

Applications of a novel algorithm for the calculation of MCSCF wavefunction: a look into possible avenues of convergence acceleration

KALYAN KUMAR DAS, DEBASHIS MUKHERJEE and
SANKAR PRASAD BHATTACHARYYA*

Department of Physical Chemistry, Indian Association for the Cultivation of Science,
Jadavpur, Calcutta 700 032, India

Abstract. The efficacy of a method based on the direct inversion in the iterative subspace (DIIS) in accelerating the approach to self consistency in the calculation of the MCSCF wavefunction using a novel algorithm developed earlier, is compared with that of a simple damping technique. Although the 'damping' turns out to be ineffective in the 'quadratic region', it accelerates remarkably in the rate of descent on the energy hypersurface in the early stages of the iterative process which leads to an impressive overall increase in the rate of approach to self consistency. The DIIS based procedure turns out to be ineffective when coupled to the present method and is plagued by ill conditioning problems. Calculations are done to compute the equilibrium geometrical parameters, charge density on different atoms, and dipole moment of HNO molecule in the lowest $1,3n\pi^*$ states at the INDO/2-MCSCF level.

Keywords. Convergence acceleration; MC-SCF theory; through damping; direct inversion in the iterative subspace; orthogonal gradient method; orthonormality constrained variation method.

1. Introduction

The single determinant SCF theory and the simple orbital picture emerging from it has undoubtedly been extremely useful in interpreting and correlating a vast body of chemical facts. This simplicity and usefulness of the independent particle model notwithstanding, there are situations when one has to abandon this simple model and look for a more appropriate one. For example, the single configuration closed shell wavefunction fails to describe correctly the dissociation of a molecule into open shell fragments. In fact, whenever one or more configurations are energetically degenerate or quasi-degenerate with the main one, the single determinant representation for the wavefunction of the system breaks down. In such a situation we are forced to adopt a many configuration description of the wavefunction where each configuration represents a Slater determinant constructed from a set of orthonormal one-electron functions (orbitals). If we now set out to make energy stationary with respect to variations in the configuration weight factors as well as in the orbital forms, we arrive at what is known as the MCSCF method in quantum chemistry parlance. The pioneering

Dedicated to Professor Sadhan Basu on the occasion of his 65th birth anniversary.

* To whom all correspondence should be addressed.

use of this method in atomic physics can be traced to the early work of Hartree and Seirles (1939). A generalized reformulation was later made by McWeeny (1955, 1956, 1960). Successful molecular calculations were made much later (Das and Wahl 1967; Clementi and Veillard 1967) and since then a lot of activity has been noted in the field of MCSCF theory (Olsen and Yeager 1983). The main thrust of this activity has been towards (i) the development of efficient techniques for solving the 'first order' MCSCF equation (ii) the development of a higher order MCSCF theory e.g. quadratically convergent MCSCF algorithms (QCMCSCF). Although considerable progress has been made in the field of QCMCSCF theory in recent years, the computational involvement in methods of this category is high enough to discourage large scale practical applications of the algorithm and even now the bulk of MCSCF calculations are based on the first order variational equation only.

Over the last few years we have developed and successfully applied in our laboratory a novel technique for solving the first order MCSCF equation (Mukherjee 1978; Adnan *et al* 1978; McWeeny and Newbould 1980; Bhattacharyya and Mukherjee 1981). The one-configuration analogue, too, has been widely tested (Bhattacharyya and Mukherjee 1979; Bhattacharyya *et al* 1982; Das *et al* 1984). The method has many desirable features e.g. smooth and fast convergence, numerical stability, ease of programming, etc. Even then, in the course of pursuing our research programme on the calculation of ground and excited state potential surfaces of small carbonyls and thiocarbonyls (Das *et al* 1986) at the INDO/2-MCSCF level, we felt that we should look for some means of accelerating the approach to self-consistency within the framework of our method even though it outperforms many of the currently used techniques of solving the first order MCSCF equation in its improved rate of approach to self-consistency. This becomes imperative just for cutting down the total computational expenses when large scale applications are envisaged. The present paper describes our experience with two such means of convergence acceleration employed by us. Besides, we present here the energy-optimized geometrical parameters of the simplest nitrosyl molecule viz, HNO in the $^1,^3n\pi^*$ state, the final state electron densities and dipole moment. The outlay of this paper is as follows: §2 discusses the general conditions of stationarity and in §3, the algorithm used by us is derived in a slightly different way. The convergence acceleration schemes are proposed in §4 while §5 presents the results.

