ACCELERATING THE CONVERGENCE OF THE COUPLED-CLUSTER APPROACH. THE USE OF THE DIIS METHOD

Gustavo E. SCUSERIA, Timothy J. LEE and Henry F. SCHAEFER III
Department of Chemistry, University of California, Berkeley, CA 94720, USA

Received 9 June 1986; in final form 7 July 1986

The direct inversion of the iterative subspace (DIIS) method is implemented into the closed-shell coupled-cluster single- and double-excitation (CCSD) model to improve the convergence of the coupled non-linear CCSD equations. As with self-consistent field and gradient methods the DIIS method proves to be quite successful. Depending on the particular case, we find that the number of iterations required to achieve convergence is generally reduced to between 30 and 70% of that required for the pure CCSD iterative procedure (without any extrapolation technique).

1. Introduction

The direct inversion of the iterative subspace (DIIS) method as introduced by Pulay [1] is a simple and efficient convergence acceleration technique which has been successfully applied to the self-consistent field (SCF) method, involving closed-shell, high-spin restricted open-shell, open-shell singlet and two-configuration wavefunctions [1–3], as well as to the problem of geometry optimization [4]. The DIIS method has also been applied to the solution of equations apearing in the coupled-pair functional method [5] and in the theory of effective Hamiltonians in Fock space [6].

The nature and convergence properties of the SCF equations have been thoroughly analyzed in the literature [7]. The standard way of solving this non-linear algebraic system is using an iterative procedure in connection with a diagonalization of the Fock matrix. Some methods which do not use diagonalization procedures are also adequate to solve the SCF equations but they are not very popular among chemists [8]. When slowly convergent cases arise or even if the iterative procedure is divergent, acceleration or extrapolation techniques which reduce the number of steps are highly desirable. Several methods have been proposed for the SCF and multiconfiguration SCF cases [9–11].

Coupled, non-linear and non-Hermitian equations occur in coupled-cluster (CC) theory [12,13]. The existence and reality of their solutions have been studied [14] but very little is known about the convergence properties.

Recently, we have developed a closed-shell CC single plus doubles (CCSD) method [15] which has proved to be efficient in computational requirements compared to the earlier and more general unrestricted Hartree–Fock (UHF) CCSD methods pioneered by Bartlett and co-workers [16]. In developing an efficient CCSD code one should optimize not only the CPU time per iteration but also the number of iterations needed to achieve convergence. Bartlett discovered some time ago that convergence of the CCSD equations, assuming the standard rearrangement involving the diagonal Fock matrix elements [16] was very slow in many cases. We have observed that as many as 91 iterations are required for H_2Be with a double-zeta (DZ) basis set in order to achieve convergence in the correlation energy to 10^{-10} au. Purvis and Bartlett developed a reduced linear equation (RLE) method [17] to deal with these slowly convergent cases and to provide generally improved convergence. Their method was used with several CC approximations and compared to Padé approximants and the perturbation theory iterative procedure [17]. However, despite the successful performance of the RLE method in these cases, calculations with a full CCSD model were not presented

0 009-2614/86 $ 03.50

and we are not aware of any other article specifically dealing with the convergence properties of the CCSD equations †.

Thus, we decided to implement the DIIS method within our closed-shell CCSD code and compare its performance with the pure CCSD iterative procedure. The choice of the DIIS method over many other available was made based on the remarkable success of DIIS in SCF procedures. It must also be mentioned that the DIIS method is related to more general procedures previously proposed to solve large eigenvalue equations [18,19].

2. Method

Detailed equations of the DIIS method have been presented by Pulay [1,2] and will not be repeated here. A more precise description of the closed-shell CCSD model is presented elsewhere [15].

It must be pointed out that the implementation of the DIIS method within the CCSD model does not require any important change in the original formulation. However, due to the specific nature of the CCSD problem some alternatives are possible.

Firstly, since the unknown t_1 and t_2 amplitudes satisfy two different (although coupled) equations [12,13], they can be collected into a unique vector or treated separately. In the former case only one direct inversion is needed for calculating the expansion coefficients of the new guess vector in terms of the previous ones, whereas in the latter the two sets of expansion coefficients (for t_1 and t_2) are not necessarily equal and two "sub DIIS" problems need to be solved separately. The computational effort of the DIIS step is extremely small so both procedures are inexpensive. However, we found that using different expansion coefficients for t_1 and t_2 does not improve the performance of the method and in some test cases the convergence was actually somewhat slower. Let X be the array containing all t_1 and t_2 amplitudes, $X=(t_1, t_2)$. Then, from the CCSD equations $X_{i+1}=f(X_i)$, we have the sequence

$$X_1\, X_2\, X_3 \ldots X_{n-1}\, X_n \rightarrow X_p\, X_{p+1} \rightarrow X_{p'}\, X_{p'+1} \ldots , \quad (1)$$

where each arrow represents a DIIS extrapolation. The error vector (which must vanish at convergence) is naturally defined as $\Delta X_i = X_{i+1} - X_i$ and it must be noticed that $X_p - X_n$ and $X_{p'} - X_{p+1}$ should not be included in the list of error vectors.

