COVER SHEET FOR PROPOSAL TO THE NATIONAL SCIENCE FOUNDATION

PROGRAM ANNOUNCEMENT/SOLICITATION NO./DUE DATE	☐ Special Exception to Deadline Date Policy	FOR NSF USE ONLY
PD 09-6881 **10/02/17**		**NSF PROPOSAL NUMBER**

FOR CONSIDERATION BY NSF ORGANIZATION UNIT(S) (Indicate the most specific unit known, i.e. program, division, etc.)

CHE - Theory, Models, Comput. Method

DATE RECEIVED	NUMBER OF COPIES	DIVISION ASSIGNED	FUND CODE	DUNS# (Data Universal Numbering System)	FILE LOCATION
				041544081	

EMPLOYER IDENTIFICATION NUMBER (EIN) OR TAXPAYER IDENTIFICATION NUMBER (TIN)	SHOW PREVIOUS AWARD NO. IF THIS IS ☒ A RENEWAL ☐ AN ACCOMPLISHMENT-BASED RENEWAL	IS THIS PROPOSAL BEING SUBMITTED TO ANOTHER FEDERAL AGENCY? YES ☐ NO ☒ IF YES, LIST ACRONYM(S)
376000511	**1361586**	

NAME OF ORGANIZATION TO WHICH AWARD SHOULD BE MADE	ADDRESS OF AWARDEE ORGANIZATION, INCLUDING 9 DIGIT ZIP CODE
University of Illinois at Urbana-Champaign	**University of Illinois at Urbana-Champaign** **SUITE A** **CHAMPAIGN, IL. 618207473**
AWARDEE ORGANIZATION CODE (IF KNOWN) **0017756000**	

NAME OF PRIMARY PLACE OF PERF	ADDRESS OF PRIMARY PLACE OF PERF, INCLUDING 9 DIGIT ZIP CODE
University of Illinois at Urbana-Champaign	**University of Illinois at Urbana-Champaign** **IL ,618207473 ,US.**

IS AWARDEE ORGANIZATION (Check All That Apply)	☐ SMALL BUSINESS	☐ MINORITY BUSINESS	☐ IF THIS IS A PRELIMINARY PROPOSAL THEN CHECK HERE
	☐ FOR-PROFIT ORGANIZATION	☐ WOMAN-OWNED BUSINESS	

TITLE OF PROPOSED PROJECT **Fast Methods for Strong and Weak Electron Correlation**

REQUESTED AMOUNT	PROPOSED DURATION (1-60 MONTHS)	REQUESTED STARTING DATE	SHOW RELATED PRELIMINARY PROPOSAL NO. IF APPLICABLE
$ **497,427**	**36** months	**06/01/18**	

THIS PROPOSAL INCLUDES ANY OF THE ITEMS LISTED BELOW

☐ BEGINNING INVESTIGATOR

☐ DISCLOSURE OF LOBBYING ACTIVITIES

☐ PROPRIETARY & PRIVILEGED INFORMATION

☐ HISTORIC PLACES

☐ VERTEBRATE ANIMALS IACUC App. Date _____

     PHS Animal Welfare Assurance Number _____

☒ TYPE OF PROPOSAL   **Research**

☐ HUMAN SUBJECTS   Human Subjects Assurance Number _____

     Exemption Subsection _____ or IRB App. Date _____

☐ INTERNATIONAL ACTIVITIES: COUNTRY/COUNTRIES INVOLVED

_____

☒ COLLABORATIVE STATUS

     Not a collaborative proposal

PI/PD DEPARTMENT **Department of Chemistry**	PI/PD POSTAL ADDRESS **600 South Mathews Avenue**
PI/PD FAX NUMBER **217-244-3186**	**Urbana, IL 61801** **United States**

NAMES (TYPED)	High Degree	Yr of Degree	Telephone Number	Email Address
PI/PD NAME **So Hirata**	**PhD**	**1998**	**217-244-0629**	**sohirata@illinois.edu**
CO-PI/PD				
CO-PI/PD				
CO-PI/PD				
CO-PI/PD				

CERTIFICATION PAGE

Certification for Authorized Organizational Representative (or Equivalent) or Individual Applicant

By electronically signing and submitting this proposal, the Authorized Organizational Representative (AOR) or Individual Applicant is: (1) certifying that statements made herein are true and complete to the best of his/her knowledge; and (2) agreeing to accept the obligation to comply with NSF award terms and conditions if an award is made as a result of this application. Further, the applicant is hereby providing certifications regarding conflict of interest (when applicable), drug-free workplace, debarment and suspension, lobbying activities (see below), nondiscrimination, flood hazard insurance (when applicable), responsible conduct of research, organizational support, Federal tax obligations, unpaid Federal tax liability, and criminal convictions as set forth in the NSF Proposal & Award Policies & Procedures Guide (PAPPG). Willful provision of false information in this application and its supporting documents or in reports required under an ensuing award is a criminal offense (U.S. Code, Title 18, Section 1001).

Certification Regarding Conflict of Interest

The AOR is required to complete certifications stating that the organization has implemented and is enforcing a written policy on conflicts of interest (COI), consistent with the provisions of PAPPG Chapter IX.A.; that, to the best of his/her knowledge, all financial disclosures required by the conflict of interest policy were made; and that conflicts of interest, if any, were, or prior to the organization's expenditure of any funds under the award, will be, satisfactorily managed, reduced or eliminated in accordance with the organization's conflict of interest policy. Conflicts that cannot be satisfactorily managed, reduced or eliminated and research that proceeds without the imposition of conditions or restrictions when a conflict of interest exists, must be disclosed to NSF via use of the Notifications and Requests Module in FastLane.

Drug Free Work Place Certification

By electronically signing the Certification Pages, the Authorized Organizational Representative (or equivalent), is providing the Drug Free Work Place Certification contained in Exhibit II-3 of the Proposal & Award Policies & Procedures Guide.

Debarment and Suspension Certification   (If answer "yes", please provide explanation.)

Is the organization or its principals presently debarred, suspended, proposed for debarment, declared ineligible, or voluntarily excluded from covered transactions by any Federal department or agency?   Yes ☐   No ☒

By electronically signing the Certification Pages, the Authorized Organizational Representative (or equivalent) or Individual Applicant is providing the Debarment and Suspension Certification contained in Exhibit II-4 of the Proposal & Award Policies & Procedures Guide.

Certification Regarding Lobbying

This certification is required for an award of a Federal contract, grant, or cooperative agreement exceeding $100,000 and for an award of a Federal loan or a commitment providing for the United States to insure or guarantee a loan exceeding $150,000.

Certification for Contracts, Grants, Loans and Cooperative Agreements

The undersigned certifies, to the best of his or her knowledge and belief, that:
(1) No Federal appropriated funds have been paid or will be paid, by or on behalf of the undersigned, to any person for influencing or attempting to influence an officer or employee of any agency, a Member of Congress, an officer or employee of Congress, or an employee of a Member of Congress in connection with the awarding of any Federal contract, the making of any Federal grant, the making of any Federal loan, the entering into of any cooperative agreement, and the extension, continuation, renewal, amendment, or modification of any Federal contract, grant, loan, or cooperative agreement.
(2) If any funds other than Federal appropriated funds have been paid or will be paid to any person for influencing or attempting to influence an officer or employee of any agency, a Member of Congress, an officer or employee of Congress, or an employee of a Member of Congress in connection with this Federal contract, grant, loan, or cooperative agreement, the undersigned shall complete and submit Standard Form-LLL, "Disclosure of Lobbying Activities," in accordance with its instructions.
(3) The undersigned shall require that the language of this certification be included in the award documents for all subawards at all tiers including subcontracts, subgrants, and contracts under grants, loans, and cooperative agreements and that all subrecipients shall certify and disclose accordingly.

This certification is a material representation of fact upon which reliance was placed when this transaction was made or entered into. Submission of this certification is a prerequisite for making or entering into this transaction imposed by section 1352, Title 31, U.S. Code. Any person who fails to file the required certification shall be subject to a civil penalty of not less than $10,000 and not more than $100,000 for each such failure.

Certification Regarding Nondiscrimination

By electronically signing the Certification Pages, the Authorized Organizational Representative (or equivalent) is providing the Certification Regarding Nondiscrimination contained in Exhibit II-6 of the Proposal & Award Policies & Procedures Guide.

Certification Regarding Flood Hazard Insurance

Two sections of the National Flood Insurance Act of 1968 (42 USC §4012a and §4106) bar Federal agencies from giving financial assistance for acquisition or construction purposes in any area identified by the Federal Emergency Management Agency (FEMA) as having special flood hazards unless the:
(1)   community in which that area is located participates in the national flood insurance program; and
(2)   building (and any related equipment) is covered by adequate flood insurance.

By electronically signing the Certification Pages, the Authorized Organizational Representative (or equivalent) or Individual Applicant located in FEMA-designated special flood hazard areas is certifying that adequate flood insurance has been or will be obtained in the following situations:
(1)   for NSF grants for the construction of a building or facility, regardless of the dollar amount of the grant; and
(2)   for other NSF grants when more than $25,000 has been budgeted in the proposal for repair, alteration or improvement (construction) of a building or facility.

Certification Regarding Responsible Conduct of Research (RCR)
(This certification is not applicable to proposals for conferences, symposia, and workshops.)

By electronically signing the Certification Pages, the Authorized Organizational Representative is certifying that, in accordance with the NSF Proposal & Award Policies & Procedures Guide, Chapter IX.B. , the institution has a plan in place to provide appropriate training and oversight in the responsible and ethical conduct of research to undergraduates, graduate students and postdoctoral researchers who will be supported by NSF to conduct research. The AOR shall require that the language of this certification be included in any award documents for all subawards at all tiers.

CERTIFICATION PAGE - CONTINUED

Certification Regarding Organizational Support

By electronically signing the Certification Pages, the Authorized Organizational Representative (or equivalent) is certifying that there is organizational support for the proposal as required by Section 526 of the America COMPETES Reauthorization Act of 2010. This support extends to the portion of the proposal developed to satisfy the Broader Impacts Review Criterion as well as the Intellectual Merit Review Criterion, and any additional review criteria specified in the solicitation. Organizational support will be made available, as described in the proposal, in order to address the broader impacts and intellectual merit activities to be undertaken.

Certification Regarding Federal Tax Obligations

When the proposal exceeds $5,000,000, the Authorized Organizational Representative (or equivalent) is required to complete the following certification regarding Federal tax obligations. By electronically signing the Certification pages, the Authorized Organizational Representative is certifying that, to the best of their knowledge and belief, the proposing organization:
(1) has filed all Federal tax returns required during the three years preceding this certification;
(2) has not been convicted of a criminal offense under the Internal Revenue Code of 1986; and
(3) has not, more than 90 days prior to this certification, been notified of any unpaid Federal tax assessment for which the liability remains unsatisfied, unless the assessment is the subject of an installment agreement or offer in compromise that has been approved by the Internal Revenue Service and is not in default, or the assessment is the subject of a non-frivolous administrative or judicial proceeding.

Certification Regarding Unpaid Federal Tax Liability

When the proposing organization is a corporation, the Authorized Organizational Representative (or equivalent) is required to complete the following certification regarding Federal Tax Liability:

By electronically signing the Certification Pages, the Authorized Organizational Representative (or equivalent) is certifying that the corporation has no unpaid Federal tax liability that has been assessed, for which all judicial and administrative remedies have been exhausted or lapsed, and that is not being paid in a timely manner pursuant to an agreement with the authority responsible for collecting the tax liability.

Certification Regarding Criminal Convictions

When the proposing organization is a corporation, the Authorized Organizational Representative (or equivalent) is required to complete the following certification regarding Criminal Convictions:

By electronically signing the Certification Pages, the Authorized Organizational Representative (or equivalent) is certifying that the corporation has not been convicted of a felony criminal violation under any Federal law within the 24 months preceding the date on which the certification is signed.

Certification Dual Use Research of Concern

By electronically signing the certification pages, the Authorized Organizational Representative is certifying that the organization will be or is in compliance with all aspects of the United States Government Policy for Institutional Oversight of Life Sciences Dual Use Research of Concern.

AUTHORIZED ORGANIZATIONAL REPRESENTATIVE	SIGNATURE	DATE
NAME		
TELEPHONE NUMBER	EMAIL ADDRESS	FAX NUMBER

B. Project Summary: *Fast Methods for Strong and Weak Electron Correlation*

Overview: Electron-correlation problems are divided into two broad categories: strong and weak correlation. Weakly correlated systems are the ones describable well by an *ab initio* mean-field theory, e.g., restricted Hartree-Fock (RHF) theory, and the remaining correlation energy can be recuperated by coupled-cluster (CC) or many-body perturbation (MP) theory. Strongly correlated systems are, in contrast, not adequately described by RHF and subsequent CC or MP series are slowly convergent (or even divergent). The structure of weakly correlated wave functions is considered to be well understood with a roadmap to their predictive calculations laid out and waiting to be deployed; the research tasks, therefore, are to enable such calculations for large or infinitely extended systems including liquids. On the other hand, strongly correlation is still poorly understood and there seems no unambiguous definition of it. Research towards understanding the fundamental aspects including the definition of strong correlation is needed. Proposed herein is a two-prong research project on both strong and weak electron correlation to broadly advance the frontiers of electronic structure theory.

