Annual Report for Blue Waters Allocation

- **Project Information**
  - Blue Water Professorship
  - So Hirata, University of Illinois at Urbana-Champaign
  - sohirata@illinois.edu

- **Executive summary (150 words)**

We have reported the first Gaussian-basis-set *ab initio* second-order many-body perturbation (MP2) simulation of a variety of properties of liquid water (including IR and Raman spectra), on an equal footing with *ab initio* calculations for molecules. We have also computationally reproduced the anomalous volume isotope effect, negative thermal expansion, and pressure-induced amorphization of ice Ih. We have more fully developed Monte Carlo many-body perturbation methods by extending them to higher orders, GPU parallelism, and explicit correlation.

- **Description of research activities and results**

  - *Key Challenges:* Our group made two breakthroughs in computational chemistry for large molecules, solids, and liquids. One is the method that allows such high-level calculations to be applied to an infinitely extended molecular crystals and molecular liquids, including nature's most important solids and liquids such as ice and liquid water. It divides these crystals and liquids into fragments embedded in their electrostatic field and treat these fragments by well-developed molecular theories and software in a highly parallel algorithm. The other weds systematic (i.e., *ab initio*) many-body electronic structure methods with quantum Monte Carlo, enabling massively parallel, systematically accurate electronic structure calculations for larger molecules and solids. With these, (1) we will predict a variety of properties of all known molecular phases of ice and dry ice to construct *ab initio* phase diagrams of these important solids. We will also compute a variety of properties of liquid water using Born-Oppenheimer molecular dynamics at an *ab initio* level that goes far beyond usual density-functional or empirical force field methods. (2) We will apply the novel quantum Monte Carlo method to predict the stacking interaction energies (important for morphology and thus functions) and optoelectronic parameters (ionization and electron attachment energies, band gaps) of conjugated organic molecular solids and supramolecular assemblies.

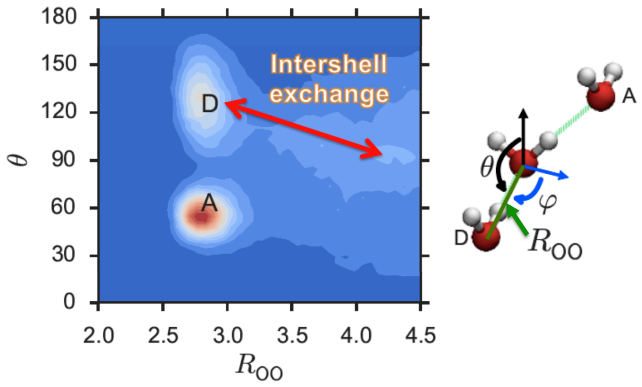
  - *Why it Matters*: Project (1) advances geochemistry, astrophysics, and planetary science where probing high-pressure phases of ices of atmospheric species on Earth or other planets are important but experimentally difficult and expensive. The systems studied by project (2) include solids that serve as bases of advanced materials such as bulk heterojunction organic solar cells, batteries, sensors, smart windows, field-effect transistors, and light emitting diodes. The optoelectronic

parameters are the quantities of prime importance determining the solids performance and functions, but the usual density-functional approximations are known to be poor for these properties. Here, our new method is uniquely useful and accurate. The impact of this research is broadly in the area of energy science and technology.

- *Why Blue Waters:* Today's workhorse computational methods for solids (density-functional methods) and liquids (classical molecular dynamics) are routine on a small computer cluster, but with limited accuracy. Therefore, the most meaningful use of Blue Waters in this area is to enhance the accuracy rather than enlarge the system size (which is already formally infinite). In electronic structure theory, this means switching from density-functional methods and empirical force fields to *ab initio* theories, which solve the fundamental equation of motion of chemistry rigorously and in systematic approximations with controlled errors. However, conventional matrix-algebra algorithms of *ab initio* theories are fundamentally non-scalable. The aforementioned combination of *ab initio* theories with quantum Monte Carlo and the embedded-fragment methods are among the few that may be realistically and usefully deployable on the large number of processors available on Blue Waters. Such calculations, in turn, can potentially directly address or answer some outstanding scientific questions of solids, liquids, or large optoelectronic materials purely computationally with high accuracy.

