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 develop and apply two highly scalable and systematically accurate electronic structure methods. One is the embedded-fragment method for various structural, thermodynamic, and spectroscopic properties of molecular crystals, amorphous solids, and liquids. The other is the Monte Carlo many-body perturbation and Green's function methods, running efficiently on thousands of CPUs or hundreds of GPUs or more. With Blue Waters, we perform grand-challenge large-scale applications of both methods.

- **Description of research activities and results**

  - *Key Challenges:* Our group have made two breakthroughs in computational chemistry for large molecules, solids, and liquids. (1) One is the method that allows such high-level calculations to be applied to an infinitely extended molecular crystals, amorphous solids, and liquids, including nature's most important solids and liquids such as ice and liquid water. It divides them into fragments embedded in their electrostatic field and treat these fragments by well-developed molecular theories and software in a highly parallel algorithm. (2) 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 large molecules and solids. With Blue Waters, we aim to perform grand-challenge applications to demonstrate these methods' full capacity. The applications may include (1) computationally explaining the thermodynamic and phase properties of crystalline and amorphous ices as well as structural and spectroscopic properties of liquid water as well as (2) near-exact calculations of ionization energies and electron affinities of large conjugated molecules, important for advanced organic materials for optoelectronic devices.

  - *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 and extremely expensive. 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, enabling formerly unthinkable predictively accurate calculations for large systems.

- *Accomplishments:* The past year's accomplishments are the following:

(1) ***Ab initio* phase diagram of ice.** We performed an *ab initio* second-order Møller–Plesset perturbation (MP2) calculation of the phase diagram covering many molecular phases of ice: Ih, II, III, VI, VII, and IX. Gibbs energy of each of these phases has been calculated in the complete-basis-set (CBS) limit in the quasiharmonic ap-proximation (QHA) for the thermal expansion effect. Calculated thermal expansion behaviors of the six phases (Figure 1) on the basis of the MP2-calculated phonon density of states (DOS) underscore the unusual nature of ice Ih; it is the only phase among the six that shows thermal contraction at low temperatures.

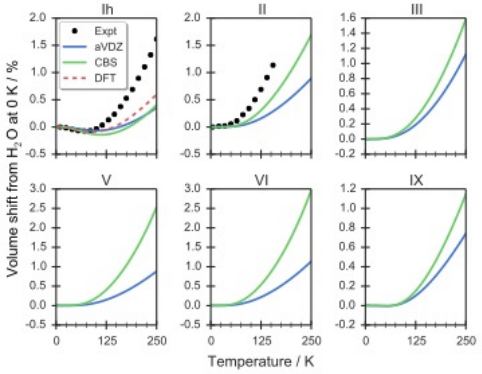

Figure 1 Thermal expansion of ice phases.

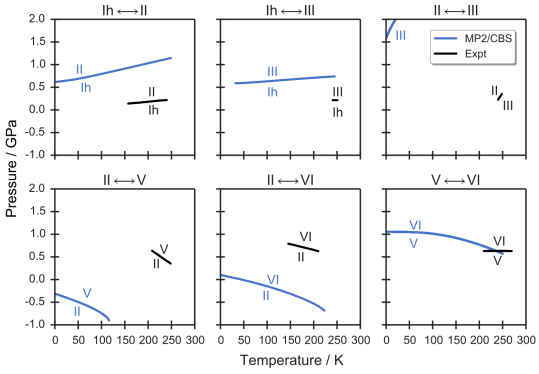

Figure 2 The *P-T* curves of Gibbs energy equivalence of ice phases.

Combining the thermal behaviors with the internal energies obtained directly from MP2, we have determined the phase boundaries of two ice phases as shown in Figure 2. The blue curves are the loci of the MP2/CBS Gibbs energy equivalence of two phases, whereas the black

curves are the experimental phase boundaries. In each case, the correct phase ordering and slope are reproduced by MP2/CBS (therefore, satisfying the Clausius–Clapeyron equation), although the transition pressures are significantly off.

