

# Semidirect MP2 Gradient Evaluation on Workstation Computers: The MPGRAD Program

Frank Haase<sup>1,2</sup> and Reinhart Ahlrichs<sup>1\*</sup>

Universität Karlsruhe, <sup>1</sup>Lehrstuhl für Theoretische Chemie, Institut für Physikalische Chemie, D-7500 Karlsruhe, Germany, and <sup>2</sup>Max-Planck-Gesellschaft, Arbeitsgruppe Quantenchemie an der Humboldt Universität Berlin, D-1199 Berlin, Germany

Received 12 October 1992; accepted 27 February 1993

A semidirect implementation of the closed-shell MP2 gradient for efficient use on workstation computers is presented. The approach is based on the algorithm proposed by Frisch and coworkers but includes several modifications to reduce disk storage requirements and exploits nonabelian point group symmetry. The performance of the resulting program MPGRAD (BIOSYM Corp., San Diego, CA) is demonstrated in applications to the molecules  $[\text{AlSi}(\text{CH}_3)_3]_4$  and ferrocene. The largest calculation involved 492 basis functions and was carried out on IBM RS/6000 workstations with memory sizes of 32 and 128 Mb. The ratio of CPU to wallclock time exceeds 90% in all typical applications. © 1993 by John Wiley & Sons, Inc.

## INTRODUCTION

The direct self-consistent field (SCF) method has proven its efficiency<sup>1,2</sup> and has become standard especially for applications on workstation computers. Because a direct SCF calculation requires the storage of only a few matrices (of size  $N^2$  with  $N$  as the number of basis functions), applications are basically CPU bound.

Attention has now focused on the extension of direct approaches to correlated methods. Because the majority of correlated methods are formulated in the molecular orbital (MO) basis, they require an integral transformation from the atomic orbital basis (AO) to the MO basis and hence need to process and store various types of MO integrals. Memory and disk storage requirements are therefore much larger than in SCF calculations, even for a second-order Møller–Plesset perturbation theory (MP2) treatment,<sup>3</sup> the simplest approach to treat electron correlation effects in a systematic manner. Because the computational requirements of correlated methods are at least  $N^5$ , correlated calculations on large molecules still represent a challenge. Direct approaches to correlated methods were initially suggested by Taylor<sup>4</sup> and later implemented for MP2 energy calculations by Head-Gordon et al.,<sup>5</sup> Sæbø et al.,<sup>6</sup> and in the TURBOMOLE program package.<sup>7</sup> Recently, direct MP2 energy calculations employing efficient integral prescreening techniques together with simple schemes to exploit symmetry of molecular systems involving more than 800 contracted basis functions were reported.<sup>8</sup>

The MP2 gradient expression and a conventional disk-based implementation was first reported by Pople et al.<sup>9</sup> A first direct approach to the MP2 gradient method was recently presented by Frisch and coworkers.<sup>10</sup> In a second article, the algorithm was extended to a semidirect approach that utilizes disk space at the expense of I/O.<sup>11</sup> Because external disks with more than 2-Gb capacity and access times of a few milliseconds at transfer rates of up to two Mb/s are now available at moderate prices, a flexible semidirect approach to the MP2 gradient method offers advantages that should be exploited. Together with tools for integral prescreening and exploitation of symmetry, this approach should allow applications to systems with several hundred basis functions.

It is our objective to demonstrate the performance of our semidirect MP2 gradient implementation for closed-shell systems on workstation computers. In the next section, we briefly describe the special features of our implementation, which is based on the semidirect approach proposed by Frisch et al.<sup>11</sup> but is further extended to reduce disk space requirements and in addition exploits nonabelian point group symmetry. Timings of two demonstrative applications to large and symmetric molecules are given in the last section.

## IMPLEMENTATION

The design of the present algorithm was derived from the requirement that applications employing a few hundred basis functions could be carried out on workstations with 32–128 Mb memory and 1 Gb to

\*Author to whom all correspondence should be addressed.

a few Gb disk space. This led to the following specifications:

1. Main memory requirement should scale with  $N^2$ .
2. Disk space requirements should be adjustable.
3. I/O is admissible as long as this overlaps sufficiently with CPU.
4. Use of all discrete point group symmetries.

