Theor Chem Acc (1997) 98:1–4

Theoretical
Chemistry Accounts

Letter

Analytical first derivatives of the energy for small CI expansions

S. Patchkovskii, W. Thiel

Organisch-chemisches Institut, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

Received: 24 April 1997 / Accepted: 10 June 1997

Abstract. Using the Z-vector formalism the analytical gradient of the energy in small CI expansions is derived and implemented for semiempirical MNDO-type methods. The computation time is shown to scale as $O(N^3)$ with the size of the system, with the memory requirements growing as $O(N^2)$. The evaluation of the analytical gradient is significantly faster than the underlying SCF and CI calculations, so that routine full geometry optimizations at the semiempirical CI level become possible for large systems.

Key words: MNDO – Configuration interaction – Analytical derivatives – Z-vector

1 Introduction

Semiempirical methods such as MNDO [1], AM1 [2], PM3 [3], and MNDO/d [4, 5] account for dynamic correlation effects in an average manner by suitable integral approximations and parametrization. Semiempirical SCF calculations therefore often provide realistic results, particularly for ground-state organic molecules [6–9]. However, an SCF description with a single determinant is no longer adequate when there are two or more interacting configurations that are degenerate or nearly degenerate (static correlation). This may occur, e.g. in reactions involving orbital crossings or in electronically excited states [10, 11]. In these cases, a limited configuration interaction (CI) treatment is mandatory even at the semiempirical level in order to ensure a qualitatively correct description. The number of active orbitals ($N_{\mathrm{CI\,MO}}$) and the number of configurations (N_{conf}) are quite small in such semiempirical CI calculations. The most typical application encountered in studies of organic reactions involves just two active orbitals (HOMO and LUMO) and three configurations.

Explorations of potential surfaces are greatly facilitated if the gradient of the energy with respect to the nuclear coordinates can be evaluated efficiently. For semiempirical closed-shell SCF wavefunctions, the computational effort for the gradient scales as $O(N^2)$ with the number of basis functions (N), both in the numerical and the analytical evaluation [12, 13], compared with an $O(N^3)$ scaling for the SCF energy. On the other hand, the work for numerical CI gradients scales as $O(N^3 N_{\mathrm{at}})$ for a molecule with N_{at} atoms, and the only published semiempirical implementation of analytical CI gradients [14] exhibits the same formal scaling. Hence, gradient evaluations and geometry optimizations at the semiempirical CI level are currently prohibitively expensive for large systems.

Analytical gradients have long been available for ab initio CI wavefunctions [15–23] (see also Refs. [24–27] for reviews). In ab initio applications, the number of CI configurations is typically very large in order to recover a significant portion of the total correlation energy. In practice, this dictates the use of direct AO-based techniques in ab initio CI programs. Semiempirical CI calculations, on the other hand, mainly address static correlation effects and thus normally include only a small number of configurations (see above). This makes the use of MO-based techniques efficient and allows significant simplifications in the computation of analytical CI gradients.

The present paper reports the derivation and implementation of the analytical gradient for small CI expansions. We adopt the restriction that it must be possible to accommodate all relevant arrays in memory, including the two-particle density matrix in the basis of the active orbitals $[O(N^4_{\mathrm{CI\,MO}})]$ as well as the CI matrix and other matrices of $O(N^2_{\mathrm{conf}})$. Under this restriction, calculations up to $N_{\mathrm{CI\,MO}} = 10$ and $N_{\mathrm{conf}} = 100$ can be done easily on current hardware which should be sufficient for most semiempirical applications. Combined with the Z-vector technique [22] this approach allows an $O(N^3)$ implementation of the analytical CI gradient which shares many features with our recent implementation of the analytical half-electron gradient [28]. Section 2 provides the necessary theoretical derivations, as

---

Correspondence to: W. Thiel

far as they are different from our previous work [28], and Sect. 3 reports on the performance obtained with the present implementation.

