JOURNAL OF COMPUTATIONAL PHYSICS **44**, 20–30 (1981)

Several Eigenvectors of a Large Matrix. II. The Generalized Eigenvalue Problem

J. A. R. COOPE AND D. W. SABO

Department of Chemistry, University of British Columbia, Vancouver, British Columbia V6T 1W5, Canada

Received May 22, 1979

1. INTRODUCTION

Recently [1], an eigenvalue independent partitioning [2] approach was proposed for calculating, simultaneously, several eigenvalues and eigenvectors of a large Hermitian matrix. Within the context of this approach, several, related, simple algorithms were considered for the ordinary eigenvalue problem, $HX = XA$, for the case where there is a large variation in diagonal elements compared to individual off-diagonal elements, and the results of calculations with these were reported. The two algorithms referred to as "generalized Nesbet" were extended to apply to the generalized eigenvalue problem

$$HX = SXA, \tag{1}$$

where H and S are Hermitian, and S positive definite. However, no calculations with these were reported. Moreover no generalization was given of the simplest, and apparently most satisfactory algorithm of those considered, that referred to as "simple diagonal Newton–Raphson" (SDNR). This note describes an appropriate generalization of this algorithm to the generalized eigenvalue problem (1), and reports some results of test calculations with this, and with the generalized Nesbet algorithms described previously. Further details of the general approach, and comparison to other approaches, can be found in [1].

2. THE GENERALIZED ALGORITHM

In eigenvalue independent partitioning the basis vectors, and the eigenvectors, are partitioned into two sets, respectively, of dimensions n_A and $n_B = n - n_A$, where n is the dimension of the eigenvalue equation. With respect to this partitioning, the matrices in (1) take the form

$$H = \begin{bmatrix} H_{AA} & H_{AB} \\ H_{BA} & H_{BB} \end{bmatrix}, \qquad S = \begin{bmatrix} S_{AA} & S_{AB} \\ S_{BA} & S_{BB} \end{bmatrix}, \qquad X = \begin{bmatrix} X_{AA} & X_{AB} \\ X_{BA} & X_{BB} \end{bmatrix}. \tag{2}$$

0021-9991/81/110020–11$02.00/0

The fact that the eigenvector matrix X, is non-singular ensures that X_{AA} and X_{BB} are non-singular for at least one partitioning of the basis, so that X can be factored in the manner

$$X = T\hat{X}, \tag{3}$$

with

$$\hat{X} = \begin{bmatrix} X_{AA} & 0 \\ 0 & X_{BB} \end{bmatrix}, \tag{4}$$

and

$$T = \begin{bmatrix} 1_A & h \\ f & 1_B \end{bmatrix}, \tag{5}$$

where $f = X_{BA}X_{AA}^{-1}$ and $h = X_{AB}X_{BB}^{-1}$. The generalized eigenvalue equation (1) can accordingly be rewritten as

$$G\hat{X} = g\hat{X}\Lambda, \tag{6}$$

where

$$G = T^{\dagger}HT, \qquad g = T^{\dagger}ST, \tag{7}$$

$T^{\dagger}$ being the adjoint.

The matrices f and h are to be determined to make the off-diagonal blocks of G and g zero:

$$\begin{aligned} G_{BA} &= H_{BA} + H_{BB}f + h^{\dagger}(H_{AA} + H_{AB}f) = 0, \\ g_{BA} &= S_{BA} + S_{BB}f + h^{\dagger}(S_{AA} + S_{AB}f) = 0, \end{aligned} \tag{8}$$

The original n-dimensional eigenvalue equation is thereby reduced to two eigenvalue equations of the smaller dimensions n_A and n_B. One of these,

$$G_A X_{AA} = g_A X_{AA} \Lambda^{(A)}, \tag{9}$$

where G_A and g_A are the diagonal blocks in (7),

$$\begin{aligned} G_A &= H_{AA} + H_{AB}f + f^{\dagger}H_{BA} + f^{\dagger}H_{BB}f, \\ g_A &= S_{AA} + S_{AB}f + f^{\dagger}S_{BA} + f^{\dagger}S_{BB}f, \end{aligned} \tag{10}$$

can be solved for the n_A eigenvalues of interest, together with the projected eigenvectors X_{AA}. The remaining components of the eigenvectors are then given, according to Eq. (3), by $X_{BA} = fX_{AA}$.