These goals can be achieved with the algorithm of Frisch et al.<sup>11</sup>—with some modifications—if combined with tools available within TURBOMOLE. To meet requirements 1 and 2, it is necessary to perform certain steps repeatedly. This especially requires a multiple-pass strategy as suggested in ref. 11: Occupied MOs are divided by the number of passes yielding batches of size  $B$  that are processed one batch at a time.

In the following, the formulation of the MP2 gradient as given by Handy et al.<sup>12</sup> will be adopted. Indices  $i, j, k, \dots$  refer to occupied MOs;  $a, b, c, \dots$  denote virtual MOs;  $p, q, r, \dots$  signify both virtual and occupied MOs; and  $\mu, \nu, \kappa, \dots$  are basis functions. In the restricted Hartree–Fock (RHF) closed-shell case, the MP2 gradient can be succinctly written as<sup>12</sup>

$$E_{\text{MP2}}^{(x)} = \sum_{pq} \mathbf{Y}_{pq} h_{pq}^x + \sum_{pq} \mathbf{W}_{pq} S_{pq}^x + \sum_{pqrs} \Gamma_{pqrs} (pq|rs)^x \quad (1)$$

where  $(x)$  denotes differentiation with respect to a Cartesian coordinate. The matrix  $\mathbf{Y}$  is the correlated density (also known as relaxed or response density) consisting of three blocks: the occupied–occupied block  $V_{ij}$ , the virtual–virtual block  $V_{ab}$ , and the occupied–virtual block  $Z_{ai}$ , which is the solution of the  $Z$ -vector equation.<sup>13</sup> The latter is a coupled perturbed Hartree–Fock (CPHF) type equation with the MP2 lagrangian  $L_{ai}$  as the corresponding right-hand side.  $\mathbf{W}$  is an intermediate matrix similar to the energy-weighted density of the SCF gradient expression. The final term contains the two-particle density matrix  $\Gamma$ , which consists of a nonseparable part  $\Gamma^{\text{NS}}$  and a separable part  $\Gamma^{\text{S}}$  and is defined through

$$\sum_{pqrs} \Gamma_{pqrs} (pq|rs)^x = -2 \underbrace{\sum_{ijab} t_{ij}^{ab} (ia|jb)^x}_{\Gamma^{\text{NS}}} + \underbrace{\sum_{pq} \mathbf{Y}_{pq} \sum_k [2(pq|kk)^x - (pk|qk)^x]}_{\Gamma^{\text{S}}} \quad (2)$$

where

$$t_{ij}^{ab} = \frac{2(ia|jb) - (ib|ja)}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} \quad (3)$$

An efficient way to obtain the MP2 gradient is to calculate the matrices  $\mathbf{Y}$ ,  $\mathbf{W}$  and the double-substitution amplitudes  $t_{ij}^{ab}$  in the MO representation, then transform them back to the AO basis, and finally contract them with derivative AO integrals. The present algorithm to evaluate these terms is sketched in Algorithm 1.

**Algorithm 1.** Structure of the present semidirect implementation. Loop over batches  $B$ .

```
(1)  $(\nu\mu|\kappa\lambda) \rightarrow (\nu\mu|ia), (\nu\mu|ij)$  and write; all  $i \in B$ 
    loop over  $i \in B$ 
      (2) read  $(\nu\mu|ia), (\nu\mu|ij) \rightarrow (jb|ia), (ab|ij), (ak|ij)$ 
          and write
      (3) read  $(ia|jb) \rightarrow t_{ij}^{ab}$  and write
      (4) increment  $V_{ij}, V_{ab}, \mathbf{W}_{ij}, \mathbf{W}_{ab}$  and  $L_{ai}$ 
      (5) read  $(\nu\mu|ia) \rightarrow (cb|ia)$  and increment  $L_{ai}$  on the fly
      (6) read  $t_{ij}^{ab} \rightarrow t_{ik}^{ac}$  and write
      (7) read  $t_{ik}^{ac} \rightarrow \Gamma_{\mu\nu, \kappa\lambda}^{\text{NS}}$  and contract with  $(\mu\nu|\kappa\lambda)^x$ 
    end loop  $i$ 
end loop  $B$ 
(8) complete  $L_{ai}$  and solve  $Z$ -vector equation for  $Z_{ai}$ 
(9) contract  $\Gamma_{\mu\nu, \kappa\lambda}^{\text{S}}$  with  $(\mu\nu|\kappa\lambda)^x$  and evaluate  $\mathbf{Y}_{\mu\nu} h_{\mu\nu}^x, \mathbf{W}_{\mu\nu} S_{\mu\nu}^x$ 
```

