Theor Chim Acta (1990) 78: 175–187

Theoretica
Chimica Acta

# A comparison of variational and non-variational internally contracted multiconfiguration-reference configuration interaction calculations

**Hans-Joachim Werner[1] and Peter J. Knowles[2]**
[1] Fakultät für Chemie, Universität Bielefeld, D-4800 Bielefeld, Federal Republic of Germany
[2] School of Chemistry and Molecular Sciences, University of Sussex, Falmer,
Brighton, BN1 9QJ, UK

Received June 14, 1990/Accepted June 15, 1990

**Summary.** The internally contracted multiconfiguration-reference configuration interaction (CMRCI) method and several non-variational variants of this method (averaged coupled pair approximation (ACPF), quasidegenerate variational perturbation theory (QD-VPT), linearized coupled pair many electron theory (LCPMET)) have been employed to compute potential energy functions and other properties for a number of diatomic molecules ($F_2$, $O_2$, $N_2$, CN, CO) using large basis sets and full valence CASSCF reference wavefunctions. In most cases the variational CMRCI wavefunctions yield more accurate spectroscopic constants than any of the employed non-variational methods. Several basis sets are compared for the $N_2$ molecule. It is found that atomic natural orbital (ANO) contractions led to significant errors in the computed $r_e$, $\omega_e$, and $D_e$ values.

**Key words:** Variational – Non-variational – Configuration interaction – Diatomic molecules

## 1. Introduction

Today, the multiconfiguration-reference configuration interaction method [1–11] (MRCI) is considered to be one of the most accurate approximations used in quantum chemistry. The main advantage of this method is that it is not restricted to specific choices of the zeroth order wavefunction and can therefore be used to compute potential energy functions and other molecular properties for ground and excited electronic states and for any nuclear geometry. A considerable number of benchmark calculations [12–21] have demonstrated that second-order MRCI wavefunctions, which include all single and double excitations out of full valence complete active space self-consistent field (CASSCF) reference wavefunctions, yield potential energy functions which more closely parallel the full CI (FCI) functions than any other approximate method tested. However, these benchmark calculations were performed with quite small orbital basis sets and therefore give no information about the absolute accuracy which can be achieved. Even though many previous applications of the MRCI method have shown that excellent agreement with experimental results can be obtained when large AO (atomic orbital) basis sets are used, it is difficult to draw general conclusions, since a large variety of basis sets and many specific choices of reference wavefunctions were employed. A systematic investigation of the effect of different selections of reference configurations was often prevented by the fact that the computational effort of the MRCI method strongly increases with the number of reference configurations. In many cases this made the simultaneous use of large basis sets and large reference spaces impossible.

This situation has changed with the development of the internally contracted MRCI method [8–11]. In this method the total wavefunction is a linear combination of the same set of configuration state functions (CSFs) as in an uncontracted calculation, but the number of variational parameters is strongly reduced by contracting subsets of the CSFs with fixed coefficients. For a given basis set, the number of variational parameters only depends on the number of correlated orbitals in the reference wavefunction; it is independent of the number of reference configurations which are constructed from this orbital set. It has been demonstrated [9, 10] that in most cases this contraction leads to only very small errors. Some results of benchmark calculations [10] are shown in Fig. 1. It is seen that the energy differences between internally contracted MRCI and uncontracted MRCI calculations are typically 30% of the difference between the uncontracted MRCI and the FCI. The absolute deviations are of the order of only 1 mH.

The price one has to pay for the reduction of the number of variational parameters is the complicated structure of the contracted $N$-electron basis functions. This complicates the calculation of the Hamiltonian matrix or, in a direct CI procedure, the evaluation of the residual vector $\boldsymbol{g} = \boldsymbol{H} \cdot \boldsymbol{c}$. However, we have recently developed very efficient new techniques which now enable us to perform internally contracted MRCI calculations with very large reference spaces and very large basis sets. Some of our previous calculations [10] included more than 3000 reference configurations with basis sets of more than 150 orbitals and were equivalent to uncontracted MRCI calculations with about $10^8$ configuration state functions.

In the present paper we first review some aspects of the internally contracted MRCI method and then present applications to the molecules $F_2$, $O_2$, $N_2$, CN, and CO. The aim of these calculations was to systematically compare the quality of calculated spectroscopic quantities for open and closed-shell molecules with single, double, and triple bonds. We have tried to make these calculations as well defined and equally based as possible. In all cases, full-valence CASSCF reference wavefunctions without any further configuration selection were used. In the special case of $F_2$ we also performed calculations which included the $2\pi_u$ and $2\pi_g$ orbitals into the active space of the reference function. We have employed the largest standard basis sets available ($13s8p3d2f1g$ contracted to $[6s5p3d2f1g]$). In the case of $N_2$ some additional calculations have been performed with other $s, p$ basis sets.