A single condition on f, namely

$$D^{(1)}(f) = H_{BA} + H_{BB}f - (S_{BA} + S_{BB}f)\,\hat{H}_A^{(1)} = 0, \tag{11}$$

where $\hat{H}_A^{(1)} = (S_{AA} + S_{AB}f)^{-1}(H_{AA} + H_{BB}f)$, can be obtained on eliminating h between the two conditions (8). The "diagonal," and "full" generalized Nesbet alorithms (DGNS[1] and FGNS), developed previously [1], are based on a similar condition, $D^{(2)}(f) = 0$, obtained by replacing $\hat{H}_A^{(1)}$ in (11), by a more accurate estimate, $\hat{H}_A^{(2)} = g_A^{-1}G_A$. However, a simpler algorithm can be based directly on the pair of conditions (8), treating both f and h as unknowns, and this is the approach developed here.

The generalized Newton–Raphson equations corresponding to Eqs. (8) are

$$\begin{aligned} \bar{H}_B^{\dagger}\delta f + \delta h^{\dagger}\bar{H}_A &= -G_{BA}, \\ \bar{S}_B^{\dagger}\delta f + \delta h^{\dagger}\bar{S}_A &= -g_{BA}, \end{aligned} \tag{12}$$

where

$$\begin{aligned} \bar{H}_B &= H_{BB} + H_{BA}h, \qquad & \bar{H}_A &= H_{AA} + H_{AB}f, \\ \bar{S}_B &= S_{BB} + S_{BA}h, \qquad & \bar{S}_A &= S_{AA} + S_{AB}f. \end{aligned} \tag{13}$$

On retaining only the diagonal parts, in order to find the individual changes $\delta f_{\sigma r}$, $\delta h_{r\sigma}$ which reduce the individual elements $G_{\sigma r}$ and $g_{\sigma r}$ approximately to zero, Eqs. (12) reduce to

$$\begin{bmatrix} (\bar{H}_B^{\dagger})_{\sigma\sigma} & (\bar{H}_A)_{rr} \\ (\bar{S}_B^{\dagger})_{\sigma\sigma} & (\bar{S}_A)_{rr} \end{bmatrix} \begin{bmatrix} \delta f_{\sigma r} \\ \delta h_{r\sigma}^{*} \end{bmatrix} = - \begin{bmatrix} G_{\sigma r} \\ g_{\sigma r} \end{bmatrix}. \tag{14}$$

Here Greek letters label basis elements in the B-space, italic letters those in the A-space. Solution of these simultaneous equations gives the iteration formulas

$$\begin{aligned} \delta f_{\sigma r} &= [G_{\sigma r}(\bar{S}_A)_{rr} - g_{\sigma r}(\bar{H}_A)_{rr}]/\Delta_{\sigma r}, \\ (\delta h^{\dagger})_{\sigma r} &= -[G_{\sigma r}(\bar{S}_B^{\dagger})_{\sigma\sigma} - g_{\sigma r}(\bar{H}_B^{\dagger})_{\sigma\sigma}]/\Delta_{\sigma r}, \end{aligned} \tag{15}$$

where

$$\Delta_{\sigma r} = (\bar{S}_B^{\dagger})_{\sigma\sigma}(\bar{H}_A)_{rr} - (\bar{H}_B^{\dagger})_{\sigma\sigma}(\bar{S}_A)_{rr}. \tag{16}$$

Only the diagonal elements of the large matrixes $\bar{H}_B$ and $\bar{S}_B$ are required. The auxiliary quantities, $\bar{H}_A$, $\bar{S}_A$, $\bar{H}_B$, $\bar{S}_B$, being linear in f or h, are easily updated as the calculation proceeds, element by element through f and h. Precise computational details are given in the appendix. The algorithm is referred to here as "Simple

[1] In [1], $f_{A\sigma}^{(\mathrm{new})\dagger}$ in Eq. (3.17) for DGN should read $-f_{A\sigma}^{(\mathrm{new})\dagger}$, and $Y_{A\sigma}^{(\mathrm{new})\dagger}$ in (4.13) for DGNS should read $-Y_{A\sigma}^{(\mathrm{new})\dagger}$. The same sign changes are required in the formula for A_{rs} in Appendices 2 and 4.

