Chemical Physics Letters 687 (2017) 163–170

Contents lists available at ScienceDirect

Chemical Physics Letters

journal homepage: www.elsevier.com/locate/cplett

Research paper

Optimization of convergence criteria for fragmentation methods

Zhaoxi Sun [a,*], Tong Zhu [b,c], Xiaohui Wang [a], Ye Mei [a,b], John Z.H. Zhang [b,c,d,*]

[a] State Key Laboratory of Precision Spectroscopy, Institute of Theoretical and Computational Science, East China Normal University, Shanghai 200062, China
[b] NYU-ECNU Center for Computational Chemistry at NYU Shanghai, Shanghai 200062, China
[c] School of Chemistry and Molecular Engineering, East China Normal University, Shanghai 200062, China
[d] Department of Chemistry, New York University, NY, NY 10003, USA

ARTICLE INFO

Article history:
Received 26 May 2017
In final form 25 August 2017
Available online 9 September 2017

ABSTRACT

Fragmentation methods reproduce the result of the exponential scaling full system calculation within an error about 1 kcal/mol. They serve as a robust linear scaling regime of quantum treatment for large molecular systems. As the response of fragmentation methods to convergence criteria in SCF iterations is different from that of full system calculation and the propagated convergence error is much smaller than the fragmentation error, the speed of fragmentation ab initio calculation can be further enhanced via loosening the convergence criterion in the calculation of each fragment without sacrificing the accuracy of the overall energy. Tests on single configuration calculations, geometry optimization and molecular dynamics simulation are performed.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

The prohibitive scaling problem of quantum methods makes them impractical to deal with large molecules, which brings the boom of various hybrid theoretical-levels methods. The hybrid ones include hybrid quantum molecular mechanics (QM/MM) [1], ONIOM (Our own N-layered Integrated molecular Orbital and molecular Mechanics) [2], and many fragment-based algorithms [3,4]. Partitioning the whole system into a series of overlapping subsystems with hydrogen or partial residual caps to be treated in a QM/MM way, fragmentation methods greatly broaden the application area of QM calculation due to their efficiency. Fragment-based methods include two main categories [3], the inclusion–exclusion principle based methods including MTA (molecular tailoring approach) [5,6], MFCC (molecular fractionation with conjugate caps) [7], SMF (Systematic Molecular Fragmentation) [8] and GEBF (generalized energy based fragmentation) [9], and many-body expansion based EE-MB (electrostatically embedded many-body) [10] and FMO (fragment molecular orbital) [11] methods. MFCC [7], being proposed in the early stage of the development of fragmentation methods, is a typical molecular tailoring scheme. Extensions and variants of it were proposed [3,12–15]. They can reproduce the full system QM energy calculation with an accuracy of about 1 kcal/mol [16–20].

Recently, *ab initio* molecular dynamic (AIMD) approach for practical dynamic simulation of large molecules has attracted significant interests [21–23]. A fragment based AIMD approach was proposed for practical application in protein dynamics study. Dynamics being described with classical mechanics, in this AIMD approach, atomic forces are calculated via EE-GMFCC (electrostatically embedded generalized molecular fractionation with conjugate caps) methods on the fly [24]. Although the efficiency of this method is much higher than that of classical QM calculation, there is still more room for efficiency enhancement.

Due to the nonlinearity of Hartree-Fock equations, we must solve them iteratively to obtain a self-consistent set of molecular orbitals. However, sometimes simple iterative procedures do not converge. A number of convergence accelerating techniques are thus introduced. The popular least-squares interpolation convergence acceleration algorithm named direct inversion in the iterative subspace (DIIS) proposed by Pulay serves as a standard acceleration method for self-consistent field [25–27]. There are several DIIS derivatives [28,29]. For instance, GDIIS [30] uses geometry displacements as error quantities and is commonly used in geometry optimization. DIIS derivatives' applications in transition state searching are also promising. Sometimes the interpolation of DIIS inherently inhibits its convergence in transition state optimization and makes the iterations get stuck on one side of the transition state [31]. Using DIIS to accelerate the convergence of the self-consistent field in quantum chemistry calculation, one observes that the energy fluctuates with the number of iterations and a truncation due to loose convergence criteria may result in unreliable and inaccurate estimates, especially when the system

