Q-Chem 2.0: A High-Performance *Ab Initio* Electronic Structure Program Package

JING KONG,[1] CHRISTOPHER A. WHITE,[2,*] ANNA I. KRYLOV,[2,†] DAVID SHERRILL,[2,‡] ROSS D. ADAMSON,[3] THOMAS R. FURLANI,[4] MICHAEL S. LEE,[2] AARON M. LEE,[3] STEVEN R. GWALTNEY,[2] TERRY R. ADAMS,[1,¶] CHRISTIAN OCHSENFELD,[5] ANDREW T. B. GILBERT,[6] GARY S. KEDZIORA,[7] VITALY A. RASSOLOV,[7] DAVID R. MAURICE,[2,#] NIKHIL NAIR,[3] YIHAN SHAO,[2] NICHOLAS A. BESLEY,[6] PAUL E. MASLEN,[2,] JEREMY P. DOMBROSKI,[1,††] HOLGER DASCHEL,[1] WEIMIN ZHANG,[1] PRAKASHAN P. KORAMBATH,[1] JON BAKER,[8] EDWARD F. C. BYRD,[2] TROY VAN VOORHIS,[2] MANABU OUMI,[2,‡‡] SO HIRATA,[2,¶¶] CHAO-PING HSU,[2] NAOTO ISHIKAWA,[9] JAN FLORIAN,[10] ARIEH WARSHEL,[10] BENNY G. JOHNSON,[1,##] PETER M. W. GILL,[6] MARTIN HEAD–GORDON,[2] JOHN A. POPLE[7]**

[1] *Q-Chem Inc, Four Triangle Lane, Suite 160, Export, Pennsylvania 15632*
[2] *Department of Chemistry, University of California, Berkeley, Berkeley, California 94720*
[3] *Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom*
[4] *Center for Computational Research, State University of New York at Buffalo, Buffalo, New York 14260*
[5] *Institut für Physikalische Chemie, Universität Mainz, D-55099 Mainz, Germany*
[6] *School of Chemistry, University of Nottingham, Nottingham NG7 2RD, United Kingdom*
[7] *Department of Chemistry, Northwestern University, Evanston, Illinois 60208*
[8] *Department of Chemistry, University of Arkansas, Fayetteville, Arkansas 72701-1201*
[9] *Department of Chemistry, Tokyo Institute of Technology, Meguro Ku, Tokyo 1528551, Japan*
[10] *Department of Chemistry, University of Southern California, Los Angeles, California 90089-1062*

[*] Present address: Bell Laboratories, Lucent Technologies, 600 Mountain Ave., Murray Hill, NJ 07974

[†] Present address: Department of Chemistry, University of Southern California, Los Angeles, CA 90089-1062

[‡] Present address: School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, GA 30332-0400

[¶] Present address: SGI Federal, 102 Central Park Square, MS 298, Los Alamos, NM 87544

[#] Present address: APAM, Inc., 201 E. Pine St., Suite 600, Orlando, FL 32801

[**] Present address: Department of Chemistry, Rutgers University, Camden, NJ 08102

[††] Present address: HortResearch, Private Bag 11030, Palmerston North, New Zealand

ABSTRACT: Q-Chem 2.0 is a new release of an electronic structure program package, capable of performing first principles calculations on the ground and excited states of molecules using both density functional theory and wave function-based methods. A review of the technical features contained within Q-Chem 2.0 is presented. This article contains brief descriptive discussions of the key physical features of all new algorithms and theoretical models, together with sample calculations that illustrate their performance.   © 2000 John Wiley & Sons, Inc.   J Comput Chem 21: 1532–1548, 2000

Keywords: quantum chemistry; electronic structure; density functional theory; computer program; computational chemistry

Introduction

A reader glancing casually at this article might suspect on the basis of its title that it is a thinly disguised piece of marketing for a program package. This is not the case. Rather, it is an attempt to document the key methodologies and algorithms of our electronic structure program package, Q-Chem 2.0, in a complete and scientifically accurate way, with full references to the original literature. This is important for two principal reasons. First, while the use of electronic structure programs is burgeoning, many users of such programs do not have much feel for the underlying algorithms that make large-scale calculations routine even on such readily available hardware as personal computers. Therefore, a link between the program package and the original literature that is written at the level of an introductory overview can be a useful bridge. Second, while citations of large-scale commercial programs in published applications are traditionally part of the conditions of use of such codes, they

are "empty citations" in that there is nothing useful that a reader may learn by looking up the citation. Other public domain electronic structure programs have provided helpful review articles (e.g., refs. 1 and 2), as have programs for molecular mechanics (e.g., ref. 3). This article will serve as the official citation for Q-Chem 2.0, and as such is intended to provide a summary of its capabilities that can provide useful information beyond a simple program citation.

As background, the development of Q-Chem commenced just before 1993, when one of the cofounders, Prof. Peter Gill (Nottingham), then a post-doctoral researcher with John Pople in Pittsburgh, began to put together parts of the program while on his Christmas vacation in New Zealand. Indeed, Q-Chem's early developers were a splinter group from the scientific collaboration responsible for the development of the Gaussian programs that grew out of the Pople laboratory. The other cofounders of Q-Chem were Dr. Benny Johnson, at that time a graduate student in the Pople group, and Dr. Carlos Gonzalez, then with the Pittsburgh Supercomputer Center. Martin Head–Gordon, who had just taken up an assistant professorship at the University of California, Berkeley, joined the development effort in May, 1993, together with his research group. The first commercial release of the code, version 1.0, occurred in mid-1997, followed by version 1.1 in December 1997 and version 1.2 in September, 1998. John Pople's research group (Northwestern) joined the development effort in January 1999. The new release described here is the joint effort of Q-Chem employees, particularly Dr. Jing Kong, numerous students and postdoctoral researchers in the Gill, Head–Gordon, and Pople groups (a number of whom now have academic positions), and several independent contributors. Their names comprise the author list of this article.

‡‡Present address: Seiko Instruments, Corporate R&D, Takatsuka Unit, 563 Takatsuka-Shinden, Matsudo-shi, Chiba 270-2222, Japan

¶¶Present address: Quantum Theory Project, University of Florida, Gainesville, FL 32611-8435

##Present address: 5275 Sardis Rd, Murrysville, PA 15668

Correspondence to: M. Head–Gordon; e-mail: mhg@cchem.berkeley.edu

To act as a bridge between readily available textbook information on electronic structure methods[4-6] and the state-of-the-art in electronic structure algorithms requires considerable heterogeneity in the level of presentation. Standard methods that are well known in textbooks are very briefly summarized, while new methods that offer novel capabilities are discussed in more depth, and in some cases selected numerical examples are presented. Similarly, it is impossible to take a comprehensive approach to referencing, because we are surveying a considerable part of electronic structure theory. The focus is, therefore, on articles directly relevant to Q-Chem's methods and algorithms. Because there is a tremendous diversity of methods included in Q-Chem, we also try to indicate very roughly their extent of applicability. We place emphasis on a variety of newly developed capabilities.

We have chosen to organize this article around the functional capabilities of the program. The very first issue is the manner in which functions of a single electron are represented, aspects of which are summarized in the next section. This defines the atomic orbital basis functions used to represent orbitals, and their matrix elements. In the following section, the methods and algorithms available for describing the ground states of molecules by self-consistent field methods (either Hartree–Fock or density functional) are discussed. The often much more expensive but systematic wave function-based methods for describing the correlations between electrons in the ground electronic state are the topic of the Wave Function-Based Treatments section. In the Excited State Methods section we turn from molecular ground states to a discussion of the methods available in Q-Chem 2.0 for describing the transitions from ground to electronically excited states. The techniques available for obtaining chemical insight from electronic structure calculations are then surveyed in Wave Function Analysis section. In the Additional Capabilities Section, a selection of important additional capabilities of Q-Chem 2.0 are summarized, including solvation modeling, relativistic energy corrections, geometry optimization techniques, and user interfaces.

Basis Functions and Their Matrix Elements

Q-Chem is entirely based on the use of contracted Gaussian atomic orbital basis sets. Gaussian basis sets are the most widely used form of spatially localized basis functions because their special properties permit analytical evaluation of all the one- and two-electron matrix elements of the Hamiltonian. A contracted Gaussian basis function is, in general, a linear combination of spherical Gaussian functions centered on each atom, multiplied by an angular function that is typically either a simple product of powers of x, y, z, or linear combinations thereof. Many standard Gaussian basis sets have been developed, and a number of general reviews of Gaussian basis sets are available.[7-9] Q-Chem contains a fairly comprehensive library of basis sets built into the program, as well as the ability to input other basis sets. In particular, all standard Pople basis sets are available,[9] including the most recent extensions of the 6-31G* basis set for all elements up to krypton.[10] The systematic sequence of correlation consistent basis sets of Dunning and coworkers[11] are available up through quadruple zeta, as well as Dunning's earlier double and triple zeta basis sets.[12] The modern Ahlrichs double and triple zeta basis sets are also available.[13] Alternatively, for the rare situations where these internal basis sets are inadequate or inappropriate, general, user-supplied basis sets can be entered as part of the job input.