2 Theory

In the general use, the matrix elements of the single-reference CI Hamiltonian can be written as [29]:

$$H^{\rm CI}_{lm} = E_{\rm ref}\delta_{lm} + \sum_{pqrs}^{\rm CI\,MOs} f^{lm}_{pqrs}(pq|rs) + \sum_{r}^{\rm CI\,MOs} e^{lm}_{r}\epsilon_r \ , \tag{1}$$

where $E_{\rm ref}$ is the energy of the reference state, $(pq|rs)$ are two-electron integrals in the MO basis, ϵ_r are orbital energies, and f^{lm}_{pqrs} and e^{lm}_r are coefficients determined by the selection of the reference state and the CI configurations. For small CI active spaces, the number of non-zero coefficients f^{lm}_{pqrs} and e^{lm}_r is rather small, so that their explicit evaluation and storage does not present a computational bottleneck. Detailed prescriptions for the determination of these coefficients are available in the literature [29–31].

The reference state in semiempirical CI is usually obtained from a closed-shell SCF calculation, or from a quasi closed-shell SCF computation with fractional occupation numbers [31]. In both cases, geometrical derivatives of the corresponding SCF energy $E_{\rm ref}$ ($E_{\rm R}$ in the notation of Ref. [28]) are well known [12, 13], and need not be discussed any further.

CI state energies ($\mathbf{E}^{\rm CI}$) and vectors ($\mathbf{A}$) are solutions of the matrix equation:

$$\mathbf{H}^{\rm CI}\mathbf{A} = \mathbf{A}\,\mathbf{E}^{\rm CI} \ , \tag{2}$$

subject to the orthonormality conditions on the state vectors:

$$\mathbf{A}^{\dagger}\mathbf{A} = \mathbf{1} \ . \tag{3}$$

Once the CI state vector $\mathbf{A}_\Lambda$ for the state Λ is available, the state energy is given by:

$$E^{\rm CI}_{\Lambda} = \mathbf{A}^{\dagger}_{\Lambda}\mathbf{H}^{\rm CI}\mathbf{A}_{\Lambda} \ . \tag{4}$$

Due to the variational nature of the CI states, the derivative of Eq. (4) with respect to an atomic coordinate τ^a contains no derivatives of the state vector [15]:

$$\frac{\partial E^{\rm CI}_{\Lambda}}{\partial \tau^a} = \mathbf{A}^{\dagger}_{\Lambda}\frac{\partial \mathbf{H}^{\rm CI}}{\partial \tau^a}\mathbf{A}_{\Lambda} \ . \tag{5}$$

Substitution of Eq. (1) for the CI matrix elements into Eq. (5) results in [16]:

$$\frac{\partial E^{\rm CI}_{\Lambda}}{\partial \tau^a} = \frac{\partial E_{\rm ref}}{\partial \tau^a} + \frac{\partial E_{\Lambda}}{\partial \tau^a} \ , \tag{6}$$

$$\frac{\partial E_{\Lambda}}{\partial \tau^a} = \sum_{pqrs}^{\rm CI\,MOs}\Gamma^{\Lambda}_{pqrs}\frac{\partial}{\partial \tau^a}(pq|rs) + \sum_{r}^{\rm CI\,MOs}\Delta^{\Lambda}_{r}\frac{\delta\epsilon_r}{\delta\tau^a} \ , \tag{7}$$

where the two-particle density Γ^{Λ} and the Lagrangian multipliers Δ^{Λ}_r corresponding to the state Λ are given by:

$$\Gamma^{\Lambda}_{pqrs} = \sum_{lm} A_{l\Lambda}A_{m\Lambda}f^{lm}_{pqrs} \ , \tag{8}$$

$$\Delta^{\Lambda}_{r} = \sum_{lm} A_{l\Lambda}A_{m\Lambda}e^{lm}_{r} \ . \tag{9}$$

Equation (7) contains derivatives of the two-electron integrals in the MO basis which involve the first-order orbital coefficients x_{ij} [27]:

$$\frac{\partial}{\partial \tau^a}(pq|rs) = (pq|rs)^a + \sum_i \{x_{ip}(iq|rs) + x_{iq}(pi|rs) + x_{ir}(pq|is) + x_{is}(pq|ri)\} \ . \tag{10}$$