Secondly, one has to decide when to start extrapolating, or more specifically, a tolerance value must be chosen such that an extrapolation will be done any time the norm of the extrapolated error vector is less than this value. After several tests we found that both large (10^{-2}) and small (10^{-7}) tolerance factors adversely affected the overall performance and we finally chose 10^{-5} as our standard value. With this tolerance the first extrapolation is usually done between the fifth and eighth iteration, depending on the particular case.

A third alternative which we have considered is letting the system "relax" for n iterations after performing an extrapolation. This means that for n consecutive iterations the pure iterative procedure is used even if the conditions for a new extrapolation are satisfied. We obtain a clear reduction in the number of iterations using $n=2$ or $n=3$ compared to the case where the system is not allowed to relax ($n=0$).

The last factor we have included in this study is the maximum number of iterations from which we are storing the information required by the DIIS technique (i.e. the error vectors and corresponding t). Our tests showed that six was too small and twelve unnecessarily large. The value we ended up adopting is a maximum of eight, which is also the value we found optimum for solving the SCF equations. This number also determines the dimension of the linear problem to be solved [1]. This point is important since for medium-size basis sets the disk file containing these data can become very large. On the other hand, we always start checking the extrapolation condition from the minimum possible number (two) of iterations. Finally, we should mention that we did not find the information coming from the first two iterations useful, probably because they contain a rather large step in the t_2 amplitudes and consequently in the correlation energy. Actually, our first iteration in the code corresponds to the second one in the sequence starting with $t_1=0$ and $t_2=0$. This is simply because

† Subsequent to our DIIS implementation we have learned (via private communication) that Bartlett and co-workers do use the RLE method for the CCSD equations and typically achieve convergence for single-point energy calculations in 10–15 iterations.

Table 1
Comparison between pure and DIIS results for the convergence of the closed-shell CCSD equations

Molecule	Basis	N [a]	N [b]	Method	E_{SCF} (au) [c]	E_{corr} (au)	Q [d]	Number of iterations
H_2 [e]	DZ	4	1	pure	−1.840873851	−0.0249062	7	14
	DZ	4	1	DIIS		−0.0249062	7	8
H_2	DZ	4	1	pure		−0.0249061772	10	21
	DZ	4	1	DIIS		−0.0249061773	10	13
HB [f]	DZ	12	3	pure	−25.113673919	−0.0730803	7	26
	DZ	12	3	DIIS		−0.0730804	7	11
HB	DZ	12	3	pure		−0.07308037245868	14	66
	DZ	12	3	DIIS		−0.07308037245869	14	27
HB+HB [g]	DZ	24	6	pure	−50.227347838	−0.1461606	7	28
	DZ	24	6	DIIS		−0.1461608	7	11
HB+HB	DZ	24	6	pure		−0.14616074491251	14	67
	DZ	24	6	DIIS		−0.14616074491936	14	27
H_2Be [h]	DZ	14	3	pure	−15.536467069	−0.0877225	7	53
	DZ	14	3	DIIS		−0.0877235	7	14
H_2Be	DZ	14	3	pure		−0.0877229611	10	91
	DZ	14	3	DIIS		−0.0877229616	10	22
H_2O [i]	DZ	14	5	pure	−76.009837603	−0.1462381	7	16
	DZ	14	5	DIIS		−0.1462381	7	11
H_2O	DZ	14	5	pure		−0.1462381254	10	26
	DZ	14	5	DIIS		−0.1462381256	10	15
H_2O [i]	DZP [j]	26	5	pure	−76.044199308	−0.2212162	7	17
	DZP	26	5	DIIS		−0.2212163	7	10
H_2O	DZP	26	5	pure		−0.2212162696	10	27
	DZP	26	5	DIIS		−0.2212162697	10	16
H_2O [l]	TZ+2P [k]	44	5	DIIS	−76.049057350	−0.2647448	7	11
H_5O_2 [l]	DZP [m]	57	10	DIIS	−152.426888390	−0.4422385	7	11

[a] Number of basis functions. [b] Number of occupied molecular orbitals.
[c] SCF energies were converged to 10^{-12} au for $Q=7$, 10 and 10^{-15} au for $Q=14$.
[d] 10^{-Q} is the tolerance factor in the correlation energy (see text).
[e] H–H = 1.4 au. [f] H–B = 2.329 au. [g] HB–HB = 90000 au.
[h] Basis set and geometry taken from ref. [16].
[i] Geometry taken from ref. [16].
[j] Polarization exponent for H is 0.75 and for O is 0.80.
[k] Polarization exponents for both H and O are 1.50 and 0.60.
[l] Optimized DZP single and double CI geometry.
[m] Polarization exponent for H is 1.00 and for O is 0.80.

the t_2 amplitudes which result from the iteration starting with nil values may be readily obtained analytically and are inserted as the initial values for the t_2 coefficients (the t_1 remain zero initially).

