**Title and Subtitle**

DEVELOPMENT OF COMPUTATIONAL CHEMISTRY SOFTWARE

**Abstract**

The objective of this work is to develop a broad array of methods for electronic structure theory, with special emphasis on improved parallel codes. A new highly scalable code for frozen core second order perturbation theory (MP2) gradients for closed shell molecules has been developed and is on-line for general use at several IBM SP2 and Cray T3E systems. The key element of this approach is our distributed data interface (DDI) that facilitates distribution of very large arrays across all available nodes. This code has been ported to the cluster environment, so it can now be used on clusters of PCs or Macs running Linux, as well as on clusters of high performance workstations. The development of analogous codes for molecules with unpaired electrons is in progress. The sequential unrestricted MP2 code has been completed, and a parallel version of this code is in progress. The derivation for the preferred restricted open shell ZAPT method is complete and a paper describing this method has been accepted for publication. The code for implementing Hartree-Fock with DDI has been completed, and a paper describing this code is in press. We are now turning to analogous developments for MCSCF and CI codes. A scalable code for the multi-reference second order perturbation theory has been completed.
1. COVER SHEET

AFOSR FINAL REPORT
10/1/97 – 1/31/02

DEVELOPMENT OF NEW SOFTWARE FOR ELECTRONIC STRUCTURE THEORY (CHSSI)

AFOSR F49620-97-1-0522

MARK S. GORDON

IOWA STATE UNIVERSITY

AMES, IA 50011
2. **Objectives:** No Change in Objectives

3. **Status of effort:** A brief statement of progress towards achieving the research objectives. (Limit to 200 words.)

   The objective of this work is to develop a broad array of methods for electronic structure theory, with special emphasis on improved parallel codes. A new highly scalable code for frozen core second order perturbation theory (MP2) gradients for closed shell molecules has been developed and is on-line for general use at several IBM SP2 and Cray T3E systems. The key element of this approach is our distributed data interface (DDI) that facilitates distribution of very large arrays across all available nodes. This code has been ported to the cluster environment, so it can now be used on clusters of PCs or Macs running Linux, as well as on clusters of high performance workstations. The development of analogous codes for molecules with unpaired electrons is in progress. The sequential unrestricted MP2 code has been completed, and a parallel version of this code is in progress. The derivation for the preferred restricted open shell ZAPT method is complete and a paper describing this method has been accepted for publication. The code for implementing Hartree-Fock with DDI has been completed, and a paper describing this code is in press. We are now turning to analogous developments for MCSCF and CI codes. A scalable code for the multi-reference second order perturbation theory has been completed. Other important developments include: the derivation of gradients for multi-reference second order perturbation theory, further developments of our effective fragment potential (EFP) method for studying solvation and liquid behavior, the development of molecular dynamics and Monte Carlo methods to facilitate the study of solvation and liquid behavior, the development and implementation of a new method for producing global potential energy surfaces from sets of ab initio points, the development and implementation of both grid-based and gridless approaches to density functional theory, and the development and implementation of several MCSCF-based approaches to spin-orbit coupling.

4. **Accomplishments/New Findings:** Describe research highlights, their significance to the field, their relationship to the original goals, their relevance to the AF's mission, and their potential applications to AF and civilian technology challenges.

   The objective of this work is to develop a broad array of methods for electronic structure theory, with special emphasis on improved parallel codes. Progress has been made in many areas.

   **Parallel Developments.** A highly scalable code for frozen core second order perturbation theory gradients for closed shell molecules has been developed and is on-line for general use at several IBM SP2 and Cray T3E systems, including the T3Es at ERDC and NAVO and the SP2 at Maui. This new code uses the Distributed Data Interface (DDI) so that the large arrays do not have to be replicated. This means that much larger molecules can be
studied using geometries based on correlated wavefunctions. The code scales very well for as many as 512 nodes and has already been applied to several challenging compounds. The development of analogous codes for molecules with unpaired electrons, using both restricted and unrestricted wavefunctions, is in progress. A paper on the restricted open-shell derivation has been accepted. A paper on a DDI implementation of the SCF method has been accepted for publication. Improved parallel MCSCF codes are under development, and an initial parallel full CI code has been implemented. The multi-reference perturbation theory code (MCQDPT) has now been implemented under DDI. Part of the effective fragment potential (EFP) code has been made scalable, and the remainder of the code will be addressed in the coming year. In an important related development, parallel GAMESS now runs on clusters of PCs and Macs (running Linux) and high end workstations. This is more challenging than self-contained massively parallel computers, since the overhead due to inter-node communications is more complex. This issue has been solved to some extent by using a Gigabit Ethernet switch with large data packets. We are also exploring alternative communications solutions, such as Myrinet and SCI. These developments will be enhanced by the construction of a new 32-node, 128-CPU IBM Power3+ cluster using funds provided by a DURIP grant. As for all other GAMESS developments, we will make our experiences in developing scalable clusters available to all users.

