

# Improvements on the Direct SCF Method

Marco Häser and Reinhart Ahlrichs

*Institut für Physikalische Chemie und Elektrochemie, Lehrstuhl f. Theoretische Chemie, Universität Karlsruhe  
Kaiserstr. 12, D-7500 Karlsruhe, West Germany (FRG)*

Received 14 March 1988; accepted 10 May 1988

Three improvements on the direct self-consistent field method are proposed and tested which together increase CPU-efficiency by about 50%: (i) selective storage of costly integral batches; (ii) improved integral bond for prescreening; (iii) decomposition of the current density matrix into a linear combination of previous density matrices—for which the two-electron contributions to the Fock matrix are available—and a remainder  $\Delta\mathbf{D}$ , which is minimized; construction of the current Fock matrix only requires processing of the small  $\Delta\mathbf{D}$  which enhances prescreening.

## I. INTRODUCTION

The self-consistent field (SCF) method<sup>1-3</sup> is a fundamental tool in molecular electronic structure theory. Its computational bottleneck is the processing of two-electron integrals

$$(\nu\mu|\kappa\lambda) = \iint \frac{\nu(1)\mu(1)\kappa(2)\lambda(2)}{r_{12}} dV_1 dV_2 \quad (1)$$

over basis functions (usually CGTFs<sup>4</sup>), which are required for the construction of the Fock matrix.

$$\mathbf{F}^{(n)} = \mathbf{h} + \mathbf{G}^{(n)}, \quad (2)$$

in each iteration (closed shell):

$$\begin{aligned} G_{\nu\mu}^{(n)} &= G_{\nu\mu}[\mathbf{D}^{(n)}] \\ &= \sum_{\kappa,\lambda} \{2(\nu\mu|\kappa\lambda) - (\nu\kappa|\mu\lambda)\} D_{\kappa\lambda}^{(n)} \end{aligned} \quad (3)$$

where  $\mathbf{D}^{(n)}$  is the density matrix calculated from the molecular orbitals (MOs) of the previous iteration, and  $\mathbf{h}$  is the one-electron hamiltonian.

The conventional approach is to compute and store the two-electron integrals, which are then retrieved and processed in each SCF iteration. However, the sheer number of two-electron integrals often prohibits their storage, especially on work station computers. To overcome this problem an alternative procedure termed direct SCF has been suggested by Almlöf, Faegri, and Korsell:<sup>5</sup> abandon integral storage and repeat the most time-consuming

step—calculation of the two-electron integrals—in each iteration ( $n$ ). Calculations on large molecules<sup>6</sup> thus become possible at the expense of increased execution times.

The additional effort is not as large as expected at first glance, since there are efficient strategies which help to keep central processing unit (CPU) time within reasonable limits:

1. Reduction of the number of iterations<sup>7,8</sup>
2. Reduction of the number of integrals to be recomputed in each iteration<sup>5,8</sup>

We have implemented a direct SCF program (DSCF) on the basis of the Karlsruhe version of the COLUMBUS program system in which integral evaluation is done by means of the HONDO procedure.<sup>9</sup> Convergence acceleration is achieved using direct inversion in the iterative subspace (DIIS).<sup>7</sup>

Additional features have been implemented and tested in our direct SCF version which mainly concern point 2 above:

1. Improved integral estimation
2. Selective storage of expensive and frequently used integrals
3. Minimization of the difference density matrix by linear combination
4. An SCF algorithm based on symmetrized basis functions for all finite point groups

In the present article we focus our attention on these features.

## II. INTEGRAL PRESCREENING

It is essential for direct SCF procedures that integral batches (NM, KL),\* which do not significantly contribute to the two-electron part of the Fock matrix  $\mathbf{G}^{(n)}$  [Eq. (3)] are identified and omitted. This integral prescreening can be enhanced by a recursive construction of  $\mathbf{G}^{(n)}$ .<sup>[5]</sup>

$$\mathbf{G}^{(n)} = \mathbf{G}^{(n-1)} + \mathbf{G}[\Delta\mathbf{D}^{(n)}], \quad (4)$$

where

$$\Delta\mathbf{D}^{(n)} = \mathbf{D}^{(n)} - \mathbf{D}^{(n-1)}. \quad (5)$$

The advantage is that the contributions of individual integral batches (NM, KL) to the Fock matrix vanish with  $\Delta\mathbf{D}^{(n)}$  as convergence is approached.

A quick and reliable estimate CON (NM, KL) for these contributions should thus be available *before* the integral batch (NM, KL) is calculated; a general form of such an estimate is:<sup>5</sup>

$$\begin{aligned} \text{CON}(\text{NM}, \text{KL}) = \max\{ & \Delta D_{KL}, \Delta D_{NM}, \frac{1}{4}\Delta D_{KM}, \\ & \frac{1}{4}\Delta D_{LM}, \frac{1}{4}\Delta D_{KN}, \\ & \frac{1}{4}\Delta D_{LN}\} \\ & * \text{EST}(\text{NM}, \text{KL}) \end{aligned} \quad (6)$$

where

$$\Delta D_{KL} := \max_{\substack{\kappa \in K \\ \lambda \in L}} \{|\Delta D_{\kappa\lambda}^{(n)}|\}, \quad (7)$$

and

$$\begin{aligned} \text{EST}(\text{NM}, \text{KL}) & \geq m(\text{NM}, \text{KL}) \\ & := \max_{\nu, \mu, \kappa, \lambda} |(\nu\mu | \kappa\lambda)| \end{aligned} \quad (8)$$

is an integral estimate ( $\nu, \mu, \dots$  run over the functions of shell N, M,  $\dots$ , respectively).