Intellectual Merit: For weak correlation, the PI will fully develop embedded-fragment methods to achieve routine ability to apply CC with singles, doubles, and perturbative triples [CCSD(T)] for molecular solids in the complete-basis-set (CBS) limit and for molecular liquids in *ab initio* molecular dynamics (MD). A specific target is a quantitative *ab initio* computational reproduction of ice phase diagram, for which second-order MP (MP2) in the CBS limit was found to fall short in our predecessor NSF project. Another is the simultaneous and *ab initio* simulations of the structural, thermodynamic, and spectroscopic properties of liquid water, the successful outcome of which will firmly establish the third era of liquid MD after empirical force field and Car-Parrinello DFT. The PI will also begin chemical applications of the embedded-fragment methods for cluster ions, developed by the predecessor NSF project.

For strong correlation, the PI will prove or disprove the hypothesis that all ground-state "strongly correlated" problems (such as bond breaking, polyradicals, metallic solids) can be reduced to weakly correlated ones by fully relaxing the RHF wave function into unrestricted HF (UHF) or complex general HF (GHF) and thereby fully minimizing the mean-field energy. The PI will develop an efficient search algorithm of a global minimum of a HF energy surface in the orbital coefficient space by borrowing the algorithmic idea from the Global Reaction Route Mapping (GRRM) method of Maeda, Ohno, and Morokuma. We will also map HF instabilities of 1D, 2D, and 3D homogeneous electron gases and of 2D hydrogen sheet of different temperature and dopant concentration as well as with phonon displacements. The latter informs about a potential mechanism of 2D high-T_c superconductivity such as in cuprates.

Broader Impacts: The PI plans to make undergraduate and graduate pedagogical innovations and a free, public release of software for molecular solids and liquids developed in this project. For the undergraduate education, the PI will continue refining the online video lectures of quantum chemistry and spectroscopies used in a highly engaging inverted course. They educate an untold number of self-motivated students in addition to registered students. For graduate education, the PI will fully develop the course for fellowship preparation using NSF GRFP as a template and intended to increase students' awareness of the importance of writing skills, cultivating a network of supporters, and growing "thick skins" in the face of defeats (in the form of rejections of fellowship applications, paper submissions, or job applications). The embedded-fragment methods have the unmatched advantage of high interoperability requiring no modification or special capabilities in the backend electronic structure software, many independent jobs of which are executed in parallel.

TABLE OF CONTENTS

For font size and page formatting specifications, see PAPPG section II.B.2.

	Total No. of Pages	Page No.* (Optional)*
Cover Sheet for Proposal to the National Science Foundation		
Project Summary (not to exceed 1 page)	1	
Table of Contents	1	
Project Description (Including Results from Prior NSF Support) (not to exceed 15 pages) **(Exceed only if allowed by a specific program announcement/solicitation or if approved in advance by the appropriate NSF Assistant Director or designee)**	15	
References Cited	7	
Biographical Sketches (Not to exceed 2 pages each)	2	
Budget (Plus up to 3 pages of budget justification)	5	
Current and Pending Support	1	
Facilities, Equipment and Other Resources	1	
Special Information/Supplementary Documents (Data Management Plan, Mentoring Plan and Other Supplementary Documents)	2	
Appendix (List below.) **(Include only if allowed by a specific program announcement/ solicitation or if approved in advance by the appropriate NSF Assistant Director or designee)**		

Appendix Items:

*Proposers may select any numbering mechanism for the proposal. The entire proposal however, must be paginated. Complete both columns only if the proposal is numbered consecutively.

C. Project Description: *Fast Methods for Strong and Weak Electron Correlation*
1. Results from Prior NSF Support: *Computational Chemistry of Clusters and Crystals*
(CHE-1361586; 6/1/14–5/31/18; $413,448)

This NSF CHE project, to which this is a renewal, has two main research objectives: methodological developments and applications. It aims at extending the embedded-fragment methods[1-3] for molecular clusters and crystals to molecular cluster ions and molecular crystals with partial charge transfer. It also plans to apply the embedded-fragment methods to *ab initio* phase diagram of ice, which amounts to a free energy calculation taking into account the effect of thermal expansion. We have accomplished both goals and additionally made a breakthrough in the form of one of the first *ab initio* molecular dynamics simulations of liquid water at the *ab initio* second-order many-body perturbation (MP2) theory level. We consider the latter to mark the arrival of a new era of liquid simulation using forces computed on the fly from a fully *ab initio* wave function theory, leaving behind the earlier eras based on fitted empirical force fields and Car-Parrinello density-functional theory (DFT) for liquids.

This NSF CHE project also made several outreach and educational contributions. They include introducing a new graduate course, enhancing undergraduate pedagogy, and postdoctoral mentorship. These are summarized in the following:

Intellectual Merit. Six peer-reviewed journal articles,[4-9] including a JCP Editors' Pick,[9] two reviews,[2,10] and a book chapter[11] resulted from this project (some of them acknowledging its predecessor NSF CHE project number). One more paper[12] is currently under review, and two other papers[13,14] are in the final stage of preparation. The following are some of the highlights:

Figure 1 MP2-caluculated and observed vibrational spectra of ice VIII.

MP2 and CCSD study on ice VIII.[4] Ice VIII is a high-pressure phase of water ice, which is proton-ordered and much denser than the usual ice Ih. It is suspected to exist in Uranus, Neptune, Jovian, and Saturnian satellites as well as in the deep ocean floor or upper mantle of the Earth. Together with ices VII and X, it occupies a large area of the phase diagram of water.

In this work supported by NSF, the structure, equation of state, IR, Raman, and inelastic neutron scattering (INS) spectra of ice VIII were computed *ab initio* at the MP2 or coupled-cluster singles and doubles (CCSD) level. Nearly all the observed features of the spectra (Figure 1) and pressure-dependence of the structure and spectra were reproduced computationally up to 60 GPa insofar as the anharmonic effects are insignificant. The calculated INS spectra corrected a systematic error in the peak positions of the observed spectra.

The calculations also shed light on the phase behavior of ice VIII and related controversies. The calculations did not support the hypothetical isostructural transition in 2-3 GPa to phase VIII′, the existence of which is a matter of controversy for over a decade. They do not contradict the interpretation of the spectral anomaly at 10-14 GPa as a precursor of the VIII-X phase transition, however.

The paper reporting this result was published in *J. Chem. Phys.*[4]

MP2 study on thermal expansion of solid carbon dioxide.[6] The embedded-fragment *ab initio* MP2 method is applied to an infinite three-dimensional crystal of CO_2 phase I using the aug-cc-pVDZ and aug-cc-pVTZ basis sets, the latter in conjunction with a counterpoise correction for the basis-set
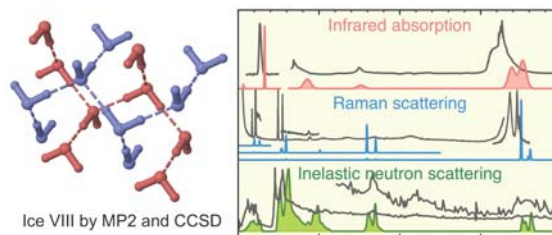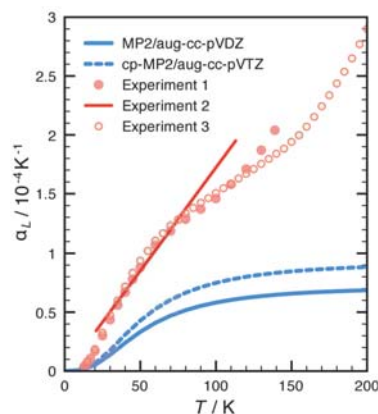

Figure 2 Linear thermal expansion coefficients of CO_2-I as a function of temperature.

superposition error. The equation of state, phonon frequencies, bulk modulus, heat capacity, Grüneisen parameter, and molar volume were reproduced accurately. We found, however, it exceedingly difficult to determine the bulk modulus correctly, the computed value of which deviates from the observed value by 50-100%. As a result, the thermal expansion coefficient calculated by MP2 is systematically too low (Figure 2), while having the correct qualitative temperature dependence. This is ascribed to the overestimation of the weak interactions by MP2.

These results were reported in *J. Chem. Theory Comput.*[6]

MP2 study on liquid water.[7] A direct, simultaneous calculation of properties of a liquid using an *ab initio* electron-correlated theory has long been unthinkable. We presented structural, dynamical, and response properties of liquid water calculated by *ab initio* molecular dynamics using the embedded-fragment spin-component-scaled MP2 (SCS-MP2) method with the aug-cc-pVDZ basis set. This level of theory is chosen as it accurately and inexpensively reproduces the water dimer potential ener-

Figure 3 The calculated and observed IR and Raman spectra of liquid water. The calculated frequencies are scaled by 0.93.

gy surface from the coupled-cluster singles, doubles, and noniterative triples [CCSD(T)] 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 experimental results (Figure 3). 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 dynamic features of the environment, which are difficult to probe experimentally, such as a surprisingly large fluctuation in the coordination number (Figure 4) and the detailed mechanism by which the hydrogen donating water molecules move across the first and second shells, thereby causing this fluctuation.
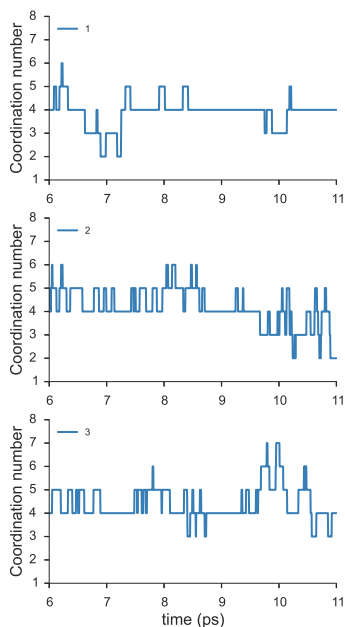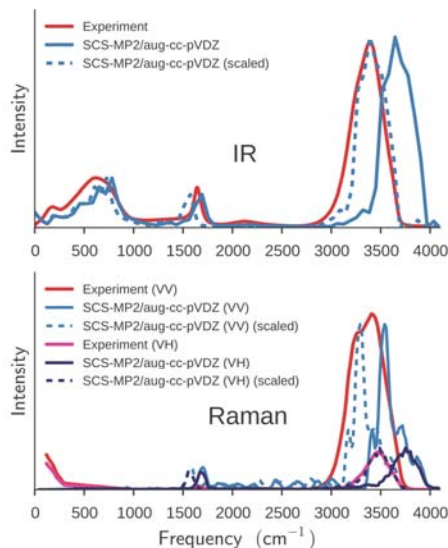
This work was published in *Sci. Rep.*[7]

Why is MP2-water "denser" and "cooler" than DFT-water?[8] DFT with a dispersion-less generalized gradient approximation (GGA) needs much higher temperature and pressure than the ambient conditions to maintain water in the liquid phase at the correct (1 g/cm^3) density. Conversely, *ab initio* MP2 calculations of liquid water require lower temperature and pressure than DFT to keep water liquid.

In this work, we presented a unifying explanation of these trends derived from classical water simulations using a polarizable force field with different sets of parameters. We showed that the different temperatures and pressures between DFT and MP2 at which the simulated water displays the experimentally observed liquid structure under the ambient conditions can be largely explained by their differences in polarizability and dispersion interaction, respectively. In DFT, the polarizability and thus the induced dipole moments and the hydrogen-bond strength are all overestimated. This hinders the rotational motion of molecules and requires a higher temperature for DFT-

Figure 4 Time evolution of the coordination number for three selected water molecules.

water to be liquid. MP2 gives a stronger dispersion interaction and thus shorter intermolecular distances than dispersion-less DFT, which is why MP2-water is denser than DFT-water under the same external pressure.

This work was published in *J. Phys. Chem. Lett.*[8]

MP2 study on ice Ih anomalies: Thermal contraction, anomalous volume isotope effect, and pressure-induced amorphization.[9] Ice Ih displays several anomalous thermo-dynamic properties such as thermal contraction at low temperatures, an anomalous volume isotope effect (VIE) rendering the volume of D_2O ice greater than that of H_2O ice, and a pressure-induced transition to the high-density amorphous (HDA) phase. Furthermore, the anomalous VIE increases with temperature in spite of its quantum-mechanical origin.

