performed simultaneously: different disciplines are generally focused on one or the other. For example large-scale simulations of biological systems in aqueous environments require millions of particles and use classical potentials. *Ab initio* simulations generally cannot be used for such applications, as they are limited to much smaller system sizes. Even under the umbrella of *ab initio* simulations, there is a hierarchy in terms of accuracy, with quantum Monte Carlo being one of the most accurate methods but limited to about 1000 particles. Despite its cost, Quantum Monte Carlo is important, since high accuracy is needed. Other methods can have errors that lead to dramatically wrong results. In high-pressure hydrogen simulations, we have recently shown that various density functionals differ significantly from benchmark quantum Monte Carlo simulations. Even widely studied systems, such as bulk water at ambient pressure and temperature currently have discrepancies with respect to experimental data; these cannot be addressed until increased accuracy and larger system sizes can be simulated.

Recent developments in the field of machine learning have made it possible that we might soon be able to perform large-scale atomistic simulations with quantum Monte Carlo accuracy. The idea is to generate the potential energy for various atomic configurations. We can then use machine learning to techniques to "learn" all the energies for these configurations, and then predict energies for new configurations at a speed that is many orders of magnitude faster than the original quantum Monte Carlo simulations.

We have a three part project which involves testing with Lennard-Jones potentials, high pressure hydrogen simulations that address shock wave experiments on the Hugoniot, and ambient water simulations. We are currently running our hydrogen simulations, and we are preparing our water simulations. During the past year we ran thousands of quantum Monte Carlo and density functional theory simulations for dense many-body hydrogen on Blue Waters. The density functional theory and neural network simulations are completely parallelized, and can be scaled up to thousands of processors. Our QMC simulations can be scaled to hundreds of thousands of processors. These QMC calculations are the most computationally expensive portion of this project, and for our current focus on hydrogen, we are looking to run, by the end of the project, on the order of 20,000 simulations. Each simulation takes roughly 2 – 10 node hours.

Plan for next year

As mentioned above we have several projects associated with the Federal Materials Genome Initiative that we will be starting. Blue Waters capability is needed to take the research from the model level to realistic description of materials where electron correlation is important. Typical materials require simulation of a unit cell with at least one thousand electrons. Just holding the single-body orbitals in memory for a realistic material requires a multiprocessor computer. Relevant accuracy for materials design will definitely require petascale computational access.

Jeongnim Kim, Ceperley and his colleagues have devised algorithms and software (QMCPACK) where the computation is shared across processors within a node using Open-MP and across nodes with MPI. In addition, parts of the calculation can be done very effectively on GPUS with speed-ups over the CPUs of more than an order of magnitude. The QMCPACK software, which was one of the benchmark algorithms for Blue Waters, is one of the few codes that can take full advantage of its architecture. Our implementation is mature, well optimized, already installed on the Blue Waters; we have demonstrated 90% parallel efficiency on up to 216,000 processor cores for production runs. Development of QMCPACK is now supported by a 5 year DOE grant mentioned above; Ceperley is one of the PIs on this grant.

We plan to continue the current project on hydrogen and initiate several new projects. We expect to submit our first paper with results from Blue Waters on hydrogen within the next few months. We will then apply the method to water. We request 240,000 node hours on Blue Waters, for the period Feb 2015-

Jan 2016. We expect to be using the time uniformly throughout the year. The data transfer, data storage and access requirements should be similar to our usage in 2014.