A disadvantage of the MRCI method is that is it not size consistent. Therefore, various non-variational modifications have been proposed, which are at least approximately size consistent. Most of these procedures are considered as approximations to a general MR-coupled cluster (CC) theory [22]. The simplest ones are the linearized CC method (MR-LCC) of Laidig and Bartlett [23, 24], the averaged coupled pair approximation (ACPF) of Gdanitz and Ahlrichs [25], the variational perturbation theory [26] (VPT, equivalent to the linear coupled electron pair approximation, CEPA(0)), and the quasidegenerate variational perturbation theory [27] (QDVPT) of Cave and Davidson. A somewhat more

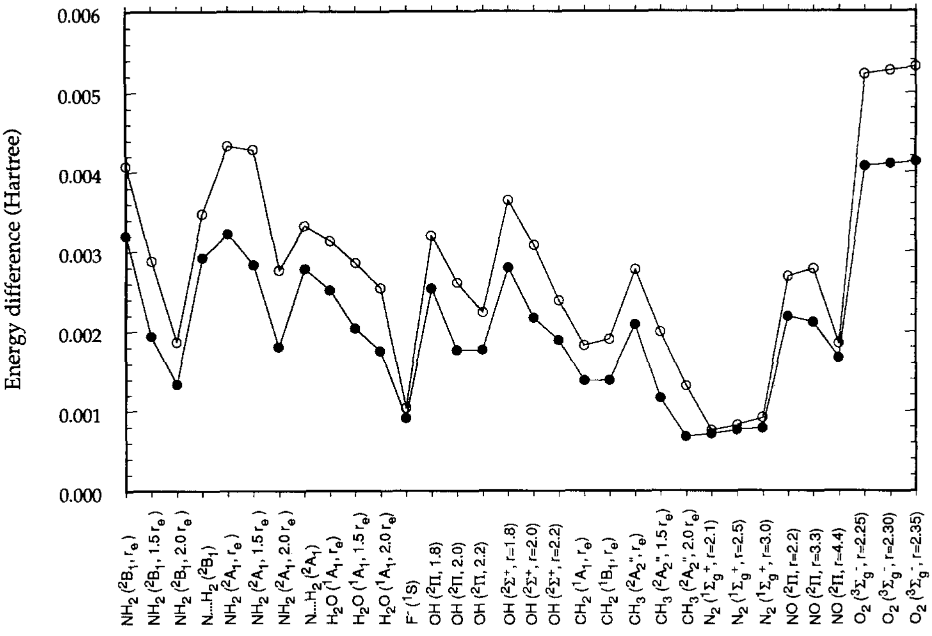

**Fig. 1.** Comparison of total energies of internally contracted (○) and uncontracted (●) MRCI calculations. The energy differences between these and the full CI results are shown

elaborate approximate MR-CC method has recently been proposed by Hoffmann and Simons [28]. This method is closely related to the internally contracted CI method, and our techniques should therefore be well suited for an efficient implementation of this method.

So far, little systematic information is available about the quality of results obtained with the above mentioned non-variational methods. In addition to the variational MRCI calculations, we have therefore performed CEPA(0), ACPF, and QDVPT calculations using the same basis sets and CASSCF reference wavefunctions. Perhaps somewhat surprisingly, we find that all non-variational methods yield substantially worse agreement with experimental data than the variational MRCI method.

The present article is organized as follows: in Sect. 2 we review the relevant aspects of the internally contracted CI method. In Sect. 3 we discuss various non-variational modifications. In Sect. 4 the results of the calculations for the electronic ground states of $F_2$, $O_2$, $N_2$, CO and CN are presented. A short summary is given in Sect. 5.

## 2. Uncontracted and internally contracted *N*-electron basis sets

Configuration state functions (CSFs) are defined by specifying the occupation number of each orbital (0, 1 or 2) and the spin coupling of the singly occupied orbitals. The number of possible spin couplings depends on the total spin quantum number and the number of singly occupied orbitals. A specific choice of occupation numbers irrespective of the spin coupling is called "orbital configuration". We denote orbital configurations by capital letters $I, J, \ldots$ and spin couplings by Greek letters $\mu, \nu, \ldots$.

It is convenient to classify the configurations according to the number of electrons in the "internal" or "external" orbitals. Internal orbitals are those that are occupied in at least one reference configuration, and are denoted by indices $i, j, k, l$. The complementary space of external orbitals is denoted by the indices $a, b, c, d$ and arbitrary orbitals by the letters $r, s, t, u$. We distinguish internal CSFs $\Phi_{I\mu}$, singly external CSFs $\Phi_{S\mu}^a$, and doubly external CSFs $\Phi_{P\mu}^{ab}$. Here, $I$ denotes an orbital configuration with $N$ electrons in the internal orbital space, $S$ denotes an orbital configuration with $N-1$ electrons in the internal space, and $P$ an orbital configuration with $N-2$ electrons in internal orbitals. Configurations with more than two electrons in external orbitals are neglected. The $N-2$ electron orbital configuration space $P$ is generated by performing all possible annihilations of two electrons from the set of reference configurations $R$. The $N-1$ electron orbital configurations $S$ are then obtained from the set $P$ by adding in all possible ways one electron to unfilled internal orbitals. Similarly, the internal configuration set $I$ is generated by adding one internal electron to the configurations $S$. In general, the reference configurations $R$ form a subset of the internal configurations $I$. For complete active space reference functions, the internal configuration space is identical to the reference configuration space.

Using these definitions an uncontracted MRCI wavefunction can be written as

$$\Psi = \sum_{I\mu} c^{I\mu}\Phi_{I\mu} + \sum_{S\mu}\sum_{a} c_a^{S\mu}\Phi_{S\mu}^{a} + \sum_{P\mu}\sum_{ab} C_{ab}^{P\mu}\Phi_{P\mu}^{ab}. \tag{1}$$

The disadvantage of this ansatz is that the number of internal orbital configurations $S$ and $P$ depends strongly on the number of reference configurations. In direct MRCI procedures, the computational effort depends most crucially on the number of $N-2$ electron functions $(P\mu)$. If this number is denoted $N_P$, the number of operations per iteration is proportional to $N_P N^4 + N_P^x N^3$, where $N$ is the number of external orbitals and $1 \leqslant x \leqslant 2$. This is the reason for the fact that it is difficult to perform uncontracted MRCI calculations with large reference configuration spaces and large basis sets.