The batch (NM, KL) can be omitted, if

$$\text{CON}(\text{NM}, \text{KL}) < \text{THR}. \quad (9)$$

The threshold THR is chosen to comply with the accuracy desired, but also depends on the quality of the integral estimate EST(NM, KL). Let us now derive an integral estimate  $Q(\text{NM}, \text{KL})$  which has advantages over existing estimates. The electrostatic interaction energy

$$\iint \frac{\rho_1(1)\rho_2(2)}{r_{12}} dV_1 dV_2$$

of two continuous charge distributions  $\rho_1$  and  $\rho_2$  satisfies the axioms of an inner

product; Roothaan<sup>1</sup> proved positive definiteness, the rest is trivial. Schwarz' inequality

$$\begin{aligned} \left| \iint \frac{\rho_1(1)\rho_2(2)}{r_{12}} dV_1 dV_2 \right|^2 \\ \leq \iint \frac{\rho_1(1)\rho_1(2)}{r_{12}} dV_1 dV_2 \\ * \iint \frac{\rho_2(1)\rho_2(2)}{r_{12}} dV_1 dV_2 \end{aligned} \quad (10)$$

then provides a separable bond for two-electron integrals  $(\nu\mu | \kappa\lambda)$ :

$$|(\nu\mu | \kappa\lambda)| \leq Q_{\nu\mu} Q_{\kappa\lambda} \quad (11)$$

with

$$Q_{\nu\mu} := (\nu\mu | \nu\mu)^{1/2} \quad (12)$$

This bound is optimal since the equality sign in (11) holds for  $\nu, \mu = \kappa, \lambda$ . This bond is easily applied to batches (NM, KL) by means of

$$Q_{NM} := \max_{\substack{\nu \in N \\ \mu \in M}} Q_{\nu\mu} \quad (13)$$

and we then put

$$\begin{aligned} \text{EST}(\text{NM}, \text{KL}) = Q(\text{NM}, \text{KL}) & := Q_{NM} Q_{KL} \\ & \geq \max |(\nu\mu | \kappa\lambda)| = m(\text{NM}, \text{KL}). \end{aligned} \quad (14)$$

We now compare the bond (14) to the integral estimates EST(NM, KL) of other authors. Cremer and Gauss<sup>8</sup> in their direct SCF program use an integral estimate based on an integral bond over s-type gaussians proposed by Ahlrichs.<sup>10</sup> As it turns out the Cremer/Gauss/Ahlrichs bound is just a special case of the Schwarz inequality [Eq. (10)]. Cremer and Gauss found the formula also useful for estimating integrals involving p and higher CGTFs, although it is then no longer a rigorous bond.

The approach by Almlöf, Faegri, and Korsell<sup>5</sup> is not related to the Schwarz inequality [Eq. (10)]. They apply radial overlap integrals

$$S_{KL} := N_K N_L \left( \frac{\pi}{\alpha_{KL}} \right)^{3/2} e^{-\frac{\alpha_K \alpha_L}{\alpha_{KL}} R_{KL}^2} \quad (15)$$

(where  $\alpha_{KL} = \alpha_K + \alpha_L$  and  $N_K, N_L$  are normalization constants)\* in estimating the magnitude of two-electron integrals:

\*N, M, K, L designate shells.

\*As far as we can see the normalization constant chosen by Almlöf et al. is  $N_K = (2\alpha_K/\pi)^{3/4}$ .

$$\text{EST}(\text{NM}, \text{KL}) = T(\text{NM}, \text{KL})$$

$$\begin{aligned} &:= S_{\text{NM}} S_{\text{KL}} \frac{\alpha_{\text{NM}} \alpha_{\text{KL}}}{\pi(\alpha_{\text{NM}} + \alpha_{\text{KL}})} |\approx \max|(\nu\mu | \kappa\lambda)| \\ &= m(\text{NM}, \text{KL}). \end{aligned} \quad (16)$$

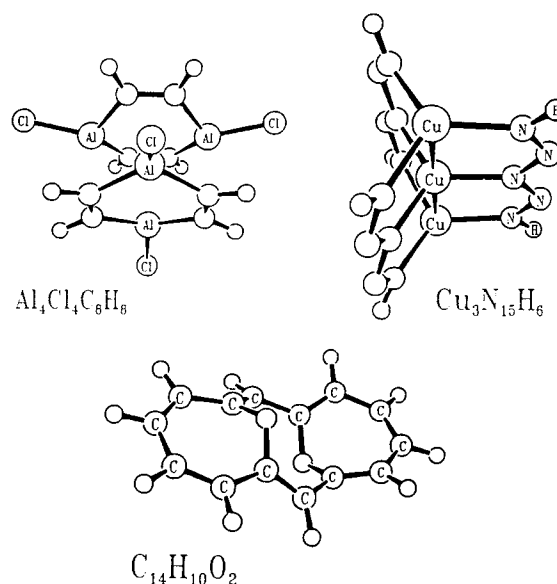
The integral estimate  $T(\text{NM}, \text{KL})$  has no product form, and for contracted GTFs the evaluation of the harmonic term  $\alpha_{\text{NM}} \alpha_{\text{KL}} / \alpha_{\text{NM}} + \alpha_{\text{KL}}$  requires additional computational labor. If the harmonic mean is substituted by a geometric mean (*which is greater than or equal to the harmonic mean*) the Almlöf formula becomes separable and can easily be applied to CGTFs:

$$m(\text{NM}, \text{KL}) \approx T'(\text{NM}, \text{KL}) := T_{\text{NM}} T_{\text{KL}}, \quad (17)$$

where

$$T_{\text{KL}} := \frac{2\alpha_K^{3/4} \alpha_L^{3/4}}{\pi^{1/2} \alpha_{\text{KL}}} e^{-(\alpha_K \alpha_L / \alpha_{\text{KL}}) R_{\text{KL}}^2}. \quad (18)$$

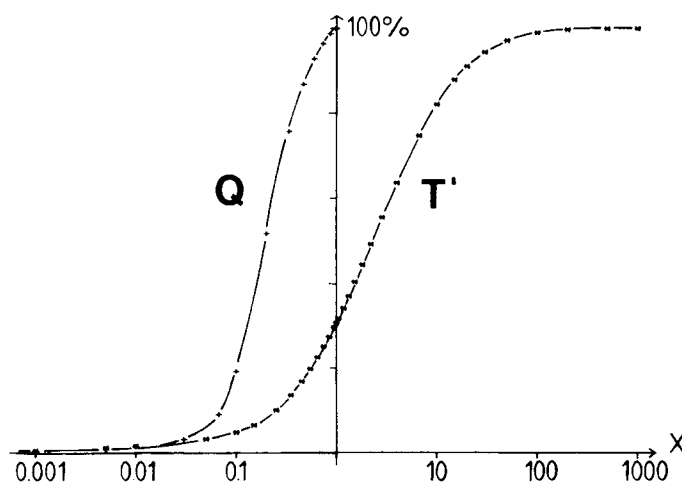
For a comparison of the integral bond  $Q(\text{NM}, \text{KL})$  [Eq. (14)] with the integral estimate  $T'(\text{NM}, \text{KL})$  [Eq. (17)], we have used the molecule  $\text{C}_{14}\text{H}_{10}\text{O}_2$  depicted in scheme 1. This molecule has previously been used to demonstrate advantages and improvements of the direct SCF approach by Cremer and Gauss.<sup>8</sup> We have used the same geometry (alternating bonds leading to  $\text{C}_s$  symmetry) but a slightly different basis set: a standard Huzinaga basis set, (8, 4)/[4, 2] for C and O, (4)/[2] for H, leading to 180 basis functions. We first tested the accuracy of approximating  $m(\text{NM}, \text{KL})$  by  $Q(\text{NM}, \text{KL})$  [Eq. (14)] and by  $T'(\text{NM}, \text{KL})$  [Eq. (18)], for all batches with



Scheme 1.

$m(\text{NM}, \text{KL}) \geq 10^{-8}$ , that is, for  $8.4 \cdot 10^6$  batches. The results in Figure 1 and Table I show the distribution function of the batches over the quotient  $m(\text{NM}, \text{KL})/\text{EST}(\text{NM}, \text{KL})$ , where EST is either  $Q$  or  $T'$ .

While  $Q$  is not only a strict bond it has also a much sharper distribution function than the  $T'$  estimate: if we ask for the size of the error within which 99% of the batches can be correctly estimated, we find about two orders of magnitude for the  $Q$  bond but twice that for the  $T'$  estimate. Note that the tail of the  $T'$  estimate, which stretches to the far right of Figure 1, and describes underestimation of  $m(\text{NM}, \text{KL})$ , is unaffected if we go from the  $T'$  estimate to Almlöf's original  $T$  estimate [Eq. (16)].



**Figure 1.** Percentage of batches (NM, KL) with  $\max|(\nu\mu | \kappa\lambda)|/\text{EST}(\text{NM}, \text{KL}) \leq x$  for  $\text{EST} = Q$  [Eq. (14)], and for  $\text{EST} = T'$  [Eq. (17)]; the depicted distribution function for the  $Q$  bond is twice as sharp as that of the  $T'$  estimate compared on logarithmic scale.

**Table I.** Percentage of batches (NM, KL) with  $m(\text{NM, KL})/\text{EST}(\text{NM, KL}) \leq X$  for the integral bond  $\text{EST} = Q$  [Eq. (14)], and the integral estimate  $\text{EST} = T'$  [Eq. (17)] (molecule  $\text{C}_{14}\text{H}_{10}\text{O}_2$ , Scheme 1).