In this study, *ab initio* MP2 theory in the quasi-harmonic approximation (QHA) is applied to the Gibbs energy of an infinite, proton-disordered crystal of ice Ih at wide ranges of temperatures and pressures. The quantum effect of nuclei moving in an-harmonic potentials is taken into account from first principles without any empirical or nonsystematic approximation to either the electronic or vibrational Hamiltonian. MP2 predicts quantitatively correctly the thermal contraction at low temperatures, which is confirmed to originate from the volume-contracting hydrogen-bond bending modes (acoustic phonons) (Figure 5). It qualitatively reproduces (but underestimates) the thermal expansion at higher tempera-tures, caused by the volume expanding hydrogen-

Figure 5 Change in the vibrationally cor-rected volumes of H_2O and D_2O ice Ih as a function of temperature.

bond stretching (and to a lesser extent librational) modes. The anomalous VIE is found to be the result of subtle cancellations among closely competing isotope effects on volume from all modes. Consequently, even *ab initio* MP2 with the aug-cc-pVDZ and aug-cc-pVTZ basis sets has difficulty reproducing this anomaly, yielding qualitatively varied predictions of the sign of the VIE depending on such computational details as the choice of the embedding field. However, the temperature growth of the anomalous VIE is repro-duced robustly and is ascribed to the librational modes (Figure 5). These solid-state MP2 calcula-tions, as well as MP2 Born–Oppenheimer molecu-lar dynamics, find a volume collapse and a loss of symmetry and long-range order in ice Ih upon pres-sure loading of 2.35 GPa or higher. Concomitantly, rapid softening of acoustic phonons is observed starting around 2 GPa. They constitute a computa-tional detection of a mechanical instability in ice Ih and the resulting pressure-induced amorphization to HDA (Figure 6).
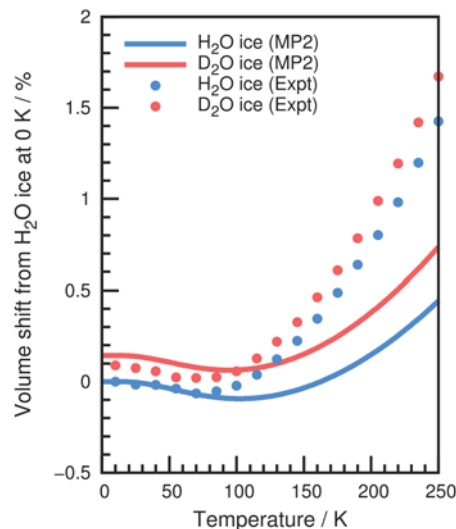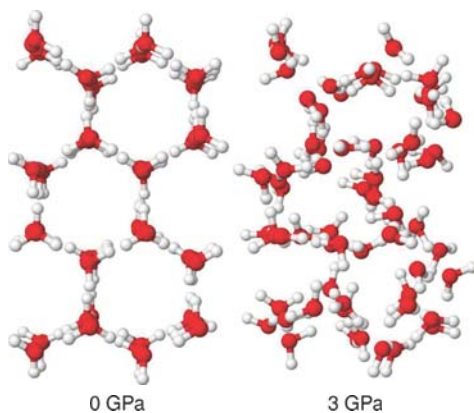

Figure 6 Snapshot structures of crystalline (left) and partially amorphous (right) ices dur-ing geometry optimization under pressure.

The paper reporting this result was published in *J. Chem. Phys.*[9] and selected as a JCP Editors' Pick.

3

Ab initio phase diagrams of ice.[13] As promised in this proposal, we performed an *ab initio* 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 approximation (QHA) for the thermal expansion effect. Calculated thermal expansion behaviors of the six phases (Figure 7) 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 (this has been thoroughly analyzed by aforementioned project,[9] also supported by NSF).
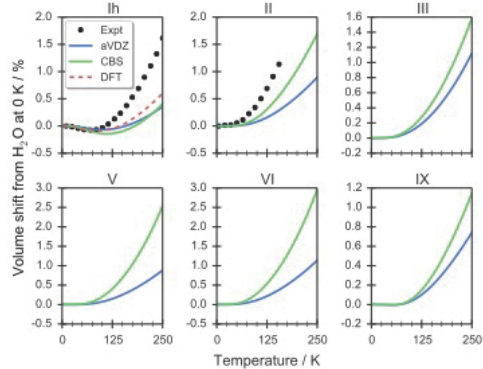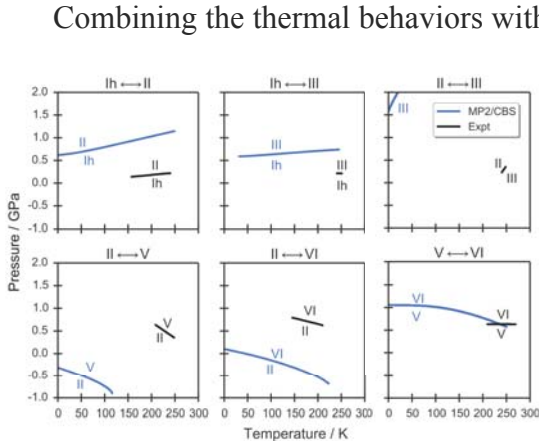

Figure 7 Thermal expansion of ice phases.

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 8. 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 9, 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 8, causing the former to be overshadowed by the latter. This along with MP2 studies of liquid water[7,8] 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.

Figure 8 The *P-T* curves of Gibbs energy equivalence of ice phases.

Embedded-fragment valence bond theory.[14] As promised in this proposal, we introduced a new embedded-fragment method for molecular cluster ions. It expresses the total wave function of such an ion as

$$|\Psi\rangle = \sum_{i=1}^{N} C_i |\Phi_i\rangle$$

Figure 9 Phase diagram of ice (observed in blue).

where $|\Phi_i\rangle$ is a many-electron state in which the charge is localized on the *i*th fragment and is composed of nonorthogonal localized molecular orbitals (MOs) of the fragment. The expansion coefficients C_i are determined as an eigenvector of the Hamiltonian matrix $\mathbf{H}$,

$$\mathbf{H} = \begin{pmatrix} \langle\Phi_1|\hat{H}|\Phi_1\rangle & \dots & \langle\Phi_1|\hat{H}|\Phi_N\rangle \\ \vdots & \ddots & \vdots \\ \langle\Phi_N|\hat{H}|\Phi_1\rangle & \dots & \langle\Phi_N|\hat{H}|\Phi_N\rangle \end{pmatrix}$$

where the diagonal and off-diagonal elements are obtained entirely by some *ab initio* MO calculations for fragments as

$$\langle \Phi_i | \hat{H} | \Phi_i \rangle = E_i^+ + \sum_{j \neq i} (E_j + {}^{\text{HF}} E_{ij}^{(i+)} - {}^{\text{HF}} E_i^+ - {}^{\text{HF}} E_j) + \sum_{j < k \neq i} (E_{jk} - E_j - E_k)$$

$$\langle \Phi_i | \hat{H} | \Phi_j \rangle = -\sqrt{(E_{ij}^+ - \epsilon_i)(E_{ij}^+ - \epsilon_j)} \quad ; \quad \epsilon_i = E_i^+ + \sum_{j \neq i} (E_j + {}^{\text{HF}} E_{ij}^{(i+)} - {}^{\text{HF}} E_i^+ - {}^{\text{HF}} E_j)$$

in the case of a cation. $E_i^+, E_j, E_{ij}, {}^{\text{HF}} E_{ij}^{(i+)}$ denote an *ab initio* electron-correlated energy of the *i*th cationic fragment, *j*th neutral fragment, *ij*th neutral fragment dimer, and a HF energy of the *ij*th cationic fragment dimer with the charge localized on the *i*th fragment. The last quantity can be obtained with a charge-local initial guess (available in many software such as NWCHEM and GAUSSIAN) and running self-consistent-field (SCF) cycle only once. All of these fragment energies are obtained in the presence of a self-consistent electrostatic embedding field. The energy of the whole cluster ion is then obtained as the lowest eigenvalue.

We showed that this ansatz works well in reproducing the brute-force *ab initio* electron-correlated calculations of cluster ions such as He_n^+, $[\text{Na}(\text{H}_2\text{O})_3]^+$, $(\text{C}_2\text{H}_4)_n^+$, $(\text{H}_2\text{O})_4^+$, 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[15,16] at a high level of *ab initio* electron-correlation theory.

Broader Impacts. Highlights of the education-outreach activities are as follows:

Undergraduate education. The PI continues to develop online video lectures on quantum chemistry and spectroscopies for senior undergraduate students to educate not only UIUC students in the inverted course format, but also any self-motivated students who have access to Internet world over. Of 37 original video lectures the PI developed, 13 have been upgraded to have more professional quality using student narrators and better visual effects. The PI has been included in the List of Teachers Ranked as Excellent by Their Students in three years in a row (2015, 2016, and 2017), which lists top 10% highest-rated instructors of UIUC. The PI received 2017 SCS Teaching Award from the School of Chemical Sciences of UIUC.

Graduate education. The PI has developed a new course that assists first and second year graduate students write competitive fellowship applications in the style of NSF GRFP. This course, modelled after a similar course designed by Professor David Masiello and coworkers of University of Washington, Seattle, has been offered three times and received well by the students.

Dr. Olaseni Sode (former graduate student supported by this NSF grant) is now a tenure-track assistant professor at The University of Tampa and recipient of 2015 Ford Foundation Postdoctoral Fellowship, 2017 Jim Talman New Investigator Award, and 2017 McKnight Junior Faculty Development Fellowship. Mr. Jacob Faucheaux (a current graduate student and NSF GRFP winner) is a 2016 winner of the Chemical Computing Group Excellence Award of the American Chemical Society Division of Computers in Chemistry. Dr. Kandis Gilliard (a former graduate student supported by this NSF grant) is slated to start tenure-track assistant professorship at University of California, Riverside, after a postdoctoral stint at University of Pennsylvania.

Postdoctoral mentorship. The group's postdoctoral researcher, Dr. Jun Zhang, is one of the inaugural MolSSI postdoctoral fellows fully funded by NSF through Virginia Tech's Molecular Science Software Institute. Dr. Zhang has written a highly efficient two-electron integral code supported by this NSF grant and also by the MolSSI fellowship and submitted a single-authored paper on it. Dr. Jinjin Li (a former postdoc supported by this NSF grant) is now a professor at Shanghai Jiaotong University.

2. Proposed Research, Education, and Outreach: *Fast Methods for Strong and Weak Electron Correlation*

2.1. Background and Significance

The prevailing paradigm of *ab initio* electronic structure theory, which aims at solving the electronic Schrödinger equation for a molecule with predictive accuracy, is a two-step process. In the first step, a mean-field theory, usually restricted Hartree-Fock (RHF) theory, is applied, capturing the vast majority of the exact total energy, say its 99%. This is followed by the second step, in which a systematic electron-correlation theory is deployed which is to recuperate the remaining 1% or so of the exact total energy, known as the correlation energy. Size-consistent electron-correlation theories include coupled-cluster (CC) and many-body perturbation (MP) theories, which form systematic series of approximations with increasing accuracy and cost.

We may then divide electronic structure problems/systems into two broad categories. One with "weak" electron correlation can be handled well by the prevailing paradigm, in the sense that the CC or MP series are rapidly convergent with an RHF reference. The other has "strong" electron correlation, for which an RHF reference is qualitatively wrong and subsequent CC series may be less rapidly convergent and MP series may be even divergent. Strongly correlated systems include molecules undergoing bond breaking, open-shell systems such as diradicals and polyradicals (i.e., magnetic molecules) as well as non-Koopmans ionized or electron-attached states, excited states, transition metal complexes, multiply covalent bonded molecules, and other systems with small HOMO-LUMO gaps as well as extended systems with high density of states near the Fermi surface including metallic and superconducting solids. Alternative terminology is "dynamic" for weak and "static" for strong electron correlation.

The goal of the proposed research is to develop fast and predictively accurate ab initio electronic structure methods for both strong and weak electron correlation.

It may be said that quantum chemists now have a firm grasp on weak correlation problems. Not only do we have the aforementioned paradigm (based on RHF, CC, and MP, e.g.) to systematically approach an exact solution efficiently for a general weak correlation problem in a relatively small molecule, but we also have several powerful algorithms we can use to deploy these methods to large or infinitely extended systems. Both CC and MP theories are diagrammatically size-consistent and maintain uniform accuracy across system size up to infinity. The algorithms in question are local correlation methods originating from Sæbø and Pulay[17,18] and their various extensions[19,20] and also emerging stochastic algorithms[21-31] for CC and MP.

Further research into weak correlation problems, therefore, may focus on algorithmic (as opposed to ansatz) enhancement to enable applications to systems that are large and complex such as solids and liquids.

The strong correlation problem, in contrast, is still wide open and considered as a holy grail of today's *ab initio* electronic structure theory. In a strongly correlated system, the usual mean-field method, typically RHF, performs extremely poorly, capturing an inadequately small portion of the exact total energy. This leaves an electron-correlation method to have to work extra hard to capture a much greater remainder of the exact total energy. There are a number of highly creative ideas proposed to tackle this problem: the density matrix renormalization group (DMRG) method[32-38] and its generalization known as tensor network theory,[39-43] projector Monte Carlo[44-46] such as full configuration-interaction quantum Monte Carlo (FCIQMC),[47-54] internally contracted multireference configuration-interaction (MRCI) method,[55] internally contracted complete-active-space self-consistent-field (CASSCF) plus a second-order correction (CASPT2),[56-62] density-matrix embedding field theory (DMET),[63-69] etc.