Steps 1 and 2, the transformation  $(\nu\mu|ia) \rightarrow (cb|ia)$  in step 5, the solution of the  $Z$ -vector equation, and step 9 are based on the MP2, CPHF, and SCF gradient routines of the TURBOMOLE program with appropriate modifications. This applies especially to the symmetry handling and the use of integral prescreening techniques.<sup>2,14</sup>

To further reduce disk storage demands and hence I/O, the semidirect approach of Frisch et al.<sup>11</sup> was modified in two ways. The maximum disk requirement of their algorithm is  $\text{BN}^3$  and appears at stage 2 of Algorithm 1. The three external integrals over MOs occupy by far the largest part of this disk space, that is,  $\frac{1}{2}BV^3$  ( $V$  is the number of virtual MOs). Because they are needed only once, to increment the MP2 lagrangian  $L_{ai}$ , they are generated on the fly and never stored in our algorithm. In contrast to ref. 11, the second half-transformation, to compute three external MO integrals (step 2), is deferred. Only when double-substitution amplitudes  $t_{ij}^{ab}$  are available, the second half-transformation is carried out together with the on-the-fly contraction of the resulting three externals with amplitudes (step 5). The basic loop structure for this task is given below ( $\mathbf{C}$  is the MO matrix).

**Algorithm 2.** Loop structure for step 5 of Algorithm 1.

```
loop  $i \in B$ , all  $a$ 
  read  $(\mu\nu|ia)$  all  $\mu, \nu$ 
  read  $t_{ji}^{ba}$  all  $j, b$ 
  loop  $b$ 
    loop  $(\nu\mu)$  pairs
       $(\mu b|ia) \leftarrow (\mu b|ia) + \mathbf{C}_{ib}(\mu\nu|ia)$ 
       $(\nu b|ia) \leftarrow (\nu b|ia) + \mathbf{C}_{ub}(\mu\nu|ia)$ 
    end loop  $(\nu\mu)$  pairs
    loop  $c \leq b$ 
      loop  $\mu$ 
         $(cb|ia) \leftarrow (cb|ia) + \mathbf{C}_{\mu c}(\mu b|ia)$ 
      end loop  $\mu$ 
      loop  $k$ 
         $L(c, k) \leftarrow L(c, k) - 4t_{ki}^{ba}(cb|ia)$ 
      end loop  $k$ 
    end loop  $c$ 
  end loop  $b$ 
end loop  $i, a$ 
```

In this way, the maximum disk requirement of the present algorithm is shifted to stage 1 of Algorithm 1 and is reduced by a factor of two ( $\frac{1}{2}BN^3$ ) compared to the original algorithm<sup>11</sup> at the only extra cost of reading the half-transformed integrals of type  $(ia|\mu\nu)$  twice and an additional reading of the amplitudes  $t_{ij}^{ab}$ . But, this is not a serious disadvantage because I/O sufficiently overlaps with the  $N^5$  arithmetic of the transformation. Besides this reduction due to the above-mentioned modification, a further saving in disk space occurs by using the full symmetry point group, as will be explained below.

The second modification concerns the back transformation of the double-substitution amplitudes from the MO to AO basis (step 6 of Algorithm 1), yielding the nonseparable part  $\Gamma^{\text{NS}}$  of the MP2 two-particle density, which is subsequently contracted with derivative integrals (step 7). We organized this back transformation in two half-transformations similar to the integral transformation.<sup>7</sup> The second half-transformation is absorbed into the subsequent contraction of the two-particle density with derivative integrals. This leads to an additional  $N^5$  step at this stage compared to the algorithm of ref. 11 but has the following advantage. It requires only a single transposition of an array of size  $\frac{1}{2}BVN^2$  (half-transformed amplitudes  $t_{ia}^{qk}$ ). In the  $N^2$  algorithm of ref. 11, an additional reordering of an array of size  $\frac{1}{2}BN^3$  is necessary, which causes additional I/O.

In the following, we comment briefly on I/O, which may have a poor performance on workstations, especially for direct access. The wallclock time of a job can then be much larger than the CPU time. Because out-of-core sort steps are unavoidable in algorithms with  $N^2$  memory requirement, the efficiency of I/O is of great importance. In our implementation, I/O is strictly  $N^4$  dependent. When sorting arrays, no second copy or indices must be held on disk. The data that have to be sorted are written once and then read several times if necessary in a certain order (like in the case of the exchange-type MO integrals or double-substitution amplitudes).