X	EST = $Q$ (%)	EST = $T'$ (%)
1/1000	0.016	0.294
5/1000	0.578	1.057
1/100	1.306	1.533
1/10	19.086	4.804
1	100.000	30.645
10	100.000	82.017
100	100.000	98.972
200	100.000	99.682
500	100.000	99.953
1000	100.000	99.991
5000	100.000	100.000

The advantage of the  $Q$  bond is confirmed by a practical test. We conducted a direct SCF calculation on the same molecule  $\text{C}_{14}\text{H}_{10}\text{O}_2$  using two versions of the DSCF program which differed only in the integral estimate  $\text{EST}$  employed. The results, collected in Table II, show that calculation 1 [ $\text{EST} = Q$  and  $\text{THR} = 10^{-9}$ , Eq. (9)] and calculation 3 ( $\text{EST} = T'$  and  $\text{THR} = 10^{-11}$ ) are of comparable accuracy. Thus, in this example, the application of the  $Q$  bond saves 18.5% of the execution time over the  $T'$  estimate. For molecules and basis sets including  $d$  functions we expect the effects to be even more pronounced.

**Table II.** Comparison of the  $Q$ -bond [Eq. (14)] and the  $T'$  estimate [Eq. (17)] in three test calculations on  $\text{C}_{14}\text{H}_{10}\text{O}_2$  (Scheme 1).

Calculation EST THR	1		2		3	
	$Q$ $10^{-9}$		$T'$ $10^{-10}$		$T'$ $10^{-11}$	
Iteration <sup>a</sup>	time <sup>b</sup>	energy <sup>c</sup>	time <sup>b</sup>	energy <sup>c</sup>	time <sup>b</sup>	energy <sup>c</sup>
1	90	-.19624377	90	-.19624204	102	-.19624370
2	31 <sup>d</sup>	-.19754769	28 <sup>d</sup>	-.19754569	41 <sup>d</sup>	-.19754764
3	26	-.19776509	22	-.19776286	34	-.19776501
4	24	-.19787187	21	-.19786943	32	-.19787175
5	23	-.19792478	20	-.19792213	— <sup>e</sup>	— <sup>e</sup>
6 <sup>f</sup>	27	-.19797799	22	-.19797482	— <sup>e</sup>	— <sup>e</sup>
7	15	-.19797808 <sup>g</sup>	15	-.19797458	— <sup>e</sup>	— <sup>e</sup>
1-4	171	—	161	—	210	—

<sup>a</sup>The energy of the start vector was -685.19419 hartree.

<sup>b</sup>CPU time per iteration in min for Siemens 7881 computer at the Rechenzentrum Karlsruhe.

<sup>c</sup>Energy + 685 hartree; the "exact" energy for the basis set chosen, obtained with  $\text{EST} = Q$  and  $\text{THR} = 10^{-10}$ , is -685.19797809 hartree.

<sup>d</sup>The marked decrease in execution time when going from the first to the second iteration is due to partial integral storage; this feature will be explained in the next section; in all three calculations the same number of integrals ( $11 \times 10^6$  out of  $39 \times 10^6$ ) has been stored under equal conditions [ $\text{THIME} = 7$ , see Eq. (20)].

<sup>e</sup>Calculation stopped after 4 iterations, because time limit (4 h) had been reached.

<sup>f</sup>DIIS convergence acceleration applied before this iteration.

<sup>g</sup>From our experience for molecules of this size the final error of the energy is less than  $10^{-6}$  hartree if the  $Q$  bond is applied with  $\text{THR} = 10^{-9}$  (and no more than about 10 iterations are necessary to reach convergence); the extreme accuracy ( $10^{-8}$  hartree) of calculation 1 is coincidental.

### III. SELECTIVE INTEGRAL STORAGE

If mass storage is not available to the extent necessary to perform a conventional SCF calculation, one can resort to the direct SCF approach, recomputing integrals as they are needed, and making little use of mass storage facilities at the expense of a considerably increased execution time. It seems straightforward then to try an approach in between in order to keep execution time to a minimum: to selectively store as many large and costly integrals as possible during the first iteration, and to retrieve them when needed later on.

For a quick selection based on the level of batches (NM, KL) one needs simple estimates of integral magnitudes  $Q(\text{NM, KL})$  [Eq. (14)], and of the CPU time per integral  $t(\text{NM, KL})$ . Thus an integral batch (NM, KL) is stored if

$$Q_{NM} Q_{KL} \geq \text{THIZE}, \quad (19)$$

and if its calculation is costly, that is, if

$$t(\text{NM, KL}) \geq \text{THIME}, \quad (20)$$

where THIZE and THIME are user-defined thresholds.

For HONDO we have found that

$$t(\text{NM, KL}) = n_N n_M n_K n_L t_l \quad (21)$$

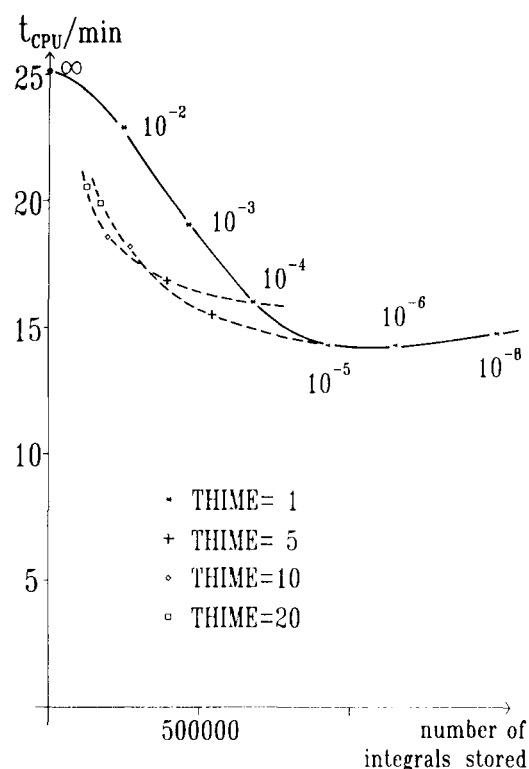
where  $n_N, n_M, \dots$  are contraction levels, and where

$$t_1 = \begin{cases} 6 & \text{if } l_N + l_M + l_K + l_L = 0 \\ 3 & \text{if } \dots = 1 \\ 2 & \text{if } \dots = 2 \\ 1 & \text{if } \dots > 2 \end{cases} \quad (22)$$