A fresh approach to this problem was proposed by Scuseria's group[70-80] at Rice. They allow the mean-field theory to break the inherent symmetry of the Hamiltonian, thereby lowering the mean-field energy considerably. This leaves an electron correlation method with an easier task to capture a smaller remainder. In other words, *by more fully optimizing the mean-*

field wave function and minimizing its energy, a strong correlation problem can be converted into a weak correlation problem, which can be handled by CC or MP; a strong correlation problem may not be a "correlation" problem but rather a mean-field problem! Additionally, Scuseria and coworkers proceed to restore symmetry before the correlation treatment, which then makes the reference wave function multi-determinantal.

We find Scuseria's overall approach (if not their implementations) particularly attractive as it potentially constitutes a universal solution to strong correlation problems. Let us further expound on this notion by taking the H_2 bond breaking as a textbook example (Figure 10). As is well known, RHF does not break the simplest covalent bond in H_2 qualitatively correctly. At a bond length greater than 1.5 bohr, the RHF potential energy curve begins to rise and resists flattening unlike the exact (CCSD) curve. Consequently, the RHF curve ends at a wrong and nonphysical dissociation limit. Whereas the RHF theory recuperates a decent >95% of the total energy near the equilibrium bond length, it captures only ~80% at 4.0 bohr and progressively less at a longer bond length. As is also well known, when one switches from RHF to unrestricted Hartree-Fock

Figure 10 H_2 bond breaking

(UHF), one can break the same bond qualitatively correctly and even exactly at the dissociation limit (capturing 100% of the exact total energy). This begs the question: *is the massive electron correlation energy (20% of the total energy) at 4.0 bohr a "strong correlation" energy or a big piece of mean-field (thus not correlation) energy left over by the poor mean-field (RHF) method?* We elect to take the latter view: by more fully optimizing the HF solution (which in this case amounts to reaching a lower-energy UHF solution), correlation is weak throughout the bond length and even zero in the dissociation limit.
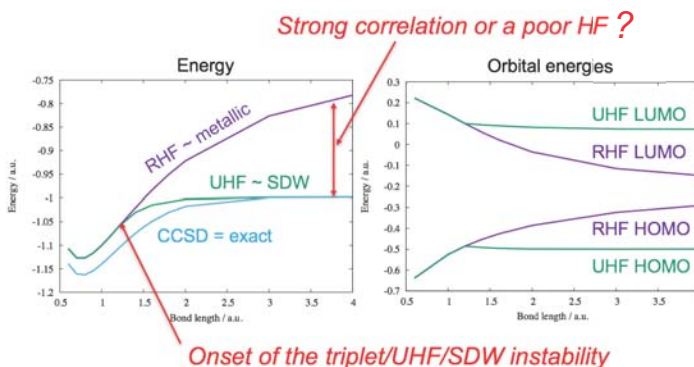
We submit a hypothesis that all strong correlation problems are illusory (let alone not unambiguously definable) in the sense that they exist only in a poor mean-field reference and can be converted into weak correlation problems (and thus generally solved) by more completely minimizing the mean-field energy.

An indication of the presence of strong correlation is a small HOMO-LUMO gap. The right panel of Figure 10 indeed shows that this gap continuously decreases with increasing bond length if and only if an RHF reference is used. However, with a UHF reference, the gap hardly varies with bond length at 1.2 bohr and remains large even in the dissociation limit. Therefore, this definition of strong correlation is rather shaky and varies depending on the reference.

Furthermore, we reported mathematical theorems[81] stating that molecular RHF solutions with form-degenerate HOMO-LUMO (including H_2 in the dissociation limit) always have certain triplet (UHF) and singlet (complex RHF) instabilities.[82-94] In other words, not only is a small HOMO-LUMO gap an unreliable indicator of strong correlation, it guarantees the presence of a lower-energy, symmetry-broken HF solution (of the UHF or complex RHF type), which compresses the correlation energy possibly down to a weak correlation regime. *These singlet and triplet instability theorems[87] provide circumstantial evidence supporting the above hypothesis for molecules with small HOMO-LUMO gaps*.

Another circumstantial supporting evidence comes from a consideration of the issue of size-consistency.[95,96] Most of the strong correlation methods are variational and lack diagrammatic size-consistency,[96] unlike CC or MP. It is hard to imagine that such methods be routinely applied to infinitely extended systems. This observation leads to a question: is strong correlation energy thermodynamically intensive or extensive? A bond-breaking and

excitation energy is clearly intensive and the correlation energy associated with them may be handled by a variational, but non-size-consistent method such as MRCI. In other words, size-intensive strong correlation problems are an easy problem and essentially solved. On the other hand, the correlation energy in a metallic solid such as a homogeneous electron gas (HEG) should be extensive, and it poses mathematical divergence problem[97,98] in perturbative, size-consistent MP and algorithmic convergence problem[99] in also perturbative, size-consistent CC. Therefore, a challenge is to answer what is a size-consistent theory for extensive strong correlation in an extended system with zero band gap such as metallic and superconducting solids? Here, like the aforementioned singlet and triplet instability theorems, a mathematical theorem known as Overhauser's theorem[100] may be instructive: it states that a complex RHF solution of a HEG at any electron density is not the lowest-energy solution of the HF problem, but there is always a lower-energy complex general Hartree-Fock (GHF) solution,[72] which is gapped and make the remaining correlation weak and potentially treatable by MP. *This theorem,[100] therefore, also seems to support the hypothesis for solids with zero band gap.*

We propose the search and development of an algorithm that fully minimizes the HF energy, whose associated wave function may be complex RHF, real UHF, complex UHF, real GHF, or complex GHF. Simple eigenvector following of each HF instability matrix will not be sufficient, according to our preliminary study. We plan to borrow ideas from the algorithms for potential energy surface scan to locate the global minimum. We then test the veracity of our hypothesis for some representative "strong correlation" problems (i.e., determine whether they all are converted into weak correlation problems with an appropriate gapped HF reference).

The distinguishing feature of our proposed research from the Scuseria group's pioneering work is that we choose to forgo symmetry restoration in favor of preserving a single-Slater-determinant structure of the HF reference, so that the it and subsequent CC or MP correction can maintain the rigorous diagrammatic size-consistency. Since non-size-consistent methods are guaranteed to deteriorate in their accuracy with system size and rendered useless for an infinitely extended system, we consider it exceedingly important that the strong correlation problem is solved by a method that is diagrammatically size-consistent. Most of the existing "strong correlation" methods lack such a provision, and is an unsatisfactory solution. We do not view symmetry restoration to outweigh the strict maintenance of size-consistency, which tends to be lost by symmetry restoration because the reference wave function becomes multi-determinantal.

Intellectual Merits of the proposed research and its successful outcome is that it advances ab initio electronic structure theory broadly and aims at benefitting calculations of all systems—strong and weak correlation in molecules, solids, and liquids, by focusing its attention to algorithmic developments and chemical applications for weak correlation problem and mean-field solution and hypothesis-testing applications for strong correlation problem. A general solution of strong correlation is considered a holy grail of today's ab initio electronic structure theory and its significance cannot be overstated.

Broader Impacts, described in the last page, include undergraduate and graduate pedagogical innovations and free, public release of software for molecular solids and liquids developed in this project. For the undergraduate education, the PI plans to continue refining the online video lectures of quantum chemistry and spectroscopies used in a highly engaging inverted course adopted by the PI and two other professors in two institutions. They educate an untold number of self-motivated students world over in addition to registered students in those institutions. For graduate education, the PI will fully develop the course for fellowship preparation using NSF GRFP as a template and intended to increase students' awareness of the importance of writing skills, cultivating a network of supporters, and growing "thick skins" in the face of defeats (in the form of rejections of fellowship applications, paper submissions, or job applications).

2.2. Proposed Research Plans and their Intellectual Merit
2.2.1. Weak Correlation: Algorithms and Chemical Applications

CCSD(T) Phase Diagram of Ice. Molecular solids and liquids include nature's most abundant and important condensed-phase systems such as ice and liquid water.[11] This is because the surface of a planet, including that of the Earth, tends to be covered by condensation of its atmospheric molecules, which are invariably molecular solids and liquids. The properties of ice and liquid water have the most decisive influence on everything from climate to geology to biology on the Earth. During the previous cycle of this NSF funding, we have made the first MP2 and one of the most comprehensive *ab initio* computational analyses of the thermodynamic properties of ice Ih,[9] specifically, its anomalous thermal expansion and volume isotope effects as well as pressure-induced amorphizaion, which may have bearing on analogous behavior of liquid water. We have also extended this work to five other molecular phases of ice and extrapolated the CBS limits of the Gibbs energies of all six phases to determine the *ab initio* MP2 phase diagram of ice.[13] These studies established a general computational procedure for *ab initio* calculations of molecular solids on an equal footing with predictive *ab initio* calculations for small gas-phase molecules, which transformed chemistry in the last century. They, however, also revealed that the simplest *ab initio* theory, MP2 (even in its CBS limit), lacks accuracy in reproducing thermal behavior and phase diagram quantitatively, which requires chemical (or even higher) accuracy in free energy differences.

In this light, <u>*we propose raising the theoretical level of the embedded-fragment calculations of ice phases to CCSD(T) in the CBS limit with the aim of reproducing the ice phase diagram quantitatively*</u>. We expect MP2/CBS accurate enough for the structure and phonon dispersion and the Gibbs energy at each geometry should be corrected by single-point CCSD(T)/CBS internal energy, which we believe is feasible on the Blue Waters supercomputer (resources made available through the Blue Waters Professorship to the PI). To make the method even more systematically convergent at exactness, we may include three-body exchange and correlation by summing trimer energies at a lower correlation and/or smaller basis-set level and replace the HF embedding field (molecular dipole or atomic partial charges) by MP2 or CCSD(T) embedding field (molecular dipole, quadrupole, etc.). <u>*Significance of this study is that its successful outcome will decisively establish the routine feasibility of predictive ab initio calculations for molecular solids, which have long been the sole domain of force field or DFT treatments; it will also actually (not theoretically) demonstrate systematic convergence of ab initio results toward exactness for solids, an unmatched advantage not shared by force field or DFT.*</u>

CCSD(T) Molecular Dynamics of Liquid Water. With support from the predecessor NSF grant, we performed one of the first *ab initio* MP2 molecular dynamics simulations of liquid water.[7] At the time of its publication, the only other group that has a similar capability is the VandeVondele group at ETH Zurich.[101,102] They implemented a fast MP2 code for a periodic system with a large unit cell by utilizing algorithmic refinements amassed in years of developments of the planewave DFT codes in the context of Car-Parrinello simulations of liquids. Our approach is based on an almost opposite idea of utilizing local basis sets that naturally exist in molecular liquids and a many-body expansion of total energy in terms of fragment energies. The VandeVondele approach has the advantage of being generally applicable (not limited to molecular solids or liquids), whereas our embedded-fragment approach is unmatched in its flexibility in the choice of *ab initio* electron-correlation theory; switching from MP2 to CCSD(T) is just a matter of changing a line in an input file, instead of years of method development or coding.

However, whether a CCSD(T) molecular dynamics of a liquid is feasible in practice is an altogether different issue. The MP2 molecular dynamics of liquid water took 6 months of continuous use of 100 CPUs, and is far from routine. CCSD(T) molecular dynamics using the same computer code is, therefore, unthinkable; a significant serial speedup and increased parallel scalability both are needed.

<u>*We propose algorithmic improvements for embedded-fragment molecular dynamics for*</u>

9

liquids so that ab initio calculations for molecular liquids become routine with a specific initial aim of reproducing the observed structural, spectral, and thermodynamic properties of liquid water by CCSD(T) with quantitative accuracy.

Concrete steps taken to achieve this goal are as follows:

(1) In an embedded-fragment calculation, a large number of small (dimer) electron-correlation calculations are performed in parallel. We found that since an electronic structure software backend that carries out each of these small calculations is not designed for such a purpose, it tends to have nonessential overheads such as many small disk I/O to frequently update human readable outputs, e.g., which add to wall time. We, therefore, map the disk I/O to virtual memory and switch to a software backend that have minimal such overheads, e.g., PSI4.[103]

(2) The coarse-grain parallelism of an embedded-fragment calculation comes from many, near-independent dimer calculations. The upper limit of scalability at about 100 CPUs in the previous liquid water simulation reflects the fact that there are roughly 100 water dimer pairs that need to be computed in each time step for a simulation cell of 32 water molecules. Hence, it is hard to increase the coarse-grain parallel scalability considerably to boost the efficiency of the calculation, if not the fidelity of the simulation (by increasing the box size). We, therefore, seek to implement dual-level parallelism[104] in which individual dimer calculations are fine-grain parallelized (e.g., over threads of a multicore processor) and they are distributed across many processors in coarse-grain parallelism. In this sense, PSI4 again seems to be an ideal backend as it supports thread-parallelism of most of its electron-correlation implementations including CCSD(T).