Diagonal Newton–Raphson with Overlap" (SDNRS). For $n_B \gg n_A$, of the order $2n_A n_B^2$ computational operations are required per iterative sweep through the elements of f and h. This is twice the number of operations required for an ordinary eigenvalue equation, but the same order as for algorithms DGNS and FGNS [1].

A refinement of the algorithm is obtained on replacing the Newton–Raphson equations (12) by the exact equations for δf and δh. These are simply

$$\begin{aligned} \bar{H}_B \delta f + \delta h^\dagger \bar{H}_A + \delta h^\dagger H_{AB} \delta f &= -G_{BA}, \\ \bar{S}_B \delta f + \delta h^\dagger \bar{S}_A + \delta h^\dagger S_{AB} \delta f &= -g_{AB}. \end{aligned} \tag{17}$$

The same diagonal approximation made before leads to a pair of simultaneous quadratic equations for $\delta h_{\sigma r}$ and $\delta h_{\sigma r}$, from which a quadratic equation for $\delta f_{\sigma r}$ is obtained

$$A \delta f_{\sigma r}^2 + B \delta f_{\sigma r} + C = 0, \tag{18}$$

with coefficients

$$\begin{aligned} A &= H_{r\sigma}(\bar{S}_B^\dagger)_{\sigma\sigma} - S_{r\sigma}(\bar{H}_B^\dagger)_{\sigma\sigma}, \\ B &= \Delta_{\sigma r} + H_{r\sigma} G_{\sigma r} - S_{r\sigma} G_{\sigma r}, \\ C &= g_{\sigma r}(\bar{H}_A)_{rr} - G_{\sigma r}(\bar{S}_A)_{rr}. \end{aligned} \tag{19}$$

The desired correction $\delta f_{\sigma r}$ is the root of smallest magnitude. Given this, the correction $\delta h_{r\sigma}$ can be found from

$$(\delta h^\dagger)_{\sigma r} = -[G_{\sigma r} + (\bar{H}_B^\dagger)_{\sigma\sigma} \delta f_{\sigma r}] / [(\bar{H}_A)_{rr} + H_{r\sigma} \delta f_{\sigma r}]. \tag{20}$$

This refinement, denoted QDNRS here (Q = quadratic), may exhibit a significantly different rate of convergence from SDNRS initially, that is when δf and δh still have large elements. The distinction disappears as the solution is approached. The smallest root of Eq. (18) then tends to $\delta f = -C/B \simeq -C/\Delta$ which, correctly, coincides with Eq. (15). For $n_B \gg n_A$, QDNRS still represents of the order of $2n_A n_B^2$ computational operations per sweep through f and h.

3. Application to a Model Problem

A series of calculations were carried out using the four algorithms SDNRS, QDNRS, DGNS, and FGNS. The matrix H in Eq. (1) was of the type used in previous calculations [1, 3], in which *all off-diagonal elements are unity*. The diagonal elements are chosen to give various distributions of eigenvalues. Permutations of these diagonal elements correspond to changes in the partitioning of the basis space. The results reported here are for relatively small matrices. Unpublished calculations with matrices up to order 12000, for $S = 1_n$, have shown

that the asymptotic convergence rates decrease only slowly for increasing n, and appear to level out for large n.

The matrix S arises as the overlap matrix (Gram matrix, metric tensor), $S_{ij} = (\phi_i, \phi_j)$, of some underlying set of basis vectors and, as such, must be positive. The following useful model matrix was used here:

$$S(\alpha) = \begin{bmatrix} 1 & \alpha & \alpha^2 & \cdots & \alpha^{n-1} \\ \alpha & 1 & \alpha & \cdots & \alpha^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ \alpha^{n-1} & \alpha^{n-2} & \alpha^{n-3} & \cdots & 1 \end{bmatrix}. \tag{21}$$

This is positive definite for all $|\alpha| < 1$. It resembles the quantum mechanical overlap matrix for a linear chain of atoms, having overlaps falling off exponentially, with distance ($S_{ij} = \alpha^{|i-j|}$), and it also serves to model a quantum mechanical configuration interaction calculation having a non-orthogonality which decreases with energy differences. For $\alpha = 0$ the orthonormal case ($S = 1_n$) is recovered, while as α approaches the maximum value unity, the eigenvalue equation becomes highly ill conditioned, becoming singular for $|\alpha| = 1$.