For heavier elements, Q-Chem contains pseudopotentials that permit only the chemically relevant valence electrons to be treated explicitly, while the chemically inert core electrons are treated implicitly. A variety of standard pseudopotentials and associated basis sets are available, as well as the ability to input custom pseudopotentials. The standard pseudopotential basis sets were obtained from the Pacific Northwest Laboratory database (as were the all-electron Dunning and Ahlrichs basis sets discussed above.[14] They are as follows: (1) Hay–Wadt Minimal Basis,[15] (2) Hay–Wadt Valence Double Zeta;[15] (3) lanl2dz (mimic of Gaussian's lanl2dz);[15] (4) Stevens–Bausch–Krauss–Jaisen–Cundari -21G;[16] (5) CRENBL–Christiansen et al. shape consistent large orbital, small core;[17] (6) CRENBS–Christiansen et al. shape consistent small basis large core;[17] (7) Stuttgart relativistic large core;[18] (8) Stuttgart relativistic small core.[19] We note that relativistic energy corrections to all-electron calculations are also available in Q-Chem and are discussed separately in the Additional Capabilities section.

One- and two-electron matrix elements involving contracted Gaussian basis functions are efficiently evaluated in Q-Chem using the most recent version[20, 21] of the PRISM methods[22] developed by Gill and coworkers. PRISM is probably the most efficient method available for evaluation of two-electron Gaussian integrals. PRISM is a family of algorithms, each of which is optimum for certain

types of integrals, depending on the angular momentum and the degree of contraction (how many primitive Gaussian integrals are combined together to make the target contracted integral). For each type of integral, the optimum method is chosen, which accounts for the adaptability and efficiency of the method. PRISM is employed both for Gaussian integrals, and also their first and second derivatives, as needed for analytical forces and vibrational frequency calculations, respectively.

Recently, Adamson and Gill[23] have extended the capabilities of PRISM to permit fully analytical treatment of integrals over pseudopotential operators of the form first proposed by Kahn et al.[24] The new fully analytical approach is in contrast to existing methods that employ numerical integration for at least some of the radial integrals that arise in pseudopotential calculations.[25] This has considerable benefits in terms of execution speed, as documented elsewhere.[23] Analytical first derivatives of pseudopotential matrix elements are also implemented using the new approach, to enable efficient force calculations for heavier elements.

Finally, we note that Q-Chem contains "fast" algorithms for linear scaling evaluation of the matrix elements that arise in practical self-consistent field calculations: they are discussed separately later. Additionally, Q-Chem contains efficient routines for transforming two-electron integrals from the atomic orbital basis to various other representations: they are discussed later in the section on local electron correlation.

Ground-State Self-Consistent Field Methods

The most inexpensive *ab initio* electronic structure methods are those that are based on only a single determinant of molecular orbitals, which may be generally described as self-consistent field (SCF) methods. After a general overview, the specialized "fast methods" available in Q-Chem for large-molecule SCF calculations are described.

CONVENTIONAL AND UNCONVENTIONAL SELF-CONSISTENT FIELD METHODS

If the single determinant is viewed as a trial wave function to be optimized via the variational principle, the result is the Hartree–Fock (HF) theory, in which each electron moves in the average field of all the others.[26] HF theory neglects the instantaneous correlations between electrons, which

can be treated by the wave function-based methods described in the following section. As a result of neglecting electron correlations, Hartree–Fock underestimates the strength of chemical bonds, and is usually poor for predicting reaction energies unless correlation energy is roughly constant between reactants and products. HF theory is much more successful for predicting molecular geometries and vibrational frequencies.[27]

A very attractive alternative is the modern Kohn–Sham density functional theory (DFT).[28, 29] In DFT, the single determinant is employed to parameterize the electron density, and to evaluate the kinetic energy. By the Hohenberg–Kohn theorem, which states that the ground-state exchange and correlation energy is a universal functional of the electron density, these terms are evaluated as complicated functional expressions. Q-Chem contains all popular standard density functionals, as well as the ability to input new ones within certain constraints on the functional form. The best common DFT methods, exemplified by the B3LYP functional,[30] yield results for relative energies in particular, but also often structural properties, which are greatly improved over HF theory. Many reviews are now available documenting the successes and limitations of modern DFT.[31] It is the most cost-effective electronic structure method available in Q-Chem, as it contains much of the physics associated with electron correlation within a framework as simple as mean field theory!

In addition to the established density functionals, Q-Chem contains several newly developed functionals. These take quite different approaches to functional design. The Empirical Density Functional 1 (EDF1) is specifically adapted to yield good results with the relatively modest-sized 6-31+G* basis set, by direct fitting to thermochemical data.[32] It has the interesting feature that exact exchange mixing was not found to be helpful with a basis set of this size. A second even more recent functional is the GG99 exchange functional,[33] which is parameter free. This is accomplished by following an approach first taken by the pioneers of the density functional theory: choosing a model density and deriving a functional that is exact for the model. The GG99 functional exactly reproduces the Hartree–Fock (HF) gradient decomposition of the energy for the model system. When coupled with the LYP correlation functional, GG99 gives a mean absolute deviation from experiment of 6.88 kcal/mol for the G2 set of atomization energies.

Q-Chem evaluates not only SCF energies, but also their analytical first and second derivatives,

as needed for geometry optimizations and vibrational frequency calculations, respectively. For large molecules, the energies and gradients employ the linear scaling methods discussed below for evaluation of all matrix elements. Matrix operations are performed conventionally via dense linear algebra routines that exhibit cubic scaling for the energy evaluation. This is asymptotically the rate-determining step for sufficiently large molecules, and effectively sets the upper limit of feasible SCF calculations in the range of several thousand basis functions, subject to the availability of sufficient memory to manipulate the matrices. In practice, the upper limit on calculations may be lower if geometry optimization proves problematical, or if SCF convergence difficulties are encountered. The linear scaling methods for Fock matrix assembly become faster than their conventional counterparts for systems of between roughly 15 and 50 first-row atoms, depending upon the geometry of the molecule, and the accuracy desired. Three-dimensional structures exhibit later crossovers because of the larger number of average neighbors. Small-gap systems lead to later crossovers for the linear scaling evaluation of exact exchange because the electronic structure is more delocalized.

Self-consistent field calculations are accelerated by the standard method of direct inversion in the iterative subspace (DIIS),[34] and good-quality initial guesses are available by superposing spherically averaged atomic densities in the target basis set. While this combination works well in most standard cases, it is necessary to have fallback strategies available for cases where convergence failures are encountered. One difficulty that is occasionally encountered is the problem of an SCF that occupies two different sets of orbitals on alternating iterations, and therefore, oscillates and fails to converge. This can be overcome by choosing orbital occupancies that maximize the overlap of the new occupied orbitals with the set previously occupied. However, this combinatorial matching problem has computational complexity that scales factorially with the number of occupied orbitals if implemented straightforwardly. Q-Chem contains the maximum overlap method (MOM),[35] which, remarkably, reduces the combinatorial problem to cubic in the number of orbitals.

Q-Chem also has other strategies available to coax recalcitrant SCF calculations to convergence. As an alternative to the conventional self-consistent field procedure, Q-Chem includes direct minimization methods that follow energy gradients to minimize the SCF energy. Additionally, if convergence

can be achieved in a smaller basis but not in a larger basis, a novel procedure for evaluating the Fock operator in a large basis set using a density matrix obtained in a small basis set is available as an alternative initial guess.[36]

Analytical frequency calculations are performed by conventional direct methods by default,[37] which require no disk storage of two-electron integrals. Still, relative to evaluating the SCF energy and gradient, the computational cost of analytical frequency calculations is high. The way in which the cost scales with system size is roughly one power of system size steeper than conventional SCF energies/gradients. Disk space and memory requirements both scale with the cube of system size, and thus the largest calculations possible are a strong function of the system resources available.

Q-Chem also contains an interesting unconventional SCF method, in which the molecular orbitals and the density matrix are not expanded directly in terms of the basis of atomic orbitals. Instead, an intermediate molecule-optimized minimal basis of polarized atomic orbitals (PAOs) is used, ref. 38. The polarized atomic orbitals are defined by an atom-blocked linear transformation from the fixed atomic orbital basis, where the coefficients of the transformation are optimized to minimize the energy, at the same time as the density matrix is obtained in the PAO representation. Thus, a PAO-SCF calculation is a constrained variational method, whose energy is above that of a full SCF calculation in the same basis. However, a molecule-optimized minimal basis is a very compact and useful representation for purposes of chemical analysis, and it also has potential computational advantages in the context of local MP2 calculations, as can be done after a PAO-HF calculation is complete to obtain the PAO-MP2 energy.