The calculated phase diagram, as depicted in Figure 3, therefore, has a limited success in reproducing the observed, even when the basis is brought to the CBS limit. Phase II is destabilized relative to phases III and V, as already seen in Figure 2, causing the former to be over-shadowed by the latter. This along with MP2 studies of liquid water clearly indicates the need to go far beyond MP2, perhaps CCSD(T) or higher, to achieve quantitative phase diagram, in which chemical accuracy in condensed-phase Gibbs energy is required. Such calculations are planned with the Blue Waters resources this year.

(2) **Embedded-fragment valence bond theory.** We introduced a new embedded-fragment method for molecular cluster ions. It expresses the total wave function of such an ion as a linear combination of wave functions of many-electron states in each of which the charge is localized on one of the fragments described by the standard *ab initio* molecular orbital (MO) theory. The expansion coefficients are determined as an eigenvector of the Hamiltonian matrix of which the diagonal and off-diagonal elements are obtained by *ab initio* MO calculations for fragments with a charge-local initial guess (available in many software such as NWCHEM and GAUSSIAN) and by running self-consistent-field (SCF) cycle only once. All of these fragment energies are computed in the presence of a self-consistent electrostatic embedding field. The energy of the whole cluster ion is then determined as the lowest eigenvalue.

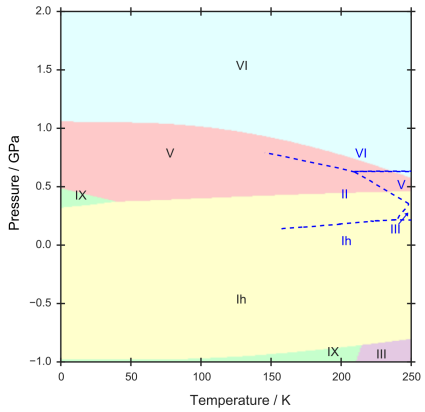

Figure 3 Phase diagram of ice (observed in blue).

We showed that this ansatz works well in reproducing the brute-force *ab initio* electron-correlated calculations of cluster ions such as He_n^+, $[Na(H_2O)_3]^+$, $(C_2H_4)_n^+$, $(H_2O)_4^+$, etc. With this good initial assessment, we believe that this newly developed method is ready for applications to cluster ions of real chemical interests such as vibrational spectra of size-selected protonated water clusters at a high level of *ab initio* electron-correlation theory. Such calculations will be performed on Blue Waters.

(3) **Universal correlation cusp.** Using the highly scalable Monte Carlo explicitly correlated MP2 method, we analyze the performance of 17 different correlation factors for total energies. We find that highly

performing correlation factors have near-universal shape and size in the short range of electron-electron distance ($0 < r_{12} < 1.5$ a.u.) (Figure 4). The long-range behavior ($r_{12} > 1.5$ a.u.) is insignificant insofar as the factor becomes near constant, leaving an orbital expansion to describe decoupled electrons. An analysis based on a low-rank Taylor expansion of the correlation factor seems limited, except that a negative second derivative with the value of around –1.3 a.u. correlates with high performance. This work was published in *Chem. Phys. Lett.* in 2017.

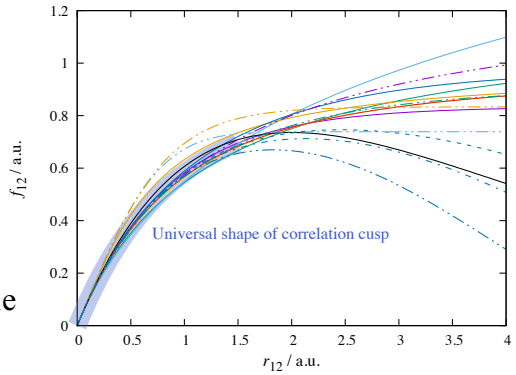

Figure 4 High-performing correlation factors sharing the same short-range shape.