The static part $(pq|rs)^a$ of these derivatives can be computed from the corresponding quantities $(\mu\nu|\lambda\sigma)^a$ in the AO basis and the LCAO coefficients $C_{\mu i}$. In MNDO approximation and standard notation [28]:

$$(pq|rs)^a = \sum_{B\neq A}\sum_{\mu\leq\nu}^{A}\sum_{\lambda\leq\sigma}^{B}(\mu\nu|\lambda\sigma)^a\left\{O^{pq}_{\mu\nu}O^{rs}_{\lambda\sigma} + O^{rs}_{\mu\nu}O^{pq}_{\lambda\sigma}\right\} \ , \tag{11}$$

$$O^{ij}_{\mu\nu} = C_{\mu i}C_{\nu j} + (1-\delta_{\mu\nu})C_{\nu i}C_{\mu j} \ . \tag{12}$$

Substitution of Eqs. (10–12) into Eq. (7) allows separation of the derivative expression into static and response parts:

$$\frac{\partial E_{\Lambda}}{\partial \tau^a} = \left(\frac{\partial E_{\Lambda}}{\partial \tau^a}\right)\Bigg|_{\rm static} + \left(\frac{\partial E_{\Lambda}}{\partial \tau^a}\right)\Bigg|_{\rm response} \ . \tag{13}$$

The static part is given by:

$$\left(\frac{\partial E_{\Lambda}}{\partial \tau^a}\right)\Bigg|_{\rm static} = \sum_{pqrs}^{\rm CI\,MOs}\Gamma^{\Lambda}_{pqrs}\,(pq|rs)^a \ , \tag{14}$$

or, equivalently, by:

$$\left(\frac{\partial E_{\Lambda}}{\partial \tau^a}\right)\Bigg|_{\rm static} = \sum_{\mu\nu\lambda\sigma}\Gamma^{\lambda}_{\mu\nu\lambda\sigma}\,(\mu\nu|\lambda\sigma)^a \ , \tag{15}$$

where the two-particle density matrix in the AO basis is defined as:

$$\Gamma^{\Lambda}_{\mu\nu\lambda\sigma} = \sum_{pqrs}^{\rm CI\,MOs}\Gamma^{\Lambda}_{pqrs}C_{\mu p}C_{\nu q}C_{\lambda r}C_{\sigma s} \ . \tag{16}$$

Both Eqs. (14) and (15) can be evaluated in an integral-driven manner, avoiding storage of either the static derivatives of two-electron MO integrals [O($N^4_{\rm CI\,MO}N_{\rm at}$), Eq. (14)] or of the two-particle density matrix in the AO basis [O(N^2) in MNDO approximation, Eq. (15)]. The MO-based approach (Eq. 14) is slightly more efficient for small CI-active spaces, and is used in the present work.

The response part of the derivative is given by:

$$\left(\frac{\partial E_{\Lambda}}{\partial \tau^a}\right)\Bigg|_{\rm response} = \sum_i\sum_p^{\rm CI\,MOs}x_{ip}q_{ip} + \sum_p^{\rm CI\,MOs}\Delta^{\Lambda}_{p}\frac{\partial\epsilon_p}{\partial\tau^a} \ , \tag{17}$$

where:

$$q_{ip} = 4 \sum_{qrs}^{\mathrm{CI\,MOs}} \Gamma^{\Lambda}_{pqrs}\,(iq|rs) \quad . \tag{18}$$

Employing the symmetry of the first-order orbital coefficients x_{ij} [28], Eq. (17) can be rewritten as:

$$\left(\frac{\partial E_\Lambda}{\partial \tau^a}\right)\Bigg|_{\mathrm{response}} = \sum_{i\leq j}^{\mathrm{active\ set}} \mathbf{Q}_{ij}\mathbf{X}^{\mathbf{a}}_{ij} = \mathbf{Q}^{\dagger}\mathbf{X}^{\mathbf{a}} \quad , \tag{19}$$

where:

$$\mathbf{Q}_{pp} = \Delta^{\Lambda}_{p} \qquad \text{if } p \in \text{CI MOs} \quad , \tag{20}$$

$$\mathbf{Q}_{ip} = q_{ip} \qquad \text{if } i \notin \text{CI MOs}, p \in \text{CI MOs, and } p > i \quad , \tag{21}$$

$$\mathbf{Q}_{pi} = -q_{ip} \qquad \text{if } i \notin \text{CI MOs}, p \in \text{CI MOs, and } p < i \quad , \tag{22}$$

$$\mathbf{Q}_{pr} = q_{pr} - q_{rp} \qquad \text{if } p, r \in \text{CI MOs and } p < r \quad . \tag{23}$$