The ill-conditioning of the equations is caused by the approach of the underlying basis vectors to linear dependence, as $|\alpha|$ increases, and is measured by the approach of the eigenvalues of S to zero. For the matrix (21), zero eigenvalues can only occur for $|\alpha| = 1$. The positive definiteness for $|\alpha| < 1$ follows by continuity from $\alpha = 0$. All but one of the eigenvalues is approaching zero, however, as $|\alpha| \to 1$, and for $|\alpha| = 1$, S is of rank 1, and the underlying n-dimensional basis space has degenerated to a one-dimensional space.

For large n, the distribution of the eigenvalues of (21) which is independent of the sign of α, is essentially the same as for the corresponding circulant matrix [5], of the same dimension. This models a quantum mechanical ring of atoms (with periodic boundary conditions) instead of a linear chain. For $n \to \infty$ these eigenvalues are given (for $|\alpha| < 1$) by

$$\begin{aligned} S_j &= 1 + 2\alpha \operatorname{Cos} \theta_j + 2\alpha^2 \operatorname{Cos} 2\theta_j + ..., \\ &= (1 - \alpha^2)/[1 - 2\alpha \operatorname{Cos} \theta_j + \alpha^2], \end{aligned} \tag{22}$$

where $\theta_j = 2\pi j/n$, $(j = 1, 2, ..., n)$. They range from $S_{\min} = (1 - |\alpha|)/(1 + |\alpha|)$, to $S_{\max} = 1/S_{\min}$, being concentrated towards the lower limit. Figure 1 illustrates the distribution for $\alpha = 0.6$.

Table 1 lists asymptotic convergence rates for a selection of these calculations, that is, the average factors by which appropriate convergence monitors decrease per iteration, once linear convergence is established.[2] The smaller these numbers the

[2] Since for $n_B \gg n_A$, a single iteration in all these algorithms represents of the order of $2n_A n_B^2$ computations, these asymptotic convergence rates represent relative cpu time requirements, and have

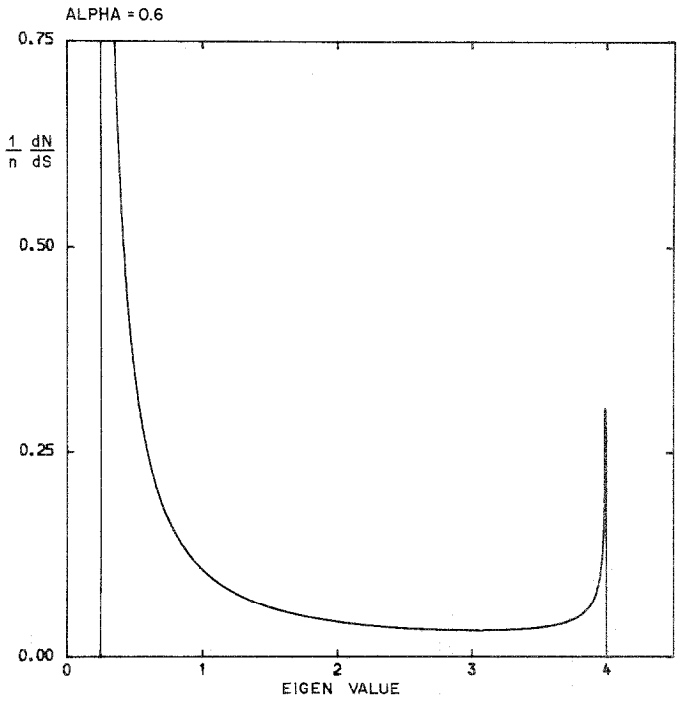

FIG. 1. Normalized distribution of eigenvalues for the overlap matrix (21), with overlap parameter $\alpha = 0.6$

faster the convergence. The l_2 convergence monitors were $\|D\|_2 = (\sum_{\sigma,r}|D_{\sigma,r}|^2)^{1/2}$ for the generalized Nesbet algorithms, and $(\|G_{BA}\|_2^2 + a\|g_{BA}\|_2^2)^{1/2}$ for SDNRS and QDNRS. The constant a, which could be considered necessary, in some applications, for dimensional consistency, was here chosen as unity throughout. No results are given in the table for QDNRS, since its performance was virtually identical to that of SDNRS in the cases presented.