2. General condition of stationarity

Let us consider a truncated set of n molecular orbitals (MO) $\Phi = \{\phi_1, \phi_2, \dots, \phi_i, \dots, \phi_n\}$ which are linear combinations of a set of m -basis functions ($m \geq n$), $\chi = (\chi_1, \chi_2, \dots, \chi_i, \dots, \chi_m)$ are related by the linear transformation:

$$\Phi = \chi T, \quad (1)$$

where T is an $(m \times n)$ matrix of the linear expansion coefficients. The variational trial function $\tilde{\psi}$ is given by:

$$\tilde{\psi} = \sum_{I=1}^N C_I \psi_I, \quad (2)$$

where the ψ_I 's are Slater determinants (with appropriate spin couplings) constructed

from the set $\Phi = \{\phi_1, \phi_2, \dots, \phi_n\}$ and the corresponding energy functional is

$$\begin{aligned} \tilde{E} &= \langle \tilde{\psi} | H | \tilde{\psi} \rangle / \langle \tilde{\psi} | \tilde{\psi} \rangle \\ &= \sum_{ij} (P_1)_{ij} \langle \phi_i | h | \phi_j \rangle + \frac{1}{2} \sum_{ijkl} (P_2)_{kl,ij} \langle kl | ij \rangle, \end{aligned} \quad (3)$$

where P_1 and P_2 are the one- and two-electron density matrices respectively. If we now proceed to make \tilde{E} stationary with respect to variations in (i) the CI coefficients $\{C_i\}$ and (ii) the orbitals, through variations δT in the linear expansion coefficients T_{ij} , we get two sets of coupled equations. Stationary conditions for the first type of variations leads to the normal CI equation

$$HC = CE, \quad (4)$$

while the condition for the stationarity with respect to variations of the type (ii) subject to the orthonormality of the orbitals, is expressed in the form

$$hTP_1 + Z = ST\varepsilon \quad (5)$$

where ε is the matrix of Lagrangian multipliers arising from the orthonormality constraints, and Z represents the electron-electron repulsion matrix with elements defined by

$$Z_{pi} = \sum_{jkl} \sum_{qrs} T_{jq}^\dagger \langle pq | rs \rangle T_{rk} T_{sl} P_{2kl,ij}.$$

One can easily show that the necessary and sufficient conditions for a given set of orbitals $\phi (= \chi T)$ to satisfy (5) is that ε must be hermitian. A variety of methods has been suggested for achieving this condition. The method adopted by us for solving (5) proceeds as follows in the next section.

3. Determination of MCSCF orbitals

Let us start by noting that the equation to be solved is $V = ST\varepsilon$ where

$$V = hTP_1 + Z, \quad (6)$$

and

$$V^\dagger = \varepsilon^\dagger T^\dagger S. \quad (7)$$

Multiplying (6) by T^\dagger from the left and (7) by T from the right we have (see Mukherjee 1978):

$$T^\dagger V = (T^\dagger ST)\varepsilon = \varepsilon; \quad (T^\dagger ST = 1), \quad (8)$$

and

$$V^\dagger T = \varepsilon^\dagger (T^\dagger ST) = \varepsilon^\dagger. \quad (9)$$

Combining (9) with (8) we have (noting that $TT^\dagger = S^{-1}$)

$$\begin{aligned} \varepsilon^\dagger \varepsilon &= (V^\dagger T T^\dagger V) \\ &= (V^\dagger S^{-1} V) \end{aligned} \quad (10)$$