PAO-SCF calculations tend to systematically underestimate binding energies (because by definition the exact result is obtained for atoms, but not for molecules). In tests on the G2 database, PAO-B3LYP/6-311+G(2df,p) atomization energies deviated from full B3LYP/6-311+G(2df,p) atomization energies by roughly 20 kcal/mol, with the error being essentially extensive with the number of bonds.[39] This deviation can be reduced to only 0.5 kcal/mol[39] with the use of a simple noniterative second-order correction for "beyond-minimal basis" effects.[40] The second-order correction is evaluated at the end of each PAO-SCF calculation, as it involves negligible computational cost. Analytical gradients are available using PAOs, to permit structure optimization. For additional discussion of

the PAO-SCF method and its uses, see the references cited above.

LINEAR SCALING METHODS

Construction of the effective Hamiltonian, or Fock matrix, has traditionally been the rate-determining step in electronic structure programs, due primarily to the cost of two-electron integral evaluation, even with the efficient methods described earlier. However, for large enough molecules, significant speedups are possible by employing recently developed linear-scaling methods for each of the nonlinear terms that can arise. Linear scaling means that if the molecule size is doubled, then the computational effort likewise only doubles. There are three computationally significant terms: electron–electron Coulomb interactions, exact exchange interactions, and exchange–correlation functional evaluation, which we discuss in turn below.

Electron–Electron Coulomb Interactions

The problem of electron–electron Coulomb interactions involves a linear number of charge distributions interacting with themselves, which conventionally requires quadratic effort via two-electron integral evaluation. However, if a given pair of charge distributions do not overlap, their interaction can be evaluated via a multipole expansion without explicitly doing the two-electron integral. Charges that are nearby can be combined together into collective multipoles to make the process of doing these long-distance interactions still more efficient. In fact, by collectivizing charge distributions into multipole and Taylor expansions based on dividing space into a binary tree structure, it is possible to evaluate all of these nonoverlapping interactions in only linear scaling time, with no loss of precision! This method for treating long-range Coulomb interactions (where bra and ket functions do not overlap) is called the Continuous Fast Multipole Method (CFMM).[41] So-called J matrix engines treat short-range terms (in the regime where the bra and ket basis functions are overlapping)[42] by directly computing elements of the J matrix. The most recent J matrix engine[43] is approximately 10 times faster than explicit integral evaluation at the level of a quartet of uncontracted d shells, yet no approximation is required. Speedups increase as angular momentum increases. Additionally, the CFMM and the J engine are also available for the Coulomb force.[44]

Exact Exchange Interactions

Hartree–Fock calculations and the popular hybrid density functionals such as B3LYP also require two-electron integrals to evaluate the exchange energy associated with a single determinant. There is no useful multipole expansion for the exchange energy, because the bra and ket of the two-electron integral are coupled by the density matrix, which carries the effect of exchange. Fortunately, density matrix elements decay exponentially with distance for systems that have a HOMO-LUMO gap.[45] The better the insulator, the more localized the electronic structure, and the faster the rate of exponential decay. Therefore, for insulators, there are only a linear number of numerically significant contributions to the exchange energy. With intelligent numerical thresholding, it is possible to rigorously evaluate the exchange matrix in linear scaling effort. For this purpose, Q-Chem contains the linear scaling K (LinK) method[46] to evaluate both exchange energies and their gradients[47] in linear scaling effort (provided the density matrix is highly sparse). The LinK method essentially reduces to the conventional direct SCF method for exchange in the small molecule limit (by adding no significant overhead), while yielding large speedups for (very) large systems where the density matrix is indeed highly sparse.

Exchange–Correlation Functional Evaluation

Density functional methods require the evaluation of exchange and correlation potentials that are spatially local. In Q-Chem, this is accomplished numerically by grid-based integration of the type pioneered by Becke;[48] these methods are intrinsically linear scaling also, and are relatively standard today.

Work is also underway to further extend the linear scaling capabilities of Q-Chem. Because, as mentioned above, there are only a linear number of numerically significant density matrix elements for sufficiently large molecules, it is possible to solve for them in linear scaling time, given the Fock matrix. This is in contrast to the cubic scaling diagonalization effort that conventional SCF requires. However, the molecule size necessary to see a crossover between cubic scaling with a small coefficient, and linear scaling with a large coefficient turns out to be quite large in practice. There is also much potential for improved methods for analytical frequency calculations, because the advantage of solving for the localized disturbance in the density matrix (instead

of a full matrix) due to each atomic displacement grows faster. New methods that solve directly for the change in the density matrix for each geometric displacement[49] are being incorporated into Q-Chem.

Wave Function-Based Treatments of Electron Correlation

While density functional methods yield a description of electronic structure that accounts for electron correlation subject only to the limitations of present-day functionals (which for example omit dispersion interactions), DFT cannot be systematically improved if the results are deficient. Wave function-based approaches for describing electron correlation[50] offer this main advantage. There are four broad classes of models for describing electron correlation that are supported within Q-Chem. The first three directly approximate the full time-independent Schrödinger equation. In order of increasing accuracy, and also increasing cost, they are: (a) perturbative treatment of pair correlations between electrons, capable of recovering typically 80% or so of the correlation energy in stable molecules. (b) Self-consistent treatment of pair correlations between electrons, capable of recovering on the order of 95% or so of the correlation energy. (c) Noniterative corrections for higher than double substitutions, which can typically account for more than 99% of the correlation energy. They are the basis of many modern methods that are capable of yielding chemical accuracy for ground state reaction energies, as exemplified by the G2[51] and G3 methods.[52] These methods are discussed in the following three subsections.

There is a fourth class of methods supported in Q-Chem, which have a different objective. This is to obtain a balanced description of electron correlation in highly correlated systems, such as biradicals, or along bond-breaking coordinates, and is discussed in the fourth part of this section.

MP2 AND LOCAL MP2 METHODS

The second-order Møller–Plesset theory (MP2)[53] is the simplest useful wave function-based electron correlation method. Revived in the mid-1970s, it remains highly popular today, because it offers systematic improvement in optimized geometries and other molecular properties relative to the Hartree–Fock (HF) theory.[54] Indeed, in a recent comparative study of small closed shell molecules,[55] MP2 outperformed much more expensive singles and doubles

coupled-cluster theory for such properties! Relative to state-of-the-art Kohn–Sham density functional theory (DFT) methods,[29] which are the most economical methods to account for electron correlation effects, MP2 has the advantage of properly incorporating long-range dispersion forces. The principal weaknesses of MP2 theory are for open shell systems, and other cases where the HF determinant is a poor starting point.

Q-Chem contains an efficient conventional semi-direct method to evaluate the MP2 energy and gradient.[56] These methods require OVN memory (O, V, N are the numbers of occupied, virtual, and total orbitals, respectively), and disk space, which is bounded from above by $OVN^2/2$. The latter can be reduced to $IVN^2/2$ by treating the occupied orbitals in batches of size I, and reevaluating the two-electron integrals O/I times. This approach is tractable on modern workstations for energy and gradient calculations of at least 500 basis functions or so, or molecules of between 15 and 30 first row atoms, depending on the basis set size. The computational cost increases between the third and fifth power of the size of the molecule, depending on which part of the calculation is time dominant.

To permit MP2 energy calculations on larger systems, Q-Chem contains newly developed local correlation methods.[57, 58] These methods involve physically motivated truncations of the full MP2 energy expression that substantially reduce the cost of very large MP2 calculations and also reduce the rate of increase of the calculations with molecular size. The answers are also slightly different, as these methods define their own theoretical model chemistries that deviate slightly from conventional MP2 theory. Below we summarize the physical content of these models, and their computational costs, and give some general guidelines to their accuracy.

MP2 theory includes the energy contributions from pair correlations between electrons, in the form of double substitutions of occupied orbitals by empty (virtual) orbitals. We describe the double substitutions in terms of a basis of nonorthogonal localized atom-centered functions to permit the development of local models. The smallest atom-centered set of functions that can span the occupied space is a minimal basis set on each atom, for which we use either polarized atomic orbitals (see earlier), or, by default, extracted polarized atomic orbitals (EPAOs),[59] projected into the occupied space. Note that the EPAOs do not alter the usual SCF energy, while as discussed above the PAOs do. The virtual space is spanned by the full set of atomic orbitals, projected into the virtual space. In terms of these
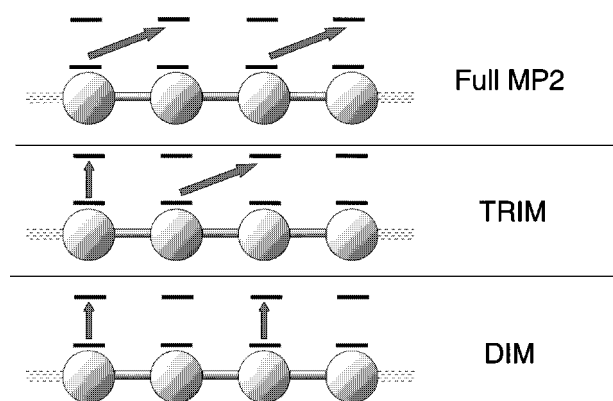
FIGURE 1. Full (untruncated) MP2 allows for double substitutions connecting up to four atoms together, because in terms of atom-centered functions, occupied orbitals on two given atoms will generally be promoted to virtual orbitals on two different atoms. Hence, the number of double substitutions rises with the fourth power of molecular size. Q-Chem's local MP2 methods restrict the double substitutions based on spatial proximity. In the triatomics in molecules (TRIM) model, one occupied to virtual substitution is required to be on a single atom, while the other can be nonlocal. Thus, the number of substitutions retained grows with the third power of molecular size. The diatomics in molecules (DIM) truncation requires both orbital substitutions to be restricted to a given atom, so that the number of retained double substitutions grows with the second power of molecular size.

atom-centered functions, a double substitution will generally couple together four different atoms, as shown in Figure 1.