For the uncontracted MRCI wavefunction, the configuration space was generated by two electron excitations from each individual reference configuration. In contrast, in the internally contracted MRCI, the configurations are generated by applying two-electron excitation operators $\hat{E}_{ri,sj}$ to the complete reference function

$$\Psi_0 = \sum_{R\mu} a^{R\mu}\Phi_{R\mu}. \tag{2}$$

In the following, we use the short-hand notation $|0\rangle \equiv \Psi_0$. The internally contracted configurations are defined as

$$\Phi_{ij}^{rs} = \hat{E}_{ri,sj}|0\rangle = (\hat{E}_{ri}\hat{E}_{sj} - \delta_{is}\hat{E}_{rj})|0\rangle \quad (i \geqslant j, \text{ all } r, s), \tag{3}$$

where $\hat{E}_{ri}$ and $\hat{E}_{ri,sj}$ are one- and two-electron excitation operators, respectively:

$$\hat{E}_{ri} = \eta_r^{\alpha\dagger}\eta_i^{\alpha} + \eta_r^{\beta\dagger}\eta_i^{\beta}, \tag{4}$$

$$\hat{E}_{ri,sj} = \hat{E}_{ri}\hat{E}_{sj} - \delta_{is}\hat{E}_{rj}. \tag{5}$$

For a given internal orbital space, the number of contracted configurations $\Phi^{rs}_{ij}$ is independent of the number of reference configurations. However, the functions $\Phi^{rs}_{ij}$ are not orthonormal and may be linearly dependent. In order to eliminate redundant functions and to orthonormalize the remaining set, we need their overlap matrix. Clearly, functions with different numbers of external orbitals are automatically orthogonal. For the doubly external configurations $\Phi^{ab}_{ij}$, the overlap is given by

$$\langle \Phi^{ab}_{ij} | \Phi^{cd}_{kl} \rangle = \delta_{ac}\delta_{bd}\langle 0|\hat{E}_{ik,jl}|0\rangle + \delta_{ad}\delta_{bc}\langle 0|\hat{E}_{il,jk}|0\rangle. \tag{6}$$

It is convenient to redefine the doubly external functions such that the external orbitals are either singlet ($p = 1$) or triplet ($p = -1$) coupled

$$\Phi_{ijp} = \tfrac{1}{2}(\Phi^{ab}_{ij} + p\Phi^{ba}_{ij}) \quad (i \geqslant j, a \geqslant b, p = \pm 1), \tag{7}$$

since then the internal and external parts in Eq. (6) factorize, and singlet and triplet pairs are orthogonal:

$$\langle \Phi^{ab}_{ijp} | \Phi^{cd}_{klq} \rangle = \tfrac{1}{2}\delta_{pq}(\delta_{ac}\delta_{bd} + \delta_{ad}\delta_{bc})\langle 0|\hat{E}_{ik,jl} + p\hat{E}_{il,jk}|0\rangle. \tag{8}$$

In order to orthonormalize the functions $\Phi^{ab}_{ijp}$ only the internal overlap matrices $S^{(p)}_{ij,kl}$ ($p = \pm 1$) are needed:

$$S^{(p)}_{ij,kl} = \langle 0|\hat{E}_{ik,jl}|0\rangle + p\langle 0|\hat{E}_{il,jk}|0\rangle. \tag{9}$$

These are given by the elements of the second-order reduced density matrix of the reference wavefunction. In principle, internally contracted singly external configurations $\Phi^{ka}_{ij}$ and internal configurations $\Phi^{kl}_{ij}$ can be defined in a similar way. This has in fact been done in the first implementation of the internally contracted MRCI method by Werner and Reinsch [8]. However, in order to orthonormalize these functions, the third- and fourth-order density matrices, respectively, are needed, and the corresponding overlap matrices can become very big. Moreover, it turns out that the calculation of the coupling coefficients in the direct CI procedure becomes quite difficult if internally contracted internal and singly external functions are used as a basis. On the other hand, the contraction of the doubly external configuration state functions presents no particular problems and leads to large savings in the cost of the calculation. We have therefore adopted a hybrid scheme [10] in which the uncontracted configuration spaces $\Phi_{I\mu}$ and $\Phi^{a}_{S\mu}$ and the internally contracted doubly external functions $\Phi^{ab}_{ijp}$ are used. The wavefunction then takes the form

$$\Psi = \sum_{I\mu} c^{I\mu}\Phi_{I\mu} + \sum_{S\mu}\sum_{a} c^{S\mu}_{a}\Phi^{a}_{S\mu} + \sum_{i\geqslant j}\sum_{p}\sum_{ab} C^{ijp}_{ab}\Phi^{ab}_{ijp} \tag{10}$$

with

$$C^{ijp}_{ab} = pC^{ijp}_{ba}. \tag{11}$$

As discussed in detail in [10], the residual vector $\boldsymbol{g} = \boldsymbol{H} \cdot \boldsymbol{c}$ is most easily calculated directly in the basis of the non-orthogonal functions $\Phi^{ab}_{ijp}$ (redundant functions are eliminated). In order to find a new expansion vector $\boldsymbol{c}$, the residual vector is transformed to an orthonormal basis $\Phi^{ab}_{Dp}$, defined by

$$\Phi^{ab}_{Dp} = \sum_{i \geqslant j} T^{(p)}_{D,ij} \Phi^{ab}_{ijp}, \tag{12}$$

