Theor Chem Acc (2007) 117: 69–72
DOI 10.1007/s00214-006-0132-2

REGULAR ARTICLE

Liat Pele · Brina Brauer · R. Benny Gerber

Acceleration of correlation-corrected vibrational self-consistent field calculation times for large polyatomic molecules

Received: 23 February 2006 / Accepted: 31 March 2006 / Published online: 30 June 2006
© Springer-Verlag 2006

Abstract Acceleration of the correlation-corrected Vibrational self-consistent field (CC-VSCF) method for anharmonic calculations of vibrational states of polyatomic molecules is described. The acceleration assumes pairwise additive interactions between different normal modes, and employs orthogonality of the single-mode vibrational wavefunctions. This greatly reduces the effort in computing correlation effects between different vibrational modes, which is treated by second order perturbation theory in CC-VSCF. The acceleration can improve the scaling of the overall computational effort from N^6 to N^4, where N is the number of vibrational modes. Sample calculation times, using semi-empirical potential surfaces (PM3), are given for a series of glycine peptides. Large computational acceleration, and significant reduction of the scaling of the effort with system size, is found and discussed.

Keywords Vibrational anharmonicity · Vibrational states · Normal modes · Vibrational self-consistent field · Correlation-corrected vibrational self-consistent field

1 Introduction

The understanding of vibrational spectral measurements is frequently aided by the use of comparison with theoretical calculations [1,2]. Until recently, such theoretical calculations were primarily performed by using harmonic approximation [3–6] and applying an empirical scaling factor in order to compensate for anharmonicity [7–9]. The advantage of harmonic frequency calculations is that they can be routinely performed on molecules of a variety of sizes, depending upon the method used to determine the molecular potential surface. For example, harmonic frequencies of molecules containing as many as ca. 50 atoms can be calculated with ab initio methods and protein frequencies may be obtained through the use of force fields [10].

Anharmonic frequency calculations, however, are required for many spectroscopic problems, since the frequencies that are obtained are generally more accurate. With the increasing availability of gas-phase measurements, such as a gas-phase protein infrared spectrum [11], accuracy in calculations is imperative. Additionally, since the scale factors used with harmonic calculations are empirical, they do not present information about the anharmonicity of the underlying potential energy surface and the nature of the anharmonic interactions. Furthermore, these scaling factors vary between different types of calculations and basis sets. Scale factors have also been found to vary within one vibrational spectrum; i.e. the scale factor for the higher frequencies differs from that of the lower frequencies [8,9].

Computation of anharmonic frequencies is complex, primarily due to non-separability of the Hamiltonian. However, during the past decade, routines for the calculation of anharmonic frequencies have become increasingly available. The vibrational self-consistent field VSCF methods [12,13] for anharmonic vibrational frequency calculations, described in this article, are based on assuming separability of the N-vibrational mode wavefunction. Together with the variational principle, this yields single mode equations that are solved self-consistently (see Sect. 2). Physically, each vibrational mode is considered in the mean field of the other vibrational modes. In the present applications, the potential is truncated by including only single and double mode terms. Accuracy can be improved by including correlation effects [using second-order perturbation theory, correlation-corrected-VSCF(CC-VSCF)], and in some cases by considering degeneracy (DPT2-VSCF [14]). The VSCF routines discussed here have been included in the GAMESS [15,16] and MOLPRO [17,18] program suites. In both

L. Pele · B. Brauer · R. B. Gerber (✉)
Department of Physical Chemistry and
Fritz Haber Research Center, The Hebrew University of Jerusalem,
Jerusalem 91904, Israel
E-mail: benny@fh.huji.ac.il
Tel.: +1-972-2-6585732
Fax: +1-972-2-6513742

R. B. Gerber
Department of Chemistry, University of California,
Irvine, CA 92697, USA

packages, grid techniques are used which permit a choice of the type of potential surface that is employed in the calculation (e.g. ab initio).

Other methods of anharmonic frequency calculations are available. For example, Barone [19] has developed an anharmonic vibrational calculation method that uses a second-order perturbation treatment based on quadratic, cubic and semidiagonal quartic force constants. His method is included in the GAUSSIAN [20] program suite. For further information about anharmonic vibrational frequency calculations the reader is referred to Ref. [21].

