An Alternative to Full Configuration Interaction Based on a Tensor Product Decomposition

Frederick A. Senese, Christopher A. Beattie,
John C. Schug, Jimmy W. Viers,
and Layne T. Watson

TR 89-14

1.

April 11, 1989

An Alternative to Full Configuration Interaction Based on a Tensor Product Decomposition

Frederick A. Senese†, Christopher A. Beattie‡, John C. Schug†, Jimmy W. Viers†,
and Layne T. Watson+
Departments of Chemistry, Mathematics, and Computer Science
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061 USA

Abstract.

A new direct full variational approach exploits a tensor (Kronecker) product decomposition of the Hamiltonian. Explicit assembly and storage of the Hamiltonian matrix is avoided by using the Kronecker product structure to form matrix-vector products directly from the molecular integrals. Computation-intensive integral transformations and formula tapes are unnecessary. The wavefunction is expanded in terms of spin-free primitive kets rather than Slater determinants or configuration state functions and is equivalent to a full configuration interaction expansion. The approach suggests compact storage schemes and algorithms which are naturally suited to parallel and pipelined machines.

1. Introduction.

A central problem in theoretical chemistry is the determination of approximate eigenvalues and eigenfunctions of spin-free Hamiltonian operators for many-electron systems. Many approaches to this problem have made use of the Hartree-Fock approximation followed by either truncated configuration interaction (C. I.), many-body perturbation theory, or a combination of the two [1–3]. However, the only post-Hartree-Fock method that is at once size consistent, extensive, and variational is a full Rayleigh-Ritz treatment [2, 4–6]. In this paper we discuss a new approach that is equivalent to full configuration interaction.

Our method is a variational treatment that is based on a tensor-product decomposition of the many-electron space [7, 8]. This leads to a computational procedure that does not require an orthogonal orbital basis. Integral transformations are avoided by using wavefunctions that are expanded in terms of primitive orbital products rather than Slater determinants or configuration state functions. Matsen's structure projectors [9] are utilized to restrict the trial vectors to physically realizable subspaces. This simultaneously reduces the dimension of the problem and suggests an efficient storage scheme for trial vectors. The method is generally applicable to ground or excited electronic states, open or closed shells, and systems of arbitrary spatial symmetry. Most importantly, the method leads to algorithms that are well-suited to parallel and vector machine architectures.

---

† Department of Chemistry
‡ Department of Mathematics
+ Department of Computer Science

2. The Tensor Product Construction.

Within the Born-Oppenheimer approximation the electronic part of the nonrelativistic Hamiltonian can be written

$$\mathbf{H} = \sum_{i=1}^{N} \hat{\mathrm{h}}(i) + \sum_{i<j}^{N} \hat{\mathrm{g}}(i,j),$$

where the one- and two-electron operators $\hat{\mathrm{h}}(i)$ and $\hat{\mathrm{g}}(i,j)$ are usually spin-free so that separation of spin and spatial variables is possible. The one-electron operators $\hat{\mathrm{h}}(i)$ include the electronic kinetic energy and nuclear attraction operators

$$\hat{\mathrm{h}}(i) = -\frac{1}{2}\nabla_i^2 - \sum_{a}^{N_a} \frac{Z_a}{|r_i - R_a|},$$

where r_i and R_a are the position vectors of the i-th electron and the a-th nucleus, respectively. The pairwise operators $\hat{\mathrm{g}}(i,j)$ are the usual Coulombic repulsion operators $|r_i - r_j|^{-1}$. If we use a basis of orbital functions $\phi_i, i = 1, \ldots, m$ that span a model space V of dimension m, the primitive N-particle functions

$$|i\rangle = |i_1, \ldots, i_N\rangle = \phi_{i_1}(1)\phi_{i_2}(2)\cdots\phi_{i_N}(N), \quad 1 \le i_j \le m,\ 1 \le j \le N,$$

span an N-fold tensor product space

$$V^{[N]} = V_1 \otimes \cdots \otimes V_N$$

of dimension m^N. $V^{[N]}$ contains all linear combinations of the elementary tensor products $c = c_1 \otimes c_2 \otimes \cdots \otimes c_N$ with $c_i \in V_i$, $i = 1, \ldots, N$.

On $V^{[N]}$ the i-th one electron operator can be written

$$\hat{I}_1 \otimes \cdots \otimes \hat{I}_{i-1} \otimes \hat{\mathrm{h}}(i) \otimes \hat{I}_{i+1} \otimes \cdots \otimes \hat{I}_N,$$

where $\hat{I}_j$ is the identity operator in V_j. The restriction of the bare nucleus Hamiltonian to $V^{[N]}$ has the matrix representation

$$h \otimes s \otimes \cdots \otimes s + s \otimes h \otimes s \otimes \cdots \otimes s + \cdots + s \otimes \cdots \otimes s \otimes h,$$