Condensed Phase Methods. We have already shown that our effective fragment potential (EFP) method for solvation is excellent for water, in a variety of applications. We are now working on extending the capabilities of the method in several ways. We are exploring several alternative approaches for incorporating dispersion and other higher order terms into the method. Such terms are particularly important for nonpolar solvents. We are also in the process of extending the model so that it is equally applicable and accurate for any solvent. Key to the success will be the derivation of general expressions for charge transfer and dispersion contributions that contain no fitted parameters. The derivation of an expression for the analytic gradient for the EFP-\textit{ab initio} interaction term is in progress. A very important new development (by co-worker Jan Jensen) is a new method for using EFPs across covalent bonds. This will facilitate the representation of large substituents, as well as the treatment of large biological molecules. We have also interfaced the EFP method with two continuum methods, the simple Onsager reaction field and the more sophisticated polarizable continuum model (PCM).

As the number of solvent molecules in the system increases, the number of configurations to be considered increases rapidly, and traditional small molecule geometry optimization methods are not effective. We are therefore developing both molecular dynamics and Monte Carlo simulation codes so that the configurational space can be probed more effectively, not only for minima, but for transition states and reaction paths as well. This is also an important first step in the development of methods based on the EFP model for predicting bulk properties and super-critical behavior. We are therefore incorporating Ewald summations into our MD method. Collaborators in this effort are Dr. Paul Day (AFRL) and Professors Teresa Head-Gordon (UC-Berkeley) and Grant Merrill (University of Missouri). Related to these dynamical methods is our development and implementation of a method for converting the large numbers of points generated in \textit{ab}
initio trajectories into a global potential energy surface, using a modified Shepard interpolation approach.

Major James Shoemaker's Ph.D dissertation in Engineering Physics at the Air Force Institute of Technology focused on the development and applications of an embedded cluster model for treating surface chemistry. A paper on the theoretical method, called SIMOMM, has been published in the Journal of Physical Chemistry. Several papers applying this method have now been published or submitted, and a manuscript that describes applications to silicon carbide surfaces is in progress. Extensions of this method to transition metal catalysts are planned; however, molecular mechanics is not likely to be a viable approach for the bulk, since electrons in such systems are too delocalized. To overcome this potential problem, we have begun to explore fast multipole methods (FMM) that scale linearly and are also highly parallel, so that the bulk part of the system can be treated by quantum mechanics. One preliminary paper has been published and the first FMM paper has been published.

Other developments There are many problems that involve species for which one cannot write one simple Lewis structure. Such species must be treated with multi-configurational (MCSCF) wavefunctions. However, the computational demands for MCSCF calculations are very high. So, we have developed a set of methods for converging such wavefunctions that range from very rapid convergence with high CPU and disk requirements to methods that converge more slowly but have very modest computational demands. The advantage of the latter is that much larger molecules than normal can be treated with this method.

For species that must be treated with MCSCF wavefunctions, non-traditional methods are also needed to recover the dynamic correlation. One method that has been implemented into GAMESS is multi-reference second order perturbation theory. The gradient formulas for this method have been derived, and the implementation into GAMESS is in progress. The key collaborator in this effort is Dr. Haruyuki Nakano (University of Tokyo).

