

Fast Density-Functional-Theory Calculation

With Q-Chem Program

Jing Kong, Shawn T. Brown, Laszlo Fusti-Monlar, Yihan Shao

Q-Chem Inc., 5001 Baum Blvd., Pittsburgh, PA 15668

www.q-chem.com

ABSTRACT

DFT (Density functional Theory) is an essential tool in the molecular modeling, especially for studies in biology and nanotechnology. A new algorithm is presented here to improve the efficiency of the computation of exchange-correlation contributions, a major time-consuming step in a DFT calculation in the Q-Chem Program. Q-Chem is a state-of-the-art commercial quantum chemistry program that is used to model atomic and molecular processes in a wide range of disciplines, including biology, chemistry, and materials science. The new method, called multiresolution exchange-correlation (mrXC), takes advantage of the variation in resolution among the Gaussian basis functions, and shifts the calculation associated with low-resolution (smooth) basis function pairs to an even-spaced cubic grid. It. Preliminary results with local density approximation have shown that mrXC yields 3 -5 times improvement in efficiency without loss of accuracy.

Keywords: density-functional theory, electronic structure, *ab initio*, molecular modeling, molecular simulations.

1 Q-CHEM PROGRAM

Q-Chem is a high-performance electronic structure program developed collaboratively between a number of university research groups and Q-Chem Inc. This synergism has proven crucial because it allows state-of-the-art quantum chemical theories and algorithms to be incorporated into a stable and usable software platform for chemical applications as well as methodological developments. Q-Chem is distributed in two ways. It is available standalone from Q-Chem Inc. (www.q-chem.com). It is also the electronic structure back-end of the Spartan modeling package from Wavefunction Inc. (www.wavefun.com).

The current release of Q-Chem is Q-Chem 2.1 [1]. It includes all of the fundamental quantum chemistry models for ground and excited states:

- Efficient, linear-scaling algorithms for large-molecule DFT and Hartree-Fock (HF) calculations
 - CFMM and FTC for fast Coulomb
 - LinK for HF exchange
 - LDA, GGA and hybrid functionals
- Extensive excited state capabilities including CIS time-dependent DFT
- Search of stable structures, transition-states and reaction coordinates in Cartesian, Z-matrix or delocalized internal coordinates
 - Impose bond angle, dihedral angle (torsion) or out-of-plane bend constraints
- Effective core potentials for heavy atoms
- Analytical second derivatives (frequency) at DFT and HF levels
- Parallel computing using MPI; memory distributed for frequency calculations
- Efficient implementation of MP2 energy and gradients with the option of local correlations
- Standard high level electron correlation methods including QCISD, QCISD(T), CCSD, CCSD(T), MP3, and MP4
- New correlation methods including OD, QCCD CCSD(2), OD(2) and active space coupled cluster methods VOD, VQCCD, VOD(2)
 - Coupled cluster excited state energies (OD, and VOD)
- Solvation with Langevin dipoles model
- Molecular properties including NMR chemical shifts.

1.1 Q-Chem 3.0

Right now, we are preparing for the release of Q-Chem 3.0. Q-Chem 3.0 is a major upgrade of Q-Chem package. It includes some of the most recent development in the field of computational chemistry, and provides the user with significant efficient gains and new computational tools. The new key features in Q-Chem 3.0 release is summarized in the following table. The technical details of the new technologies are presented in the paper appearing in Physical Chemistry Chemical Physics [2].

Key Features	Innovation	Benefit
NMR Chemical Shifts	1st linear-scaling NMR	Feasible for hundreds of atoms
Fast DFT Calculations	New FTC and DFT methods	Up to 3 times faster
Fast MP2 Calculations	RI-MP2 (with gradients) RI-TRIM local MP2	3 to 10 times faster
Scaled MP2 Schemes	New SOS-MP2, MOS-MP2 schemes with RI, gradients	More accurate and faster
Coupled Cluster Methods	EOM-(EE, IP, EA, SF)-CCSD with gradients	Robust treatment of radicals, bond-breaking and symmetry-breaking problems
Valence Correlation Models	SSG, PP, IP, RP methods with gradients	Inexpensive alternative to GVB/CASSCF-type models
QM/MM Hybrid Methods	ONIOM Energy, Gradient, Frequency	Suitable for huge structures
Anharmonic Frequencies	New TOSH Model	Direct prediction of IR peaks
Continuum Solvation Model	SS(V)PE Electronic Model	Improved accuracy for ions
Parallel DFT/HF	Energies, gradients and frequencies (with distributed memory)	Scalable to many CPUs
Dual Basis Methods	HF, DFT, MP2 energies	About 10 times faster for large basis sets
Transition Structure Finder	Growing String Method	Improved search Tool
Direct Dynamics	Extended Lagrangian and Born-Oppenheimer	Enables trajectory studies
Graphic User Interface	WebMO on distribution CD	Basic builder and viewer

2 NEW DFT ALGORITHM: MULTIRESOLUTION EXCHANGE- CORRELATION

Numerous efforts have been made in the last decade to reduce the computational cost of the DFT calculation, since it is still quite computationally demanding for a system with a hundred atoms or more. In the past, majority of the efforts have been concentrated on the development of efficient algorithms for Coulomb and HF exchange. For example, we have implemented a Coulomb method in Q-Chem called the Fourier Transform Coulomb (FTC) method [3]. It has been shown through a series of benchmark calculations that FTC provides several-fold improvement in efficiency over the most efficient existing Coulomb scheme imbedded in Q-Chem.