(4) **Monte Carlo explicitly correlated GF2.** Green's function (GF) theory can directly compute electron-detachment and attachment energies of a molecule and quasiparticle energy bands of a solid, the key parameters of many chemical and materials processes. Its *ab initio* implementations, such as second-order GF (GF2) method, however, suffer from an extremely slow convergence of the results with the size of one-electron basis sets. We thus developed a Monte Carlo explicitly correlated GF2 method, in which a basis function (correlation factor) that depends on two-electron positions, or inter-electronic distance is additionally used to vastly accelerate the convergence. Its stochastic algorithm based on Metropolis Monte Carlo integrations makes the method highly scalable with respect to both computer size and system size. We have completed the formulation and coding as well as initial tests of the algorithm and we will perform large-scale parallel calculations of ionization energies of C_{60}, C_{70}, etc., using Blue Waters.

- **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, *The Journal of Chemical Physics* **144**, 204503 (2016), "Ice Ih anomalies: Thermal contraction, anomalous volume isotope effect, and pressure-induced amorphization."

S. Hirata, K. Gilliard, X. He, M. Keçeli, J. Li, M. A. Salim, O. Sode, and K. Yagi, an invited book chapter in *Fragmentation: Toward Accurate Calculations on Complex Molecular Systems* edited by M. S. Gordon (Wiley, Chichester, 2017), "*Ab initio* ice, dry ice, and liquid water."

A. E. Doran and S. Hirata, *Journal of Chemical Theory and Computation* **12**, 4821-4932 (2016), "Monte Carlo MP2 on many graphical processing units."

C. M. Johnson, A. E. Doran, J. Zhang, E. F. Valeev, and S. Hirata, *The Journal of Chemical Physics* **145**, 154115 (2016), "Monte Carlo explicitly correlated second-order many-body perturbation theory."

C. M. Johnson, S. Hirata, and S. Ten-no, *Chemical Physics Letters* [A. H. Zewail Commemorative Issue] **683**, 247-252 (2017), "Correlation factors in explicitly correlated methods."

M. A. Salim and S. Hirata, "Ab initio phase diagram of ice Ih, II, III, V, VI, and IX," in preparation (2018).

M. A. Salim and S. Hirata, "Embedded-fragment VB theory," in preparation (2018).

C. M. Johnson, A. E. Doran, S. Ten-no, and S. Hirata, "Monte Carlo explicitly correlated second-order many-body Green's function theory," in preparation (2018).

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); Low-Scaling and Unconventional Electronic Structure Techniques Conference (2016); Molecular Quantum Mechanics 2016; International Symposium of Theoretical Chemical Physics (2016); Molecular Electronic Structure in Buenos Aires (2016); Southeast Regional ACS Meeting (2016); Seminar, Institute for Molecular Science (2016); Seminar, Kobe University (2016); Seminar, Waseda University (2016); Seminar, Keio University (2016); Sanibel Symposium (2017); ACS National Meeting in San Francisco (2017); Seminar, BU-Harvard-MIT (2017); Molecular Science of Ice Symposium (2017); Department of Energy Chemical Theory and Computation Meeting (2017); Canadian Society of Chemistry National Meeting (2017); ACS National Meeting in Washington DC (2017); WATOC (2017); Waterloo Chemical Physics Symposium (2017); XVI Reunion Mexicana de Fiscoquímica Téorica (2017).

- **Plan for next year**

  We request 200,000 node-hours (100,000 node-hours on XE and 100,000 node-hours on XK).

  Grand-challenge applications of Monte Carlo MP and Green's function (GF) calculations for large molecules will be performed with the 100,000 XK node-hours using the GPU-based algorithm developed in this project with a demonstrated speedup in excess of 2,700-fold on 256 XK nodes relative to 1 XE node. We will also use the additional 50,000 XE node-hours for calculations of newer Monte Carlo MP/GF methods that are to be developed this year. We will use the remaining 50,000 XE node-hours on embedded-fragment calculations on ices and liquid water at the coupled-cluster levels, and for other molecular liquids and the development of multi-time-step approach.

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