The present work is primarily focused on the CC-VSCF routine as implemented in GAMESS, although the acceleration is suitable for other computational applications. With the original implementation of CC-VSCF, anharmonic frequency calculations could easily be performed on molecules with up to 30 atoms, using semi-empirical potential surfaces. However,the use of VSCF and CC-VSCF for calculation of protein frequencies is desirable, and although they have been used for a small protein calculation [22], this remains computationally expensive. Additionally, this protein calculation was performed using an empirical force field, and ideally, a more accurate potential surface would be desired. Reduction in the computational effort could help in making anharmonic calculations for large molecules routine. Until now, efforts in reduction of calculation time have primarily focused on decreasing the computational effort used in describing the potential surface [18,23–25]. However, even with a fast potential surface calculation, such as with semi-empirical calculations, solution of the CC-VSCF and DPT2-VSCF equations can cause a computational bottleneck. The present work describes a shortcut in the CC-VSCF and DPT2-VSCF calculations that can accelerate the computation by an order of magnitude.

Relevant aspects of the VSCF theory and their application to the GAMESS program are discussed in Sect. 2, results are presented in Sect. 3, and conclusions are given in Sect. 4.

2 Theory

VSCF equations and their computational solution

As described earlier, VSCF equations for anharmonic vibrational frequency calculations are based on assuming separability of the N-vibrational mode wavefunction:

$$\Psi_n(Q_1, \ldots, Q_N) = \prod_{j=1}^{N} \Psi_j^{(n)}(Q_j) \tag{1}$$

where $Q_1, Q_2, \ldots, Q_N$ are mass-weighted normal coordinates (for a more detailed description of VSCF, see Ref. [21] and References therein). Equation (1) is inserted into the vibrational Schrödinger equation:

$$\left[-\frac{1}{2} \sum_{j=1}^{N} \frac{\partial^2}{\partial Q_j^2} + V(Q_1, \ldots, Q_N) \right] \Psi_n(Q_1, \ldots, Q_N) = E_n \Psi_n(Q_1, \ldots, Q_N) \tag{2}$$

where V is the potential function of the system, n is the state number, and N is the number of vibrational modes. Using the variational principle leads to VSCF Eqs. (3) and (4), containing the effective potential $\bar{V}_j^{(n)}(Q_j)$, which are solved self-consistently:

$$\left[-\frac{1}{2} \frac{\partial^2}{\partial Q_j^2} + \bar{V}_j^{(n)}(Q_j) \right] \Psi_j^{(n)}(Q_j) = \varepsilon_j^{(n)} \Psi_j^{(n)}(Q_j) \tag{3}$$

$$\overline{V_j^n}(Q_j) = \left\langle \prod_{l \neq j}^{N} \Psi_l^{(n)}(Q_l) \middle| V(Q_1, \ldots, Q_N) \middle| \prod_{l \neq j}^{N} \Psi_l^{(n)}(Q_l) \right\rangle. \tag{4}$$

The potential can be approximated by using one-mode and two-mode terms:

$$V(Q_1, \ldots, Q_N) = \sum_{j=1}^{N} V_j^{\mathrm{diag}}(Q_j) + \sum_{i} \sum_{j>i} W_{ij}^{\mathrm{coup}}(Q_i, Q_j) \tag{5}$$

CC-VSCF is an improvement on VSCF by the use of perturbation theory to include correlation effects; the difference between the VSCF Hamiltonian and the true Hamiltonian is treated as a perturbation [13,26].

$$H = H^{\mathrm{SCF},(n)} + \Delta V(Q_1, \ldots, Q_N) \tag{6}$$

where $H^{\mathrm{SCF},(n)}$ is the Hamiltonian in the VSCF approximation.

In Eq. (6), ΔV is given by:

$$\Delta V(Q_1, \ldots, Q_N) = V(Q_1, \ldots, Q_N) - \sum_{j=1}^{N} \overline{V_j^{(n)}}(Q_j) \tag{7}$$

Correlation effects are included in ΔV, the difference between the correct Hamiltonian and the VSCF one, and the effective potential $\overline{V}_j^{(n)}(Q_j)$for the mode Q_j is given in Eqs. (3) and (4).

Proceeding in analogous fashion to the second order Møller–Plesset method in the electronic structure theory (Ref. [26] and references therein) gives:

$$E_n^{\mathrm{cc}} = E_n^{\mathrm{VSCF}} + \sum_{m \neq n} \frac{\left| \left\langle \prod_{j=1}^{N} \Psi_j^{(n)}(Q_j) | \Delta V | \prod_{j=1}^{N} \Psi_j^{(m)}(Q_j) \right\rangle \right|^2}{E_n^{(o)} - E_m^{(o)}} \tag{8}$$

where E_n^{cc} is the correlation-corrected energy of state n. In the case of the pairwise approximation, the perturbation is