Advances in the Coulomb and HF exchange computation have made more pronounced the determination of exchange-correlation (XC) contribution. In particular, we have shown [3] that the advances in the Coulomb computation have made the calculation of the XC contribution the most time-consuming step in a DFT calculation with GGA functionals. At this conference, we will present a new scheme for the efficient evaluation of the XC contribution called the multiresolution exchange-correlation (mrXC) method. It is based on the well-known observation that a Gaussian function can be evaluated accurately on a numerical grid with sufficient resolution, and if a Gaussian is more diffuse (or smoother) than another, the resolution (or the density) of the grid for the former can be lower than that for the latter while retaining the same accuracy. Some preliminary results will be given to demonstrate the effectiveness of this method. The details of the algorithm and some preliminary results will be published in the Journal of Chemical Physics [4].

In practical DFT calculations, the forms of modern XC functionals are complicated and integrals over the functionals are generally evaluated numerically. A typical numerical quadrature for XC is composed of atom-centered grids (ACG's). In such a grid, the density of the grid points, or the resolution of the points, is high near the nucleus because the electron density changes rapidly in this region, and becomes increasingly lower away from the nucleus as the electron density changes slower. The grid density can vary by several orders of magnitude between the high and low density regions. In the current implementation, a basis function is evaluated at each grid point that it overlaps with, regardless of the grid density. While this is necessary for functions with high resolutions, it is unnecessary for smooth functions with low resolutions because such a function changes slowly in space. Instead, one can just evaluate the function values at an appropriate interval in the 3-dimensional space (i.e. forming a cubic grid) and interpolate the values at the required points through those at the nearby vertices on the even-spaced cubic grid.

The feasibility of the interpolation was recently studied [5]. Divided difference polynomial interpolation was explored as a means of translating the electron density and its gradients from the cubic grid to points on the atom-centered grid. Aspects of accuracy, error control through the use of the grid density were studied, and it is shown that majority of the effective basis function pairs (up to 80%) can be evaluated on the cubic grid accurately.

2.1 Theory

To demonstrate the mrXC algorithm, we will first start with the simple local density approximation, i.e. the XC functional is a function of only the electron spin density, $f \equiv f(\rho_\alpha, \rho_\beta)$. Examples of such functionals include the Slater exchange functional [6] and the VWN(Vosko-Wilk-Nusair) correlation functional [7].

The current computation of XC contribution is typically done in an ACG with grid points centered around each atom. The two time-consuming steps are:

1. Evaluate the spin-densities on the grid:

$$\rho^\alpha(i) = \sum_{\mu\nu} P_{\mu\nu}^\alpha \phi_{\mu\nu}(i) \quad (1)$$

where i is the index for the grid points and $\phi_{\mu\nu}$ is the product of two basis functions, i.e. $\phi_{\mu\nu}(i) = \phi_\mu(i)\phi_\nu(i)$.

In general, $\phi_{\mu\nu}$ is a linear combination of Gaussian functions at different centers.

2. Evaluate the XC matrix, the exchange-correlation contribution to the Fock matrix:

$$F_{\mu\nu}^\alpha = \sum_i f^{\rho^\alpha}(i) \phi_{\mu\nu}(i) \quad (2)$$

To speed up the calculation, we recognize that for a well-behaved function like a Gaussian, the value of the function at any point in the space (e.g. a point in the ACG) can be interpolated accurately using the values of the function at the grid points of an even-spaced cubic grid, especially for the purpose of integration, when the cubic grid has sufficient density. That is:

$$\phi_{\mu\nu,i} = \sum_m C_{im} \phi_{\mu\nu,m} \quad (3)$$

It is then straightforward to obtain the smooth part of the density and the Fock matrix on the ACG through interpolation for the BFPs that are smooth:

$$\rho_i = \sum_{\mu\nu} \sum_m C_{im} P_{\mu\nu} \phi_{\mu\nu,m} = \sum_m C_{im} \rho_m \quad (4)$$

$$F_{\mu\nu} = \sum_i \sum_m f_i^\rho C_{im} \phi_{\mu\nu,m} = \sum_m \tilde{f}_m^\rho \phi_{\mu\nu,m} \quad (5)$$