* Corresponding authors at: State Key Laboratory of Precision Spectroscopy, Institute of Theoretical and Computational Science, East China Normal University, Shanghai 200062, China (Zhaoxi Sun) and (J.Z.H. Zhang).

E-mail addresses: proszx@163.com (Z. Sun), zhzhang@phy.ecnu.edu.cn (J.Z.H. Zhang).

http://dx.doi.org/10.1016/j.cplett.2017.08.059
0009-2614/© 2017 Elsevier B.V. All rights reserved.

is large or contains unusual species or geometries. However, in this work we prove that loosening the convergence criteria properly in fragmentation methods still ensures an accurate and efficient estimate. A benchmark test on the influences of the SCF convergence criteria on the single point fragmentation calculations, geometry optimization and AIMD simulations is performed.

The paper is organized as follows: Section 2 provides an overview of the methodology and presents some discussions about the errors involved in the fragmentation method. Section 3 gives an outline of the computational details of the investigation of the effects of convergence criterion on single point calculations, dynamics simulations and geometry optimization of several protein systems. Then Section 4 presents results and discussions.

2. Theory and methodology

2.1. Brief review of the EE-GMFCC method

The EE-GMFCC method was proposed to accurately calculate the potential energy of protein. In this method, a protein is decomposed into fragments which are embedded in a field of point charges representing the remaining parts of protein. Two-body QM interaction energies between short-range non-neighboring fragment interactions are also included. Thus the gas-phase expression of energy is given by,

$$E_{EE-GMFCC} = \sum_{i=2}^{N-1} \tilde{E}(Cap_{i-1}^{*} A_i Cap_{i+1}) - \sum_{i=2}^{N-2} \tilde{E}(Cap_i^{*} Cap_{i+1}) + \sum_{\substack{i,j>i+2 \\ |R_i - R_j| \leqslant \lambda}} (\tilde{E}_{ij} - \tilde{E}_i - \tilde{E}_j) - \left\{ \sum_{i,j} \sum_{m,n} \frac{q_{m(i)} q_{n(j)}}{R_{m(i)n(j)}} - \sum_{\substack{i,j>i+2 \\ |R_i - R_j| \leqslant \lambda}} \sum_{m',n'} \frac{q_{m'(i)} q_{n'(j)}}{R_{m'(i)n'(j)}} \right\}, \quad (1)$$

where i and j represent the indexes of the ith and jth residue, respectively, $\tilde{E}$ is the energy of the each QM area, q_i denotes the atomic partial charge of atom i and R_{ij} is the distance between atom i and atom j. Variables of m and n are the indexes of residues which are used as back ground charges in the QM calculation of the jth and ith residue respectively. More thorough descriptions of the method can be found elsewhere. [24] Basis Set Superposition Error (BSSE) is relatively small in the fragmentation method of GEBF due to error cancellation in the combination and assembly. [32] Due to the similarity between these two fragmentation methods, BSSE would be small in our method. Thus no BSSE correction is calculated. In the following discussion we use MFCC to represent EE-GMFCC for brevity.