(3) Most importantly, we consider it essential to adopt the state-of-the-art molecular dynamics algorithms especially when *ab initio* force calculations are inevitably expensive. Specifically, we plan to implement the multiple-time-step approach of Steele and coworkers[105-107] at Utah, in which a lower-level electronic structure theory (such as HF) is used to compute forces at every time step, whereas a med-level theory (such as MP2) once every several time steps, and finally a high-level theory, say CCSD(T), occurs at an even longer time interval. The mathematics underlying this mixed use of different levels has been sorted out[105-107] with a demonstrated increased efficiency. Since a higher-level theory calculation costs orders of magnitude CPU time than a lower-level theory calculation, decreasing the frequency of the former calculations directly and greatly impact the serial efficiency.

The foregoing paragraph corresponds to coarse graining in the time domain. We can also incorporate coarse graining in the distance domain, that is, we use the high-level theory only for the nearest neighbor dimers, while the rest are treated by mid-level or even low-level theory as the dimer becomes longer. This will also significantly decrease the number of expensive high-level theory calculations without impairing the accuracy.

With these algorithmic improvements, the following applications are planned:

We will compute the structural, thermodynamic, and spectroscopic properties of liquid water. From our previous MP2 simulation[7] and subsequent analysis,[8] we know that MP2 overestimates weak interactions, causing the liquid to contract homogenously and to have too high a density. To make liquid water have the correct density, the pressure needed to be negative. We also found that MP2 underestimates polarizability and thus directional hydrogen-bond strength, causing the water molecule to rotate more freely breaking hydrogen bonds. This means that the MP2 liquid water exists at a nonphysically low temperature. DFT have the exact opposite trends (too large polarizability and too high temperature; too small dispersion and too small density or large positive pressure). Although MP2 and DFT both have their shortcomings, the former has the advantage of being a part of converging series of size-consistent approximations and systematically improvable. We consider it important to actually demonstrate that by going from MP2 to CCSD and to CCSD(T) we can reproduce even liquid properties systematically accurately.

We will also apply embedded-fragment molecular dynamics to high- and low-density

amorphous ices. Although they are molecular solids in the time span of our simulations or most experiments, they are better treated as dynamical systems. Observed Raman spectra are available for both phases[108-115] and can be simulated as Fourier transform of time correlation functions of polarizabilities.

Ultimately, however, the true utility of the developed method does not lie in the applications to neat liquid water or ice because for such systems an overall more efficient research protocol exists: a polarizable force field is fitted against accurate [e.g., CCSD(T)] calculations for water clusters and then applied to various phases at maximum efficiency. Rather, *ab initio* molecular dynamics has a unique advantage over this conventional force field molecular dynamics in applications to systems for which fitting of force fields is too expensive or difficult such as multi-component molecular liquids or liquids whose intermolecular interactions are expressly quantum mechanical (such as conjugation). As an example, we plan to apply the method at MP2 or higher for liquid hydrogen fluoride and an aqueous solution of hydrogen fluoride,[116] in which chains or rings of various lengths with conjugated hydrogen fluoride molecules are suspected to exist.

Significance of this project is that its successful outcome will firmly establish the third era of liquid simulations, after force field (the first era) and Car-Parrinello DFT (the second era) simulations, bringing the same kind of computation revolution in chemical research in the last century to liquids. The algorithmic improvements will benefit molecular solids and amorphous calculations also.

CCSD(T) Spectra of Protonated Water Clusters. In the previous funding cycle of this NSF grant, we have developed a critical extension of the embedded-fragment method to molecular cluster ions and molecular clusters with partial charge transfer.[14] These systems may be easily divided into well-defined fragments, but it is not possible to assign an integer electron count to each fragment. We, therefore, introduced a new formalism in which the wave function of the whole cluster is a superposition of various valence bond (VB) configurations with well-defined integer electron counts to fragments. The superposition makes the time-averaged electron count of each fragment to be noninteger. We found that this ansatz works well in reproducing the potential energy curves of various molecular cluster ions. Figure 11 shows such agreement between the embedded-fragment VB method and brute-force calculation of the entire cluster ion, in this case, $[Na(H_2O)_3]^+$.

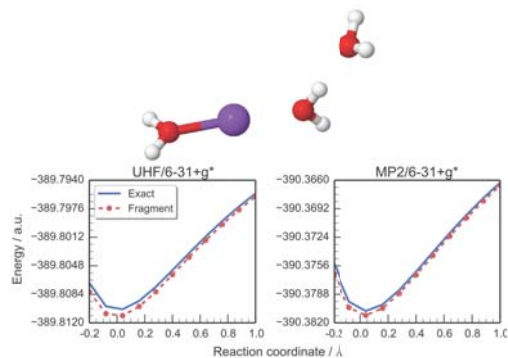**Figure 11** $[Na(H_2O)_3]^+$ potential energy curves

We will apply this new embedded-fragment method to molecular cluster ions such as protonated water clusters. Vibrational spectra have been reported for them with various sizes.[15,16] In both reports in 2004 and 2014, computational analyses were done with DFT. We plan to raise the theoretical levels to MP2, CCSD, and CCSD(T) in the calculations of optimized structures and vibrational spectra.

2.2.2. Strong Correlation: General Theory and Test Applications

Complex GHF Global Minimum Search. *We will prove or disprove the following hypothesis as a general solution to the notorious "strong correlation" problem: All strongly correlated problems are reduced to weakly correlated ones (treatable by CC or even MP) by fully minimizing the mean-field energy, which can be associated with the most general form of a HF solution, i.e., complex GHF wave function.*[72] Complex GHF orbitals are a linear combination of α-spin and β-spin components, whose spatial parts are complex-valued. It is analogous to a two-component relativistic mean-field wave function. Our preliminary study (discussed in the next paragraph) showed that such a global-minimum solution of the HF problem cannot necessarily be reached by following an eigenvector corresponding to the negative eigenvalue of the HF stability matrix.[117] In other words, the positive definiteness of the stability

matrix merely tests whether the solution is a minimum or any other type of stationary point (such as a saddle point), but cannot tell if it is a local minimum or the global minimum. This means that we need a more powerful, brute-force, blanket search algorithm scanning a broad area of the high-dimensional HF solution space efficiently to locate all local and global minima.

We plan to borrow algorithmic ideas from geometry optimization and potential energy surface (PES) scan, which suffer from a similar difficulty. We will use the Global Reaction Route Mapping (GRRM) method[118-120] of Maeda, Ohno, and Morokuma, which has been highly successful in locating all important points in the PES and reaction paths. We will apply GRRM to the HF PES in the space of MO coefficients (not in the space of geometry), by replacing geometrical coordinates with atomic orbital (AO) bases. Specifically, we adopt their ADDF (Anharmonic Downward Distortion Following) algorithm, designed originally to find reaction coordinates for isomerizations (and other reactions). On a hyperellipsoid centered at a local minimum, it identifies the direction in which the rate of the increase in the hyperellipsoid-PES is the smallest by virtue of strong anharmonicity, which is followed to reach a transition state and then to another local minimum (Figure 12).

Starting from a local minimum in the MO space, typically a real RHF or UHF solution, we perform the HF stability analysis by diagonalizing the relevant stability matrix. This determines the "normal coordinates" (as eigenvectors) and "harmonic frequencies" (as eigenvalues) of the HF PES in the MO space (as opposed to in the geometrical space). If the lowest "harmonic frequency" (eigenvalue) is negative, the initial solution is not even a local minimum, but a saddle point, and we follow the corresponding eigenvector to reach a true local minimum.[117] If all eigenvalues are nonnegative, the initial solution is a local minimum. Then, we perform the

Figure 12 The ADDF algorithm for search for isomerization reaction path and local minimum (Taken from Maeda, Ohno, and Morokuma, PCCP, 2013, **15**, 3683).

ADDF algorithm using the normal coordinates and harmonic frequencies in the MO space in the complete analogy to the same in the original geometrical context of Maeda and coworkers. We hope to develop an algorithm as powerful as Maeda's original one for reaction paths but for general HF solutions. We note that Thom and Head-Gordon proposed and implemented an alternative algorithm[121] to locate HF solutions also inspired by a PES scan (specifically, metadynamics), which can serve as "plan B," in the event that GRRM is found not to work for general HF solutions.

With this algorithm, we investigate whether "strong correlation" problems can be reduced to weak correlation problems by fully minimizing the HF energy and, if so, determine what the nature of the true HF solution is: Are the orbitals real- or complex-valued? Is it UHF, RHF, or GHF? Are the orbitals spatially localized or delocalized and symmetry-adapted? As the target problems, we consider various covalent bond breaking including the notorious Cr_2.[60,122,123] We note that Jiménez-Hoyos, Henderson, and Scuseria[72] have indeed found that GHF breaks bonds correctly for O_2 and CO_2, which is in favor of our hypothesis. We will also study the H_4 system[124-127] and H_n chains[128,129] as well as Cu_2O_2 "torture track."[130] When these test calculations support our hypothesis, we proceed to apply it to strongly correlated problems of real chemical interests, such as single-molecule magnets,[131-135] polyradicals,[136,137] and transition metal complexes.[138,139]

Significance of this research is that it proves or disproves the hypothesis that long-standing strong correlation problems can be solved altogether at an appropriate mean-field (HF) theory level. It also provides a robust algorithm to fully optimize HF solution beyond the usual instability analysis, regardless of the veracity of the hypothesis. If the hypothesis were proven correct, this algorithm could potentially serve as a practical computational basis to nearly all electronic structure problems, both strongly and weakly correlated, upon which CC or even MP can be applied to converge at exactness rapidly.

HF Instabilities of HEG. In parallel with the aforementioned research for molecules, we will study the instabilities of infinitely extended systems, strong correlation in which is the most challenging problem. *We will numerically determine the stability of complex RHF solutions of 1D, 2D, and 3D homogeneous electron gas (HEG) as a function of its density, which serves as a model of metallic or superconducting solids in the respective dimensions*. It has a strong bearing on the 2D high-T_c superconductivity discussed in the next subsection.

In our previous study,[81] we proved theorems stating that a real RHF solution of a molecule with degenerate HOMO and LUMO is always unstable in the space of real or complex UHF and complex RHF solutions. We were, however, unable to prove the same for an infinite gapless system. It is interesting and important to determine whether these theorems can be extended to infinite systems. An argument in favor of this extensibility is provided by
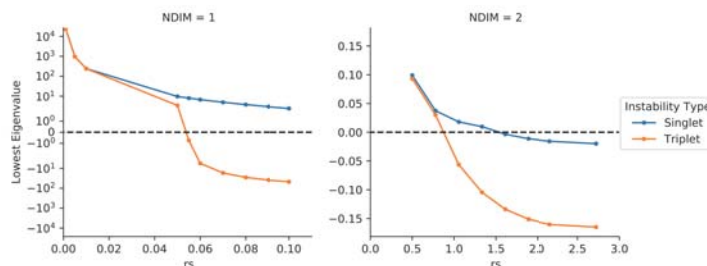

Figure 13 The lowest eigenvalue of the HF singlet and triplet instability of 1D and 2D HEG as a function of average electron-electron distance (r_s). The negative lowest eigenvalue (seen at larger r_s or lower density) implies an instability.

Overhauser's theorem[100,140] stating that there is always a "spirial spin density wave (SDW)" solution (which is a complex GHF solution) with energy lower than that of a complex RHF solution for a 1D or 3D HEG at any density. This suggests (but not implies) an analogous instability theorem for infinite systems in general (not just for HEG as in the case of Overhauser's theorem). We will answer this question by actually building the instability matrices of 1D, 2D, and 3D HEG. Our preliminary calculations (which are only preliminary because convergence with respect to the number of wave vector sampling points is yet to be convincingly shown) indicated that complex RHF solutions of 1D and 2D HEG are *stable* against complex UHF or complex GHF instability at higher densities (Figure 13). This is not contradictory with Overhauser's theorem because the presence of instability merely means that the solution is not a minimum and is a sufficient (but not necessary) condition for the existence of a lower-energy (global) minimum solution. We will actively construct a complex GHF solution (spiral SDW or other types) with a lower energy at higher densities to understand the structure of such solutions and come up with an algorithm to find them in an infinite metallic system.

Significance of this study is that it bridges the gap in theoretical understanding of strong correlation and HF instability between finite systems and infinite systems, the latter including metallic and superconducting solids with zero band gaps (in a finite system, HOMO-LUMO gap is almost never zero). It also sheds light on the fundamental physics question concerning the size dependence (thermodynamic extensive or intensive nature) of "strong correlation" energy, if it can be unambiguously defined.

High-T_c Superconductivity UHF Modeling. A typical, albeit simplified phase diagram[141] (as a function of dopant concentration / pressure and temperature) of 2D high-T_c superconductor (such as cuprate or iron arsenide) is drawn in Figure 14. At low dopant concentration or low pressure (corresponding to a small r_s in Figure 13) and low temperature, the system undergoes a spontaneous symmetry breaking in electronic structure to form an insulating SDW or antiferromagnetic phase (also called Mott-Hubbard insulator as opposed to band insulator[141]). As the dopant concentration or pressure increases, SDW is
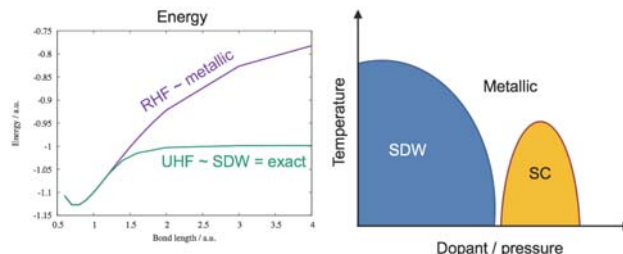

Figure 14 Phase diagram of cuprate high-T_c superconductor.

resolved and periodic symmetry of orbitals is restored, making the system metallic. Around the area of this electronic phase transition is the superconducting (SC) phase.