where

$$T^{(p)} = (S^{(p)})^{-1/2}. \tag{13}$$

## 3. Non-variational MRCI approximations

In order to achieve approximate size extensivity, Gdanitz and Ahlrichs suggested the minimization of the energy functional [25]

$$E = E_0 + \frac{\langle \Psi_0 + \Psi_c | H - E_0 | \Psi_0 + \Psi_c \rangle}{1 + g_a \langle \Psi_a | \Psi_a \rangle + g_e \langle \Psi_e | \Psi_e \rangle}, \tag{14}$$

where $\Psi_0$ is the normalized reference wavefunction (Eq. (2)), $\Psi_c$ is the correlation function

$$\Psi_c = \Psi - \langle \Psi | \Psi_0 \rangle \Psi_0 = \Psi_a + \Psi_e, \tag{15}$$

and $\Psi_a$ and $\Psi_e$ are the mutually orthogonal internal and external parts of $\Psi_c$. Equation (15) implies $\langle \Psi_c | \Psi_0 \rangle = 0$. For our internally contracted MRCI wavefunction the internal and external correlation functions take the form

$$\Psi_a = \sum_{I\mu} c^{I\mu} \Phi_{I\mu} - \langle \Psi | \Psi_0 \rangle \Psi_0, \tag{16}$$

$$\Psi_e = \sum_{S\mu} \sum_{a} c^{S\mu}_{a} \Phi^{a}_{S\mu} + \sum_{i \geqslant j} \sum_{p} \sum_{ab} C^{ijp}_{ab} \Phi^{ab}_{ijp}. \tag{17}$$

In Eq. (14), $g_a$ and $g_e$ are numerical factors, which in the ACPF method of Gdanitz and Ahlrichs are chosen to be [25]

$$g_a = 1, \qquad g_e = 2/n, \tag{18}$$

where $n$ is the number of electrons correlated. Different choices of $g_a$ and $g_e$ lead to other approximations, e.g.,

$$g_a = 1, \qquad g_e = 1 \tag{19}$$

yields the variational MRCI solution,

$$g_a = 0, \qquad g_e = 0 \tag{20}$$

corresponds to variational perturbation theory [26] (VPT) or CEPA(0), and

$$g_a = 1, \qquad g_e = 0 \tag{21}$$

is equivalent to Cave and Davidson's quasi-degenerate variational perturbation theory [27] (QDVPT). The multireference linear coupled cluster (MR-LCC) method of Laidig and Barlett [23, 24] is similar to VPT and CEPA(0), but does not include the orthogonal complement of the reference function in the internal space. In the early version [23] of MR-LCC the important semi-internal configurations were also omitted.

Varying the energy functional (14) with respect to the coefficients $c^{I\mu}$, $c_a^{S\mu}$, and $C_{ab}^{ijp}$ yields the variational conditions

$$0 = \langle \Psi_{I\mu} | H - \varepsilon_a | \Psi \rangle - a^{I\mu}(E - E_0)(1 - g_a)\langle \Psi_0 | \Psi \rangle, \tag{22}$$

$$0 = \langle \Psi_{S\mu}^{a} | H - \varepsilon_e | \Psi \rangle, \tag{23}$$

$$0 = \langle \Psi_{ijp}^{ab} | H - \varepsilon_e | \Psi \rangle, \tag{24}$$

where

$$\varepsilon_a = E_0 + g_a(E - E_0), \tag{25}$$

$$\varepsilon_e = E_0 + g_e(E - E_0). \tag{26}$$

and $a^{I\mu}$ are the coefficients of the reference wavefunction. In the direct CI procedure, these equations are solved iteratively. Since for the non-variational methods Eqs. (22)–(26) represent a system of non-linear (ACPF, QDVPT) or linear (CEPA(0)) equations, some modifications of the usual Davidson procedure [29] are necessary.

While in most previous approximate MRCC methods uncontracted $N$-electron basis sets were used, similar excitation operators and internally contracted configurations as those defined in Eq. (3) were employed in the approximate linear MRCC method of Hoffmann and Simons [28]. Compared to other linear MRCC approximations, Hoffmann and Simons did not neglect some terms which involve higher order excitations as intermediate states. However, this method was implemented like a conventional (non-direct) CI, and therefore allowed only the use of quite small configuration spaces. It appears straightforward to implement their method more efficiently using the same techniques [10, 11] as in our internally contracted CI method.

## 4. Benchmark calculations for diatomics

The availability of an efficient implementation of the internally contracted MRCI method described in Sect. 2 has opened up the possibility of routinely using large CASSCF reference spaces, rather than being forced into the difficult problem of selecting a few appropriate reference configurations. In particular, for nearly all diatomic molecules, it is now feasible to perform CMRCI calculations which are nearly "black box" in nature, using complete valence reference spaces, with large standard basis sets. In this section, we present the results of such calculations for the series of first row diatomics $F_2$, $O_2$, $N_2$, CN, CO, and compare the resulting spectroscopic properties with those derived from experiment. This comparison then acts as a benchmark by which we can estimate the reliability of this computational ansatz in future studies. The series of molecules chosen gives a reasonably comprehensive test, since it spans a range of different bond lengths, number of electrons, and degree of multiple bonding.