The stationarity condition demands that

$$\varepsilon^\dagger = \varepsilon,$$

so that at convergence

$$\varepsilon^\dagger \varepsilon = \varepsilon^2 = (V^\dagger S^{-1} V), \quad (11)$$

i.e.

$$\varepsilon = (V^\dagger S^{-1} V)^{1/2}. \quad (12)$$

Away from the stationary point (12) provides an 'estimate' of the ε matrix ($\tilde{\varepsilon}$, say). At the n th iterative stage (7) with ε replaced by $\tilde{\varepsilon}$ reads as follows:

$$V_n = S T_n \tilde{\varepsilon}_n,$$

or,

$$T_n = S^{-1} V_n \tilde{\varepsilon}_n^{-1} = S^{-1} V (V^\dagger S^{-1} V)^{-1/2}. \quad (13)$$

At the true stationary point $\varepsilon^\dagger = \varepsilon$ and (13) does not update the coefficient matrix T any further—otherwise, the right hand side of (13) provides an updated coefficient matrix. This leads us to the following iterative sequence for the self consistent determination of T (or ϕ):

$$T_n \rightarrow T_{n+1} = S^{-1} V_n (V_n^\dagger S^{-1} V_n)^{-1/2}. \quad (14)$$

If $\varepsilon^\dagger = \varepsilon$ within a preassigned limit, iterations are terminated (a corresponding check can also be made on the electronic energy itself).

4. Convergence acceleration

Any iterative scheme for solving (5) would converge (when it does) at a rate guided by the structure of the algorithm itself. We may call it the 'intrinsic convergence characteristic' which is found to be very good for the present method. However, the determination of MCSCF wavefunction requires solution of the CI equation (4) and equation (5) as coupled tasks and convergence is found to be very slow with many of the available techniques for solution of (5). Normally, one resorts to sophisticated damping and interpolation techniques to achieve fast and smooth convergence. One wonders whether the adoption of a 'convergence-aid' of this kind could improve the efficiency of our algorithm further. In what follows we present our findings in this respect.

4.1 An interpolation scheme

An interpolation method based on direct inversion in the iterative subspace (DIS) (Pulay 1980, 1982; Csazar and Pulay 1984) has been shown to enhance the rate of approach to self-consistency in one-configuration SCF calculation, the enhancement being particularly noticeable in the quadratic region. Surprisingly, the method, in spite of the promise it holds out, has never been applied in the context of MCSCF calculation. To adapt the DIS procedure to the framework of our algorithm for solving the orbital equation we have to define first a suitable error vector (e_i) at the i th stage of the iterative process. This vector should be a faithful measure of the 'distance' of the MCSCF orbitals

in the i th iteration from the converged solution. The most natural choice would be to take

$$e_i = \varepsilon_i - \varepsilon_i^\dagger = [T_i^\dagger V_i - V_i^\dagger T_i], \quad (15)$$

Since $\varepsilon_i^\dagger = \varepsilon_i$ at convergence, $e_i \rightarrow 0$ (null vector) as iterations proceed to self-consistency. One can safely assume that the changes in the MCSCF orbitals in two consecutive iterations ($i, i+1$) are *small in the quadratic* region. Therefore, the error vector e_i at the i th iterative stage may be taken to be a linear function of the parameters $\{\beta_{pq}\}$ characterizing the trial wavefunction. One can then search for a linear interpolation of the consecutive parameter vectors as follows:

$$\tilde{\beta} = \sum_i a_i \beta^i. \quad (16)$$

The linear interpolation in (16) should be such that the corresponding linear interpolation of the consecutive error vectors ($e_1, e_2, \dots, e_i, \dots$) approach the null vector as closely as possible i.e.,

$$\tilde{e} = \sum_i a_i e_i = 0 \quad (17)$$

Obviously, (17) can be satisfied only in the least squares sense. Thus minimizing the norm of \tilde{e} with a suitable defined metric subject to the constraint