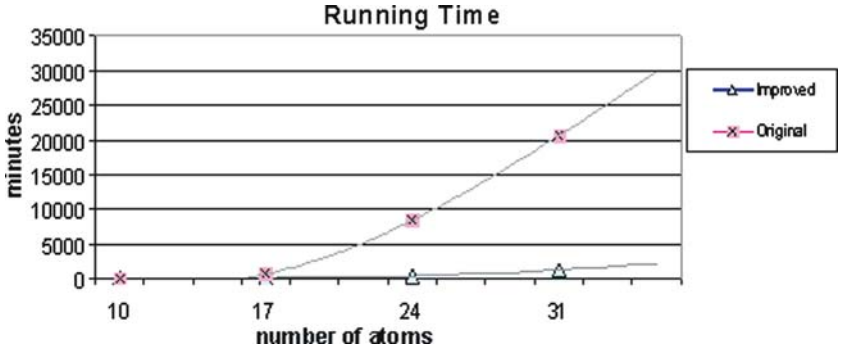

Fig. 1 Running times for VSCF/CC-VSCF in GAMESS as a function of number of atoms in glycine peptides

given by

$$\Delta V(Q_1, \ldots, Q_N) = \sum_{j=1}^{N} V_j^{\mathrm{diag}}(Q_j) + \sum_i \sum_{j>i} W_{ij}^{\mathrm{coup}}(Q_i, Q_j) - \sum_{j=1}^{N} \overline{V_j^{(n)}}(Q_j). \quad (9)$$

Substitution of (9) into the numerator of (8), leads to the cancellation of the diagonal terms, and thus yields:

$$\left\langle \prod_{j=1}^{N} \psi_j^{(n)}(Q_j) \,|\Delta V|\, \prod_{j}^{N} \psi_j^{(m)}(Q_j) \right\rangle = \sum_{j>i} \int \psi_i^{(n)}(Q_i)\psi_j^{(n)}(Q_j) W_{ij}^{\mathrm{coup}}(Q_i, Q_j) \times \psi_i^{(m)}(Q_i)\psi_j^{(m)}(Q_j)\mathrm{d}Q_i\mathrm{d}Q_j \quad (10)$$

where the integration is over the normal modes i and j, and m labels the excited vibrational states ($m \neq n$). In the derivation of (10), we have also used the orthonormality of the single-mode wavefunctions. [Note that strictly speaking, $\psi_i^{(n)}(Q_i)$ and $\psi_i^{(m)}(Q_i)$ are only very nearly orthogonal, since the effective potentials $\overline{V}^{(n)}(Q_i)$ and $\overline{V}^{(m)}(Q_i)$ are not strictly the same However, orthogonality holds in this case to very good accuracy. In fact, wavefunction overlaps between different vibrational states were found to be of the order of 1×10^{-3} in the case of H_2O.] By testing the condition for orthogonality first, time is saved by not calculating wavefunction overlap terms that will be eliminated. The implementation in the GAMESS code is done via a subroutine that is also called from the VCI (first-order degenerate perturbation theory VSCF) routine, where an analogous correction to the Hamiltonian is considered; this adaptation could improve calculational running times in any similar situation.

3 Results

The VSCF code change was tested on a series of glycine peptides, using a PM3 potential surface. A PM3 surface can be calculated quickly with the original version of VSCF, since it is a semi-empirical method, and the remainder of the calculational time is due to the solution of the VSCF equations. In order to study the effect of the change in the VSCF routine, runtimes were checked computationally using the GAMESS runtime subroutines on an Intel Pentium 4 CPU 3.00 GHz computer with 2 GB memory. Additionally, runtime profiling was performed for each run, using a profiler that was developed by us, based on Linux system time calls. The extent of the testing is limited to tetraglycine by the size of the original CC-VSCF calculations. All frequency results were identical in both versions of the code.

In the original version of VSCF, the evaluation of Eq. (10) took 85–95% of the total run time, while in the accelerated version the evaluation of Eq. (10) takes only 0.5–2.5% of the total run time. The bottleneck in the new version is in evaluation to the potential energy grid, rather than in the solution of the VSCF equations. Runtimes for the series of peptides are given in Table I and in Fig. 1. The results indicate that the improvement in running time is generally larger for larger molecules; for glycine the improvement is a factor of 5.9, for triglycine it is 20, and for tetraglycine it is 16.5. The dependence of the runtime on the number of the normal modes was analyzed. The perturbation theory section of the code was originally roughly $O(N^4)$, and with the improve-

Table 1 Runtimes of VSCF/CC-VSCF in GAMESS (non-degenerate)

Molecule	Original (min)	Improved (min)
Glycine (10 atoms)	49.8	8.4
Diglycine (17 atoms)	622.3	74.5
Triglycine (24 atoms)	8,411.5	416.5
Tetraglycine (31 atoms)	20,662.9	1,255.1

Table 2 Runtimes of VSCF/DPT2-VSCF in GAMESS (degenerate)

Molecule	Original (min)	Improved (min)
Glycine (10 atoms)	1,062.5	119.7
Diglycine (17 atoms)	22,588 (~17 days)	1,476.4 (~1 day)

ment it is $O(N^2+(c*N^4))$ where c is a very small constant that could be lowered further to zero. The overall runtime is thus lowered from $O(N^6)$, with the original VSCF code, to practically $O(N^4)$. Use of a different potential surface with fewer potential points, such as VSCF/2MR-QFF[25], would further lower the runtime by a factor of ca. 10. Improvements in running times with degenerate perturbation theory are given in Table 2.