For each molecule we have used the largest van Duijneveldt [31] *sp* basis, 13*s*8*p*, with the innermost 8*s*, 4*p* contracted together according to the atomic SCF 1*s* orbital. Polarization and angular correlation effects are represented using Dunning's [32] 3*d*2*f*1*g* primitive gaussian set. For the CI calculations, the reference space consisted of all configurations obtained by distributing the valence electrons amongst those orbitals corresponding to atomic 2*s* and 2*p* orbitals, giving in most cases several hundred configurations. This choice allows for the $2s$ correlation effects which are particularly important in N and C, and is in general sufficiently flexible to account for most significant non-dynamical correlation effects. Amongst the series of molecules studied, the exception is $F_2$, where, in order to describe both neutral and anionic character, it is desirable to include two sets of $p$ orbitals in the active space. Inclusion of six additional orbitals is impractical, and following previous studies [33] we use reference spaces which arise from the $2p$ orbitals together with the $2\pi_u$ and $2\pi_g$ orbitals. In order to generate reference orbitals for the CI treatment, CASSCF calculations with the reference configuration sets were performed with inner shell orbitals optimized but not correlated.

Tables 1–5 give the computed spectroscopic constants for each molecule, derived by fitting the calculated energies at ten bond distances to a polynomial of eighth degree. The errors in $r_e$, $\omega_e$, and $D_e$ relative to the experimental values are shown in Fig. 2. The stability of the spectroscopic constants with respect to variations of the fit was carefully checked. In the case of $F_2$, the $2s2p$ reference results are compared with $2p2\pi_u$ and $2p2\pi_u2\pi_g$ reference results. The agreement of the MRCI results with experiment is seen to be significantly superior than that of all the non-variational methods for the near equilibrium properties ($r_e$, $\omega_e$, etc.). The errors of the CMRCI equilibrium distances vary between 0.001 and 0.003 Å, while the harmonic frequencies are accurate to within 5–10 $\mathrm{cm}^{-1}$. For $D_e$, it is difficult to draw significant conclusions; here size consistency is an important issue, and one expects the non-variational methods to be an improvement over variational MRCI, but this is not the case for $N_2$, CN and CO. The errors in the MRCI dissociation energies amount to about 0.2 eV in all cases.

For the case of $F_2$, despite the known problem of inadequate simultaneous description of F and $F^-$, the $2s2p$ active space results display an accuracy similar

**Table 1.** Calculated[a] and experimental spectroscopic constants[b] for $F_2$ ($X\,^1\Sigma_g^+$)

| | Energy | $r_e$ | $B_e$ | $\alpha_e$ | $\omega_e$ | $\omega_e x_e$ | $D_e$ |
|---|---|---|---|---|---|---|---|
| Exp.[d] | | 1.4119 | 0.890 | 0.0138 | 916.6 | 11.2 | 1.66 |
| MRCI[e] | −199.3116873 | 1.4155 | 0.886 | 0.0119 | 909.75 | 18.01 | 1.46 |
| ACPF[e] | −199.3430536 | 1.4158 | 0.885 | 0.0113 | 924.39 | 17.30 | 1.55 |
| QDVPT[e] | −199.3494482 | 1.4161 | 0.885 | 0.0112 | 927.36 | 17.16 | 1.57 |
| CEPA(0)[e,f] | −199.3510003 | 1.4211 | 0.879 | 0.0116 | 911.45 | 17.70 | |
| MRCI[g] | −199.3226396 | 1.4130 | 0.889 | 0.0109 | 934.43 | 16.57 | 1.65 |
| ACPF[f,g] | −199.3469359 | 1.4167 | 0.884 | 0.0112 | 924.38 | 17.14 | |
| QDVPT[f,g] | −199.3517274 | 1.4176 | 0.883 | 0.0113 | 922.12 | 17.27 | |
| CEPA(0)[f,g] | −199.3526138 | 1.4172 | 0.884 | 0.0123 | 925.41 | 19.59 | |
| MRCI[h] | −199.3316887[i] | | | | | | 1.62 |

[a] For basis set, see text

[b] $B_e$, $\alpha_e$, $\omega_e$, $\omega_e x_e$ in $\mathrm{cm}^{-1}$, $r_e$ in Å, $D_e$ in eV

[c] Energy at $r_e$

[d] [34]

[e] $2s2p$ reference space

[f] Because of intruder state problems, no convergence at large distance

[g] $2p2\pi_u$ reference space

[h] $2p2\pi_u2\pi_g$ reference space, see text

[i] Energy at $r = 1.4120$ Å

**Table 2.** Calculated[a] and experimental spectroscopic constants[b] for $O_2$ ($X\,^3\Sigma_g^-$)

| | Energy | $r_e$ | $B_e$ | $\alpha_e$ | $\omega_e$ | $\omega_e x_e$ | $D_e$ |
|---|---|---|---|---|---|---|---|
| Exp.[d] | | 1.2075 | 1.446 | 0.0159 | 1580.2 | 12.0 | 5.21 |
| MRCI | −150.1400817 | 1.2085 | 1.443 | 0.0159 | 1585.3 | 11.9 | 5.00 |
| ACPF | −150.1650993 | 1.2113 | 1.437 | 0.0158 | 1573.8 | 11.8 | 5.06 |
| QDVPT | −150.1712708 | 1.2121 | 1.435 | 0.0158 | 1569.9 | 11.8 | 5.08 |
| CEPA(0)[e] | −150.1721413 | 1.2130 | 1.433 | 0.0159 | 1564.2 | 11.8 | |

[a] For basis set, see text

[b] $B_e$, $\alpha_e$, $\omega_e$, $\omega_e x_e$ in $cm^{-1}$, $r_e$ in Å, $D_e$ in eV

[c] Energy at $r_e$

[d] [34]