3. Results

In table 1 results obtained for a set of test cases are presented. All DIIS results were obtained using only one set of expansion coefficients and setting the parameter $n=3$, as explained above. The test cases presented in table 1 were chosen from well-behaved calculations (H_2 and H_2O) and slowly convergent cases (HB, HB+HB and H_2Be). The basis sets used are the Huzinaga–Dunning double zeta [20,21] and triple zeta [22] with the hydrogen exponents scaled by a factor of 1.2. The DIIS CCSD results are compared to "pure" CCSD approach wherein no extrapolation procedure is employed. The procedures were considered converged when the correlation energy difference (in au) between two successive iterations was less than a certain tolerance (10^{-Q}). Three different cases were analyzed: $Q=7$, $Q=10$ and $Q=14$. These values represent the three different ranges of precision usually imposed for single-point energies, gradients and very accurate calculations, respectively. We als carried out very precise computations to check the size extensivity of the CCSD model $E(\mathrm{HB+HB})=2E(\mathrm{HB})$.

A similar convergence criterion can be chosen by requiring the norm of X_n-X_{n-1} to be smaller than a certain tolerance ($10^{-Q'}$). Using the Euclidean norm $X^2=\sum x_i^2$, we found that typically $Q=7$ implies $Q'=6$, which means that the t coefficients satisfied the CCSD equation to an accuracy of 10^{-6}.

Results given in table 1 show that in well-behaved cases the DIIS method decreases the number of iterations by 30%. In the slowly convergent cases the acceleration is even more remarkable, and between 1/3 and 1/4 the number of iterations of the pure iterative procedure are required to reach convergence. For single-point energy calculations ($Q=7$) between eight and eleven iterations are typically needed to achieve convergence and, in the majority of cases, having extrapolated only once.

In conclusion, we found that the DIIS method is well suited to accelerate the solution of the CCSD equations and we hope that these results will be useful to other researchers.

Acknowledgement

GES wishes to acknowledge the financial support by CONICET, the National Research Council of Argentina. This research was supported by the US National Science Foundation, Chemistry Division, Grant CHE-8218785.

References

[1] P. Pulay, Chem. Phys. Letters 73 (1980) 393.
[2] P. Pulay, J. Comput. Chem. 3 (1982) 556.
[3] T.P. Hamilton and P. Pulay, J. Chem. Phys. 84 (1986) 5728.
[4] P. Csaszar and P. Pulay, J. Mol. Struct. 114 (1984) 31.
[5] R. Ahlrichs, P. Scharf and C. Ehrhardt, J. Chem. Phys. 82 (1985) 890.
[6] W. Kutzelnigg and S. Koch, J. Chem. Phys. 79 (1983) 4315.
[7] M.A. Natiello and G.E. Scuseria, Intern. J. Quantum Chem. 26 (1984) 1039, and references therein.
[8] P.-G. Reinhard and R.Y. Cusson, Nucl. Phys. A378 (1982) 418.
[9] G.B. Bacskay, Chem. Phys. 65 (1982) 383.
[10] R.N. Camp and H.F. King, J. Chem. Phys. 77 (1982) 3056.
[11] M. Page and J.W. McIver, J. Chem. Phys. 79 (1983) 4985.
[12] J. Čížek, Advan. Chem. Phys. 14 (1969) 35.
[13] R.J. Bartlett, Ann. Rev. Phys. Chem. 32 (1981) 359; in: Advanced theories and computational approaches to the electronic structure of molecules, ed. C.E. Dykstra (Reidel, Dordrecht, 1984) pp. 127–159.
[14] T.P. Zivkovic, Intern. J. Quantum Chem. Symp. 11 (1977) 413;
T.P. Zivkovic and H.J. Monkhorst, J. Math. Phys. 19 (1978) 1007.
[15] G.E. Scuseria, A.C. Scheiner, J.E. Rice, T.J. Lee and H.F. Schaefer III, to be published.
[16] G.D. Purvis and R.J. Bartlett, J. Chem. Phys. 76 (1982) 1910.
[17] G.D. Purvis and R.J. Bartlett, J. Chem. Phys. 75 (1981) 1284.
[18] R.J. Bartlett and E. Brändas, J. Chem. Phys. 56 (1972) 5467; 59 (1973) 2032.
[19] P.E.S. Wormer, F. Visser and J. Paldus, J. Comput. Phys. 48 (1982) 23.
[20] S. Huzinaga, J. Chem. Phys. 42 (1965) 1293.
[21] T.H. Dunning, J. Chem. Phys. 55 (1971) 716.
[22] T.H. Dunning and P.J. Hay, in: Modern theoretical chemistry, Vol. 3, ed. H.F. Schaefer III (Plenum Press, New York, 1977).