$$\sum_i a_i = 1, \quad (18)$$

we arrive at the following set of linear simultaneous equations to be solved for the determination of the interpolation coefficients $\{a_i\}$

$$\begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} & 1 \\ b_{21} & b_{22} & \dots & b_{2n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \quad (19)$$

where λ is the Lagrangian multiplier that takes care of the constraint (18), $b_{ij} = \text{Tr}(e_i e_j^\dagger)$. Once the coefficients $\{a_i\}$ are known we can determine the interpolated orbitals T (or any other parameter vector suitably chosen) as follows:

$$\tilde{T}_{n+1} = \sum_{i=1}^n a_i T_i.$$

In practice, however, such an interpolation is not entirely suitable as it does not preserve the orthonormality of T . Instead, it is better to perform the interpolation on the MCSCF operator (matrix) V itself as follows

$$\tilde{V}_{n+1} = \sum_{i=1}^n a_i V_i.$$

A similar scheme has been successfully adopted in one-configuration work (Pulay 1980).

4.2 A damping technique

Since its introduction by Hartree, the use of damping of some kind or the other has been quite common in SCF or MCSCF calculations. What we have adopted in the course of our calculations is essentially a kind of Hartree damping applied to the MCSCF operator V directly. Let V_{i-1} and V_i be the MCSCF operator (matrix) in any two consecutive iterations. The updating of the orbitals (coefficient matrix T) is then done after replacing V_i by \tilde{V}_i where

$$\tilde{V}_i = V_{i-1}(1 - \lambda) + \lambda V_i \quad (0 < \lambda \leq 1).$$

The conventional undamped iteration scheme corresponds to the choice $\lambda = 1$ which should be restored as soon as the iterations move into the quadratic region. Otherwise, the descent rate becomes too slow. To judge whether damping should be withdrawn or continued, one should ideally calculate $(\partial^2 E_i / \partial \lambda^2)$ at that particular iterative stage. If $(\partial^2 E_i / \partial \lambda^2) > 0$ damping is continued, and otherwise withdrawn.

5. Results and discussion

5.1 Relative performance of the convergence acceleration scheme (damping versus DIIS)

The analysis has been carried out with H_2CO and HNO molecules in the $1,3n\pi^*$ states. The trial wavefunction ($\tilde{\psi}$) is of the form

$$\tilde{\psi} = \frac{1}{\sqrt{2}} [|\phi_1 \bar{\phi}_1 \dots \phi_i \bar{\phi}_j \dots \phi_n \bar{\phi}_n| \pm |\phi_1 \bar{\phi}_1 \dots \phi_j \bar{\phi}_i \dots \phi_n \bar{\phi}_n|]$$

so that only the orbital forms are to be varied, CI coefficients being symmetry determined. The molecules are taken in their respective ground state equilibrium geometries as calculated under the standard INDO/2 approximations. The MCSCF calculations, too, are carried out at the same level of approximation (viz INDO/2).

Table 1 displays convergence profiles for H_2CO in the $1n\pi^*$ states for (i) unaided iterations (ii) damped iterations (iii) iterations with interpolation done under the DIIS procedure. In each case, the iterations were started with the ground state vectors and the convergence criterion was set equal to 10^{-5} a.u. in energy. The damping factor λ was chosen to be equal to 0.75 and was introduced right at the beginning. The damping was withdrawn when $\Delta E = |E_{n+1} - E_n| \leq 0.0002$ a.u. subject to at least 10 iterations being carried out. These parameters were determined through a number of exploratory calculations and were later found to be more or less optimal for different systems. This saved additional computational labour required to compute $(\partial^2 E_i / \partial \lambda^2)$ to determine the point of withdrawal of damping. The DIIS procedure was switched on only after 10 iterations. A perusal of table 1 clearly indicates that although convergence could be achieved in all the cases, the fastest convergence was achieved with damped iterations. The DIIS procedure is found to be ineffective so far as acceleration of convergence is concerned. Moreover, the DIIS procedure suffers from a severe ill-conditioning problem and could not be implemented without special measures being adopted to eliminate the 'ill conditioning' of the b -matrix in (19) (viz the error vector of the largest norm was eliminated, if unsuccessful with that, the next one was eliminated, and so on till the ill conditioning disappeared). This result is a little curious in view of the reported efficacy