[e] No convergence at large distance

**Table 3.** Calculated[a] and experimental spectroscopic constants[b] for $N_2$ ($X\,^1\Sigma_g^+$)

| | Energy[c] | $r_e$ | $B_e$ | $\alpha_e$ | $\omega_e$ | $\omega_e x_e$ | $D_e$ |
|---|---|---|---|---|---|---|---|
| Exp.[d] | | 1.0977 | 1.998 | 0.0173 | 2358.6 | 14.3 | 9.91 |
| MRCI | −109.3862961 | 1.1008 | 1.987 | 0.0171 | 2350.7 | 13.9 | 9.72 |
| ACPF | −109.3984824 | 1.1016 | 1.984 | 0.0172 | 2343.7 | 14.0 | 9.68 |
| QDVPT | −109.4020138 | 1.1019 | 1.983 | 0.0172 | 2341.4 | 14.0 | 9.66 |
| CEPA(0) | −109.4021260 | 1.1020 | 1.983 | 0.0172 | 2340.8 | 14.0 | 9.66 |

[a] For basis set, see text

[b] $B_e$, $\alpha_e$, $\omega_e$, $\omega_e x_e$ in $cm^{-1}$, $r_e$ in Å, $D_e$ in eV

[c] Energy at $r_e$

[d] [34]

**Table 4.** Calculated[a] and experimental spectroscopic constants[b] for CO ($X\,^1\Sigma^+$)

| | Energy[c] | $r_e$ | $B_e$ | $\alpha_e$ | $\omega_e$ | $\omega_e x_e$ | $D_e$ |
|---|---|---|---|---|---|---|---|
| Exp.[d] | | 1.1283 | 1.931 | 0.0175 | 2169.81 | 13.29 | 11.36 |
| MRCI | −113.1689381 | 1.1319 | 1.919 | 0.0173 | 2164.80 | 13.1 | 11.14 |
| ACPF | −113.1811840 | 1.1330 | 1.915 | 0.0173 | 2156.68 | 13.1 | 11.08 |
| QDVPT | −113.1847670 | 1.1334 | 1.914 | 0.0173 | 2153.91 | 13.1 | 11.06 |
| CEPA(0)[e] | −113.1851947 | 1.1336 | 1.914 | 0.0173 | 2152.41 | 13.1 | |

[a] For basis set, see text

[b] $B_e$, $\alpha_e$, $\omega_e$, $\omega_e x_e$ in $cm^{-1}$, $r_e$ in Å, $D_e$ in eV

[c] Energy at $r_e$

[d] [34]

[e] No convergence at large distances

Table 5. Calculated[a] and experimental spectroscopic constants[b] for CN ($X\,^2\Sigma^+$)

| | Energy[c] | $r_e$ | $B_e$ | $\alpha_e$ | $\omega_e$ | $\omega_e x_e$ | $D_e$ |
|---|---|---|---|---|---|---|---|
| Exp.[d] | | 1.1718 | 1.900 | 0.0174 | 2068.59 | 13.09 | 7.78<br>7.89 |
| MRCI | −92.5782757 | 1.1754 | 1.888 | 0.0173 | 2060.52 | 13.0 | 7.70 |
| ACPF | −92.5878526 | 1.1762 | 1.886 | 0.0173 | 2055.47 | 13.1 | 7.67 |
| QDVPT | −92.5910468 | 1.1765 | 1.885 | 0.0173 | 2053.58 | 13.1 | 7.66 |

[a] For basis set, see text

[b] $B_e$, $\alpha_e$, $\omega_e$, $\omega_e x_e$ in cm$^{-1}$, $r_e$ in Å, $D_e$ in eV

[c] Energy at $r_e$

[d] [34]

to that obtained in the other systems. The inclusion of the $2\pi_u$ orbital ($2p2\pi_u$ active space) apparently improves the $D_e$ and $r_e$ values, but actually leads to a worse, and overestimating, error in $\omega_e$. The good $D_e$ value is partly due to a compensation of basis set shortcomings and an overestimate of the correlation effect. The inclusion of only the $2\pi_u$ orbital in the reference wavefunction improves the calculation at $r_e$, but has no effect at large $r$. For a balanced calculation it is therefore necessary to include excitations into the $2\pi_g$ orbital in the reference wavefunctions as well. In this case it is not possible, however, to use a CASSCF wavefunction as reference in the MRCI. To make the MRCI calculation tractable, we used a reference consisting of all possible configurations in the $2p$ space, plus all single and double excitations out of this space into the $2\pi_u$ and $2\pi_g$ orbitals. This resulted in 278 reference CSFs. The orbitals were taken from a CASSCF with the $2p$, $2\pi_u$, $2\pi_g$ orbitals active (2584 CSFs). The dissociation energy obtained in this case (1.62 eV) is in fact lower than without the $2\pi_g$ orbitals (1.65 eV). The importance of the $2\pi_u$ and $2\pi_g$ orbitals is particularly apparent in the CASSCF dissociation energies. The corresponding $D_e$ values for the $2s2p$, $2p$, $2p2\pi_u$ and $2p2\pi_u2\pi_g$ active spaces are 0.80 eV, 0.74 eV, 1.68 eV, and 1.51 eV, respectively (all values calculated as difference of the energies at 0.706 Å and 20 Å).