The two local models arise by making atomic truncations of the full set of double substitutions as described below, and as shown in Figure 1. Note that by the nature of an atomic truncation, these local models yield smooth, globally defined potential energy surfaces. (1) Triatomics in molecules (TRIM): The gentlest atomic truncation is to force one substitution to be restricted to a single atom, while no restriction is placed on the other. This means that only one electron can be transferred with a double substitution. The number of amplitudes is reduced to cubic, as only three separate atoms can be coupled by TRIM double substitution. The TRIM models recovers around 99.7% of the MP2 correlation energy for covalent bonding.[58] The performace for relative energies is very robust, as shown in ref. 58 for the challenging case of torsional barriers in conjugated molecules. (2) Diatomics-in-molecules (DIM): a stronger truncation is to force each substitution to be restricted to a single atom, so that only pairs of atoms are coupled by the double

substitution. The long-range dispersion is still correctly recovered, but nonlocal charge transfer is no longer allowed. We find that typically 95% of the correlation energy associated with covalent bonding is recovered, but the performance for hydrogen bonding is much poorer.

Finally, for comparison, TRIM comfortably exceeds the accuracy of the widely used Pulay–Saebo[60] local MP2 method (roughly 98–99% correlation energy recovery), while DIM is below it. Note, however, that the Pulay–Saebo method as conventionally defined does not yield strictly continuous potential energy surfaces.

The computational advantage associated with the local MP2 methods varies, depending upon the size of molecule and the basis set. As a very rough general estimate, TRIM-MP2 calculations are feasible on molecule sizes about twice as large as those for which conventional MP2 calculations are feasible on a given computer, and this is their primary advantage. Q-Chem's implementation is well suited for large basis set calculations, because the memory requirement for the integral transformation does not exceed OON, and is thresholded so that it asymptotically grows linearly with molecule size. Additional memory of approximately $32N^2$ is required to complete the local MP2 energy evaluation. The disk space requirement is only about $8OVN$, but is not thresholded. DIM-MP2 calculations are faster than TRIM-MP2, and do not require disk storage, but have similar memory requirements.

COUPLED-CLUSTER SINGLES AND DOUBLES METHODS

The standard approach for treating pair correlations self-consistently are coupled-cluster methods where the cluster operator contains all single and double substitutions,[61] abbreviated as CCSD. Quadratic configuration interaction with singles and doubles (QCISD)[62] is a widely used alternative, which is probably best viewed as an approximation to CCSD. These methods yield results that are only slightly superior to MP2 for structures and frequencies of stable closed-shell molecules. However, they are far superior for reactive species, such as transition structures and radicals, for which the performance of MP2 is quite erratic. Q-Chem supports energy evaluation for both CCSD and QCISD. In addition, the MP3 and MP4 energies are also available.

There is an alternative to CCSD and QCISD available in Q-Chem that has some additional advantages. This is the optimized orbital CCSD method (OO-CCD), which we normally refer to as sim-

ply optimized doubles (OD).[63] In the OD method, there are no single substitutions in the cluster operator. Instead, the orbitals that define the mean field reference are optimized to minimize the total energy (defined in the usual nonvariational cluster fashion). This can be viewed as an alternative definition of approximate Brueckner orbitals. The OD method has the advantage of formal simplicity (orbital variations and single substitutions are redundant variables). In cases where Hartree–Fock theory performs poorly (e.g., artifactual symmetry breaking), it is also practically advantageous to use the OD method, where the HF orbitals are not required, rather than CCSD or QCISD. Q-Chem supports both energies and analytical gradients using the OD method. The computational cost for the OD energy is roughly twice that of the CCSD or QCISD method, but the total cost of energy plus gradient is similar.

The implementation of the QCISD, CCSD, and OD methods in Q-Chem requires substantial disk space, which grows as the number of atomic orbitals to the fourth power. The computational effort increases with the sixth power of molecule size, or the fourth power of basis set size, for fixed molecule size. These are the conventional scalings associated with these methods, and result from the use of delocalized occupied and virtual orbitals. In other words, the local methods described for MP2 theory in the previous section have not yet been implemented for these higher correlation methods. Finally, we note that the internal details of our implementation make our code readily extensible to new correlated techniques relative to traditional designs. The code is fully object oriented, and the top-level calls closely resemble spin-orbital algebra.

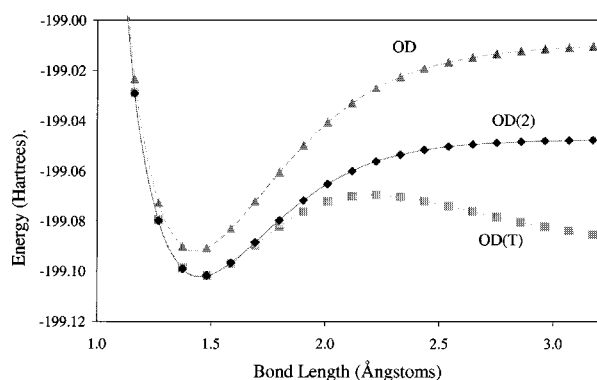### NONITERATIVE CORRECTIONS DUE TO HIGHER CORRELATION EFFECTS

To approach chemical accuracy in reaction energies and related properties, it is necessary to account for electron correlation effects that involve three electrons simultaneously, as represented by triple substitutions relative to the mean field single determinant reference, which arise in MP4. The best standart methods for including triple substitutions are the CCSD(T)[64] and QCISD(T) methods.[62] The accuracy of these methods is well documented for many cases,[65] and, in general, is a very significant improvement relative to the starting point (either CCSD or QCISD). The cost of these corrections scales with the seventh power of molecule size (or the fourth power of the number of basis functions for fixed molecule size), although no additional disk resourses are required relative to the starting coupled cluster calculation. Q-Chem supports the evaluation of CCSD(T) and QCISD(T) energies, as well as the corresponding OD(T) correction to the optimized doubles method discussed in the previous subsection. As discussed in the next section, a local triples method is also being developed.

While the (T) corrections have been extraordinarily successful, there is nonetheless still room for improvement. They contain judiciously chosen terms from fourth- and fifth-order Møller–Plesset perturbation theory, as well as higher order terms that result from the fact that the converged cluster amplitudes are employed to evaluate the fourth- and fifth-order terms. The correction, therefore, depends upon the bare reference orbitals and orbital energies, and in this way its effectiveness still depends on the quality of the reference determinant. Because we are correcting a coupled cluster solution rather that a single determinant, this is an aspect of the (T) corrections that can be improved.

Such an improvement has recently been reported,[66] and Q-Chem contains this new method. The new correction is a true second-order correction to a coupled cluster starting point, and is, therefore, denoted as (2). It is available for the three cluster methods discussed above, as OD(2), CCSD(2), and QCISD(2) energies.[66, 67] The basis of the (2) method is to partition not the regular Hamiltonian into perturbed and unperturbed parts, but rather to partition a similarity-transformed Hamiltonian, defined as $\hat{\bar{H}} = e^{-\hat{T}} \hat{H} e^{\hat{T}}$. In the truncated space (call it the p-space) within which the cluster problem is solved (e.g., singles and doubles for CCSD), the coupled cluster wavefunction is a true eigenvalue of $\hat{\bar{H}}$. Therefore, we take the zero-order Hamiltonian, $\hat{\bar{H}}^{(0)}$, to be the full $\hat{\bar{H}}$ in the p-space, while in the space of excluded substitutions (the q-space) we take only the one-body part of $\hat{\bar{H}}$ (which can be made diagonal). The fluctuation potential describing electron correlations in the q-space is $\hat{\bar{H}} - \hat{\bar{H}}^{(0)}$, and the (2) correction then follows from second-order perturbation theory.