We can then submit the following hypothesis for the mechanism of the high-T_c superconductivity (Figure 14): The metallic phase corresponds to the complex RHF solution, whose orbitals are periodic-symmetry-adapted. The gapped SDW phase corresponds to the complex UHF solution, whose orbitals spontaneously break the underlying periodic symmetry of the Hamiltonian. The SC phase occurs right around the onset of the UHF instability, where the electronic structure is fragile and easily affected by a minute perturbation (phonons) to bring about a large symmetry-breaking change, and, therefore, there is an unusually large electron-phonon coupling. In other words, the simple phonon-based Bardeen-Cooper-Schrieffer (BCS) theory may explain the high-T_c superconductivity with the strong electron-phonon coupling caused by the symmetry-breaking instability of electronic structure. It is speculated that enlarging the SDW phase leads to the SC phase and higher T_c can be achieved.[142,143]

Figure 15 A symmetry-broken SDW (Mott-Hubbard) insulator phase found by UHF in 2D hydrogen sheet.

We will determine if this hypothetical mechanism is plausible. We propose the following research plan. *We will map the HF SDW-metal (if not SC directly) phase diagram of a 2D hydrogen sheet as a function of dopant concentration or pressure and temperature*. We have already developed a periodic RHF / UHF method at a finite temperature (a finite-temperature extension of a mean-field theory is well established and trivially implemented), finding a symmetry-broken SDW-Mott-Hubbard-insulator solution of the HF problem of a 2D hydrogen sheet (Figure 15). The effect of dopants can be simulated by varying the Fermi level (and concomitantly scaling nuclear charges to maintain the unit cell charge neutrality). The effect of pressure is simply the change in the inter-nuclear distances. Our preliminary study shows that this "doped" HF theory indeed can reproduce the basic structure of the SDW-metal phase diagram of 2D hydrogen sheet, in which a SDW phase (which roughly but not exactly corresponds to a gapped, insulating phase) appears at a low temperature, low dopant concentration (Figure 16). *We will find the condition (dimensionality, structure, stoichiometry) that maximizes the SDW phase as a potential proxy of high T_c. We will then determine which phonon varies the gap of the SDW phase most greatly, which likely serves as the Cooper pair glue*.

Figure 16 Phase diagram of 2D hydrogen sheet.

We note that this hypothetical mechanism (assuming the phonon-mediated Cooper-pair formation based on the conventional BCS theory) goes against the prevailing view assuming a strongly correlated, magnetic-interaction-mediated Cooper-pair formation of theoretical physicists on high-T_c superconductivity. We consider it both interesting and worthwhile to inject a fresh *ab initio* viewpoint of quantum chemists (based largely on HF theory as opposed to DFT of physicists which are more immune to instabilities) into this long-standing problem. We also note that our hydrogen-sheet model is admittedly crude but maintains atomic identities and positions and thus generates distinct phonons unlike the Hubbard model.

Significance of this research is that it tests a HF-based, weakly correlated, BCS picture of high-T_c superconductivity, as opposed to the prevailing strongly correlated, non-BCS view. It may also shed light on the type of phonons that serve as a Cooper pair glue and the structure-T_c relationship and thus a general guideline for finding and designing higher-T_c materials.
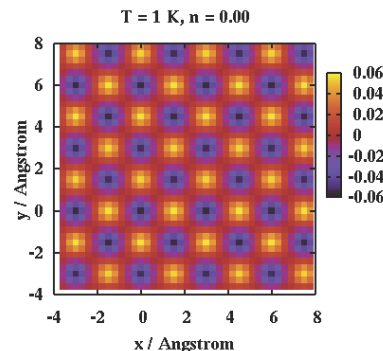
2.3. Education and Outreach and their Broader Impacts

Undergraduate Instructional Innovation. With partial support from NSF, the PI has developed 37 video lectures on quantum chemistry and spectroscopies, which are used in an inverted course. In this course, a student is asked to watch the video lecture and solve the matching problems beforehand. During a class, randomly chosen students are asked to present their solutions on the blackboard and explain them to the class. This modern format of instructions, which more fully engage students during and outside the classroom, was found to be highly effective and also well received by students. The PI's particular implementation was adopted by his colleague, Professor Sharon Hammes-Schiffer, who refined the format further, with success. It was also experimentally used by Professor Francesco Evangelista at Emory University. For this and other educational efforts, the PI received the 2017 SCS Teaching Award from the School of Chemical Sciences of UIUC and was included in the List of Teachers Ranked as Excellent by Their Students in 2015, 2016, and 2017.

Of 37 original video lectures (powerpoint presentations with the PI's voice-only narration), 13 have been upgraded with more professional versions featuring a student playing the part of an instructor (http://butane.chem.illinois.edu/sohirata/). The paid student staff were recruited from the highest achieving undergraduate students who took the course in the previous semesters. The PI will continue upgrading the remaining 24 video lectures, at the completion of which we may consider broadcasting them through higher-profile, but free-of-charge platforms (such as YouTube or iTune) in addition to the UIUC website where they are already available to anyone who has access to the Internet. *We expect that these resources have an even greater lasting educational value than popular textbooks on the same subject, as unlike the latter, they are free, readily available, and effective, and signal the future mode of pedagogy even at such high levels of materials as quantum chemistry and spectroscopies. Unlike similar efforts elsewhere, our online video lectures are presented by recent former undergraduate students of the course (not by a professor), making these difficult subjects more accessible psychologically to the listener students.*

Graduate Instructional Innovation. The PI will continue and refine his newly designed graduate course on fellowships. The unique feature of this course is that it not only gives students a training on general grant writing, but it does so in such timing and format that they can submit what they write to NSF GRFP (submission is optional but encouraged), serving as an incentive for students to get started with the lifelong pursuit of perfecting their writing skills. The tasks the students are asked to perform have focused objectives and hard deadlines with potentially great tangible benefits to the students' career. NSF GRFP is chosen as a convenient template for its inclusiveness of all chemistry areas, but the course's benefit transcends this particular program and includes students' increased awareness of the significance of writing skills and cultivating the network of supporters who would provide them with recommendation letters. It also helps students become more immune to inevitable defeats (i.e., not winning NSF fellowships) and grow "thick skins" necessary to survive and thrive in competitive intellectual professions.

Software and Technology Transfer. We plan to release software of some of the more fully developed methods to general public. This serves as a dual purpose of providing the most direct means of verification of published results and of having interested colleagues an access to the methods for applications. Specifically, we plan to package embedded-fragment software for molecular solids and liquids with documentations and quality assurance tests, which can be interfaced with a number of backend molecular electronic structure software such as PSI4 and NWCHEM. We note that these backends need not be included in or modified for the embedded-fragment software, as the communications between these two software modules occur through files.

D. References Cited: *Fast Methods for Strong and Weak Electron Correlation*

(1) Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. Fragmentation methods: a route to accurate calculations on large systems. *Chem Rev* **2012,** *112* (1), 632.

(2) Hirata, S.; Gilliard, K.; He, X.; Li, J. J.; Sode, O. Ab Initio Molecular Crystal Structures, Spectra, and Phase Diagrams. *Accounts Chem Res* **2014,** *47* (9), 2721.

(3) Beran, G. J.; Hartman, J. D.; Heit, Y. N. Predicting Molecular Crystal Properties from First Principles: Finite-Temperature Thermochemistry to NMR Crystallography. *Acc Chem Res* **2016,** *49* (11), 2501.

(4) Gilliard, K.; Sode, O.; Hirata, S. Second-order many-body perturbation and coupled-cluster singles and doubles study of ice VIII. *J Chem Phys* **2014,** *140* (17), 174507.

(5) Hirata, S.; Sode, O.; Keeli, M.; Yagi, K.; Li, J. J. Response to "Comment on 'Fermi resonance in solid CO2 under pressure'" [J. Chem. Phys. 140, 177101 (2014)]. *J Chem Phys* **2014,** *140* (17), 177102.

(6) Li, J. J.; Sode, O.; Hirata, S. Second-Order Many-Body Perturbation Study on Thermal Expansion of Solid Carbon Dioxide. *J Chem Theory Comput* **2015,** *11* (1), 224.

(7) Willow, S. Y.; Salim, M. A.; Kim, K. S.; Hirata, S. Ab initio molecular dynamics of liquid water using embedded-fragment second-order many-body perturbation theory towards its accurate property prediction. *Sci Rep-Uk* **2015,** *5*, 14358.

(8) Willow, S. Y.; Zeng, X. C.; Xantheas, S. S.; Kim, K. S.; Hirata, S. Why Is MP2-Water "Cooler" and "Denser" than DFT-Water? *J Phys Chem Lett* **2016,** *7* (4), 680.

(9) Salim, M. A.; Willow, S. Y.; Hirata, S. Ice Ih anomalies: Thermal contraction, anomalous volume isotope effect, and pressure-induced amorphization. *J Chem Phys* **2016,** *144* (20), 204503.

(10) Hirata, S.; He, X.; Hermes, M. R.; Willow, S. Y. Second-Order Many-Body Perturbation Theory: An Eternal Frontier. *J Phys Chem A* **2014,** *118* (4), 655.

(11) Hirata, S.; Gilliard, K.; He, X.; Keçeli, M.; Li, J. J.; Salim, M. A.; Sode, O.; Yagi, K. In *Fragmentation: Toward Accurate Calculations on Complex Molecular Systems*; Gordon, M. S., Ed.; Wiley: Chichester, 2017.

(12) Zhang, J. LIBRETA: Computerized Optimization and Code Synthesis for Electron Repulsion Integral Evaluation. *J Chem Theory Comput* **2017,** (under review).

(13) Salim, M. A.; Hirata, S. Ab initio phase diagram of ice Ih, II, III, V, VI, and IX. **2017,** (in preparation).

(14) Salim, M. A.; Hirata, S. Embedded-fragment VB theory. **2017,** (in preparation).

(15) Shin, J. W.; Hammer, N. I.; Diken, E. G.; Johnson, M. A.; Walters, R. S.; Jaeger, T. D.; Duncan, M. A.; Christie, R. A.; Jordan, K. D. Infrared signature of structures associated with the $H^+(H_2O)_n$ ($n = 6$ to 27) clusters. *Science* **2004,** *304* (5674), 1137.

(16) Fournier, J. A.; Wolke, C. T.; Johnson, C. J.; Johnson, M. A.; Heine, N.; Gewinner, S.; Schollkopf, W.; Esser, T. K.; Fagiani, M. R.; Knorke, H.; Asmis, K. R. Site-specific vibrational spectral signatures of water molecules in the magic $H_3O^+ (H_2O)_{20}$ and $Cs^+ (H_2O)_{20}$ clusters. *Proc Natl Acad Sci U S A* **2014,** *111* (51), 18132.

(17) Saebo, S.; Pulay, P. Local Configuration-Interaction - an Efficient Approach for Larger Molecules. *Chem Phys Lett* **1985,** *113* (1), 13.

(18) Saebo, S.; Pulay, P. Local Treatment of Electron Correlation. *Annu Rev Phys Chem* **1993,** *44*, 213.

(19) Werner, H. J.; Schutz, M. An efficient local coupled cluster method for accurate thermochemistry of large systems. *J Chem Phys* **2011,** *135* (14), 144116.

(20) Pinski, P.; Riplinger, C.; Valeev, E. F.; Neese, F. Sparse maps: A systematic infrastructure for reduced-scaling electronic structure methods. I. An efficient and simple linear scaling local MP2 method that uses an intermediate basis of pair natural orbitals. *J Chem Phys* **2015,** *143* (3), 034108.