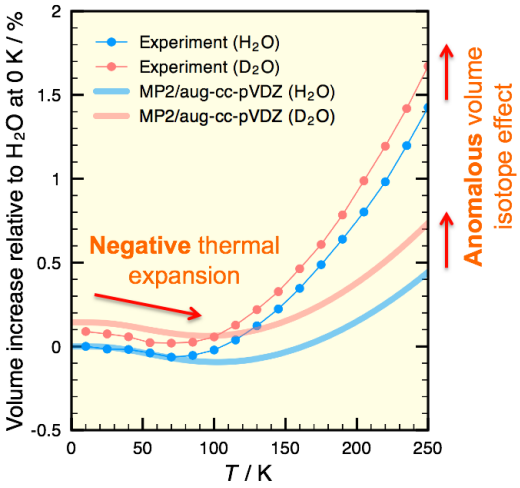
- *Accomplishments:* There are three main accomplishments this year, in all of which Blue Waters has played an essential role:

  (1) ***Ab initio* MP2 molecular dynamics of liquid water.** In paper #2 published in *Sci. Rep.*, we reported structural, dynamical, and response properties of liquid water calculated by *ab initio* molecular dynamics using the forces evaluated by the embedded-fragment spin-component-scaled second-order many-body perturbation method with the aug-cc-pVDZ basis set. The level of theory is chosen as it accurately and inexpensively reproduces the water dimer potential energy surface from the coupled-cluster singles, doubles, and noniterative triples with the aug-cc-pVQZ basis set, which is nearly exact. The calculated radial distribution function, self-diffusion coefficient, coordinate number, and dipole moment, as well as the IR and Raman spectra are in excellent agreement with the experimental results. The shapes and widths of the OH stretching bands in the IR and Raman spectra and their



isotropic-anisotropic Raman noncoincidence, which reflect the diverse local hydrogen-bond environment, are also reproduced computationally. The simulation also reveals intriguing dynamical features of the environment, which are difficult to probe experimentally, such as a surprisingly large fluctuation in the coordination number and the detailed mechanism by which the hydrogen donating water molecules move across the first and second solvation shells, thereby causing this fluctuation.

(2) ***Ab initio* MP2 study of ice anomalies.** Ice Ih displays unusual thermodynamic properties such as negative thermal expansion at low temperatures and the anomalous volume isotope effect, which renders the volume of D_2O ice greater than that of H_2O ice. The former is related to the well-known negative slope of water's melting curve and the pressure-induced collapse to an amorphous solid at low temperatures. We use the embedded-fragment second-order many-body perturbation method and the quasiharmonic approximation to determine the crystal structures that minimize the Gibbs energy of proton-disordered ice Ih at the temperatures and pressures spanning the entire domain of its stability. The calculations correctly predict the thermal contraction at low temperatures followed by expansion at higher temperatures, confirming this behavior originates from the negative Gruüneisen parameters of acoustic phonons. They also reproduce the anomalous volume isotope effect, but only when atomic partial charges (but not molecular dipole moments) are used in the embedding field. They confirm that it is caused by the large volume- contracting effect of the OH/OD stretching phonons, but reveal that its sign and magnitude are the result of subtle cancellation among closely competing effects of all phonons. The equally unusual temperature dependence of the volume isotope effect, on the other hand, is reliably reproduced by theory and is shown to originate from the librational phonons. We furthermore observe discontinuous changes in the volume and the ratio of lattice constants as well as softening of acoustic phonons at ca. 1.5 GPa, computationally detecting the mechanical instability of ice Ih that causes pressure-induced amorphization.



(3) **Explicitly correlated Monte Carlo MP2.** *Ab initio* methods for electron correlation, such as MP2, are well known to be slowly convergent with respect to the size of one-electron basis set. The slow convergence is due to the inability of such a basis set to describe the cusped form of the wave

function at electron-electron coalescence. The cure is to include two-electron basis functions, particularly, a function that explicitly depends on interelectronic distance (R12), which can describe the cusps exactly. The problem of the methods that include such basis functions, known as explicitly correlated or R12 methods, are that the formalism now involves high-dimensional integrals, that, if such high-dimensional integrals are avoided by the resolution-of-the-identity (RI) approximation, a large auxiliary basis set becomes necessary, that one has to derive and implement analytical Gaussian integrals of R12 basis function, which is difficult and possible only for the simplest such functions, and that their matrix-algebra algorithm is not scalable. The Monte Carlo MP algorithm solves all of these problems. A Metropolis Monte Carlo integration performs particularly well (relative to cubature) as the dimension is higher and hence RI with an auxiliary basis is unnecessary; it works for any functional form of R12 basis; it is fundamentally more scalable than matrix-algebra algorithms with respect to both computer size and system size. We have implemented explicitly correlated Monte Carlo MP2, which has at least one-rank lower cost scaling with system size ($O(n^4)$ as opposed to usual $O(n^5)$ where n is the number of orbitals) and naturally parallel. The method gives the MP2 reaction enthalpies that are within a few kJ/mol of the complete basis set limits with just the aug-cc-pVDZ basis set.