The computation cost can be further reduced when one recognizes that using Fourier functions instead of local polynomial ones will require less dense of a cubic grid. The following picture completes mrXC algorithm:

$$\rho_n \xrightarrow[1]{FFT} \rho_k \xrightarrow[2]{K_{\max}^n \rightarrow K_{\max}^m} \rho_k \xrightarrow[3]{FFT} \rho_m \xrightarrow[4]{DD} \rho_i$$

$$f_i^\rho \xrightarrow[5]{DD} \tilde{f}_m^\rho \xrightarrow[6]{FFT} \tilde{f}_k^\rho \xrightarrow[7]{K_{\max}^m \rightarrow K_{\max}^n} \tilde{f}_k^\rho \xrightarrow[8]{FFT} \tilde{f}_n^\rho$$

The fast Fourier transformations (FFTs) above are between two cubic grids with different grid densities. The major load of the work is actually carried out on the coarse grid which involves a loop over basis function pairs.

2.2 Illustration of accuracy of mrXC

To demonstrate the accuracy of mrXC, we have calculated the DFT energies with the Slater functional for a chain of 5 Alanine units with 53 atoms. Two Pople-type basis sets are used, 6-31G(df,pd) and 6-31G(2df,2pd). The quadrature for the XC numerical integration is SG-1, a widely used efficient grid that yields good accuracy. The calculation results are listed in Table 1.

Table 1. The difference in the DFT energy (ΔE) from mrXC, in comparison with the error of the SG-1 grid measured in $\mu\text{h}/\text{atom}$.

Molecule/Basis sets	Error of SG-1	ΔE of mrXC
(Alanine) ₅ /6-31G(df,pd)	2.9	0.1
(alanine) ₅ /6-31G(2df,2pd)	2.7	0.2

To estimate the errors of SG-1 for the testing calculations, we have calculated the DFT energies with a very fine ACG grid with 99 radial points and 590 angular points. The differences between SG-1 and the very fine grid are listed in Table 1 as the error of SG-1 without using mrXC. The next column lists the differences introduced by mrXC. It is apparent that mrXC does not change the accuracy of the calculations. SG-1 gives errors about a few microhartrees (μh) per atom and using mrXC only changes the errors by a few percent, which has little statistical significance.

2.3 Illustration of efficiency gain

Taxol, a drug molecule with 113 atoms has been chosen to demonstrate the effectiveness of mrXC. Two basis sets are used – 6-31G(df,pd) and 6-31G(2df,2pd). The computer is an AMD 2.1 MHz Athlon MP. The results are listed in Table 2.

As elucidated in the **Theory** section, the explicit evaluation of density values and the elements of XC matrix on the ACG (atom-centered grid) count for almost all of the computational cost (c.a. 97 %) of the XC part in

the current scheme. MrXC is designed to replace the explicit evaluation of smooth functions with the calculations on an even-spaced cubic grid.

Table 2 lists the CPU time of the explicit evaluations and that of the mrXC scheme for the comparison of the efficiency of the two approaches for different combinations of molecules and basis sets. It shows that using mrXC results in 3-5 times of speed-up for these molecules of moderate size. Also the speed-up increases when valence functions are added to the basis sets, since those valence functions have relatively small exponents and are therefore smooth.

Table 2. The speed up of CPU time of the electron density and XC matrix formations on the SG-1 quadrature grid with the original standard implementation and mrXC.

Molecule/Basis sets	Number of basis functions	Speed-up
Taxol/6-31G(df,pd)	1925	3.92
Taxol/6-31G(2df,2pd)	2450	4.94

REFERENCES

- [1] J. Kong, C. A. White, et al, "Q-Chem 2.0: A high-performance *ab initio* electronic structure program", *J. Comput. Chem.* **21**, 1532-1548 (2000).
- [2] Y. Shao, L. Fusti-Molnar, et al, "Q-Chem 3.0: A high-performance quantum chemistry program package for molecular electronic structure calculations", *Phys. Chem. Chem. Phys.*, accepted.
- [3] L. Fusti-Molnar, J. Kong, "Fast Coulomb Calculations with Gaussian Functions", *J. Chem. Phys.*, **122**, 074108 (2005).
- [4] J. Kong, S. T. Brown, L. Fusti-Molnar, "Efficient Computation of the Exchange-Correlation Contribution in the Density Functional Theory Through Multiresolution", *J. Chem. Phys.*, in press.
- [5] S. T. Brown, L. Fusti-Molnar, J. Kong, "Interpolation density values on a cartesian grid: Improving the efficiency of Lebedev based numerical integration in Kohn-Sham density functional algorithms", *Chem. Phys. Lett.*, **418**, 490 (2006).
- [6] P. A. M. Dirac, *Proc. Cam. Phil. Soc.*, 26, 376 (1930).
- [7] S. H. Vosko, L. Wilk and M. Nursair, *Can. J. Phys.*, **58**, 1200 (1980).