The elements of the vector $\mathbf{X^a}$ are [28]:

$$\mathbf{X}^{\mathbf{a}}_{ij} = x_{ij} \quad \text{if } i < j \quad , \tag{24}$$

$$\mathbf{X}^{\mathbf{a}}_{ii} = \frac{\partial \epsilon_i}{\partial \tau^a} \quad . \tag{25}$$

Since the response part of the derivative is linear in $\mathbf{X^a}$ [Eq. 19], the Z-vector technique [22] can be applied to improve the performance of the gradient computation. An efficient symmetric formulation of the Z-vector method suitable for semiempirical applications has been discussed in detail previously [28] and is directly applicable to the CI gradients.

Implementation issues for the response part of the CI gradient are essentially the same as in the half-electron case [28], with one important distinction. Whereas the number of redundant CPHF variables in half-electron computations is small (typically zero or one), there are generally $\mathrm{O}(N_{\mathrm{CI\,MO}}N)$ redundant variables in the CI gradient expression. Hence, the MO-based transformation of the redundant parts [28] scales as $\mathrm{O}(N_{\mathrm{CI\,MO}}N^4)$ for the CI gradient and becomes inefficient. The AO-based procedure [28] remains efficient with a formal $\mathrm{O}(N^3)$ scaling and is therefore used in the CI case.

Analytical CI gradients have been implemented in the MNDO94 program [32]. Since the rate-determining part of the code (i.e. the solution of the Z-vector CPHF equations) is essentially the same as in our previous implementation [13, 28], the program structure and the detailed distribution of the computational effort need not be discussed here.

3 Results and discussion

In order to assess the overall performance of the present implementation relative to numerical techniques and to the previous analytical implementation [14], Cartesian energy gradients were computed for a series of D_{2h} condensed aromatic hydrocarbons in the highest singlet state of the minimal 3×3 CI involving HOMO and LUMO of the closed-shell reference configuration. The convergence criterion for the SCF energy was 10^{-7} eV (ca. 4×10^{-9} au), while the CPHF equations were required to converge within 10^{-6} au. Although less stringent convergence criteria are usually sufficient for SCF and half-electron gradients, the CI results appear to be more sensitive to the quality of the orbitals. With these convergence criteria, numerical and analytical gradients agree within 0.1 kcal mol^{-1} Å^{-1} (ca. 10^{-4} au), and need not be shown here.

Table 1. Comparison of computation times[a] (in seconds)

System[b]	Orbitals	SCF + CI	Numerical gradient	Analytical gradient, Ref. [14][c]	Analytical gradient, present work
$C_{10}H_8$	48	0.76	49.69	15.80	0.59
$C_{32}H_{14}$	142	15.53	2458.19	1227.99	8.18
$C_{66}H_{20}$	284	165.80	37833.31	24049.04	62.33
$C_{112}H_{26}$	474	1094.41	[d]	[d]	288.81
$C_{170}H_{32}$	712	7778.39	[d]	[d]	1122.22

[a] On an SGI Indigo^2 with a 100 MHz R4000 CPU and 112 MB of main memory

[b] At an idealized D_{2h} geometry with $R_{\mathrm{CC}} = 1.4066$ Å and $R_{\mathrm{CH}} = 1.0904$ Å

[c] As implemented in MOPAC7 [33]

[d] Not computed, estimated execution time too large

As can be seen from the execution times given in Table 1, gradient evaluation requires a small fraction of the time necessary for an energy evaluation (SCF + CI), except for very small molecules, where the overhead of setting up the analytical computation is significant. The analytical CI gradient becomes progressively more efficient compared to SCF + CI for larger systems, which reflects a more localized memory access pattern in the CI gradient program.