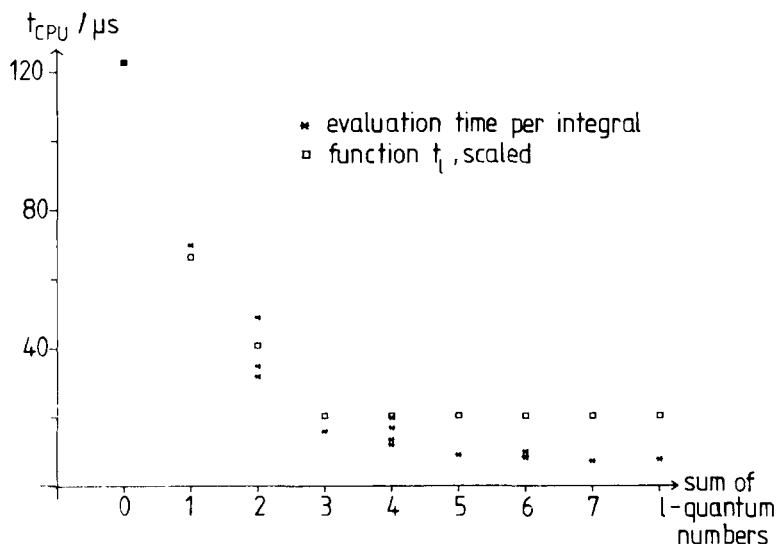
complies with the CPU times we measured (see Fig. 2). A more thorough analysis upon different integral evaluation algorithms is given by Hegarty<sup>11</sup> and van der Velde.<sup>12</sup>

The example  $C_{14}H_{10}O_2$  already demonstrated the importance of partial integral storage (roughly a quarter of the integrals needed during the first iteration were stored to restrict the execution time to 4 h).

In order to monitor the dependence of execution time versus THIZE and THIME, we have performed test calculations on dibenzene (Fig. 3). The direct SCF calculation, namely all integrals recomputed in each iteration (THIME =  $+\infty$ , THIZE =  $+\infty$ ), required 25 min 8 s CPU time.\* The choice THIZE =  $10^{-4}$  to  $10^{-5}$  and THIME = 5 to 10 reduces the CPU time to 15½–18½ min\* (corresponding to 60–75%) with still modest storage requirements (10% to 35% of a "conventional" SCF calculation, THIME =



**Figure 3.** CPU time vs. number of stored integrals for direct SCF calculations on dibenzene with different choices of THIME and THIZE [Eqs. (19) and (20)]; the solid line connects points with constant THIME = 1 (no selection of integrals with respect to evaluation cost) while the dashed lines connect points with constant THIZE. The calculations were done in  $D_{6h}$  symmetry with (8, 4)/[4, 2] and (4)/[2] Huzinaga basis sets (144 CGTFs); the ring-ring distance was 8 a.u., 8 iterations were necessary to attain convergence ( $|\Delta E| \leq 10^{-7}$ ) from start MOs corresponding to a ring distance of 100 a.u..



**Figure 2.** Measured and estimated dependence of CPU time (of the Siemens 7881 computer at the Rechenzentrum Karlsruhe) for batch-wise integral evaluation (HONDO) upon the  $l$ -quantum numbers of the shells (contraction level is 1).

\*Siemens 7881 computer at the Rechenzentrum Karlsruhe.

1, THIZE =  $10^{-8}$ ). This holds true for bigger calculations: in the course of a geometry op-

timization on  $C_8H_8Al_4Cl_4$  depicted in Scheme 1 (320 CGTFs including one  $d$  shell at each C, Al, and Cl atom) we have done a single SCF calculation first with THIME = 4 and THIZE =  $7.5 \times 10^{-6}$ , and then with THIME = 1 and THIZE  $\approx 4 \times 10^{-3}$ . Both choices lead to  $14 \times 10^6$  integrals being stored ( $94 \times 10^6$  integrals are processed in the first iteration), the first choice, however, saves 25% of the execution time<sup>†</sup> as compared to the second choice (9 iterations were needed to reach convergence). This is the benefit of storing not only the bigger integrals, but also the more time-consuming ones.