The new partitioning of terms between the perturbed and unperturbed Hamiltonians inherent in the (2) correction leads to a correction that shows both similarities and differences relative to the existing (T) corrections. There are two types of higher correlations that enter at second order: not only triple substitutions, but also quadruple substitutions. The quadruples are treated with a factoriza-

FIGURE 2. Comparison of calculated potential curves for breaking the F_2 bond using restricted (spin-pure orbitals) using the optimized orbital coupled cluster doubles (OD) method, and two noniterative corrections for higher than double substitutions. The basis set is the cc-pVDZ basis. OD itself is qualitatively correct, but makes the well depth far too deep because it cannot recover a uniform fraction of the correlation energy as the bond is stretched. The first noniterative correction is the standard triples correction, (T), which exhibits an unphysical hump, and clearly is diverging as the bond length is stretched. Second is the new (2) correction, which corrects the well depth, and is stable even as the bonding and antibonding σ orbitals approach degeneracy.

tion ansatz, that is exact in fifth-order Møller–Plesset theory,[68] to reduce their computational cost from N^9 to N^6. For large basis sets this can still be larger than the cost of the triples terms, which scale as the seventh power of molecule size, with a factor twice as large as the usual (T) corrections. These corrections are feasible for molecules containing between 4 and 10 first-row atoms, depending on computer resources, and the size of the basis set chosen. There is early evidence that the (2) corrections are superior to the (T) corrections for highly correlated systems.[66] This shows up in improved potential curves, particularly at long range, as illustrated in Figure 2, and may also extend to improved energetic and structural properties at equilibrium in problematical cases. It will be some time before sufficient testing on the new (2) corrections has been done to permit a general assessment of the performance of these methods, but they are clearly very promising.

LOCAL TRIPLE EXCITATION METHODS

Work is underway to incorporate local correlation methods for triple substitutions,[69, 70] for closed-shell molecules only. Triple substitutions are essential for predicting chemical reaction energies

to 1 kcal/mol or better, and are employed in the widely used G2 and G3 thermochemical methods, where the rate-determining step is usually the triples contribution to the fourth-order Møller–Plesset energy: MP4(T). The cost of methods including triple substitutions, such as the widely used MP4(T) and CCSD(T) methods, scales as the seventh power of molecule size, and this limits their application to roughly 10 first row atoms. The local triples algorithm reduces the cost from seventh order to fifth order, and thereby enables calculations on molecules two to three times larger than previously feasible. It crosses over with the conventional triples algorithm around 25 occupied orbitals, or roughly 150 basis functions. The local algorithm requires all the doubles amplitudes to be stored on disk, and also requires disk storage for a cubic number of local integrals. A relatively small (quadratic) amount of memory is required.

The basis of the local triples method is a generalization of the TRIM method discussed above for MP2 theory. Three electrons are simultaneously promoted in a triple substitution, and in the local model, two of those promotions are restricted to a single atom, while the third is permitted to be nonlocal. This reduces the number of triple substitutions from rising with the sixth power of molecule size to rising with the fourth power. Benchmark thermochemical calculations on 105 molecules in the G2/97 database[52] indicate that local truncation recovers at least 95% of the untruncated triples energy. The local error introduced into the G3 binding energies is typically 0.1–0.2 kcal/mol, with a maximum error of 0.26 kcal/mol. This error is small enough for most applications. If higher accuracy is required, the error can be reduced by employing a less severe local truncation.

VALENCE ACTIVE SPACE METHODS

Electron correlation effects can be qualitatively divided into two classes. The first class is static or nondynamical correlation: long wavelength low-energy correlations associated with other electron configurations that are nearly as low in energy as the lowest energy configuration. These correlation effects are important for problems such as bond breaking, and are the hardest to describe because by definition the single configuration Hartree–Fock description is not a good starting point. The second class is dynamical correlation: short wavelength high-energy correlations associated with atomic-like effects. Dynamical correlation is essential for quantitative accuracy.

In the methods discussed in the previous several subsections, the objective was to approximate the total correlation energy. However, in some cases, it is useful to instead directly model the nondynamical and dynamical correlation energies separately. The reasons for this are pragmatic: with approximate methods, such a separation can give a better balanced treatment of electron correlation along bond-breaking coordinates, or reaction coordinates that involve biradicaloid intermediates. The nondynamical correlation energy is conveniently defined as the solution of the Schrödinger equation within a small basis set composed of valence bonding, antibonding, and lone-pair orbitals: the so-called full valence active space. Solved exactly, this is the so-called full valence complete active space SCF (CASSCF),[71] or equivalently, the fully optimized reaction space (FORS) method.[72]

Full-valence CASSCF and FORS involve computational complexity, which increases exponentially with the number of atoms, and is thus unfeasible beyond systems of only a few atoms, unless the active space is further restricted on a case-by-case basis. Q-Chem includes an economical method that directly approximates these theories using a truncated coupled cluster doubles wave function with optimized orbitals.[73] This represents a generalization of the optimized doubles method (OD) discussed above such that double excitations are only permitted within the valence active space, and the orbitals describing this active space are simultaneously optimized. This method may be fully specified as valence optimized orbital coupled cluster doubles (VOO-CCD), or, more concisely as valence optimized doubles (VOD). VOD calculations have computational complexity that scales similarly to OD, although with a smaller prefactor, and with far lower disk storage requirements. It is thus an economical and size-consistent approximation to full valence CASSCF, which can be applied to medium-sized organic molecules with a full valence active space.

The performance of VOD for bond-breaking processes may be briefly summarized as follows.[73] For processes that are essentially single bond breaking in nature, VOD appears to give essentially quantitative accuracy in comparison with full valence CASSCF. Caution is required in the description of multiple bond breaking, because if restricted orbitals are enforced at long bond lengths, VOD can exhibit nonvariational collapse to energies well below the true values (like any conventional coupled cluster method). For a discussion of these limitations, see refs. 50 and 74, and we note that efforts are currently under way to develop modified coupled cluster methods that perform more satisfactorily for multiple bond breaking.[74, 75] Both energies and analytical gradients are available for VOD in Q-Chem. We note that convergence of VOD calculations is generally poorer than that of OD calculations, because of the difficulty of optimizing the active space orbitals.

VOD energies can be perturbatively corrected for dynamical correlation effects by an appropriate generalization of the OD(2) model discussed in the previous section. Q-Chem contains an implementation of the corresponding VOD(2) method,[76] for energies only. VOD(2) includes the dynamical correlation effects associated with the single and double substitutions omitted in VOD, as well as a subset of triple and quadruple substitutions. The latter allow a consistent treatment of correlation effects from *all* determinants in the VOD reference, as is required when the HF determinant is a poor starting point. Use of the VOD(2) correction yields relative energies that are greatly improved relative to VOD.[76] In summary, VOD(2) is a true second-order perturbation correction to the VOD reference for dynamical correlation, analogous to the manner in which MP2 is a second-order correction to the Hartree–Fock reference for all correlation effects. However, it is of much higher quality than MP2 because the perturbation is much smaller, because VOD includes the leading nondynamical correlations.

Excited State Methods

The development of effective approaches to modeling electronic excited states has historically lagged behind advances in treating the ground state. In part, this is because of the much greater diversity in the character of the wave functions for excited states, making it more difficult to develop broadly applicable methods without molecule-specific or even state-specific specification in the form of the wave function. Broadly speaking, Q-Chem contains methods that are capable of giving qualitative agreement, and in many cases quantitative agreement with experiment for lower optically allowed states. The situation is less satisfactory for states that involve bi-electronic excitations, although even here reasonable results can sometimes be obtained. The discussion of specific methods below is divided into two sections. The first part describes excited state versions of single configuration treatments of the ground state via Hartree–Fock theory, and in particular, Kohn–Sham density functional theory, which

are applicable to quite large molecules. The second part discusses more accurate, and also more expensive excited state methods based upon coupled cluster theory, which are only feasible for smaller molecules.

EXCITED STATE METHODS BEGINNING FROM SELF-CONSISTENT FIELD GROUND STATES

Here we consider the ground state to be given as a single determinant: either the Hartree–Fock determinant for wave function-based methods, or the Kohn–Sham determinant, which parameterizes the ground-state electron density in the density functional theory (DFT). The single determinant is optimized by minimizing the total energy, which is equivalent via the Brillouin theorem to obtaining a set of occupied orbitals whose matrix elements with all single substitutions are zero. Therefore, single substitutions are Hamiltonian noninteracting with the ground state, and form a natural basis for describing excited states at a level of theory roughly similar to the ground state.

Beginning from the Hartree–Fock theory, this leads to excited states that are a mixture of all such single substitutions (the CIS method).[77] Q-Chem contains efficient direct implementations of CIS not only for closed- and open-shell energies, but also for analytical first and second derivatives.[78] For closed-shell molecules, CIS yields qualitatively correct descriptions of one-electron excited states, although due to neglect of electron correlation, the excitation energies cannot generally be expected to be reliable to within more than roughly 2 eV. Optimized geometries and vibrational frequencies are of reasonably good quality,[79] roughly similar to HF theory for ground states. Conventional CIS is qualitatively unsatisfactory for radicals, but the extended CIS (XCIS) method is available for doublet and quartet excitation energies,[80] and with comparable performance to CIS for closed-shell molecules. To correct CIS excitation energies for the leading effects of electron correlation, Q-Chem also includes a perturbative doubles correction CIS(D).[81] CIS(D) is roughly an excited state analog of MP2 theory, and typically reduces errors in CIS excitation energies by a factor of two or more, while the computational cost per state is roughly similar to an MP2 calculation.