Table 1. MCSCF convergence profiles in (i) unaided (ii) damped (iii) interpolated (by DIIS) iterations (convergence criterion $\epsilon = 10^{-5}$ a.u.). The system is H_2CO in $^1n\pi^*$ state.

Number of iterations	Electronic energy (a.u.)		
	Without damping	With damping	With DIIS*
0	-42-885845	-42-885845	-42-885845
1	-42-899554	-42-899554	-42-899554
2	-42-908499	-42-908499	-42-908499
3	-42-915320	-42-917523	-42-915326
4	-42-920630	-42-925188	-42-920630
5	-42-924665	-42-929959	-42-924665
6	-42-927694	-42-932496	-42-927694
7	-42-929915	-42-933646	-42-929915
8	-42-931517	-42-934046	-42-931517
9	-42-932653	-42-934246	-42-932653
10	-42-933447	-42-934809	-42-933447
11	-42-933996	-42-935008	-42-933996
12	-42-934372	-42-935111	-42-934372
13	-42-934628	converges at 14th iteration	-42-934519
14	-42-934800		-42-934665
15	-42-934916		-42-934777
	converges at 21st iteration		converges at 21st iteration

* DIIS procedure switched on after 10 iterations.

of the DIIS procedure (Pulay 1980) in SCF calculations with the traditional repeated diagonalization scheme. It may be that a different choice error vector would be more appropriate for our method. The same conclusion could also be drawn from the calculation on the $^3n\pi^*$ state of H_2CO .

In table 2 convergence parameters for the damped and undamped iterations in MCSCF calculations performed on the $^1,^3n\pi^*$ states of H-N-O are displayed. The damping parameters are just the same as used in the case of H_2CO . The case of HNO is particularly interesting as it provides us with a typical example of a slowly converging

Table 2. Convergence parameters obtained in MCSCF calculations on the $^1,^3n\pi^*$ states of HNO with and without damping.

Molecule	Electronic state	Number of iterations required for convergence	
		With damping	Without damping
HNO	$^1n\pi^*$	29	44
	$^3n\pi^*$	28	38

Damping parameter $\lambda = 0.75$ and convergence criterion in energy $\epsilon = 10^{-5}$ a.u.

MCSCF iterations within the framework of our method. A perusal of table 2 clearly indicates the efficacy of the simple Hartree type of damping in MCSCF calculations. It also proves that the damping parameters chosen by us are system independent and are more or less optimal although minor variations from system to system can not be ruled out.

5.2 Properties of HNO in $^{1,3}n\pi^*$ states

In table 3 we present the fully optimized geometrical parameters of HNO in the ground and $^{1,3}n\pi^*$ states. Experimental or *ab initio* theoretical data have been included in the table wherever such data are available. Our computed results are in fairly good agreement with these results. One interesting point is the predicted opening of the H-N-O angle in the $^3n\pi^*$ state and the narrowing of it in the $^1n\pi^*$ state with respect to the corresponding angle in the ground state. Similarly, increase in the computed N-O bond length is slightly larger in the $^1n\pi^*$ state than in the corresponding triplet.

In table 4 the computed net charges on different atoms of HNO in the ground and $^{1,3}n\pi^*$ states at the respective equilibrium geometry are displayed. The calculated net dipole moments are also given. There appears to be a significant increase in the dipole moment following $n \rightarrow \pi^*$ transition. Experimental data, however, is not available for confirming this prediction. Studies in photochemical reactions of this molecule and related species at the INDO/2-MCSCF level are under way.

Table 3. Equilibrium geometrical parameters in the ground and $^{1,3}n\pi^*$ states of HNO as computed by MCSCF-INDO/2 calculations.