#### IV. MINIMIZATION OF THE DIFFERENCE DENSITY MATRIX BY LINEAR COMBINATION

The advantages of a recursive construction of the Fock matrix [Eqs. (4) and (5)] are well known.<sup>5,8</sup> This approach of enhancing integral prescreening can be improved if "minimized" density differences, rather than simple density differences [Eq. (5)], are derived for the construction of  $G^{(n)}$ . For this purpose we approximate  $D^{(n)}$  as a linear combination of density matrices  $\Delta D^{(j)}$  for which  $G[\Delta D^{(j)}]$  [Eq. (3)] has already been computed, such that the residual  $\Delta D^{(n)}$  is minimized:

$$\Delta D^{(n)} := D^{(n)} - \sum_{j=1}^{n-1} x_j^{(n)} \Delta D^{(j)} \quad (23)$$

$$\Delta G^{(n)} := G[\Delta D^{(n)}] \quad (24)$$

$$G^{(n)} = \Delta G^{(n)} + \sum_{j=1}^{n-1} x_j^{(n)} \Delta G^{(j)} \quad (25)$$

The vector of the linear coefficients  $\mathbf{x} := (x_1^{(n)}, \dots, x_{n-1}^{(n)})^T$  is determined by

$$\text{tr} \{ \Delta D^{(n)} \Delta D^{(n)} \} = \min \quad (26)$$

This constraint leads to the equation

$$\mathbf{A}\mathbf{x} = \mathbf{b} \quad (27)$$

where

$$A_{ij} := \text{tr} \{ \Delta D^{(i)} \Delta D^{(j)} \}, \quad (28)$$

$$b_i := \text{tr} \{ \Delta D^{(i)} D^{(n)} \}. \quad (29)$$

$\mathbf{A}$  is symmetric and positive definite (if the density matrices of all iterations are linearly independent). The unique solution of

Eq. (27) thus corresponds to the global minimum of Eq. (26). Furthermore  $\mathbf{A}$  is diagonal if the prior difference density matrices  $\Delta D^{(j)}$  have been optimized according to Eqs. (26) and (27), as is easily shown.

If a direct SCF calculation is close to convergence, however, the optimal  $\mathbf{x}$  with respect to Eq. (26) may not be optimal in terms of the desired accuracy of  $G^{(n)}$ . We now present a way to control the *amplification* of errors while still assuring the benefits of minimized difference density.

The main source of errors is integral prescreening:  $\Delta G^{(n)}$  does *not* incorporate *all* two-electron integrals. The integral neglect threshold  $\text{THR} = \text{THR}(n)$  (in general a function of the iteration number  $n$ ) and the magnitude of the difference density matrix  $\Delta D^{(n)}$  both influence which integral batches will be neglected [Eqs. (6)–(9)].

With  $\Delta G_{\text{ex}}^{(n)}$  being the exact difference Fock matrix (built with the complete list of two-electron integrals) we define the error matrix  $\Delta \mathbf{e}^{(n)}$  of the actually calculated  $\Delta G^{(n)}$  as

$$\Delta \mathbf{e}^{(n)} := \Delta G_{\text{ex}}^{(n)} - \Delta G^{(n)}. \quad (30)$$

From Eq. (25), the similarly defined error matrix  $\mathbf{e}^{(n)}$  of  $G^{(n)}$  then turns out to be

$$\mathbf{e}^{(n)} = \Delta \mathbf{e}^{(n)} + \sum_{j=1}^{n-1} x_j^{(n)} \Delta \mathbf{e}^{(j)}. \quad (31)$$

For an estimate of this error we are interested in the quantity

$$\begin{aligned} E(n)^2 &:= \text{tr} \{ \mathbf{e}^{(n)} \mathbf{e}^{(n)} \} = \text{tr} \{ \Delta \mathbf{e}^{(n)} \Delta \mathbf{e}^{(n)} \} \\ &+ 2 \sum_{j=1}^{n-1} x_j^{(n)} \text{tr} \{ \Delta \mathbf{e}^{(n)} \Delta \mathbf{e}^{(j)} \} \\ &+ \sum_{i,j=1}^{n-1} x_i^{(n)} x_j^{(n)} \text{tr} \{ \Delta \mathbf{e}^{(i)} \Delta \mathbf{e}^{(j)} \}. \end{aligned} \quad (32)$$

The most reasonable assumption which can be made about the error matrices is, that

$$\text{tr} \{ \Delta \mathbf{e}^{(i)} \Delta \mathbf{e}^{(j)} \} \sim \text{THR}(i) \text{THR}(j) \delta_{ij}. \quad (33)$$

As a justification we note that the  $\Delta D^{(j)}$  are orthogonal if optimized ( $A_{ij} = A_{ii} \delta_{ij}$ ), and the integrals neglected behave rather statistically. One then obtains

$$\begin{aligned} E(n)^2 &\sim E'(n)^2 := \text{THR}(n)^2 \\ &+ \sum_{j=1}^{n-1} x_j^{(n)2} \text{THR}(j)^2. \end{aligned} \quad (34)$$

<sup>†</sup>Silicon Graphics Workstation IRIS-4D/60T.