In order to investigate whether the contraction of the $s, p$ basis set has a significant effect on the computed spectroscopic constants and the dissociation energy, some additional test calculations were performed for the $N_2$ molecule. In these calculations the orbitals were determined in a CASSCF with only the six MOs (molecular orbitals) which correlate with the atomic $2p$ orbitals active. This has virtually no effect on $r_e$ and $\omega_e$, but gives slightly better total energies and $D_e$ values (at large distance the CASSCF(10) and CASSCF(6) calculations yield identical results). In the MRCI all ten valence electrons were correlated and a full valence CI (10 electrons active) was used as reference function. In Table 6 the results obtained with two different segmented basis sets are compared with two ANO (atomic natural orbital) basis sets [35] and with the new generally contracted valence quadruple (V4Z) and pentuple (V5Z) basis sets of Dunning [32]. In calculations not shown in Table 6, it was found that the larger segmented basis set (SEG-86) yields dissociation energy identical to that of the entirely contracted basis. The stronger contraction in the SEG-65 basis leads to a slight overestimate of $D_e$ and $\omega_e$. The generally-contracted V4Z and V5Z basis sets yield very similar results to the segmented basis. As expected, the total energies

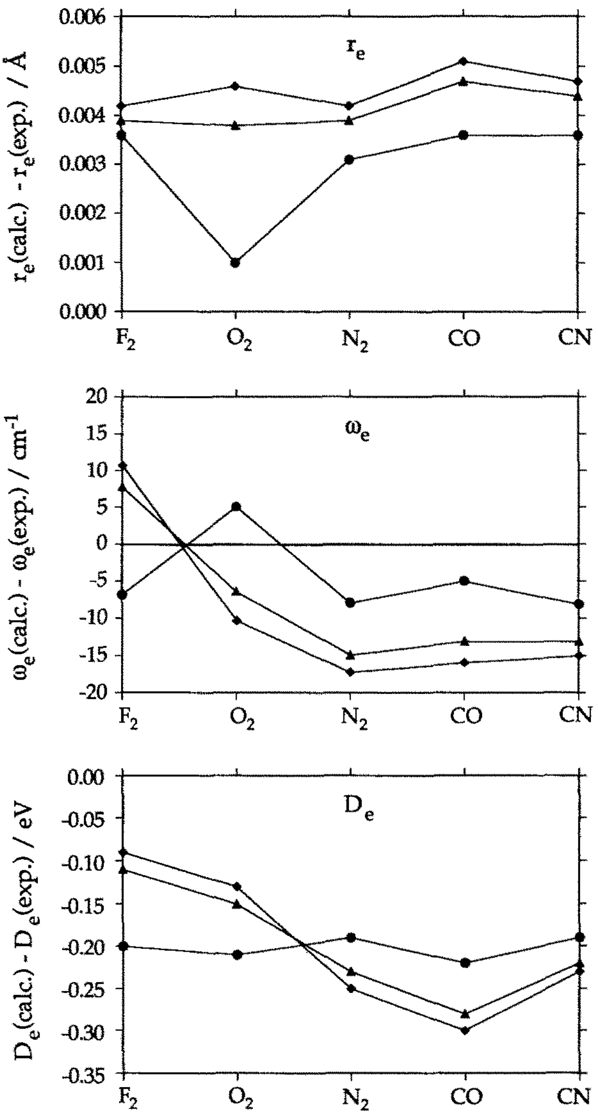

**Fig. 2.** Comparison of computed and experimental spectroscopic constants for various molecules. In all cases the absolute errors relative to the experimental data are plotted. ◆, QDVPT; ▲, ACPF; ●, MRCI

of the V4Z and V5Z bases are lower than that of the SEG-65 set and almost as good as those of the SEG-86 basis. The ANO basis sets also yield very good total energies. However, the spectroscopic constants obtained with these bases show significant errors: For the ANO-54 basis, $r_e$ is 0.0016 Å too long, and $\omega_e$ 7 $\text{cm}^{-1}$ too small. Even more significant is the error in $D_e$. The value for the ANO-54 basis is 0.047 eV (1.08 kcal/mole) two low; for the ANO-65 basis set the error still amounts to 0.48 kcal/mole. Contraction errors of this size are quite unacceptable for calculations of the present accuracy. In a recent MRCI study, Almlöf et al. [36] addressed the problem of calculating the dissociation energy of $N_2$ with "chemical accuracy". In their best calculation, which employed an

**Table 6.** Calculated[a] and spectroscopic constants[b] for $N_2$ ($X\ ^1\Sigma_g^+$) for MRCI with different basis sets

| Basis[c] | $E(R_e)$ | $E(\infty)$ | $r_e$ | $B_e$ | $\alpha_e$ | $\omega_e$ | $\omega_e x_e$ | $D_e$ |
|---|---|---|---|---|---|---|---|---|
| ANO[54] | −109.3907768 | −109.03548622 | 1.1019 | 1.983 | 0.0171 | 2343.8 | 13.9 | 9.682 |
| ANO[65] | −109.3921873 | −109.03542221 | 1.1008 | 1.987 | 0.0171 | 2347.9 | 13.9 | 9.708 |
| V4Z[54] | −109.3897586 | −109.03264645 | 1.1006 | 1.988 | 0.0171 | 2349.6 | 13.9 | 9.718 |
| V5Z[65] | −109.3927441 | −109.03535487 | 1.1004 | 1.989 | 0.0172 | 2350.2 | 13.9 | 9.725 |
| SEG-65 | −109.3865110 | −109.02863835 | 1.1002 | 1.989 | 0.0171 | 2352.5 | 13.9 | 9.738 |
| SEG-86 | −109.3929710 | −109.03545165 | 1.1003 | 1.989 | 0.0172 | 2350.6 | 13.9 | 9.729 |
| Exp. | | | 1.0977 | 1.998 | 0.0173 | 2358.6 | 14.3 | 9.904 |