Excited states may be obtained from density functional theory by time-dependent density functional theory,[82, 83] which calculates poles in the response of the ground-state density to a time-varying applied electric field. These poles are Bohr frequen-cies or excitation energies, and are available in Q-Chem,[84] together with the CIS-like Tamm–Dancoff approximation.[85] TDDFT is becoming very popular as a method for studying excited states because the computational cost is roughly similar to the simple CIS method (scaling as roughly the square of molecular size), but a description of differential electron correlation effects is implicit in the method. The excitation energies for low-lying valence excited states of molecules (below the ionization threshold) are often remarkably improved relative to CIS, with an accuracy of roughly 0.3 eV being observed with either gradient corrected or local density functionals.

However, standard density functionals do not yield a potential with the correct long-range Coulomb tail (due to the so-called self-interaction problem), and therefore, excited states that sample this tail (e.g., diffuse Rydberg states, and some charge transfer excited states) are not given accurately.[86, 87] Hence it is advisable to only employ TDDFT for low-lying valence excited states that are below the first ionization potential of the molecule. This makes radical cations a particularly favorable choice of system, as exploited in ref. 88. TDDFT for low-lying valence excited states of radicals is in general a remarkable improvement relative to CIS, including some states, that, when treated by wave function-based methods can involve a significant fraction of double excitation character.[84]

EXCITED STATE METHODS BASED ON COUPLED-CLUSTER GROUND STATES

It is possible to obtain a description of electronic excited states at a level of theory similar to that associated with the coupled-cluster theory for the ground state, by applying either the linear response theory[89] for equations of motion methods.[90] A number of groups have demonstrated that excitation energies based on a coupled-cluster singles and doubles ground state are generally very accurate for states that are primarily single electron promotions. The errors observed in calculated excitation energies to such states is approximately 0.3 eV, including both valence and Rydberg excited states. This, of course, assumes that a basis set large and flexible enough to describe valence and Rydberg states is employed. The accuracy of the excited state coupled-cluster methods is much lower for excited states that involve a substantial component of double excitation character, where errors may be 1 eV or even more. Such errors arise because the description of electron correlation is better in the ground state

than for an excited state with substantial double excitation character.

Q-Chem includes coupled-cluster methods for excited states based on the optimized orbital coupled-cluster doubles (OD) method, described earlier. OD excitation energies are essentially identical in numerical performance to CCSD excited states, as has been recently demonstrated.[91] This method, while far more computationally expensive than TDDFT, is nevertheless useful as a proven high accuracy method for the study of excited states of small molecules. Also, when studying a series of related molecules it can be very useful to compare the performance of TDDFT and coupled-cluster theory for at least a small example to understand its performance. Along similar lines, the CIS(D) method described earlier as an economical correlation energy correction to CIS excitation energies is, in fact, an approximation to coupled-cluster excitation energies. It is useful to assess the performance of CIS(D) for a class of problems by benchmarking against the full coupled-cluster treatment. Finally, Q-Chem also includes excited states by the equation of motion version of the valence optimized doubles (VOD) method (see above), whose validity and use is fully discussed in ref. 91.

Wave Function Analysis

Wave functions are not easy to interpret, and this has continued to fuel an interest in the extraction of simpler quantities from them. The most basic tools that have been developed are the familiar Mulliken and Lowdin population analyses,[26] but these attempt to reduce electronic structure to a set of numbers, and this is often not entirely satisfactory. A more sophisticated approach to population analysis the widely used Natural Bond Order (NBO) approach, developed by Weinhold and coworkers.[92, 93] Q-Chem contains a basic interface to this package,[94] although not all of its functionality is currently supported. Alternatively, instead of numbers, one may seek simple *functions* that capture the key properties of the wave function. Accordingly, it has now become commonplace for researchers to study systems of interest by plotting the electron density, key molecular orbitals, electrostatic potential, electric field, or other such functions. Q-Chem is well equipped for such calculations.

Q-Chem can also compute a number of less familiar functions and four classes of these are described in the following paragraphs:

STEWART ATOMS

Stewart atoms are the unique nuclear-centered spherical functions whose sum best fits a molecular density in a least-squares sense.[95] Their usefulness lies in the fact that they recover atomic identity from a molecular density, and yield a much simplified (although approximate) description of the electronic density. This second fact has ramifications for the rapid calculation of Coulomb energies, and for the evaluation of integrals arising in DFT without recourse to numerical quadrature grids. Although several methods have been developed for calculating Stewart atoms,[96] the most conceptually straightforward approach is based on resolution of the identity (RI), which expands the Stewart atoms in a radial basis located on each center.[97] The expansion coefficients are determined via a least-squares fitting procedure, and this has been fully implemented within the Q-Chem package. Work is also being carried out on an integral equation formulation of the Stewart theory in which the Stewart atoms are obtained by convolving the molecular density with solution kernels.[98] This method avoids the need for an auxiliary expansion basis, and thus circumvents the basis set convergence problems that affect the RI expansion.

MOMENTUM DENSITIES

The electron density of a molecule is a familiar quantity, and plots of such densities adorn many modern chemistry textbooks. It is useful because it reveals the positions at which an electron is most likely to be found. Furthermore, because it can be measured by X-ray crystallography, the electron density provides an important bridge between theory and experiment. The less familiar momentum density is also illuminating, because it reveals the momenta that an electron is most likely to possess. Momentum densities also offer a bridge to experiment through Compton scattering measurements. Electron densities and momentum densities thus provide complementary information about a chemical system and, taken together, provide a more detailed picture of electronic structure than emerges from either separately.

INTRACULES

Electron and momentum densities are one-electron functions, which means that they provide information about the probability that one electron will be found in a certain position or with a certain momentum. However, because the interactions

between *pairs* of electrons are so important in chemistry, it is also beneficial to generate and study the analogous two-electron functions. The position intracule, $P(u)$, and momentum intracule, $\hat{P}(v)$, are the probability distributions of the relative position $u = |\mathbf{r}_1 - \mathbf{r}_2|$ and relative momentum $v = |\mathbf{p}_1 - \mathbf{p}_2|$ of two electrons. An intracule derived from a Hartree–Fock wave function can be separated into its Coulomb and exchange components, and these sometimes yield more information than their sum. Q-Chem can separately calculate the $J(u)$, $K(u)$, $\hat{J}(v)$, and $\hat{K}(v)$ intracules.[99–101]

ATTACHMENT AND DETACHMENT DENSITIES

While an electronic transition is a collective rearrangement of all electrons in a molecule it is very desirable to have a one-electron picture of the main changes in charge density. In attachment–detachment density analysis,[102] the difference density matrix between the ground state and the excited state of interest is decomposed into a detachment density that is rearranged upon excitation into an attachment density. The remaining electron density is unaffected. This procedure is often helpful for visualizing electronic transitions and classifying them as valence, or Rydberg, or mixed, and examining whether or not there is charge-transfer character. Attachment–detachment analysis has been successfully employed in a number of chemical applications.[103]

Additional Capabilities

USER INTERFACE

Like traditional electronic structure programs, Q-Chem is fundamentally a backend compute engine. It contains a simple text-based input mechanism, including scripts that permit emulation of common Gaussian-style keywords. On some platforms this release of Q-Chem is bundled as a back end with the Spartan user interface from Wavefunction, Inc., which provides a versatile and sophisticated platform for building molecules, and visualizing the results of the calculations.

SOLVATION MODELING

Q-Chem contains two solvation models.[104] First, the simple spherical Onsager reaction field model, and second, the considerably more sophisticated Langevin dipoles model developed by Warshel and Florian.[105] The Langevin dipoles approach is also

a continuum solvation model, but it more realistically treats effects such as dielectric saturation, and local solvation, by incorporating a layer of dipoles around the Van der Waals surface of the solute, and this dipole layer is, in turn, surrounded by continuum solvent. The individual dipoles orient themselves self-consistently with the solute charge distribution (while having constant magnitude) to determine the short-range part of the reaction field. This approach has been successfully applied in a number of chemical applications.[106]

WALKING ON POTENTIAL ENERGY SURFACES

Q-Chem's optimization capabilities were developed by Dr. Jon Baker, and are based on the eigenvector following method.[107] This is capable of searching for both minima and transition structures, using either gradients, or gradients and hessians as input. It employs redundant internal coordinates[108] to ensure good convergence even when an initial estimate of the force constant matrix is not available. Additionally, a sophisticated treatment of optimization with constraints[109] is available, in which constraints can be specified either as frozen internal coordinates (bond lengths, angles, etc.), or frozen Cartesian coordinates.