In all cases reported, the starting approximation was $f = h = 0$. The use, instead, of the exact f and h for an m-dimensional truncated eigenvalue problem ($n_A < m \ll n$), as the starting approximation for the n-dimensional problem, was of little if any value in improving convergence. The use of the exact $f(=-h^\dagger)$ for the eigenvalue equation without overlap, i.e., for $S = 1_n$, was also of little value. A detailed analysis [4] of the convergence is consistent with these observations, predicting that the accuracy of the starting approximation normally has little bearing on the final convergence rate

intrinsic significance. Absolute cpu times are not given since, with a variety of computers, programming languages, and compilers now in use, timings for a particular program on a particular machine have relative meaning only. They are subject to non-quantifiable influences such as programming style, compiler design, and machine architecture, which have litle relation to the underlying mathematical features of the calculation being done.

TABLE I

Linear[a] Convergence Rates of the Algorithms in Selected Calculations[b]

Diagonal matrix elements in order	α[c]	Method[d] SDNRS $_*$		DGNS $_*$		FGNS $_*$	
1, 3, 5, 7,..., 39	0.0	0.29		0.37		0.29	
1, 3, 5, 7,..., 39	0.1	0.29	0.18	0.22	0.13	0.28	0.17
1, 3, 5, 7,..., 39	0.2	0.46	0.22	0.57	0.27	0.33	0.22
1, 3, 5, 7,..., 39	0.4	0.71	0.33	0.90	0.37	0.98[e]	0.32
1, 3, 5, 7,..., 39	0.6	f	0.41	0.90[e]	0.60	g	0.59
1, 3, 5, 7,..., 39	0.7	f	0.52	h	0.89	f	0.88
1, 3, 5, 7,..., 39	0.8	f	0.51	f	0.99[i]	f	0.99[i]
1, 3, 5, 7,..., 39	0.9	f	f	f	f	f	1.00[i]
1, 3,..., 11, 9,..., 37, 39	0.2	0.34	0.26[j]	0.72[j]	0.27[j]	0.63[j]	0.27[j]
1, 3,..., 11, 9, 7, 13,..., 37, 39	0.2	f	0.70[k]	1.00[i]	0.95[k]	0.97[k]	0.95[k]
1, 3,..., 11, 13, 7, 9,..., 37, 39	0.2	0.45[l]	0.34[m]	0.97[m]	0.80[m]	0.97[m]	0.73[m]
1, 3,..., 13, 11, 9, 7,..., 37, 39	0.2	f	0.28[m]	0.95[m]	0.63[m]	0.95[m]	0.57[m]
1, 1.1, 1.2, 1.3, 1.4, 11, 13,...,	0.2	0.18	0.12	0.14	0.11	0.14	0.11
1, 1.1, 1.2, 1.3, 1.4, 11, 13,...,	0.4	0.24	0.11	0.13	0.10	0.13	0.10
1, 1.1, 1.2, 1.3, 1.4, 3, 5, 7,...,	0.2	0.40	0.15	0.43	0.20	0.33	0.20
1, 1.1, 1.2, 1.3, 1.4, 3, 5, 7,...,	04	0.36	0.18	0.28	0.16	0.31	0.16
1, 1.1, 1.2, 1.3, 1.4, 1.5, 13, 15,...,	0.2	0.19	0.16	f	0.16	0.93	0.16
1, 1.1, 1.2, 1.3, 1.4, 1.5, 13, 15,...,	0.4	0.34	0.25	0.33	0.25	0.94[n]	0.25
1, 1.1, 1.2, 1.3, 1.4, 1.5, 3.5,...,	0.2	0.48	0.25	0.95[o]	0.25	p	0.25
1, 1.1, 1.2, 1.3, 1.4, 1.5, 3.5,...,	0.4	0.57	0.38	0.50	0.38	n	0.38

[a] From a least-squares calculation of slope of logarithm of the cnvergence monitor as a function of iteration number. These numbers are the average factor by which the convergence monitor decreases per iteration.

[b] All off-diagonal elements of H are unity. Unless otherwise indicated, convergence is to the eigenspace of the n_A lowest eigenvalues. In all these calculations $n_A = 5$, $n = 20$.

[c] The elements of the metric matrix in the generalized eigenvalue equation are $S_{ij} = \alpha^{|i-j|}$.

[d] In the calculations denoted by *, the iterative evaluation of f and h was done after a 10×10 prediagonalization.