Besides the usual double buffering of all I/O operations, some help comes from the dynamic memory allocation used throughout in the program. It immediately frees previously allocated memory if it is no longer in use. The main memory thus made available to the operating system (UNIX) is used as (additional) internal buffer that facilitates overlapping of CPU and channel activities. Finally, for the largest files—the half-transformed integrals and amplitudes—the I/O is coded in C, which is slightly faster than FORTRAN.

The scheme implemented for exploiting molecular point group symmetry is based on the formalism used throughout the TURBOMOLE program system.<sup>7</sup> It is due to Häser and is outlined elsewhere in detail.<sup>15</sup> The algorithm applies projection operators that project quadruples of MO onto their totally sym-

metric part and in this way allows to identify those MO integrals or amplitudes that are zero or equal to others for symmetry reasons. Further, the factorization of such projectors provides a means of packing and unpacking the above-mentioned quantities, leading to a symmetry speed-up and a reduction of disk space. For quantities in the MO basis, the disk requirement is reduced by a factor of  $g^{-1}$  if  $g$  is the order of the molecular point group. However, for half-transformed quantities like the integrals  $(\nu\mu|ia)$  and amplitudes  $t_{ia}^{qk}$  the saving factor is  $g^{-1}$  at best and may deteriorate if atoms are located on symmetry elements as in planar and linear molecules.

After the submission of the article, a revision of the present algorithm was completed concerning the solution of the Z-vector equation (step 8 of Algorithm 1). Although the original MO-based algorithm is fast, it requires the evaluation and storage of an additional type of half-transformed integrals  $(\nu\mu|ij)$  in stage 1 and Coulomb-type MO integrals  $(ab|ij)$  in stage 2. The algorithm was substituted by the direct inversion of iterative subspace (DIIS) method.<sup>9</sup> The Z-vector equation is now solved directly in the AO basis,<sup>16</sup> similar to a direct SCF calculation. While the overall computation time is increased marginally, disk storage requirements are reduced significantly for larger molecules at least (see Table III).

## REPRESENTATIVE TIMINGS

In Table I, timings and disk space requirements for two symmetric molecules are listed that illustrate the savings due to symmetry. The CPU times for

**Table I.** Performance of MP2 gradient calculations using different point groups. Times refer to an IBM RS6000/320H unless otherwise indicated.

	(AlH) <sub>6</sub>			N <sub>4</sub> (CH <sub>2</sub> ) <sub>6</sub>		
	138			210		
Number of basis functions <sup>a</sup>						
Point group	<i>D</i> <sub>3d</sub>	<i>D</i> <sub>3</sub>	<i>C</i> <sub>2h</sub>	<i>T</i> <sub>d</sub>	<i>D</i> <sub>2d</sub>	<i>D</i> <sub>2</sub>
CPU time <sup>b</sup> (min)	45	80	130 <sup>c</sup>	154	373	688 (737) <sup>d</sup>
CPU/wallclock (%)	90	90	94	92	93	96
Disk space (Mb)	64	103	180 <sup>c</sup>	166	353	670 (286) <sup>d</sup>

<sup>a</sup>A basis set of SVP type<sup>17</sup> was employed.

<sup>b</sup>All calculations performed with a single pass if not otherwise indicated.

<sup>c</sup>For comparison, Gaussian 92<sup>18</sup> calculations were carried out under the same conditions (semidirect  $N^2$  algorithm, 20-Mb memory, full correlation, same symmetry and number of passes). The following keywords were specified: %MEM=2621440; MP2=(SEMIDIRECT,FULL,MAXDISK=227840000); SCF=(DIRECT) GEN FORCE. The CPU time of the Gaussian 92 MP2 gradient modules (links 906, 1002, 701, 702, 703, 716) was 317 min and the disk requirement (rwf-file) was 1259 Mb.

<sup>d</sup>CPU time and disk requirement of an MPGRAD calculation on an IBM RS6000/350 with three passes given in parentheses. The same calculation with the Gaussian 92 MP2 gradient required 1456 min CPU time and 1735-Mb disk space.