There are many problems that involve multiple electronic states on multiple potential energy surfaces. Since such surfaces can cross and therefore induce radiationless transitions, it is important to have the capability to predict non-adiabatic coupling among these surfaces. One origin of non-adiabatic interactions among surfaces is spin-orbit coupling. Spin-orbit coupling matrix elements must be calculated using multi-reference wavefunctions, but until recently, there have been no generally available codes for the computation of these matrix elements using all electrons and any arbitrary MCSCF wavefunction for any arbitrary set of spin states. We have now developed and implemented such a method. It will be particularly valuable for the study of high energy density materials and for the study of any processes involving transition metals, since both types of species involve multiple electronic states. Since such all-electron calculations are very demanding, two alternative approaches to this problem have also been develop and implemented into GAMESS for use with larger molecules. One of these is called the $Z_{\text{eff}}$ method, in which the two electron term is ignored and a fitted
empirical parameter is introduced into the one-electron part of the Hamiltonian. While this is a very efficient approach, its accuracy is limited. However, there is an intermediate method, in which only those matrix elements connecting orbitals in the active space with each other are neglected. We have shown that in this case excellent agreement with the all electron method is obtained, without any need for fitted parameters. We have also implemented the RESC (relativistic elimination of the small component) method for calculating spin-orbit coupling interactions, complete with analytic gradients.

Density functional theory (DFT) has become extraordinarily popular in recent years, because it has the same computational requirements as Hartree-Fock theory, but often has a level of accuracy that is comparable to second order perturbation theory. We have developed and implemented a gridless approach to DFT, by using the resolution of the identity to evaluate the complicated integrals. This procedure requires the use of auxiliary basis sets to resolve the identity. After extensive analysis of the convergence of several molecular properties as a function of basis set, we have recommended a series of auxiliary basis sets for use with the first row in the periodic table. The next step will be to extend this basis set development to heavier elements. We have also implemented a more traditional grid-based DFT method, including all of the most popular functionals.

Another approach that is very important for large molecules, especially molecules that contain heavier elements, is the effective core potential (ECP) method. This method replaces inner shell electrons by model potentials, making the calculations much more efficient. These potentials include relativistic corrections for the heavier elements. To make these calculations more useful, we have developed much more efficient methods for obtaining the first and second energy derivatives; that is, the gradients and hessians. The increase in speed is a factor of 5-10, and the method has been implemented to take advantage of parallel computers.

As calculations on larger and larger species are made feasible by advances in parallel programs and other innovative algorithms, it becomes increasingly important to develop graphical methods for interpreting the results of complex calculations. This has led to the continuing development of MacMolPlt, a graphical interface for GAMESS. MacMolPlt provides the ability to visualize (and rotate in real time) orbitals, wavefunctions, electron densities and electrostatic potentials in two and three dimensions. We are now in the process of developing a Windows based version of this code.

We have also been exploring alternatives and extensions to the popular G2 and G3 methods developed by Pople and co-workers for the accurate prediction of such thermodynamic properties as heats of formation, ionization potentials and electron affinities. One limitation of these methods is that they are applicable only to species that are adequately described by single configuration wavefunctions. This eliminates many transition metal compounds, diradicals and most transition states. We have therefore initiated, in collaboration with the Radom group, the development of multi-reference analogs of the G2 and G3 methods. The multi-reference methods are based on CASSCF wavefunctions, instead of Hartree-Fock, followed by multi-reference perturbation theory.
(instead of MP2) and finally multi-reference CI. The initial paper describing this method has been published.

5. **Personnel Supported**: List professional personnel (Faculty, Post-Docs, Graduate Students, etc.) supported by and/or associated with the research effort.

   Dr. Michael W. Schmidt, Senior Research Associate  
   Mr. Yuri Alexeev, graduate student  
   Dr. Pradipta Bandyopadhyay, postdoctoral research associate  
   Dr. Brett Bode, graduate student, postdoctoral research associate  
   Dr. Galina Chaban, graduate student, postdoctoral research associate  
   Dr. Cheol Ho Choi, postdoctoral research associate  
   Mr. Dmitri Fedorov, graduate student  
   Dr. Graham Fletcher, postdoctoral research associate  
   Dr. Kurt Glaesemann, graduate student, postdoctoral research associate  
   Professor Jan Jensen, University of Iowa  
   Professor Shiro Koseki, Mie University  
   Professor Takako Kudo, Gunma University  
   Dr. Haruyuki Nakano, Tokyo University  
   Mr. Ryan Olson, graduate student  
   Dr. Walter Stevens, NIST

6. **Publications**: Peer-reviewed publications submitted and/or accepted during the 12-month period starting the previous 1 October.