[e] Convergence is apparently to the four lowest, and the highest eigenvalue, but is very slow. The behaviour is as if the largest eigenvalue of S had passed through infinity and became negative.

[f] Calculation is divergent, the convergence monitor increases in most of the first 50 iterations.

[g] The iterative calculation is oscillatory, no convergence in 50 iterations.

[h] The iterative calculation is divergent even if started with the exact solution ($\|D^{(2)}\| < 10^{-12}$).

[i] Calculations in wich these convergences rates are near 1.00 are for practical purposes non-convergent, since hundreds or thousands of iterations may be required to obtain 2 or 3 figure accuracy. Also, these results are based only on the first 50 iterations, so that long shallow oscillation can show up here as initial slow convergence.

[j] Converges to eigenvalues 1, 2, 3, 4, and 6.

[k] Apparent convergence 1, 2, 3, 5, and 6, but very slow.

[l] Linear convergence was established in this case only after 31 iterations.

[m] Apparent convergence to eigenvalues 1, 2, 3, 6, and 7.

[n] Apparent convergence to eigenvalues 1, 2, 3, 4, and 7.

[o] Apparent convergence to eigenvalues 1, 2, 3, 4, and 9, but very slow.

[p] Apparent convergence to eigenvalues 1, 2, 3, 4, and 8, but very slow.

established. Calculation 6 of Table I gives an instance in which DGNS is divergent even when started with $\|D\|_2 < 10^{-12}$, i.e., virtually at the exact solution, indicating that the norm of the first order error constant matrix in this case, and for the basis used, is greater than unity at the exact solution.

On the other hand significant improvement in convergence may be obtained by a preliminary partial diagonalization. In contrast to a mere change of starting approximation, this is a change of basis, in which m basis vectors are replaced by the eigenvectors of an m-dimensional truncated eigenvalue problem. The m-dimensional blocks of H and S include, but are larger than, H_{AA} and H_{BB}. These calculations are listed under an asterisk in Table I. Theoretical considerations suggest that improvement in asymptotic convergence rates is likely only if m is a substantial fraction of n. However, the initial behaviour of the calculation may be greatly improved even for small m. For $n_A \ll m \ll n_B$ the transformation to partially diagonal form, and subsequent back transformation of f, requires only of the order of $m^2 n_B$ computational operations, a fraction of the number for one iteration $(n_A n_B^2)$, so that it is practicable to take $m \sim (n_A n_B)^{1/2}$.

While rates of convergence decrease only slowly with increasing n_A and n, they decrease rather rapidly with increasing overlap parameter α. For calculations of the lowest n_A eigenvalues, the effective upper limit for convergence to occur, without preliminary partial diagonalization, is $\alpha = 0.6$. Sets of n_A eigenvalues other than the n_A lowest or highest, can be obtained by appropriate partitioning of the basis, equivalent here to a reordering of the diagonal elements, but as α increases beyond 0.2 the convergence of these calculations becomes problematical, especially when non-consecutive subsets of eigenvalues are sought (calculations 9–12 in Table I). Preliminary partial diagonalization is particularly useful in these cases in improving convergence rates, and especially initial behaviour.

The last eight calculations in Table 1 illustrate the point that near degeneracy of eigenvalues presents no problem whatever, provided all nearly degenerate, or degenerate eigenvalues are included in the set of n_A eigenvalues calculated (or in the complement). In contrast, the corresponding calculations with Nesbet's algorithm for the *single* lowest eigenvalue, to which the generalized Nesbet algorithms reduce in the case $n_A = 1$, do not converge. The practical advantages of determining an entire set of nearly equal eigenvalues simultaneously, rather than one at a time successively, have been well illustrated for the ordinary eigenvalue problem as well, both through our approach [1] and in recent work of Liu [6].

When convergence is straightforward, that is, when α is small, and the eigenvalues to be calculated are well separated from the other eigenvalues, these several algorithms exhibit similar rates of convergence. However, for the calculations in Table 1 it is seen that the algorithm outlined in this note (SDNRS) is generally more reliable than the other two. One reason is probably the more frequent updating of the very simple quantities appearing in the iteration formulas. While SDNRS is simple and requires little storage, it is clear that more powerful algorithms are possible, within he general paritioning approach since the knowledge of the successive calculated correction "vectors" δf and δh or, alternatively, knowledge of the

successive calculated residual "vectors" G_{BA} and g_{BA}, is not being fully exploited in the algorithms considered here. The exploitation of the knowledge of successive vectors is a feature of Davidson's [6, 7] powerful method. We have formulated a number of such extrapolation strategies, but comparative tests of various alternatives have not yet been carried out.