(AlH)<sub>6</sub> scale as 1.6 and 2.9 if symmetry is increased from C<sub>2h</sub> to D<sub>3</sub> and D<sub>3d</sub>, in close agreement with the ratio of the number of symmetry operations (1.5 and 3, respectively). In the case of N<sub>4</sub>(CH<sub>2</sub>)<sub>6</sub>, the situation is less satisfactory but the values of 1.8 and 4.5 do not differ much from the corresponding theoretical factors 2 and 6. The savings in disk space are similar to those in CPU times. It should be mentioned that for those terms of the gradient expression that are fully evaluated in the MO representation (perturbed density Y) symmetry savings by a factor of 1/*g*<sup>2</sup> occur. But, this hardly affects the total CPU time because in the dominating steps (steps 1, 2, 5, and 7 of Algorithm 1) the symmetry speed-up scales usually as 1/*g*.

A sensitive indicator for the efficiency of the I/O implementation is the ratio of the CPU to wallclock time of a calculation. As can be seen from Table I, it exceeds 90% in all cases. This shows that already for such small calculations (note that I/O scales as *N*<sup>4</sup> and arithmetic as *N*<sup>5</sup>) the program performs well on a workstation computer.

In the footnotes of Table I, we also document the performance of Gaussian 92<sup>18</sup> in comparison to MPGRAD as requested by one of the referees. In either case, the semidirect *N*<sup>2</sup> algorithm (which would be selected automatically under the given memory restrictions) required about twice the CPU time although a low (abelian) symmetry group was used. Disk space requirements were roughly six times larger for the two test cases considered.

MP2 gradient calculations on [AlSi(CH<sub>3</sub>)<sub>3</sub>]<sub>4</sub> and ferrocene were carried out to assess the performance of MPGRAD for large calculations. The technical data are summarized in Table II. The former molecule consists of 56 atoms, is highly symmetric (point group *T<sub>d</sub>*), and thus should provide a good test case, especially for the efficiency of the symmetry scheme implemented. A salient feature of the semidirect algorithm is its capability of adapting to the respective resources of a computer. This has been carried to the extreme with the first calculation in Table II, where only 32 Mb memory and a mass storage capacity of no more than 1.3 Gb were made available. Three passes and hence three integral and four derivative evaluations were necessary under these con-

ditions. The main memory requirement was 45 Mb and therefore significant paging to virtual memory was to be expected. However, the CPU performance of 83% is satisfactory under the given circumstances.

As a second application, an MP2 gradient calculation on ferrocene was performed. This molecule was recently investigated by Park and Almlöf<sup>19</sup> using the modified coupled-pair functional (MCPF) and MP2 methods. Because the authors had to resort to single-point calculations on the iron—cyclopentadienyl distance and employed small basis sets, it was our goal to show that the analytic MP2 optimization with a sufficiently large basis set is no longer prohibitive with the present program. A detailed investigation of ferrocene on the MP2 level including the full optimization will be presented in a separate article. Because of the memory requirement of 94 Mb, we used a 128-Mb memory workstation. Thus problems due to paging were avoided and it was possible to assess the I/O performance. The CPU/wallclock time ratio exceeded 90%, as for the smaller calculations, although a much larger data transfer to disk was necessary.

In a series of calculations with different numbers of passes, we wanted to examine what one has to pay for a significant reduction of the disk storage requirement. From inspection of Algorithm 1, it follows that each additional pass through all *N*<sup>5</sup>-dependent terms inevitably leads to an additional integral and derivative evaluation. Because in the present implementation the integral and derivative evaluation scales with the order of the molecular point group, two cases were considered: one with high symmetry and one with moderate symmetry. Again, [AlSi(CH<sub>3</sub>)<sub>3</sub>]<sub>4</sub> in *T<sub>d</sub>* symmetry was chosen together with the hexamethylenetetramine molecule in *D<sub>2</sub>* symmetry. To get a more detailed insight into the individual steps of the algorithm, Table III contains additional timings of all steps of Algorithm 1. Let us first consider the data for the calculation on [AlSi(CH<sub>3</sub>)<sub>3</sub>]<sub>4</sub>. Comparing the total CPU times of the three different calculations, we see a surprisingly small increase when increasing the number of passes. A calculation with three passes (two passes) requires 66% (44%) less disk space but only 3.4% (2.3%) more CPU time than a run with a single pass. This additional expense is mostly caused by the integral derivative generation, while additional integral computations are less expensive. This is due to the high symmetry, which leads to a speed-up of the integral and derivative computation by a factor of 24. Thus, these two tasks have only a minor influence on the total CPU time. The situation changes for lower symmetry. In the case of N<sub>4</sub>(CH<sub>2</sub>)<sub>6</sub>, the savings in disk space when increasing the number of passes from one to eight (four) are 68% (58%), but now one must accept an increase in CPU time of 188% (80%). All that can be gained by symmetry is a factor of 2 (if point group *D<sub>2</sub>* is imposed), and therefore integral