For a constant integral neglect threshold  $\text{THR}(j) = \text{THR}$  and proceeding as in Eqs. (4) and (5) one finds

$$E'(n) = \sqrt{n} \text{THR}. \quad (35)$$

Now that we have a simple measure  $E'(n)$  of the error accumulated in  $\mathbf{G}^{(n)}$ , we want a simple tool for reducing  $E'(n)$  to a user-defined error bound  $B(n)$  if necessary.  $B(n)$  may be chosen as constant or as a stepwise decreasing function as convergence is approached.\* Usually  $E'(n) < B(n)$  (in fact, error control is not necessary if the integral neglect thresholds  $\text{THR}(j)$  are chosen small enough), but close to convergence it may happen that  $E'(n) > B(n)$  for the optimal  $\mathbf{x}$  [solution of Eq. (27)]. Then the minimal difference density matrix [Eq. (26), *not* (27)] must be found under the constraint  $E'(n) = B(n)$  by solving the equation

$$(\mathbf{A} - \mu\mathbf{T})\mathbf{x} = \mathbf{b}, \quad (36)$$

where  $T_{ij} = \delta_{ij} \text{THR}(i)^2$ , and  $\mu$  is a Lagrange multiplier corresponding to the constraint. There is one and only one solution  $\mu \leq 0$ , while solutions with  $\mu > 0$  are less optimal with respect to Eq. (26).

A calculation which utilizes minimized difference density matrices is shown in Table III. The molecule is  $\text{Cu}_3\text{N}_{15}\text{H}_6$  ( $\text{D}_{3h}$ -

symmetry, see Scheme 1), the basis set is a (14, 11, 6)/[8, 6, 4] for Cu, (8, 4, 1)/[4, 2, 1] for N, and (4)/[2] for H. Parameters for selective integral storage are  $\text{THIME} = 4$ ,  $\text{THIZE} = 10^{-5}$  ( $26 * 10^6$  integrals stored,  $180 * 10^6$  integrals processed in the first iteration). The total CPU time amounts to 35.4 h,\* while the same calculation<sup>†</sup> utilizing simple difference densities [Eqs. (4) and (5)] takes 43.3 h.\* Execution time is reduced by 18% at the same degree of accuracy. Time savings are less pronounced for smaller calculations: 8% for  $\text{C}_{14}\text{H}_{10}\text{O}_2$  (as mentioned before, depicted in Scheme 1).

## V. USE OF SYMMETRY

In the DSCF program symmetry is used

1. For the construction of a skeleton Fock matrix in each iteration using a "petit list" of integral batches over cartesian contracted GTFs according to Dupuis and King;<sup>13</sup>
2. For a symmetrizing transformation of this skeleton Fock matrix to symmetry-adapted basis functions.

The first point guarantees a fast assembly of the skeleton Fock matrix,<sup>13-15</sup> whereas the

**Table III.** Test calculation on  $\text{Cu}_3\text{N}_{15}\text{H}_6$  (375 contracted GTFs) utilizing minimized difference density matrices  $\Delta\mathbf{D}^{(n)}$  [eq. (23)].

Iteration	$10^{10}\text{THR}(n)$	$10^{10}E'(n)$	$\ \mathbf{D}^{(n)} - \mathbf{D}^{(n-1)}\ ^a$	$\ \Delta\mathbf{D}^{(n)}\ ^a$	Number of batches included in $\Delta\mathbf{G}^{(n)}$	CPU time <sup>b</sup>	Energy in hartree
1 <sup>c</sup>	1.00	1.00	10.3	10.3	4052507	5.00	-5735.2790092
2	0.66	1.16	2.14	2.08	3924700	3.24	-5736.3051323
4	0.43	1.62	$2.02 \cdot 10^{-1}$	$2.62 \cdot 10^{-2}$	3171290	2.48	-5736.3972022
6	0.32	2.51	$8.78 \cdot 10^{-2}$	$2.77 \cdot 10^{-3}$	2722035	2.02	-5736.4143577
8	0.25	3.90	$4.25 \cdot 10^{-2}$	$1.17 \cdot 10^{-4}$	2014553	1.39	-5736.4179374
10	0.21	6.51	$2.15 \cdot 10^{-2}$	$1.33 \cdot 10^{-5}$	1600121	1.16	-5736.4187748
12	0.17	11.7	$1.07 \cdot 10^{-2}$	$1.51 \cdot 10^{-6}$	1171392	0.93	-5736.4189969
14	0.14	19.8	$5.46 \cdot 10^{-3}$	$1.00 \cdot 10^{-6}$	1125174	0.90	-5736.4190628
16	0.13	30.4	$3.00 \cdot 10^{-3}$	$2.61 \cdot 10^{-7}$	834017	0.74	-5736.4190857
18 <sup>d</sup>	0.29	3.01 <sup>e</sup>	$8.18 \cdot 10^{-3}$	$1.06 \cdot 10^{-1}$	3541225	2.84	-5736.4191006
20	0.29	3.01 <sup>e</sup>	$8.76 \cdot 10^{-5}$	$1.46 \cdot 10^{-4}$	2102412	1.42	-5736.4191006

<sup>a</sup> $\|\mathbf{X}\| := (\text{tr}\mathbf{X}\mathbf{X})^{1/2}$ .

<sup>b</sup>CPU time in hours on Silicon Graphics workstation IRIS-4D/60T.

<sup>c</sup>Start MOs generated by basis set expansion 8 from MOs of a prior calculation with smaller basis set.

<sup>d</sup>DIIS convergence acceleration performed before this iteration.

<sup>e</sup> $E'(n) = B(n) = B = 3 * 10^{-10}$  enforced in iterations 18-20.

\*We have chosen  $B(n) \geq B_{\min} = 3 \text{THR}(1)$  and for  $\text{THR}(n)$  we have chosen a monotonic decreasing function of  $\sum_{j=1}^{n-1} x_j^{(n)2} \text{THR}(j)^2$  (the very details are insignificant), see Table III for an example.