RELATIVISTIC ENERGY CORRECTION

A relativistic energy of the order $1/c^2$, where c is the speed of light, is calculated automatically each time a Hartree–Fock frequency calculation is requested. This is an additive correction to the Hartree–Fock energy that approximately accounts for the increase of the electron mass as the electron velocity approaches the speed of light, which occurs near an atomic nucleus. It is the first-order contribution of the stationary direct perturbation theory expansion[110] of the Dirac–Fock energy. The Dirac–Fock method is a many-electron generalization of the Dirac equation,[111] and includes four components for each molecular orbital. Dirac–Fock is accurate for heavy-atom-containing molecules, but it is too costly, and there are numerical problems with convergence. See ref. 112 for a review of relativistic methods including Dirac–Fock. The stationary direct perturbation expansion in powers of $1/c^2$ gives an order-by-order approximation to the Dirac–Fock energy, where the zero-order energy is the nonrelativistic Hartree–Fock energy, and the infinite order energy is the Dirac–Fock energy. In Q-Chem, real UHF orbitals are used as the zero-order orbitals in conjunction with the first-order

energy formula.[113, 114] This results in a scalar relativistic correction (that does not include spin-orbit effects) which is a good approximation to the Dirac–Fock energy minus the spin-orbit contribution for molecules containing the atoms hydrogen through krypton and perhaps heavier atoms.

DIAGONAL ADIABATIC CORRECTION

The commonly used Born–Oppenheimer approximation separates the motion of electrons from that of nuclei. For a total molecular energy, this approximation is correct up to the order of $O(m/M)$, where m/M is the ratio of the electronic to the nuclear masses. The leading contribution to this molecular energy *beyond* the Born–Oppenheimer approximation is described by the so-called diagonal Born–Oppenheimer correction (DBOC),[115] which has been proven to be rigorous even for the molecules with separated center-of-mass motion.[116] Q-Chem is capable of calculating the DBOC for self-consistent field wave functions. Additionally, Q-Chem can evaluate the leading correlation correction to the DBOC, based on the Møller–Plesset perturbation expansion of the correlated wave function.[117]

PARALLEL COMPUTING

One way to extend the applicability of quantum chemistry methods to larger molecules is to take advantage of the growing power and availability of parallel computers. Q-Chem recognizes this need, and has been making systematic effort to parallelize the most computationally demanding steps in the code. Parallelized features in the current version of Q-Chem includes HF and DFT models for single-point and gradient calculations.[118] A dynamic load-balancing scheme is used to achieve optimal parallel efficiency and the Message-Passing Interface (MPI) is used for portability. The speedups obtained in benchmark calculations are 12–14 on 16 nodes for single-point and 40–50 on 64 nodes for energy gradient evaluation (SGI Origin 2000 and Cray T3E timings). An IBM SP2 version is now being completed, and efforts are being made in development of parallel linear-scaling methods.

Conclusions

The new release of the Q-Chem program described in this article represents the second generation of this package. This article has described the key features contained in the program with particular emphasis on theories and algorithms that are new and not available elsewhere. As the result of collaboration between a software company and several academic research groups, such features represent one of the strongest aspects of Q-Chem.

References

1. Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. J Comput Chem 1993, 14, 1347.

2. Kendall, R. A.; Apra, E.; Bernholdt, D. E.; Bylaska, E. J.; Dupuis, M.; Fann, G. I.; Harrison, R. J.; Ju, J.; Nichols, J. A.; Nieplocha, J.; Straatsma, T. P.; Windus, T. L.; Wong, A. T. Comput Phys Commun 2000, 128, 260.

3. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. J Comput Chem 1983, 4, 187.

4. Levine, I. N. Quantum Chemistry; Prentice Hall: Englewood Cliffs, NJ, 2000.

5. Jensen, F. Introduction to Computational Chemistry; Wiley: New York, 1999.

6. Simons, J.; Nichols, J. Quantum Mechanics in Chemistry; Oxford: New York, 1997.

7. Davidson, E. R.; Feller, D. Chem Rev 1986, 86, 681.

8. Helgaker, T.; Taylor, P. R. In Modern Electronic Structure Theory; Yarkony, D. R., Ed.; World Scientific: Singapore, 1995, p. 725.

9. Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. Ab Initio Molecular Orbital Theory; Wiley: New York, 1986.

10. Rassolov, V. A.; Pople, J. A.; Ratner, M. A.; Windus, T. L. J Chem Phys 1998, 109, 1223; Rassolov, V. A.; co-workers, in preparation.

11. Dunning, T. H., Jr. J Chem Phys 1989, 90, 1007. The full set of correlation-consistent basis sets are available for download at the Pacific Northwest Laboratory web site (http://www.emsl.pnl.gov).

12. Dunning, T. H., Jr.; Hay, P. J. In Methods of Electronic Structure Theory; Schaefer, H. F., III, Ed.; Plenum Press: New York, 1977.

13. Schafer, A.; Horn, H.; Ahlrichs, R. J Chem Phys 1992, 97, 2571.

14. Basis sets were obtained from the Extensible Computational Chemistry Environment Basis Set Database, Version, as developed and distributed by the Molecular Science Computing Facility, Environmental and Molecular Sciences Laboratory, which is part of the Pacific Northwest Laboratory, P.O. Box 999, Richland, WA 99352, and funded by the U.S. Department of Energy. The Pacific Northwest Laboratory is a multiprogram laboratory operated by Battelle Memorial Institute for the U.S. Department of Energy under contract DE-AC06-76RLO 1830. Contact David Feller or Karen Schuchardt for further information.

15. Hay, P. J.; Wadt, W. R. J Chem Phys 1985, 82, 270; ibid 284, ibid 299.

16. See, for example, Stevens, W. J.; Basch, H.; Krauss, M. J Chem Phys 1984, 81, 6026; Stevens, W. J.; Krauss, M.; Basch, H.; Jasien, P. G. Can J Chem 1992, 70, 612; Cundari, T. R.; Stevens, W. J. J Chem Phys 1993, 98, 5555.

17. See, for example, Hurley, M. M.; Pacios, L. F.; Christiansen, P. A.; Ross, R. B.; Ermler, W. C. J Chem Phys 1986, 84, 6840; Lajohn, L. A.; Christiansen, P. A.; Ross, R. B.; Atashroo, T.; Ermler, W. C. J Chem Phys 1987, 87, 2812.

18. See, for example, Bergner, A.; Dolg, M.; Kuechle, W.; Stoll, H.; Preuss, H. Mol Phys 1993, 80, 1431.

19. See, for example, Dolg, M.; Stoll, H.; Preuss, H.; Pitzer, R. M. J Phys Chem 1993, 97, 5852.

20. Adams, T. R.; Adamson, R. D.; Gill, P. M. W. J Chem Phys 1997, 107, 124.

21. Gill, P. M. W.; Gilbert, A. T. B; Adams, T. R. J Comput Chem.

22. Gill, P. M. W. Adv Quantum Chem 1994, 25, 141.

23. Adamson, R. D.; Gill, P. M. W. J Chem Phys, in preparation.

24. Kahn, L. R.; Goddard, W. A., III J Chem Phys 1972, 56, 2685.

25. See, for example, Skylaris, C.-K.; Cagliardi, L.; Handy, N. C.; Ioannou, A. G.; Spencer, S.; Willets, A.; Simper, A. M. Chem Phys Lett 1998, 296, 445.

26. For a general introduction to Hartree–Fock theory, see Szabo, A.; Ostlund, N. S. Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory; Dover: New York, 1996.

27. A comprehensive discussion of the performance of Hartree–Fock theory, and also second-order Møller–Plesset theory for chemical applications is available in ref. 9.

28. Parr, R. G.; Yang, W. Density Functional Theory of Atoms and Molecules; Oxford: New York, 1989.

29. Kohn, W.; Becke, A. D.; Parr, R. G. J Phys Chem 1996, 100, 12974.

30. Becke, A. D. J Chem Phys 1993, 98, 5648.

31. See, for example, Ziegler, T. Chem Rev 1991, 91, 651; Chemical Applications of Density Functional Methods; Laird, B. B.; Ross, R. B.; Ziegler, T., Eds.; American Chemical Society: Washington, DC, 1996.