Since the focus of the present work is computational runtimes, accuracy of the frequencies, which of course depends on the quality of the potential, was not of concern here. The PM3 potential employed here is not very accurate. Anharmonic frequency calculations on diglycine using the much superior MP2/DZP potential surface have been carried out by G.M. Chaban (personal communication).

4 Conclusions

For the CC-VSCF method, using pairwise potentials, the present results yield results identical to those of previous codes, but much faster. Basically, the acceleration comes from the consideration of simplifications due to orthogonality relations in the framework of the pairwise mode coupling approximation. This results in major speedups which become increasingly important as the number of vibrational modes (N) increases. Indeed, the computational effort of the CC-VSCF calculation is reduced from a $O(N^6)$ to practically a $O(N^4)$ type of behavior. The practical significance of this result is obvious. The current improvement has been introduced into the GAMESS suite of programs.

Acknowledgments We wish to thank Dr. G. M. Chaban for her helpful comments. The help of Yifat Miller and Dorit Shemesh is gratefully acknowledged. We thank Shai Meir and Tom Rosenfeld for their computational assistance. This research was supported by a grant from the US-Israel Binational Science Foundation (BSF - 2004009).

References

1. Wilson EB, Jr, Decius JC, Cross PC (1955, reprinted 1980) Molecular Vibrations. The theory of infrared and Raman vibrational spectra. Dover, New York.
2. Herzberg G (1945, reprinted 1991) Molecular spectra and molecular structure. vol. II. Infrared and Raman spectra of polyatomic molecules. Krieger Publishing Company, Malabar
3. Gwinn WD (1971) J Chem Phys 55:477–481
4. Boatz JA, Gordon MS (1989) J Phys Chem 93:1819–1826
5. Blum CE**,** Altona C, Oskam A (1977) Mol Phys, 34:557–571
6. Fogarasi G, Pulay P (1984) Ann Rev Phys Chem 35:191–213
7. Scott AP, Radom L (1996) J Phys Chem 100:16502–16513
8. Halls MD, Velkovski J, Schlegel HB (2001) Theor Chem Acc 105:413–421
9. Sinha P, Boesch SE, Gu C, Wheeler RA, Wilson AK (2004) J Phys Chem A 108:9213–9217
10. Brooks B, Karplus M (1983) Proc Natl Acad Sci USA 80:6571–6575
11. Oomens J, Polfer N, Moore DT, van der Meer L, Marshall AG, Eyler JR, Meijer G, von Helden G (2005) Phys Chem Chem Phys 7:1345–1348
12. Chaban GM, Jung JO, Gerber RB (1999) J Chem Phys 111:1823–1829
13. Jung JO, Gerber RB (1996) J Chem Phys 105:10332–10348
14. Matsunaga N, Chaban GM, Gerber RB (2002) J Chem Phys 117:3541–3547
15. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JJ, Koseki S, Matsunaga N, Nguyen KA, Su S, Windus TL, Dupuis M, Montgomery JA (1993) J Comput Chem 14:1347–1363
16. http://www.msg.ameslab.gov/GAMESS/GAMESS.html
17. Werner H-J, Knowels PJ et al (2002) MOLPRO version 2002.6, a package of ab initio programs, Birmingham
18. Rauhut G (2004) J Chem Phys 121:9313–9322
19. Barone V (2005) J Chem Phys 122:014108
20. Frisch MJ et al (2003) Gaussian 03, revision, A.1, Gaussian Inc., Pittsburgh,
21. Gerber RB, Chaban GM, Brauer B, Miller Y (2005) In:Dykstra CE, Frenking G, Kim K, Suseria G, (eds) Theory and applications of computational chemistry: the first 40 years, Chapter 9, pp 165–193
22. Roitberg AE, Gerber RB, Elber R, Ratner MA (1995) Science 268:1319–1322
23. Benoit DM (2004) J Chem Phys 120:562–573.
24. Yagi K, Taketsugu T, Hirao K, Gordon MS (2000) J Chem Phys 113:1005–1017
25. Yagi K, Hirao K, Taketsugu T, Schmidt MW, Gordon MS (2004) J Chem Phys 121:1383–1389
26. Norris LS, Ratner MA, Roitberg AE, Gerber RB (1996) J Chem Phys 105:11261–11267