Reaction	MP2/aDZ	***MC-MP2-F12/aDZ***	MP2/CBS
$CO + H_2 \rightarrow CH_2O$	−17.9	***−23.6 (2.3)***	−28.2
$N_2 + 3H_2 \rightarrow 2NH_3$	−10.0	***−25.2 (1.4)***	−25.2
$C_2H_2 + H_2 \rightarrow C_2H_4$	11.5	***12.7 (1.1)***	14.9
$CO_2 + 4H_2 \rightarrow CH_4 + 2H_2O$	12.4	***0.6 (2.4)***	−1.3
$CH_2O + 2H_2 \rightarrow CH_4 + H_2O$	−2.8	***−12.5 (1.4)***	−10.9
$CO + 3H_2 \rightarrow CH_4 + H_2O$	−20.7	***−36.1 (1.5)***	−39.1
$HCN + 3H_2 \rightarrow CH_4 + NH_3$	14.8	***4.9 (1.2)***	6.4
$HNO + 2H_2 \rightarrow NH_3 + H_2O$	−4.4	***−14.0 (1.4)***	−15.6
$C_2H_2 + 3H_2 \rightarrow 2CH_4$	19.5	***20.5 (1.1)***	21.9
$CH_2 + H_2 \rightarrow CH_4$	−66.4	***−76.2 (0.6)***	−76.3
$2CH_2 + H_2 \rightarrow C_2H_4$	−140.7	***−160.1 (1.2)***	−159.5
$F_2 + H_2 \rightarrow 2HF$	8.0	***−0.8 (2.3)***	−1.1

in kJ/mol

We also note that considerable progress has been made to render Monte Carlo MP2 execute efficiently on GPU. It has achieved ca. 200-fold speedup on one GPU on the XK node relative to one CPU on the XE node.

- **List of publications and presentations associated with this work**

S. Y. Willow, M. R. Hermes, K. S. Kim, and S. Hirata, *Journal of Chemical Theory and Computation* **9**, 4396-4402 (2013), "Convergence acceleration of parallel Monte Carlo second-order many-body perturbation calculations using redundant walkers."

S. Y. Willlow, M. A. Salim, K. S. Kim, and S. Hirata, *Scientific Reports* **5**, 14358 (2015) (14 pages), "*Ab initio* molecular dynamics of liquid water using embedded-

fragment second-order many-body perturbation theory towards its accurate property prediction."

M. A. Salim, S. Y. Willow, and S. Hirata, (in preparation, 2016), "Second-order many-body perturbation study of thermal expansion in ice Ih."

S. Hirata *et al.*, an invited book chapter in *Fragmentation: Toward Accurate Calculations on Complex Molecular Systems* edited by M. S. Gordon (in preparation, 2016), "*Ab initio* calculations of ice, dry ice, and liquid water."

C. Johnson *et al*., (in preparation, 2016), "Explicitly correlated Monte Carlo second-order many-body perturbation method."

A. Doran *et al*., (in preparation, 2016), "Monte Carlo second-order many-body perturbation method on graphical processing units."

Work benefitted from Blue Waters has been highlighted in invited talks by the PI in the following conferences and university seminars: American Conference on Theoretical Chemistry (2014); Low-Scaling and Unconventional Electronic Structure Techniques Conference (2014); International Symposium on "Quantum Chemistry for Extended Systems" in honor of Professor Jean-Marie André (2014); Seminar, Department of Chemistry and Biochemistry, University of Arkansas (2014); Seminar, Department of Chemistry, Virginia Tech (2014); Blue Waters Symposium (2014); Sanibel Symposium (2015); CECAM meeting (2015); Blue Waters Symposium (2015); RAEST (2015); Pacifichem 2015 (2015).

- **Plan for next year**

We request 200,000 node-hours (100,000 node-hours on XE and 100,000 node-hours on XK). The software for the Monte Carlo MP2 method has been rewritten for Blue Waters running on either XE and XK nodes. Its scalability is measured and being further improved. In 2015, we have used many of the node-hours for the development and scalability measurements of Monte Carlo MP2. In 2016, we will begin production calculations on large systems, which are outside the applicability of other existing algorithms of MP. This project will use the entirety of the 100,000 node-hours on XK and 50,000 node-hours on XE. The remaining 50,000 node-hours will be reserved for *ab initio* simulation of molecular liquids with the embedded-fragment method at *ab initio* levels of theory. We plan to upgrade the embedded-fragment method by introducing a multitimestep algorithm and its software by eliminating all I/O from the backend electronic structure code. Once these are complete, we expect the code to be more scalable on Blue Waters and execute for a variety of molecular liquid simulations. We note that the production run of the liquid water simulation described above took 400K CPU-hours on a supercomputer in Korea.

The usage schedule is, Q1: 25%, Q2: 25%, Q3: 25%, Q4: 25%.