**PUBLICATIONS**


15. D.G. Fedorov, M. Evans, Y. Song, M.S. Gordon and C.Y. Ng, "An Experimental and Theoretical Study of the Spin-orbit Interaction for CO$^+$($A^2\Pi_{\frac{3}{2},\frac{1}{2}}$, $v^*=0-41$) and O$_2^+$ ($X^2\Pi_{\frac{3}{2},\frac{1}{2}}$, $v^*=0-38$)", J. Chem. Phys., 111, 6413 (1999).


7. **Interactions/Transitions:**

Participation/presentations at meetings, conferences, seminars, etc.

**TALKS AT PROFESSIONAL MEETINGS**


5. “The Role of Theory in Chemical Vapor Deposition”, Electrochemical Society Meeting (INVITED), Boston, November 1998


11. “Structure and Energetics of Silicon Carbide Clusters, \( \text{Si}_n\text{C}_m, n + m = 4 \)” , presented by Jamie Rintelman, , American Conference on Theoretical Chemistry, Boulder, CO, June 1999


19. “Cycloaddition reactions of 1,3-cyclohexadiene on Silicon (001) surface using SIMOMM”, presented by Cheol Ho Choi, National ACS Meeting, New Orleans, LA, August 1999

20. “Direct recurrence relations for the rapid and stable determination of rotation matrices between complex spherical harmonics”, presented by Cheol Ho Choi, National ACS Meeting, New Orleans, LA, August 1999


23. Excited State Potential Energy Surfaces from Diatomics to Biomolecules, National ACS Meeting, San Diego, April 2001 (INVITED)


25. “The mechanism for formation of Al₂O₃ from Al and O₂” AFOSR HEDM Contractors’ Meeting, October 2000, Michael Pak presenting.


INVITED LECTURES

1. "Ubiquitous Water", Advanced Supercomputer Center, Wright-Patterson AFB, October 1997

2. "Ubiquitous Water", Coulson Lecture, University of Georgia, October 1997

3. "Ubiquitous Water", University of Utah, November 1997


5. "Ubiquitous Water", Melbourne University, February 1998


7. "Ubiquitous Water", Roseman Lecture, Johns Hopkins University, April 1998


9. “Ubiquitous Water”, University of Nebraska, September 1998


11. “Ubiquitous Water”, Marquette University, February 1999
Consultative and advisory functions to other laboratories and agencies, especially Air Force and other DoD laboratories. Provide factual information about the subject matter, institutions, locations, dates, and name(s) of principal individuals involved.

- Served as co-advisor to Major James Shoemaker, who received his Ph.D. from the Department of Engineering Physics, Air Force Institute of Technology, October 1997. Several papers will result from this work; one is already in print [J. Phys. Chem. A, 103, 3245 (1999)], one is submitted to J. Chem. Phys, and a third is in preparation.
• Presented CHSSI workshop on GAMESS at the Army Research Laboratory, June 1998 and July 1999 (presented by Jerry Boatz).

• Collaborating with the Group of Dr. Ruth Pachter (AFRL) to develop new codes for molecular dynamics and Monte Carlo simulations (ongoing).

• Collaborating with Dr. Ruth Pachter to develop density-functional-based effective fragment potentials.

c. Transitions. Describe cases where knowledge resulting from your effort is used, or will be used, in a technology application. Transitions can be to entities in the DOD, other federal agencies, or industry. Briefly list the enabling research, the laboratory or company, and an individual in that organization who made use of your research.

Both GAMESS and MacMolPlt have an increasing number of users in universities, industry, and government laboratories. The latest version of parallel GAMESS is on-line at all DoD Major Shared Resource Centers.

8. New discoveries, inventions, or patent disclosures. (If none, report None.) None

9. Honors/Awards: List honors and awards received during the grant/contract period. List lifetime achievement honors such as Nobel prize, honorary doctorates, and society fellowships prior to this effort.

‖ Coulson Lecturer, University of Georgia, October, 1997

‖ Roseman Lecturer, Johns Hopkins University, April, 1998

‖ Named Distinguished Professor of Chemistry, May, 1998

‖ Named to list of 100 most cited chemists, 1981-2000.

10. Markings: In order to ensure prompt receipt and acceptance, mark the outside of the package clearly to indicate that it is a progress report.