The speedup over previously available techniques is substantial, e.g. more than two orders of magnitude for $C_{66}H_{20}$. Combined with memory requirements comparable to those of the SCF procedure [13, 28], this allows routine geometry optimizations and vibrational analysis of large molecules at the CI level. For example, the MNDO vibrational analysis of a C_{60} dimer (480 orbitals, C_{2h} symmetry) at the 3×3 CI level (via finite differences of analytical CI gradients) took about 17 h on an SGI PowerChallenge using one 90 MHz R8000 CPU. With previously available techniques, such computations would not have been practical.

4 Conclusion

The Z-vector formulation of the analytical gradient of the energy in small CI expansions provides a substantial increase in the computational efficiency over both numerical and previous analytical [14] implementations. The computation time formally scales as $\mathrm{O}(N^3)$ with the size of the system, and is typically a small fraction of the SCF computation time, with memory requirements growing as $\mathrm{O}(N^2)$.

Acknowledgements. This work was supported by the Schweizerischer Nationalfonds.

References

1. Dewar MJS, Thiel W (1977) J Am Chem Soc 99:4899, 4907
2. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) J Am Chem Soc 107:3902
3. Stewart JJP (1989) J Comp Chem 10:209, 221
4. Thiel W, Voityuk AA (1992) Theor Chim Acta 81:391, 93:315
5. Thiel W, Voityuk AA (1996) J Phys Chem 100:616
6. Dewar MJS (1985) J Phys Chem 89:2145
7. Thiel W (1988) Tetrahedron 44:7393
8. Stewart JJP (1990) In: Lipkowitz JB, Boyd DB (eds) Reviews in computational chemistry, vol 1. VCH, New York, pp 45–81
9. Thiel W (1996) Adv Chem Phys 93:703
10. Dewar MJS, Olivella S, Stewart JJP (1986) J Am Chem Soc 108:5771
11. Schweig A, Thiel W (1981) J Am Chem Soc 103:1425
12. Dewar MJS, Yamaguchi Y (1978) Comput Chem 2:25
13. Patchkovskii S, Thiel W (1996) J Comp Chem 17:1318
14. Dewar MJS, Liotard DA (1990) J Mol Struct (Theochem) 206:123
15. Tachibana A, Yamashita K, Yamabe T, Fukui K (1978) Chem Phys Lett 59:255
16. Brooks BR, Laidig WD, Saxe P, Goddard JD, Yamaguchi Y, Schaefer III HF (1980) J Chem Phys 72:4652
17. Krishnan R, Schlegel HB, Pople JA (1980) J Chem Phys 72:4654
18. Brooks BR, Laidig WD, Saxe P, Handy NC, Schaefer III HF (1980) Phys Scr 21:312
19. Dupuis M (1981) J Chem Phys 74:5758
20. Osamura Y, Yamaguchi Y, Schaefer III HF (1981) J Chem Phys 75:2919
21. Osamura Y, Yamaguchi Y, Schaefer III HF (1982) J Chem Phys 77:383
22. Handy NC, Schaefer III HF (1984) J Chem Phys 81:5031
23. Rice JE, Amos RD, Handy NC, Lee TJ, Schaefer III HF (1986) J Chem Phys 85:963
24. Amos RD (1987) Adv Chem Phys 67:99
25. Pulay P (1987) Adv Chem Phys 69:241
26. Jørgensen P, Simons J (eds) (1985) Geometrical derivatives of energy surfaces and molecular properties. Reidel, Dordrecht
27. Yamaguchi Y, Osamura Y, Goddard JD, Schaefer III HF (1994) A new dimension to quantum chemistry: analytic derivative methods in ab initio molecular electronic structure theory. Oxford University Press, New York
28. Patchkovskii S, Thiel W (1996) Theor Chim Acta 93:87
29. Shavitt I (1977) In: Schaefer III HF (ed) Methods of electronic structure theory. Plenum, New York, pp 189–275
30. Roos BO, Siegbahn PEM (1977) In: Schaefer III HF (ed) Methods of electronic structure theory. Plenum, New York, pp 277–318
31. Armstrong DR, Fortune R, Perkins PG, Stewart JJP (1972) J Chem Soc Faraday Trans II 68:1839
32. Thiel W (1994) Program MNDO94, version 4.0
33. Stewart JJP (1993) MOPAC7 Rev. 1: a public-domain general molecular orbital package Quantum Chemistry Program Exchange (QCPE 455 7.0); see QCPE Bull 13:42