[a] Spectroscopic constants from polynomial fit of 8th degree to 10 energies (1.6–3.4 au). The orbitals were determined in a CASSCF with the $6p$ electrons active. In the MRCI a full valence CI (10 electrons active) was used as reference function

[b] $B_e$, $\alpha_e$, $\omega_e$, $\omega_e x_c$ in $\text{cm}^{-1}$, $r_e$ in Å, $D_e$ in eV

[c] In all cases uncontracted $3d$, $ef$, $1g$ functions optimized by Dunning are used. The ANO, SEG-65, and SEG-86 $s, p$ bases use van Duijneveldt's $13s$, $8p$ primitive functions. In the segmented [65] basis the innermost $8s$, $4p$ are contracted. In the SEG-86 basis the innermost $6s$, $3p$ are contracted. The latter basis yields a dissociation energy identical to that of the entirely uncontracted $13s$, $8p$ basis. The ANO contraction coefficients are from natural orbits of an SDCI for N($^4S$) with $13s$, $8p$, $2d$, $1f$ functions

ANO-65 basis, the remaining error in $D_e$ was about 2.4 kcal/mole, after correcting for core-valence correlation and the effect of $i$ functions. It appears that a significant fraction of this error is due to the ANO contraction of the basis set. These findings have led us to perform a more detailed study of basis set and correlation effects on the dissociation energy and spectroscopic constants of $N_2$, which will be published elsewhere [37].

## 5. Conclusions

In this paper, we have demonstrated the viability of complete active space multireference CI calculations for diatomics, using CMRCI wavefunctions. The accuracy of computed potentials appears to be rather consistent over a wide range of different systems, and since large reference spaces can be handled rather economically, the difficult and potentially dangerous task of selecting reference configurations is eliminated. For the $N_2$ molecule it was demonstrated that small ANO basis sets lead to quite significant errors in the spectroscopic constants and $D_e$, while the energy optimized generally-contracted basis sets of Dunning of the same size yield excellent agreement with essentially uncontracted basis sets.

## References

1. Buenker RJ, Peyerimhoff SD (1974) Theor Chim Acta 35:33
2. Buenker RJ, Peyerimhoff SD, Butscher W (1978) Mol Phys 35:771
3. Siegbahn PEM (1980) Int J Quantum Chem 18:1229

4. Liu B, Yoshimine M (1981) J Chem Phys 74:612
5. Lischka H, Shepard R, Brown FB, Shavitt I (1981) Int J Quantum Chem Symp 15:91
6. Saxe P, Fox DJ, Schaefer III HF, Handy NC (1982) J. Chem Phys 77:5584
7. Saunders VR, van Lenthe JH (1983) Mol Phys 48:923
8. Werner H-J, Reinsch EA (1982) J Chem Phys 76:3144
9. Werner H-J, Reinsch EA (1984) In: Dykstra CE (ed) Advanced theories and computational approaches to the electronic structure of molecules. Reidel, Dordrecht
10. Werner H-J, Knowles PJ (1988) J Chem Phys 89:5803
11. Knowles PJ, Werner H-J (1988) Chem Phys Lett 145:514
12. Bauschlicher CW, Taylor PR, Handy NC, Knowles PJ (1986) J Chem Phys 85:1469
13. Bauschlicher CW, Taylor PR (1986) J Chem Phys 85:2779
14. Bauschlicher CW, Taylor PR (1986) J Chem Phys 85:6510
15. Bauschlicher CW, Langhoff SR, Taylor PR (1987) J Chem Phys 87:387
16. Bauschlicher CW, Taylor PR (1987) J Chem Phys 86:2844
17. Bauschlicher CW, Taylor PR (1987) J Chem Phys 86:5600
18. Bauschlicher CW, Langhoff SR (1987) J Chem Phys 87:4665
19. Bauschlicher CW, Langhoff SR (1987) J Chem Phys 87:5595
20. Langhoff SR, Bauschlicher CW, Taylor PR (1987) J Chem Phys 86:6992
21. Bauschlicher CW, Taylor PR (1987) J Chem Phys 86:5600
22. Jeziorski B, Paldus J (1989) J Chem Phys 90:2714
23. Laidig WD, Bartlett RJ (1984) Chem Phys Lett 104:424
24. Laidig WD, Saxe P, Bartlett RJ (1987) J Chem Phys 86:887
25. Gdanitz RJ, Ahlrichs R (1988) Chem Phys Lett 143:413
26. Cave RJ, Davidson ER (1988) J Chem Phys 88:5770
27. Cave RJ, Davidson ER (1988) J Chem Phys 89:6798
28. Hoffmann MR, Simons J (1988) J Chem Phys 88:993
29. Davidson ER (1975) J Comput Phys 17:87
30. Malmqvist PA, Roos BO (1989) Chem Phys Lett 155:189
31. van Duijneveldt FB (1971) IBM Research Report RJ945
32. Dunning TH (1989) J Chem Phys 90:1007
33. Langhoff SR, Bauschlicher CW, Taylor PR (1987) Chem Phys Lett 135:543
34. Huber KP, Herzberg G (1979) Constants of diatomic molecules. Van Nostrand Reinhold
35. Almlöf J, Taylor PR (1987) J Chem Phys 86:4070
36. Almlöf J, Deleeuw BJ, Taylor PR, Bauschlicher CW, Siegbahn PEM (1989) Int J Quantum Chem Sympos 23:345
37. Werner H-J, Knowles PJ (1990) J Chem Phys (submitted for publication)