(21)   Thom, A. J.; Alavi, A. Stochastic perturbation theory: a low-scaling approach to correlated electronic energies. *Phys Rev Lett* **2007,** *99* (14), 143001.
(22)   Thom, A. J. Stochastic coupled cluster theory. *Phys Rev Lett* **2010,** *105* (26), 263004.
(23)   Willow, S. Y.; Kim, K. S.; Hirata, S. Stochastic evaluation of second-order many-body perturbation energies. *J Chem Phys* **2012,** *137* (20), 204122.
(24)   Willow, S. Y.; Kim, K. S.; Hirata, S. Stochastic evaluation of second-order Dyson self-energies. *J Chem Phys* **2013,** *138* (16), 164111.
(25)   Neuhauser, D.; Rabani, E.; Baer, R. Expeditious Stochastic Approach for MP2 Energies in Large Electronic Systems. *J Chem Theory Comput* **2013,** *9* (1), 24.
(26)   Willow, S. Y.; Hirata, S. Stochastic, real-space, imaginary-time evaluation of third-order Feynman-Goldstone diagrams. *J Chem Phys* **2014,** *140* (2), 024111.
(27)   Willow, S. Y.; Kim, K. S.; Hirata, S. Brueckner-Goldstone quantum Monte Carlo for correlation energies and quasiparticle energy bands of one-dimensional solids. *Phys Rev B* **2014,** *90* (20), 201110.
(28)   Willow, S. Y.; Zhang, J. M.; Valeev, E. F.; Hirata, S. Communication: Stochastic evaluation of explicitly correlated second-order many-body perturbation energy. *J Chem Phys* **2014,** *140* (3), 031101.
(29)   Johnson, C. M.; Doran, A. E.; Zhang, J. M.; Valeev, E. F.; Hirata, S. Monte Carlo explicitly correlated second-order many-body perturbation theory. *J Chem Phys* **2016,** *145* (15), 154115.
(30)   Franklin, R. S.; Spencer, J. S.; Zoccante, A.; Thom, A. J. Linked coupled cluster Monte Carlo. *J Chem Phys* **2016,** *144* (4), 044111.
(31)   Spencer, J. S.; Thom, A. J. W. Developments in stochastic coupled cluster theory: The initiator approximation and application to the uniform electron gas. *J Chem Phys* **2016,** *144* (8), 084108.
(32)   Chan, G. K. An algorithm for large scale density matrix renormalization group calculations. *J Chem Phys* **2004,** *120* (7), 3172.
(33)   Chan, G. K.; Kallay, M.; Gauss, J. State-of-the-art density matrix renormalization group and coupled cluster theory studies of the nitrogen binding curve. *J Chem Phys* **2004,** *121* (13), 6110.
(34)   Marti, K. H.; Ondik, I. M.; Moritz, G.; Reiher, M. Density matrix renormalization group calculations on relative energies of transition metal complexes and clusters. *J Chem Phys* **2008,** *128* (1), 014104.
(35)   Zgid, D.; Nooijen, M. The density matrix renormalization group self-consistent field method: orbital optimization with the density matrix renormalization group method in the active space. *J Chem Phys* **2008,** *128* (14), 144116.
(36)   Kurashige, Y.; Yanai, T. High-performance ab initio density matrix renormalization group method: applicability to large-scale multireference problems for metal compounds. *J Chem Phys* **2009,** *130* (23), 234114.
(37)   Yanai, T.; Kurashige, Y.; Neuscamman, E.; Chan, G. K. Multireference quantum chemistry through a joint density matrix renormalization group and canonical transformation theory. *J Chem Phys* **2010,** *132* (2), 024105.
(38)   Knecht, S.; Legeza, O.; Reiher, M. Communication: four-component density matrix renormalization group. *J Chem Phys* **2014,** *140* (4), 041101.
(39)   Jiang, H. C.; Weng, Z. Y.; Xiang, T. Accurate determination of tensor network state of quantum lattice models in two dimensions. *Phys Rev Lett* **2008,** *101* (9), 090603.
(40)   Zhou, H. Q.; Orus, R.; Vidal, G. Ground state fidelity from tensor network representations. *Phys Rev Lett* **2008,** *100* (8), 080601.
(41)   Beri, B.; Cooper, N. R. Local tensor network for strongly correlated projective States. *Phys Rev Lett* **2011,** *106* (15), 156401.
(42)   Nakatani, N.; Chan, G. K. Efficient tree tensor network states (TTNS) for quantum chemistry: generalizations of the density matrix renormalization group algorithm. *J Chem Phys* **2013,** *138* (13), 134113.
(43)   Evenbly, G.; Vidal, G. Tensor Network Renormalization. *Phys Rev Lett* **2015,** *115* (18), 180405.

(44) Ohtsuka, Y.; Nagase, S. Projector Monte Carlo method based on configuration state functions. Test applications to the H_4 system and dissociation of LiH. *Chem Phys Lett* **2008,** *463* (4-6), 431.

(45) Ohtsuka, Y.; Nagase, S. Projector Monte Carlo method based on Slater determinants: Test application to singlet excited states of H_2O and LiF. *Chem Phys Lett* **2010,** *485* (4-6), 367.

(46) Ohtsuka, Y.; Nagase, S. Projector Monte Carlo method based on Slater determinants: a new sampling method for singlet state calculations. *Theor Chem Acc* **2011,** *130* (2-3), 501.

(47) Booth, G. H.; Thom, A. J. W.; Alavi, A. Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space. *J Chem Phys* **2009,** *131* (5), 054106.

(48) Cleland, D.; Booth, G. H.; Alavi, A. Communications: Survival of the fittest: Accelerating convergence in full configuration-interaction quantum Monte Carlo. *J Chem Phys* **2010,** *132* (4), 041103.

(49) Booth, G. H.; Alavi, A. Approaching chemical accuracy using full configuration-interaction quantum Monte Carlo: A study of ionization potentials. *J Chem Phys* **2010,** *132* (17), 174104.

(50) Ohtsuka, Y.; Ten-no, S. Explicitly correlated projector Monte Carlo method based on Slater determinants (PMC-SD-F12 method). *AIP Conf Proc* **2012,** *1456*, 97.

(51) Booth, G. H.; Cleland, D.; Alavi, A.; Tew, D. P. An explicitly correlated approach to basis set incompleteness in full configuration interaction quantum Monte Carlo. *J Chem Phys* **2012,** *137* (16), 164112.

(52) Shepherd, J. J.; Booth, G. H.; Alavi, A. Investigation of the full configuration interaction quantum Monte Carlo method using homogeneous electron gas models. *J Chem Phys* **2012,** *136* (24), 244101.

(53) Booth, G. H.; Gruneis, A.; Kresse, G.; Alavi, A. Towards an exact description of electronic wavefunctions in real solids. *Nature* **2013,** *493* (7432), 365.

(54) Ohtsuka, Y.; Ten-no, S. A study of potential energy curves from the model space quantum Monte Carlo method. *J Chem Phys* **2015,** *143* (21), 144116.

(55) Shamasundar, K. R.; Knizia, G.; Werner, H. J. A new internally contracted multi-reference configuration interaction method. *J Chem Phys* **2011,** *135* (5), 054101.

(56) Andersson, K.; Malmqvist, P. A.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. Second-Order Perturbation-Theory with a CASSCF Reference Function. *J Phys Chem-Us* **1990,** *94* (14), 5483.

(57) Andersson, K.; Malmqvist, P. A.; Roos, B. O. Second-Order Perturbation-Theory with a Complete Active Space Self-Consistent Field Reference Function. *J Chem Phys* **1992,** *96* (2), 1218.

(58) Nakano, H. Quasi-Degenerate Perturbation-Theory with Multiconfigurational Self-Consistent-Field Reference Functions. *J Chem Phys* **1993,** *99* (10), 7983.

(59) Finley, J.; Malmqvist, P. A.; Roos, B. O.; Serrano-Andres, L. The multi-state CASPT2 method. *Chem Phys Lett* **1998,** *288* (2-4), 299.

(60) Pulay, P. A Perspective on the CASPT2 Method. *Int J Quantum Chem* **2011,** *111* (13), 3273.

(61) Granovsky, A. A. Extended multi-configuration quasi-degenerate perturbation theory: the new approach to multi-state multi-reference perturbation theory. *J Chem Phys* **2011,** *134* (21), 214113.

(62) Shiozaki, T.; Gyorffy, W.; Celani, P.; Werner, H. J. Communication: extended multi-state complete active space second-order perturbation theory: energy and nuclear gradients. *J Chem Phys* **2011,** *135* (8), 081106.

(63) Knizia, G.; Chan, G. K. Density matrix embedding: a simple alternative to dynamical mean-field theory. *Phys Rev Lett* **2012,** *109* (18), 186404.

(64) Knizia, G.; Chan, G. K. Density Matrix Embedding: A Strong-Coupling Quantum Embedding Theory. *J Chem Theory Comput* **2013,** *9* (3), 1428.

(65) Dresselhaus, T.; Neugebauer, J.; Knecht, S.; Keller, S.; Ma, Y.; Reiher, M. Self-consistent embedding of density-matrix renormalization group wavefunctions in a density functional environment. *J Chem Phys* **2015,** *142* (4), 044111.

18

(66) Tsuchimochi, T.; Welborn, M.; Van Voorhis, T. Density matrix embedding in an antisymmetrized geminal power bath. *J Chem Phys* **2015,** *143* (2), 024107.

(67) Hedegard, E. D.; Reiher, M. Polarizable Embedding Density Matrix Renormalization Group. *J Chem Theory Comput* **2016,** *12* (9), 4242.

(68) Pernal, K. Reduced density matrix embedding. General formalism and inter-domain correlation functional. *Phys Chem Chem Phys* **2016,** *18* (31), 21111.

(69) Wouters, S.; Jimenez-Hoyos, C. A.; Sun, Q.; Chan, G. K. A Practical Guide to Density Matrix Embedding Theory in Quantum Chemistry. *J Chem Theory Comput* **2016,** *12* (6), 2706.

(70) Tsuchimochi, T.; Scuseria, G. E. Strong correlations via constrained-pairing mean-field theory. *J Chem Phys* **2009,** *131* (12), 121102.

(71) Tsuchimochi, T.; Scuseria, G. E. Constrained active space unrestricted mean-field methods for controlling spin-contamination. *J Chem Phys* **2011,** *134* (6), 064101.

(72) Jimenez-Hoyos, C. A.; Henderson, T. M.; Scuseria, G. E. Generalized Hartree-Fock Description of Molecular Dissociation. *J Chem Theory Comput* **2011,** *7* (9), 2667.

(73) Scuseria, G. E.; Jimenez-Hoyos, C. A.; Henderson, T. M.; Samanta, K.; Ellis, J. K. Projected quasiparticle theory for molecular electronic structure. *J Chem Phys* **2011,** *135* (12), 124108.

(74) Jimenez-Hoyos, C. A.; Henderson, T. M.; Tsuchimochi, T.; Scuseria, G. E. Projected Hartree-Fock theory. *J Chem Phys* **2012,** *136* (16), 164109.

(75) Samanta, K.; Jimenez-Hoyos, C. A.; Scuseria, G. E. Exploring Copper Oxide Cores Using the Projected Hartree-Fock Method. *J Chem Theory Comput* **2012,** *8* (12), 4944.

(76) Jimenez-Hoyos, C. A.; Rodriguez-Guzman, R.; Scuseria, G. E. Multi-component symmetry-projected approach for molecular ground state correlations. *J Chem Phys* **2013,** *139* (20), 204102.

(77) Garza, A. J.; Jimenez-Hoyos, C. A.; Scuseria, G. E. Capturing static and dynamic correlations by a combination of projected Hartree-Fock and density functional theories. *J Chem Phys* **2013,** *138* (13), 134102.

(78) Henderson, T. M.; Scuseria, G. E. Linearized Jastrow-style fluctuations on spin-projected Hartree-Fock. *J Chem Phys* **2013,** *139* (23), 234113.

(79) Jimenez-Hoyos, C. A.; Rodriguez-Guzman, R.; Scuseria, G. E. Excited electronic states from a variational approach based on symmetry-projected Hartree-Fock configurations. *J Chem Phys* **2013,** *139* (22), 224110.

(80) Rivero, P.; Jimenez-Hoyos, C. A.; Scuseria, G. E. Predicting Singlet-Triplet Energy Splittings with Projected Hartree-Fock Methods. *J Phys Chem A* **2013,** *117* (33), 8073.

(81) Yamada, T.; Hirata, S. Singlet and triplet instability theorems. *J Chem Phys* **2015,** *143* (11), 114112.

(82) Cizek, J.; Paldus, J. Stability Conditions for Solutions of Hartree-Fock Equations for Atomic and Molecular Systems . Application to Pi-Electron Model of Cyclic Polyenes. *J Chem Phys* **1967,** *47* (10), 3976.

(83) Paldus, J.; Cizek, J. Stability conditions for the solutions of the Hartree-Fock equations for the simple open-shell case. *Chem Phys Lett* **1969,** *3* (1), 1.

(84) Cizek, J.; Paldus, J. Stability Conditions for Solutions of Hartree-Fock Equations for Atomic and Molecular Systems .3. Rules for Singlet Stability of Hartree-Fock Solutions of Pi-Electronic Systems. *J Chem Phys* **1970,** *53* (2), 821.

(85) Paldus, J.; Cizek, J. Stability Conditions for the Solutions of the Hartree-Fock Equations for Atomic and Molecular Systems. VI. Singlet-Type Instabilities and Charge-Density-Wave Hartree-Fock Solutions for Cyclic Polyenes. *Phys Rev A* **1970,** *2* (6), 2268.

(86) Paldus, J.; Cizek, J. Stability Conditions for Solutions of Hartree-Fock Equations for Atomic and Molecular Systems .2. Simple Open-Shell Case. *J Chem Phys* **1970,** *52* (6), 2919.

(87) Cizek, J.; Paldus, J. Stability Conditions for Solutions of Hartree-Fock Equations for Atomic and Molecular Systems .5. Nonanalytic Behavior of Broken-Symmetry Solutions at Branching Point. *Phys Rev A* **1971,** *3* (2), 525.