**Table II.** Technical data of MP2 gradient calculations on [AlSi(CH<sub>3</sub>)<sub>3</sub>]<sub>4</sub> and ferrocene.

	[AlSi(CH <sub>3</sub> ) <sub>3</sub> ] <sub>4</sub> <sup>a</sup>	Fe(C <sub>5</sub> H <sub>5</sub> ) <sub>2</sub> <sup>b</sup>
Number of basis functions	492 <sup>c</sup>	483 <sup>d</sup>
Point group	<i>T<sub>d</sub></i>	<i>D<sub>5h</sub></i>
Number of passes	3	2
CPU time (min)	5061	3342
CPU/wallclock (%)	83	92
Disk space (Mb)	1221	1235

<sup>a</sup>Timings refer to an IBM RS6000/320H (32-Mb memory).

<sup>b</sup>Timings refer to an IBM RS6000/550 (128-Mb memory).

<sup>c</sup>A basis set of SVP type<sup>17</sup> was employed.

<sup>d</sup>Basis set of Fe, [8s7p4d2f]; C, [6s4p3d]; H, [3s2p].

**Table III.** Timings (min) of individual steps of MP2 gradient calculations on  $[\text{AlSi}(\text{CH}_3)_3]_4$  and  $\text{N}_4(\text{CH}_2)_6$  with different numbers of passes.

	$[\text{AlSi}(\text{CH}_3)_3]_4^a$			$\text{N}_4(\text{CH}_2)_6^b$		
Number of basis functions <sup>c</sup>	492			210		
Point group	$T_d$			$D_2$		
Number of passes	1	2	3	1	4	8
First half-transformation, step 1	174	194	210	97	216	379
Second half-transformation, step 2	654	644	630	31	31	31
Increment <b>Y</b> , <b>W</b> , <b>L</b> , steps 3 and 4	163	164	160	17	17	18
Increment <b>L</b> with $(cb ia)$ , step 5	897	892	878	67	69	69
$t_{ij}^{qb}$ backtransformation, step 6	322	320	317	35	36	36
$\Gamma_{\mu\nu,\kappa\lambda}^{\text{NS}} (\mu\nu \kappa\lambda)^r$ evaluation, step 7	273	332	382	289	707	1280
Solution of Z-vector equation, step 8	62	63	62	10	11	11
$\Gamma_{\mu\nu,\kappa\lambda}^{\text{S}} (\mu\nu \kappa\lambda)^r$ evaluation, step 9	35	35	35	107	111	119
Total CPU time	2643	2703	2732	688	1233	1978
CPU/wallclock (%)	93	95	92	96	92	91
Disk space (Mb)	3598	2007	1221	670	283	215
Disk space (Mb) <sup>d</sup>	2680	1357	912	620	246	180

<sup>a</sup>Timings refer to an IBM RS6000/550 (128-Mb memory).<sup>b</sup>Timings refer to an IBM RS6000/320H (32-Mb memory).<sup>c</sup>A basis set of SVP type<sup>17</sup> was employed.<sup>d</sup>Disk space requirement of the current version as explained in text.

and derivative computation becomes much more expensive and, further, more passes are necessary to achieve a disk space reduction comparable to the high-symmetry calculations. Besides this, steps 1 and 7 of Algorithm 1 are now dominant because hexamethylenetetramine has only 38 occupied MOs (108 occupied MOs were involved in the  $[\text{AlSi}(\text{CH}_3)_3]_4$  calculation), which considerably reduces the computational effort for  $N^5$ -dependent steps.

## CONCLUSION

We have shown that a semidirect implementation of the MP2 gradient method together with an efficient scheme to exploit molecular symmetry and techniques for integral prescreening provides a powerful tool in examining structures of large and symmetric molecules at the MP2 level of theory. Applications employing a few hundred basis functions are in this way feasible even on computers with limited resources. In cases of high symmetry, the extra expense in CPU time to drastically reduce the disk storage requirements of a calculation is small. For cases with lower symmetry, this is no longer true. However, although such calculations may require CPU times of several days the CPU/wallclock ratio still exceeds 90%.