2.2. Convergence criterion and DIIS

EE-GMFCC is combined with Gaussian 09 package [33] in the current work. Note that our discussion about the influence of SCF convergence criteria on the accuracy of the EE-GMFCC method can be extended to other QM packages as well as other fragment-based methods. (If the number of atoms included in each fragment is no more than that in our method, 50 atoms approximately.) There are a number of convergence criteria available. In the implementation of Gaussian 09, SCF convergence requires both $<10^{-N}$ RMS density matrix change (RMSDP) and $<10^{-(N-2)}$ maximum change (MaxDP) in the density matrix, while energy changes (energy differences, DE) are not used to check the convergence. N is a positive integer, which normally ranges from 4 to 8. The tested convergence criterion of energy is $10^{-(N-2)}$, which is only used for testing of convergence rather than determining it. Typically an SCF 10^{-N} RMS change in the density matrix corresponds to a 10^{-2N} change in energy in atomic units. As RMSDP is the most significant criterion, in the following discussion its magnitude is used to represent the magnitude of all criteria.

The SCF convergence is strongly related to stability conditions and sometimes a convergence is difficult to reach [34–36]. Near degeneracy between two states and small energy difference between occupied and virtual orbitals compared with 2 electron contributions to Hamiltonian incur large changes in the Fock matrix [27]. Hence convergence acceleration algorithms such as DIIS are introduced.

Basic concepts and properties of the convergence acceleration algorithm of DIIS are illustrated as follows. Basic iterative methods always contain an update step $x_{n+1} = x_n + e_n$, where x_n is the value of the function to be minimized in the nth-step iteration and the correction term e_n is the derivative or a variant of it. In most algorithms only the information in the current step is used [37]. while DIIS exploits the information in the current step as well as in several previous iterative steps. It defines a trial move $\tilde{x}_{n+1} = x_n + e_n$ and computes an improved iterates $\tilde{x}_{n+1} = \sum_{i=1}^{n} c_i \tilde{x}_i with \sum_{i=1}^{n} c_i = 1$ by minimizing the least square functional $J_{DIIS} = \frac{1}{2} \| \sum_{i=ini}^{n} c_i e_i \|^2$.

Using a combination of EDIIS (energy-DIIS) and DIIS to further accelerate the convergence [38], however, may lead to an increase of the absolute value of the energy change, which will not happen with Bacskay's quadratically convergent SCF method as a result of different natures of the two methods [39,40]. (A quadratically convergent SCF involves (approximate) linear searches and calculation of the exact Hessian matrix to determine the direction for optimizing, thus being much slower.)

2.3. Uncertainty of MFCC estimates

MFCC and various fragmentation methods are subject to both convergence errors and fragmentation errors. The convergence errors represent the differences between the estimates with infinite precision and those under the convergence criterion of 10^{-N}. Normally the most precise estimate is substituted with the results under the convergence criterion of 10^{-8}. When the convergence criterion is infinitely small, the fragmentation estimates still mismatch their full system components. Such differences are named fragmentation errors.

As the convergence is determined by the changes in the density matrix, energy or other quantities between consecutive SCF cycles, a small enough RMSDP or DE does not ensure small convergence errors or uncertainty of QM results. Thus small delta values are always recommended in quantum mechanics calculations for safety. As is done in normal QM and QM/MM calculations to ensure reliable estimates for various properties of the system, fragmentation methods always use the tight convergence criterion of 10^{-8}.

As noted above, in MFCC the estimations of energy and the other properties are decomposed into the calculations of the fragments. Thus by neglecting the cross correlations of various properties between fragments, the variances (convergence errors) of the these properties computed via MFCC are the summations of the variances of all fragments and caps,

$$\sigma_A^2 = \sum_{i=1}^{N_{fragments}+N_{caps}} \sigma_{A,i}^2 \quad (2)$$

Theoretically the value of convergence criterion is not exactly the same with the convergence errors. Convergence is measured by the changes of quantities (such as energy) between successive iterations. When the convergence is well achieved, the convergence criterion can give a trial estimate of convergence errors. Thus if the target precision of full system quantum mechanics calculation is 10^{-N} and the number of the fragments and caps is less than

Download English Version:

https://daneshyari.com/en/article/5377507

Download Persian Version:

https://daneshyari.com/article/5377507

Daneshyari.com