\*Silicon Graphics Workstation IRIS-4D/60T.

<sup>†</sup> $\text{THR}(n) = 10^{-10}$  was set constant.

second point assures numeric stability and reduces memory requirements (symmetry blocking) for the subsequent processing of the Fock matrix.

The input necessary is the Schönfliess symbol (plus occupation numbers); all finite-point groups can be exploited.

The formula for the symmetrizing transformation is

$$\mathbf{X}^{(\Gamma)} = \frac{1}{\dim(\Gamma)} \sum_{\gamma=1}^{\dim(\Gamma)} \mathbf{V}^{(\gamma, \Gamma)\dagger} \mathbf{X}_{\text{skel}} \mathbf{V}^{(\gamma, \Gamma)} \quad (37)$$

where  $\mathbf{X}_{\text{skel}}$  is the (Hermitian) skelton matrix,  $\mathbf{V}^{(\gamma, \Gamma)}$  is the matrix of the symmetrized basis functions (which belong to the  $\gamma$ -th row of the real irreducible representation  $\Gamma^*$ ) expressed as column vectors in the cartesian basis, and  $\mathbf{X}^{(\Gamma)}$  is the corresponding (sub-)block of the symmetrized matrix. Note that for degenerate irreducible representations  $\Gamma$  only one out of  $\dim(\Gamma)$  identical subblocks  $\mathbf{X}^{(\Gamma)}$ , which are related to the first, second, ...,  $\dim(\Gamma)$ -th row of  $\Gamma$ , need to be calculated and stored.<sup>†</sup>

## VI. SUMMARY

The techniques to improve upon the direct SCF approach proposed in this article increase efficiency by varying degrees:

1. The more precise integral bond employed for prescreening reduces CPU time by about 15%.
2. Selective storage of two-electron integral batches allows for a continuous transition between standard and direct SCF techniques. This flexibility enables optimal usage of available computer hardware. CPU times can be typically reduced by 25–50%.
3. Minimization of  $\Delta D$  reduces CPU times by 5–18%. Our algorithm to control error accumulation stabilizes the calculation. We now routinely perform calculations

<sup>\*</sup>If the group contains complex conjugate representations we work with the reducible E-representations; formula (37) has to be modified in these cases to account for skew-symmetric off-diagonal subblocks, which relate the first with the second row of  $\Gamma$ .

<sup>†</sup>If  $\Gamma$  is reducible to two complex conjugate representations there is an extra off-diagonal block, which is skew-symmetric.

with 20 iterations (if necessary, e.g., for the first run of a series of calculations to optimize geometries), whereas other authors have recommended an update after about 10 iterations.<sup>8</sup>

The above estimates are conservative. The three effects enhance one another.

Finally, we want to comment on the comparison between standard SCF (based on a supermatrix formulation) and the present implementation of direct SCF. In test runs on  $\text{P}_5\text{H}_5$  in  $C_1$  symmetry (basis [6, 4, 1; 3, 1], a total of 145 CGTFs) the DSCF required 30% more CPU time; IO charge, however, was reduced by a factor of 9 since only 12% of the integrals were stored. These runs required 7 iterations which is typical in geometry optimizations where good starting MOs are available.

The authors thank Stefan Brode, Christoph Kölmel, and Dr. David Swanton for valuable help. The financial support by the Fonds der Chemischen Industrie is gratefully acknowledged.

In reply to the suggestion of a referee we point out that the Cremer/Gauß estimate<sup>8</sup> is a bond only for s-type CGTFs. In this case it is identical to our bond [eq. (10)] and to the original bond by Ahlrichs.<sup>10</sup>

## References

1. C. C. J. Roothaan, *Rev. Mod. Phys.*, **23**, 69 (1951).
2. G. G. Hall, *Proc. Roy. Soc.*, **A205**, 541 (1951).
3. C. C. J. Roothaan, *Rev. Mod. Phys.*, **32**, 179 (1960).
4. S. Huzinaga, *J. Chem. Phys.*, **42**, 1293 (1965).
5. J. Almlöf, K. Faegri Jr., and K. Korsell, *J. Comput. Chem.*, **3**, 385 (1982).
6. H. P. Luethi and J. Almlöf, *Chem. Phys. Lett.*, **135**, 357 (1987).
7. P. Pulay, *Chem. Phys. Lett.*, **73**, 393 (1980).
8. D. Cremer and J. Gauss, *J. Comput. Chem.*, **7**, 274 (1986).
9. M. Dupuis, J. Rys, and H. F. King, *J. Chem. Phys.*, **65**, 111 (1976).
10. R. Ahlrichs, *Theor. Chim. Acta*, **33**, 157 (1974).
11. D. Hegarty, in *Advanced Theories and Computational Approaches to the Electronic Structure of Molecules*, C. E. Dykstra, Ed., D. Reidel Publishing Company, Dordrecht, Holland 1984.
12. D. Hegarty and G. van der Velde, *Int. J. Quant. Chem.*, **23**, 1135 (1983).
13. M. Dupuis and H. F. King, *Int. J. Quant. Chem.*, **11**, 613 (1977).
14. P. D. Dacre, *Chem. Phys. Lett.*, **7**, 47 (1970).
15. M. Elder, *Int. J. Quant. Chem.*, **7**, 75 (1973).