The authors thank J. Gauss and M. Häser for several valuable hints given during the development of the program. They are grateful to H. Horn, C. Kölmel, and H. Weiss for many helpful discussions. Financial support by the Stipendienstiftung of the Fonds der Chemischen Industrie via a postdoctorate stipendium for one of us (F.H.) is gratefully acknowledged. Finally, F.H. thanks the TURBOMOLE

team for the hospitality and exuberant but stimulating atmosphere during his stay in Karlsruhe.

## REFERENCES

1. J. Almlöf, K. Fægri Jr., and K. Korsell, *J. Comp. Chem.*, **3**, 385 (1982); D. Cremer and J. Gauss, *J. Comp. Chem.*, **7**, 274 (1986).
2. M. Häser and R. Ahlrichs, *J. Comp. Chem.*, **10**, 104 (1989).
3. C. Møller and M.S. Plesset, *Phys. Rev.*, **46**, 618 (1934).
4. P.R. Taylor, *Int. J. Quantum Chem.*, **31**, 521 (1987).
5. M. Head-Gordon, J.A. Pople, and M.J. Frisch, *Chem. Phys. Lett.*, **153**, 281 (1988).
6. S. Sæbø and J. Almlöf, *Chem. Phys. Lett.*, **154**, 83 (1989).
7. R. Ahlrichs, M. Bär, M. Häser, H. Horn, and C. Kölmel, *Chem. Phys. Lett.*, **162**, 165 (1989).
8. V. Parasuk, J. Almlöf, and M.W. Feyereisen, *J. Am. Chem. Soc.*, **113**, 1049 (1991); M. Häser, J. Almlöf, and G.E. Scuseria, *Chem. Phys. Lett.*, **181**, 497 (1991); G.E. Scuseria, *Chem. Phys. Lett.*, **176**, 423 (1991).
9. J.A. Pople, R. Krishnan, H.B. Schlegel, and J.S. Binkley, *Int. J. Quantum Chem. Symp.*, **13**, 225 (1979).
10. M.J. Frisch, M. Head-Gordon, and J.A. Pople, *Chem. Phys. Lett.*, **166**, 275 (1990).
11. M.J. Frisch, M. Head-Gordon, and J.A. Pople, *Chem. Phys. Lett.*, **166**, 281 (1990).
12. N.C. Handy, R.D. Amos, J.F. Gaw, J.E. Rice, and E.D. Simandiras, *Chem. Phys. Lett.*, **120**, 151 (1985); E.D. Simandiras, R.D. Amos, and N.C. Handy, *Chem. Phys. Lett.*, **114**, 9 (1987).
13. N.C. Handy and H.F. Schaefer, *J. Chem. Phys.*, **81**, 5031 (1984).
14. H. Horn, H. Weiss, M. Häser, M. Ehrig, and R. Ahlrichs, *J. Comp. Chem.*, **12**, 1058 (1991).
15. M. Häser, Dissertation, Universität Karlsruhe, 1990; M. Häser, J. Almlöf, and M.W. Feyereisen, *Theor. Chim. Acta*, **79**, 115 (1991); M. Häser, *J. Chem. Phys.*, **95**, 8259 (1991).

16. Y. Osamura, Y. Yamaguchi, and H.F. Schaefer III, *J. Chem. Phys.*, **77**, 383 (1982).
17. A. Schäfer, H. Horn, and R. Ahlrichs, *J. Chem. Phys.*, **97**, 2571 (1992).
18. M.J. Frisch, G.W. Trucks, M. Head-Gordon, P.M.W. Gill, M.W. Wong, J.B. Foresman, B.G. Johnson, H.B. Schlegel, M.A. Robb, E.S. Replogle, R. Gomperts, J.L. Andres, K. Raghavachari, J.S. Binkley, C. Gonzalez, R.L. Martin, D.J. Fox, D.J. Defrees, J. Baker, J.J.P. Stewart, and J.A. Pople, Gaussian 92, revision C, Gaussian, Inc., Pittsburgh, PA, 1992.
19. C. Park and J. Almlöf, *J. Chem. Phys.*, **95**, 1829 (1991).