(88) Paldus, J.; Cizek, J. Stability Conditions for Solutions of Hartree-Fock Equations for Atomic and Molecular Systems .4. Study of Doublet Stability for Odd Linear Polyenic Radicals. *J Chem Phys* **1971,** *54* (6), 2293.

(89) Ostlund, N. S. Complex and Unrestricted Hartree-Fock Wavefunctions. *J Chem Phys* **1972,** *57* (7), 2994.

(90) Fukutome, H. Unrestricted Hartree-Fock Theory of Chemical Reactions .1. Electronic Instabilities in Chemical Reactions and Solutions of Unrestricted Scf Lcao Mo Equation for Homopolar Two-Center Two-Electron System. *Prog Theor Phys* **1972,** *47* (4), 1156.

(91) Fukutome, H. Unrestricted Hartree-Fock Theory of Chemical Reactions III. Instability Conditions for Paramagnetic and Spin Density Wave States and Application to Internal-Rotation of Ethylene. *Prog Theor Phys* **1973,** *50* (5), 1433.

(92) Fukutome, H. Theory of Unrestricted Hartree-Fock Equation and Its Solutions III. Classification of Instabilities and Interconnection Relation between 8 Classes of Uhf Solutions. *Prog Theor Phys* **1974,** *52* (6), 1766.

(93) Fukutome, H. Theory of Unrestricted Hartree-Fock Equation and Its Solutions IV. Behavior of UHF Solutions in Vicinity of Interconnecting Instability Threshold. *Prog Theor Phys* **1975,** *53* (5), 1320.

(94) Seeger, R.; Pople, J. A. Self-Consistent Molecular-Orbital Methods XVIII. Constraints and Stability in Hartree-Fock Theory. *J Chem Phys* **1977,** *66* (7), 3045.

(95) Hirata, S. Thermodynamic limit and size-consistent design. *Theor Chem Acc* **2011,** *129* (6), 727.

(96) Hirata, S.; Grabowski, I. On the mutual exclusion of variationality and size consistency. *Theor Chem Acc* **2014,** *133* (3).

(97) Hirata, S.; He, X. On the Kohn-Luttinger conundrum. *J Chem Phys* **2013,** *138* (20), 204112.

(98) He, X.; Ryu, S.; Hirata, S. Finite-temperature second-order many-body perturbation and Hartree-Fock theories for one-dimensional solids: An application to Peierls and charge-density-wave transitions in conjugated polymers. *J Chem Phys* **2014,** *140* (2), 024702.

(99) Hermes, M. R.; Hirata, S. Finite-temperature coupled-cluster, many-body perturbation, and restricted and unrestricted Hartree-Fock study on one-dimensional solids: Luttinger liquids, Peierls transitions, and spin-and charge-density waves. *J Chem Phys* **2015,** *143* (10), 102818.

(100) Overhauser, A. W. Spin Density Waves in an Electron Gas. *Phys Rev* **1962,** *128* (3), 1437.

(101) Del Ben, M.; Hutter, J.; VandeVondele, J. Electron Correlation in the Condensed Phase from a Resolution of Identity Approach Based on the Gaussian and Plane Waves Scheme. *J Chem Theory Comput* **2013,** *9* (6), 2654.

(102) Del Ben, M.; Schonherr, M.; Hutter, J.; VandeVondele, J. Bulk Liquid Water at Ambient Temperature and Pressure from MP2 Theory. *J Phys Chem Lett* **2013,** *4* (21), 3753.

(103) Parrish, R. M.; Burns, L. A.; Smith, D. G. A.; Simmonett, A. C.; DePrince, A. E.; Hohenstein, E. G.; Bozkaya, U.; Sokolov, A. Y.; Di Remigio, R.; Richard, R. M.; Gonthier, J. F.; James, A. M.; McAlexander, H. R.; Kumar, A.; Saitow, M.; Wang, X.; Pritchard, B. P.; Prakash, V.; Schaefer, H. F.; Patkowski, K.; King, R. A.; Valeev, E. F.; Evangelista, F. A.; Turney, J. M.; Crawford, T. D.; Sherrill, D. PSI4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Interoperability. *J Chem Theory Comput* **2017,** *13* (7), 3185.

(104) Doran, A. E.; Hirata, S. Monte Carlo MP2 on Many Graphical Processing Units. *J Chem Theory Comput* **2016,** *12* (10), 4821.

(105) Steele, R. P. Communication: Multiple-timestep ab initio molecular dynamics with electron correlation. *J Chem Phys* **2013,** *139* (1).

(106) Fatehi, S.; Steele, R. P. Multiple-Time Step Ab Initio Molecular Dynamics Based on Two-Electron Integral Screening. *J Chem Theory Comput* **2015,** *11* (3), 884.

(107) Steele, R. P. Multiple-Timestep ab Initio Molecular Dynamics Using an Atomic Basis Set Partitioning. *J Phys Chem A* **2015,** *119* (50), 12119.

(108) Mazzacurati, V.; Nardone, M. Raman-Scattering from Amorphous Ice Films. *Chem Phys Lett* **1975,** *32* (1), 99.

(109) Klug, D. D.; Mishima, O.; Whalley, E. Raman-Spectrum of High-Density Amorphous Ice. *Physica B & C* **1986,** *139* (1-3), 475.

(110) Klug, D. D.; Mishima, O.; Whalley, E. High-Density Amorphous Ice IV. Raman-Spectrum of the Uncoupled O-H and O-D Oscillators. *J Chem Phys* **1987,** *86* (10), 5323.

(111) Devlin, J. P. Polarized Raman-Spectra for the Full Range of Isotopic Dilution for Ice Ic and Amorphous Ice - Mixtures of Intact H_2O and D_2O. *J Chem Phys* **1989,** *90* (3), 1322.

(112) Kanno, H.; Tomikawa, K.; Mishima, O. Raman spectra of low- and high-density amorphous ices. *Chem Phys Lett* **1998,** *293* (5-6), 412.

(113) Suzuki, Y.; Takasaki, Y.; Tominaga, Y.; Mishima, O. Low-frequency Raman spectra of amorphous ices. *Chem Phys Lett* **2000,** *319* (1-2), 81.

(114) Suzuki, Y.; Tominaga, Y. Polarized Raman spectroscopic study of relaxed high density amorphous ices under pressure. *J Chem Phys* **2010,** *133* (16), 164508.

(115) Yoshimura, Y.; Stewart, S. T.; Somayazulu, M.; Mao, H. K.; Hemley, R. J. Convergent Raman Features in High Density Amorphous Ice, Ice VII, and Ice VIII under Pressure. *J Phys Chem B* **2011,** *115* (14), 3756.

(116) Galindo, A.; Burton, S. J.; Jackson, G.; Visco, D. P.; Kofke, D. A. Improved models for the phase behaviour of hydrogen fluoride: chain and ring aggregates in the SAFT approach and the AEOS model. *Mol Phys* **2002,** *100* (14), 2241.

(117) Toth, Z.; Pulay, P. Finding symmetry breaking Hartree-Fock solutions: The case of triplet instability. *J Chem Phys* **2016,** *145* (16), 164102.

(118) Maeda, S.; Ohno, K.; Morokuma, K. An Automated and Systematic Transition Structure Explorer in Large Flexible Molecular Systems Based on Combined Global Reaction Route Mapping and Microiteration Methods. *J Chem Theory Comput* **2009,** *5* (10), 2734.

(119) Maeda, S.; Ohno, K.; Morokuma, K. Automated global mapping of minimal energy points on seams of crossing by the anharmonic downward distortion following method: a case study of H_2CO. *J Phys Chem A* **2009,** *113* (9), 1704.

(120) Maeda, S.; Ohno, K.; Morokuma, K. Systematic exploration of the mechanism of chemical reactions: the global reaction route mapping (GRRM) strategy using the ADDF and AFIR methods. *Phys Chem Chem Phys* **2013,** *15* (11), 3683.

(121) Thom, A. J.; Head-Gordon, M. Locating multiple self-consistent field solutions: an approach inspired by metadynamics. *Phys Rev Lett* **2008,** *101* (19), 193001.

(122) Purwanto, W.; Zhang, S.; Krakauer, H. An auxiliary-field quantum Monte Carlo study of the chromium dimer. *J Chem Phys* **2015,** *142* (6), 064302.

(123) Ma, Y. J.; Knecht, S.; Keller, S.; Reiher, M. Second-Order Self-Consistent-Field Density-Matrix Renormalization Group. *J Chem Theory Comput* **2017,** *13* (6), 2533.

(124) Paldus, J.; Wormer, P. E. S.; Benard, M. Coupled-Pair Theories and Davidson-Type Corrections for Quasidegenerate States - the H_4 Model Revisited. *Collect Czech Chem C* **1988,** *53* (9), 1919.

(125) Paldus, J.; Piecuch, P.; Pylypow, L.; Jeziorski, B. Application of Hilbert-Space Coupled-Cluster Theory to Simple $(H_2)_2$ Model Systems - Planar Models. *Phys Rev A* **1993,** *47* (4), 2738.

(126) Piecuch, P.; Tobola, R.; Paldus, J. Approximate Account of Connected Quadruply Excited Clusters in Multireference Hilbert-Space Coupled-Cluster Theory - Application to Planar H_4 Models. *Chem Phys Lett* **1993,** *210* (1-3), 243.

(127) Piecuch, P.; Paldus, J. Application of Hilbert-Space Coupled-Cluster Theory to Simple $(H_2)_2$ Model Systems II. Nonplanar Models. *Phys Rev A* **1994,** *49* (5), 3479.

(128) Sinitskiy, A. V.; Greenman, L.; Mazziotti, D. A. Strong correlation in hydrogen chains and lattices using the variational two-electron reduced density matrix method. *J Chem Phys* **2010,** *133* (1), 014104.

(129) Stella, L.; Attaccalite, C.; Sorella, S.; Rubio, A. Strong electronic correlation in the hydrogen chain: A variational Monte Carlo study. *Phys Rev B* **2011,** *84* (24), 245117.

(130) Cramer, C. J.; Wloch, M.; Piecuch, P.; Puzzarini, C.; Gagliardi, L. Theoretical models on the Cu_2O_2 torture track: mechanistic implications for oxytyrosinase and small-molecule analogues. *J Phys Chem A* **2006,** *110* (5), 1991.

(131) Pederson, M. R.; Porezag, D. V.; Kortus, J.; Khanna, S. N. Theoretical calculations of magnetic order and anisotropy energies in molecular magnets. *J Appl Phys* **2000,** *87* (9), 5487.

(132) Pederson, M. R.; Kortus, J.; Khanna, S. N. Electronic-structure-based investigation of magnetism in the Fe-8 molecular magnet. *J Appl Phys* **2002,** *91* (10), 7149.

(133) Pederson, M. R.; Liu, A. Y.; Baruah, T.; Kurmaev, E. Z.; Moewes, A.; Chiuzbaian, S.; Neumann, M.; Kmety, C. R.; Stevenson, K. L.; Ederer, D. Electronic structure of the molecule-based magnet $Mn[N(CN)_2]_2$ from theory and experiment. *Phys Rev B* **2002,** *66* (1), 014446.

(134) Ribas-Arino, J.; Baruah, T.; Pederson, M. R. Density-functional study of two Fe_4-based single-molecule magnets. *J Chem Phys* **2005,** *123* (4).

(135) Postnikov, A. V.; Kortus, J.; Pederson, M. R. Density functional studies of molecular magnets. *Phys Status Solidi B* **2006,** *243* (11), 2533.

(136) Hachmann, J.; Dorando, J. J.; Aviles, M.; Chan, G. K. The radical character of the acenes: a density matrix renormalization group study. *J Chem Phys* **2007,** *127* (13), 134309.

(137) Mizukami, W.; Kurashige, Y.; Yanai, T. More pi Electrons Make a Difference: Emergence of Many Radicals on Graphene Nanoribbons Studied by Ab Initio DMRG Theory. *J Chem Theory Comput* **2013,** *9* (1), 401.

(138) Hachmann, J.; Frazier, B. A.; Wolczanski, P. T.; Chan, G. K. A theoretical study of the 3d-$M(smif)_2$ complexes: structure, magnetism, and oxidation states. *ChemPhysChem* **2011,** *12* (17), 3236.

(139) Kurashige, Y.; Chan, G. K. L.; Yanai, T. Entangled quantum electronic wavefunctions of the Mn4CaO5 cluster in photosystem II. *Nat Chem* **2013,** *5* (8), 660.

(140) Overhauser, A. W. Giant Spin Density Waves. *Phys Rev Lett* **1960,** *4* (9), 462.

(141) Lee, P. A.; Nagaosa, N.; Wen, X. G. Doping a Mott insulator: Physics of high-temperature superconductivity. *Rev Mod Phys* **2006,** *78* (1), 17.

(142) Phillips, P. Mottness. *Ann Phys-New York* **2006,** *321* (7), 1634.

(143) Phillips, P.; Choy, T. P.; Leigh, R. G. Mottness in high-temperature copper-oxide superconductors. *Rep Prog Phys* **2009,** *72* (3), 036501.