Molecule	Electronic state	Computed geometrical parameters		
		rN-H(Å)	rN-O(Å)	H-N-O (Degrees)
HNO	$^1A'$ (Ground)	1.090 [1.063] ^a	1.190 [1.212] ^a	111.2 [108.6] ^a
	$^1A''$ ($^1n\pi^*$)	1.075 [1.06] ^a	1.218 [1.21] ^a	107.1 [108.5] ^a
	$^3A''$ ($^3n\pi^*$)	1.072 [1.03] ^b	1.205 [1.24] ^b	115.3 [116.7] ^b

Bracketted quantities refer to experimental or *ab initio* theoretical results.

^a Experimental results (Dalbey 1958); ^b Theoretical results (Bruna and Marian 1979; Momura 1980).

Table 4. Computed net charges on different atoms in the ground and $^{1,3}n\pi^*$ states of HNO. Calculated dipole moments in these states are also included (calculations refer to equilibrium geometry in each state).

Molecule	State	Computed net charges on different atoms			Dipole moment (Debye)
		H	N	O	
HNO	$^1A'$	0.0169	0.0970	-0.1138	1.487
	$^1A''$	0.1147	-0.0496	-0.0651	1.823
	$^3A''$	0.1124	-0.0323	-0.0801	1.806

6. Conclusions

The INDO/2-MCSCF calculations appear to be quite useful for studying molecular geometry in excited electronic states. The algorithm used by us for solving the MCSCF equation is fast. The speed of convergence can, moreover, be significantly improved by using a simple damping technique. The DIIS procedure when coupled to our algorithm proved to be ineffective.

Acknowledgements

The authors wish to thank the DST, New Delhi for a generous research grant. One of us (SPB) wishes to thank Dr. Gabor Pongr, Lorand Eötvös University, Budapest, Hungary, for a discussion on the DIIS procedure.

References

- Adnan S S Z, Bhattacharyya S P and Mukherjee D 1978 *Int. J. Quantum Chem.* **14** 289
Bhattacharyya S P and Mukherjee D 1979 *Chem. Phys. Lett.* **66** 511
Bhattacharyya S P and Mukherjee D 1981 *Int. J. Quantum Chem.* **20** 1165
Bhattacharyya S P, Mukherjee D, Bhattacharyya D M and Khan P 1982 *Int. J. Quantum Chem.* **22** 753
Bruna P J and Marian C M 1979 *Chem. Phys. Lett.* **67** 109
Clementi E and Veillard A 1967 *Theor. Chim. Acta* **7** 133
Császár P and Pulay P 1984 *J. Mol. Struct.* **114** 31
Dalbey F W 1958 *Can. J. Phys.* **36** 1336
Das K K, Bhattacharyya S P and Mukherjee D 1986 *Int. J. Quantum Chem.* (to be published)
Das K K, Mukherjee D, Bhattacharyya S P, Bhattacharyya D M and Khan P 1984 *Int. J. Quantum Chem.* **25** 809
Das G and Wahl A C 1967 *J. Chem. Phys.* **47** 2934
Hartree D R and Seirles B 1939 *Philos. Trans. R. Soc.* **A238** 229
McWeeny R 1955 *Proc. R. Soc. London* **A232** 114
McWeeny R 1956 *Proc. R. Soc. London* **A235** 496
McWeeny R 1960 *Rev. Mod. Phys.* **32** 335
McWeeny R and Newbould L M 1980 *Mol. Phys.* **39** 353
Mukherjee D 1978 *Proc. Indian Acad. Sci.* **A87** 37
Momura O 1980 *Int. J. Quantum Chem.* **18** 143
Olsen J and Yeager D L 1983 in *Advances in chemical physics* (New York: John Wiley and Sons) vol. 54, pp. 1-176, and references therein
Pulay P 1980 *Chem. Phys. Lett.* **73** 393
Pulay P 1982 *J. Comput. Chem.* **3** 556