APPENDIX: COMPUTATIONAL DETAILS

Simple Diagonal Newton–Raphson Algorithm with Overlap (SDNRS)

Initialization:

$$f = 0, \qquad h^{\dagger} = 0,$$
$$\bar{H}_A = H_{AA}, \qquad \bar{S}_A = S_{AA},$$
$$\bar{H}_B^{\dagger \mathrm{diag}} = H_{BB}^{\mathrm{diag}}, \qquad \bar{S}_B^{\dagger \mathrm{diag}} = S_{BB}^{\mathrm{diag}}.$$

Then:

$$\left.\begin{aligned}
G_{\sigma r} &= H_{\sigma r} + \sum_{\rho=1}^{n_B} H_{\sigma\rho} f_{\rho r} + \sum_{s=1}^{n_A} h_{\sigma s} (\bar{H}_A)_{sr}, \\
g_{\sigma r} &= S_{\sigma r} + \sum_{\rho=1}^{n_B} S_{\sigma\rho} f_{\rho r} + \sum_{s=1}^{n_A} h_{\sigma s}^{\dagger} (\bar{S}_A)_{sr}, \\
\Delta_{\sigma r} &= (\bar{S}_B^{\dagger})_{\sigma\sigma} (\bar{H}_A)_{rr} - (\bar{H}_B^{\dagger})_{\sigma\sigma} (\bar{S}_A)_{rr}, \\
\delta f_{\sigma r} &= [G_{\sigma r} (\bar{S}_A)_{rr} - g_{\sigma r} (\bar{H}_A)_{rr}]/\Delta_{\sigma r}, \\
\delta h_{\sigma r}^{\dagger} &= -[G_{\sigma r} (\bar{S}_B^{\dagger})_{\sigma\sigma} - g_{\sigma r} (\bar{H}_B^{\dagger})_{\sigma\sigma}]/\Delta_{\sigma r}.
\end{aligned}\right\} \quad \begin{aligned} r &= 1,\ldots, n_A, \\ \sigma &= 1,\ldots, n_B. \end{aligned}$$

up date:

$$\left.\begin{aligned}
(\bar{H}_A)_{sr} &\to (\bar{H}_A)_{sr} + H_{s\sigma}\, \delta f_{\sigma r} \\
(\bar{S}_A)_{sr} &\to (\bar{S}_A)_{sr} + S_{s\sigma}\, \delta f_{\sigma r}
\end{aligned}\right] s = 1,\ldots, n_A,$$
$$(\bar{H}_B^{\dagger})_{\sigma\sigma} \to (\bar{H}_B^{\dagger})_{\sigma\sigma} + \delta h_{\sigma r}^{\dagger} H_{r\sigma},$$
$$(\bar{S}_B^{\dagger})_{\sigma\sigma} \to (\bar{S}_B^{\dagger})_{\sigma\sigma} + \delta h_{\sigma r}^{\dagger} S_{r\sigma},$$
$$f_{\sigma r} \to f_{\sigma r} + \delta f_{\sigma r},$$
$$h_{\sigma r}^{\dagger} \to h_{\sigma r}^{\dagger} + \delta h_{\sigma r}^{\dagger}.$$

ACKNOWLEDGMENT

The support of the National Research Council of Canada is greatfully acknowledged.

REFERENCES

1. J. A. R. COOPE AND D. W. SABO, *J. Comp. Phys.* **23** (1977), 404.
2. J. A. R. COOPE, *Mol. Phys.* **18** (1970), 571.
3. R. K. NESBET, *J. Chem Phys.* **43** (1965), 311.
4. J. A. R. COOPE AND D. W. SABO, to appear.
5. D. E. RUTHERFORD, *Roy. Soc. of Edinburgh, Proc. A.* **63** (1949/52), 232.
6. B. LIU, NRC Report LBL-8158/UC/CONF-780878, "Numerical Algorithms in Chemistry: Algebraic Methods," Lawrence Berkeley Laboratory, University of California, Berkeley, p. 49.
7. E. R. DAVIDSON, *J. Comp. Phys.* **17** (1975), 87.