32. Adamson, R. D.; Gill, P. M. W.; Pople, J. A. Chem Phys Lett 1998, 284, 6.

33. Gilbert, A. T. B.; Gill, P. M. W. Chem Phys Lett 1999, 312, 511.

34. Pulay, P. J Comput Chem 1982, 3, 556.

35. Gilbert, A. T. B.; Gill, G. B. W.; Gill, P. M. W., in preparation.

36. Kong, J.; Johnson, B. G., unpublished.

37. Frisch, M. J.; Head–Gordon, M.; Pople, J. A. Chem Phys 1990, 141, 189.

38. Lee, M. S.; Head–Gordon, M. J Chem Phys 1997, 107, 9085.

39. Lee, M. S.; Head–Gordon, M., to appear.

40. Lee, M. S.; Head–Gordon, M. Comp Chem 2000, 24, 295.

41. White, C. A.; Johnson, B. G.; Gill, P. M. W.; Head–Gordon, M. Chem Phys Lett 1994, 230, 8; 1996, 253, 268.

42. White, C. A.; Head–Gordon, M. J Chem Phys 1996, 104, 2620.

43. Shao, Y.; Head–Gordon, M. Chem Phys Lett 2000, 323, 425.

44. Shao, Y.; Head–Gordon, M. Chem Phys Lett, in preparation.

45. Schwegler, E.; Challacombe, M.; Head–Gordon, M. J Chem Phys 1997, 106, 9708.

46. Ochsenfeld, C.; White, C. A.; Head–Gordon, M. J Chem Phys 1998, 109, 1663.

47. Ochsenfeld, C. Chem Phys Lett 2000, 327, 216.

48. Becke, A. D. J Chem Phys 1988, 88, 2547.

49. Ochsenfeld, C.; Head–Gordon, M. Chem Phys Lett 1997, 270, 399.

50. For a tutorial introduction to electron correlation methods based on wave functions, see Bartlett, R. J.; Stanton, J. F. In Reviews in Computational Chemistry; Lipkowitz, K. B.; Boyd, D. B., Eds.; VCH: New York, 1994, p. 65, vol. V.

51. Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. J Chem Phys 1991, 94, 722.

52. Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. J Chem Phys 1998, 109, 7764.

53. Møller, C.; Plesset, M. S. Phys Rev 1934, 46, 618.

54. Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. Ab Initio Molecular Orbital Theory; Wiley: New York, 1986.

55. Helgaker, T.; Gauss, J.; Jorgensen, P.; Olsen, J. J Chem Phys 1997, 106, 6430.

56. Head–Gordon, M. Mol Phys 1999, 96, 673.

57. Head–Gordon, M.; Lee, M. S.; Maslen, P. E. In Simulation and Theory of Electrostatic Interactions in Solution; Pratt, L. R.; Hummer, G., Eds.; AIP Conference Proceedings 492; AIP: New York, 1999, p. 301.

58. Lee, M. S.; Maslen, P. E.; Head–Gordon, M. J Chem Phys 2000, 112, 3592.

59. Lee, M. S.; Head–Gordon, M. Int J Quantum Chem 2000, 76, 169.

60. Saebo, S.; Pulay, P. Annu Rev Phys Chem 1993, 44, 213.

61. Purvis, G. D.; Bartlett, R. J. J Chem Phys 1982, 76, 1910.

62. Pople, J. A.; Head–Gordon, M.; Raghavachari, K. J Chem Phys 1987, 87, 5968.

63. Sherrill, C. D.; Krylov, A. I.; Byrd, E. F. C.; Head–Gordon, M. J Chem Phys 1998, 109, 4171.

64. Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head–Gordon, M. Chem Phys Lett 1989, 157, 479.

65. See, for example, Lee, T. J.; Scuseria, G. E. In Quantum Mechanical Calculations with Chemical Accuracy; Langhoff, S. R., Ed.; Kluwer: Dordrecht, 1995, p. 47.

66. Gwaltney, S. R.; Head–Gordon, M. Chem Phys Lett 2000, 323, 21.

67. Gwaltney, S. R.; Head–Gordon, M., in preparation.

68. Kucharski, S. A.; Bartlett, R. J. J Chem Phys 1998, 108, 5243; 1998, 108, 9221.

69. Maslen, P. E.; Lee, M. S.; Head–Gordon, M. Chem Phys Lett 2000, 319, 205.

70. Maslen, P. E.; Lee, M. S.; Head–Gordon, M., in preparation.

71. For an introduction to the CASSCF method, see Roos, B. O. Adv Chem Phys 1987, 69, 399.

72. See, for example, Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M.; Elbert, S. T. Chem Phys 1982, 71, 49; 1982, 71, 51.

73. Krylov, A. I.; Sherrill, C. D.; Byrd, E. F. C.; Head–Gordon, M. J Chem Phys 1998, 109, 10669.

74. Van Voorhis, T.; Head–Gordon, M. J Chem Phys, submitted.

75. Van Voorhis, T.; Head–Gordon, M. Chem Phys Lett, submitted; Van Voorhis, T.; Byrd, E. F. C.; Head–Gordon, M., work in progress.

76. Gwaltney, S. R.; Sherrill, C. D.; Head–Gordon, M.; Krylov, A. I., J Chem Phys 2000, 113, 3548.

77. Foresman, J. B.; Head–Gordon, M.; Pople, J. A.; Frisch, M. J. J Phys Chem 1992, 96, 137.

78. Maurice, D.; Head–Gordon, M. Mol Phys 1999, 96, 1533.

79. Stanton, J. F.; Gauss, J.; Ishikawa, N.; Head–Gordon, M. J Chem Phys 1995, 103, 4160.

80. Maurice, D.; Head–Gordon, M. J Phys Chem 1996, 100, 6131.

81. Head–Gordon, M.; Rico, R. J.; Oumi, M.; Lee, T. J. Chem Phys Lett 1994, 219, 21.

82. Runge, E.; Gross, E. K. U. Phys Rev Lett 1984, 52, 997.

83. Casida, M. E. In Recent Advances in Density Functional Methods, Part I; Chong, D. P., Ed.; World Scientific: Singapore, 1995, p. 155.

84. Hirata, S.; Head–Gordon, M. Chem Phys Lett 1999, 302, 375.

85. Hirata, S.; Head–Gordon, M. Chem Phys Lett 1999, 314, 291.

86. Casida, M. E.; Jamorski, C.; Casida, K. C.; Salahub, D. R. J Chem Phys 1998, 108, 4439.

87. Tozer, D. J.; Handy, N. C. J Chem Phys 1998, 109, 10180.

88. Hirata, S.; Lee, T. J.; Head–Gordon, M. J Chem Phys 1999, 111, 8904.

89. See, for example, Koch, H.; Jorgensen, P. J Chem Phys 1999, 93, 3333.

90. Stanton, J. F.; Bartlett, R. J. J Chem Phys 1993, 98, 7029.

91. Krylov, A. I.; Sherrill, C. D.; Head–Gordon, M. J Chem Phys, to appear.

92. Reed, A. E.; Weinstock, R. B.; Weinhold, F. J Chem Phys 1985, 83, 735.

93. Reed, A. E.; Curtiss, L. A.; Weinhold, F. Chem Rev 1988, 88, 899.

94. Weinhold, F. NBO 4.0 Program Manual, University of Wisconsin Theoretical Chemistry Institute, Madison, 1994.

95. Gill, P. M. W. J Phys Chem 1996, 100, 15421.

96. Gilbert, A. T. B.; Lee, A. M.; Gill, P. M. W. Mol Struct (Theochem) 2000, 500, 363.

97. Lee, A. M.; Gill, P. M. W. Chem Phys Lett 1998, 286, 226.

98. Gillbert, A. T. B.; Taylor, S. W.; Gill, P. M. W. Proceedings of the 2000 Molecular Densities and Distributions Workshop, submitted.

99. Lee, A. M.; Gill, P. M. W. Chem Phys Lett 1999, 313, 271.

100. Gill, P. M. W.; Lee, A. M.; Nair, N.; Adamson, R. D. J Mol Struct (Theochem) 2000, 506, 303.

101. Besley, N. A.; Lee, A. M.; Gill, P. M. W., in preparation.

102. Head–Gordon, M.; Graña, A. M.; Maurice, D.; White, C. A. J Phys Chem 1998, 99, 14261.

103. See, for example, Oumi, M.; Maurice, D.; Head–Gordon, M. Spectrochim Acta 1999, 55A, 525.

104. For a general introduction to continuum solvation models, see, for example, Tomasi, J.; Persico, M. Chem Rev 1994, 94, 2027.

105. Florian, J.; Warshel, A. J Phys Chem B 1997, 101, 5583.

106. See, for example, Florian, J.; Sponer, J.; Warshel, A. J Phys Chem B 1999, 103, 884.

107. Baker, J. J Comput Chem 1986, 7, 385.

108. Baker, J.; Kessi, A.; Delley, B. J Chem Phys 1996, 105, 192.

109. Baker, J. J Comput Chem 1997, 18, 1079.

110. Kutzelnigg, W.; Ottschofski, E.; Franke, R. J Chem Phys 1995, 102, 1740.

111. Dirac, P. A. M. Proc R Soc (London) A 1928, 117, 610.

112. Heß, B. A. Ber Bunsenges Phys Chem 1997, 101, 1.

113. Kedziora, G. S.; Pople, J. A.; Rassolov, V. A.; Ratner, M. A.; Redfern, P. C.; Curtiss, L. A. J Chem Phys 1999, 110, 7123.

114. Kedziora, G. S., manuscript in preparation.

115. Handy, N. C.; Yamaguchi, Y.; Schaefer, H. F., III J Chem Phys 1986, 84, 4481.

116. Kutzelnigg, W. Mol Phys 1997, 90, 909.

117. Rassolov, V. A.; et al., in preparation.

118. Furlani, T. R.; Kong, J.; Gill, P. M. W. Comput Phys Commun 2000, 128, 170.