where h is the one-electron core Hamiltonian matrix $[\langle\phi_i \mid \hat{\mathrm{h}} \mid \phi_j\rangle]$ and s is the one-electron overlap matrix $[\langle\phi_i \mid \phi_j\rangle]$. Similarly each two-electron operator $\hat{\mathrm{g}}(i,j)$ may be first restricted to $V^{[N]}$ and then viewed as an operator on $V_i \otimes V_j$. For example, $\hat{\mathrm{g}}(1,2)$ when restricted to functions in $V^{[N]}$ can be represented as

$$\hat{\mathrm{g}}(1,2) \otimes \hat{I}_3 \otimes \cdots \otimes \hat{I}_N,$$

where $\hat{\mathrm{g}}(1,2)$ acts nontrivially only in the space $V_1 \otimes V_2$. If we define g to be the matrix of two-electron integrals $\langle ij \mid kl\rangle$, with ϕ_i, ϕ_k in V_1 and ϕ_j, ϕ_l in V_2, then the restriction of $\hat{\mathrm{g}}(1,2)$ to $V^{[N]}$ has the matrix representation

$$[g \otimes s \otimes \cdots \otimes s],$$

where $N-2$ copies of s appear in the tensor product. To model the general two-electron operators $\hat{\mathrm{g}}(k,l)$ the unitary mapping P_{kl} is introduced for notational convenience [10]:

$$P_{kl}\,\phi_{i_1}(1)\phi_{i_2}(2)\cdots\phi_{i_N}(N) = \phi_{i_k}(1)\phi_{i_l}(2)\phi_{i_1}(3)\phi_{i_2}(4)\cdots$$
$$\cdots\phi_{i_{k-1}}(k+1)\phi_{i_{k+1}}(k+2)\cdots\phi_{i_{l-1}}(l)\phi_{i_{l+1}}(l+1)\cdots\phi_{i_N}(N).$$

Then as before, on $V^{[N]}$ $\hat{\mathrm{g}}(k,l)$ can be written

$$P_{kl}^{-1}\left[\hat{\mathrm{g}}(1,2)\otimes\hat{I}_3\otimes\cdots\otimes\hat{I}_N\right]P_{kl},$$

and the matrix representation of its restriction to $V^{[N]}$ is

$$P_{kl}^{-1}\left[g\otimes s\otimes\cdots\otimes s\right]P_{kl},$$

where P_{kl} in this context is the permutation matrix corresponding to the previously defined unitary mapping. The complete tensor product construction [8] for H is

$$H=\sum_{i=1}^{N}s^{[i-1]}\otimes h\otimes s^{[N-i]}+\sum_{i<j}^{N}P_{ij}^{-1}(g\otimes s^{[N-2]})P_{ij},$$

and the N-particle overlap matrix is

$$S=s^{[N]}=\bigotimes_{i=1}^{N}s.$$

3. Eigenfunction Structure and Storage.

The Hamiltonian matrix has been constructed in $V^{[N]}$ without assuming orthonormality or symmetry adaptation of the basis set. The full space $V^{[N]}$ spanned by the N-th rank tensors generated from the basis of m orbital functions has dimension m^N. While the dimension of the model Hamiltonian within this reducible representation space is very large, the extreme simplicity of the matrix structure is a potent analog of sparsity and can be exploited to good effect in computational algorithms used to resolve the resultant large matrix eigenvalue problem.

Attempting to store all ket expansion coefficients leads to storage requirements that grow *exponentially* in the number of electrons. Thus, full storage of trial vectors used in most iterative computational schemes is impractical for systems with extended basis sets. This provides strong motivation for attempting to reduce the working dimension of the problem by exploiting internal symmetries and restricting consideration to physically allowable subspaces of $V^{[N]}$.

Physically realizable eigenstates can be obtained by restricting the trial vectors to the alternating irreducible representation spaces, using Matsen's structure projection methods [9]. The restriction to physically relevant subspaces results in a pronounced reduction in the dimension of the problem and also suggests an efficient storage scheme for the trial vectors.

According to the Pauli principle, systems of fermions are described by state functions which belong to the antisymmetric (alternating) representations of the symmetric group $\mathcal{S}_N$. Consequently all naturally occurring permutation states must have cyclic structures characterized by partitions $[\lambda]$ of the integers $1, \ldots, N$ which satisfy

$$[\lambda] = [\overbrace{\{i_1, i_2\}, \ldots, \{i_{2p-1}, i_{2p}\}}^{p \text{ groups of } 2}, \overbrace{\{i_{2p+1}\}, \ldots, \{i_N\}}^{N-2p \text{ groups of } 1}],$$

where $p = N/2 - S$ is the permutation quantum number and S is the spin quantum number. Equivalently, the physically relevant irreducible representation subspaces of $V^{[N]}$ are labelled by Young frames with one or two columns.

A Young tableau can be constructed by placing each of the first N integers in a box of a Young frame. The tableau is called *standard* if the integers always increase from left to right and from top to bottom; the number of standard tableaux associated with a frame gives the dimension of the corresponding irreducible representation. These tableaux are concise sources of information about the irreducible representations; for example, they can be used to build the matrix representations of $\mathcal{S}_N$ and can be used to form *structure projectors* [9] which project primitive N-particle functions into the physically relevant subspaces. The structure projector associated with a standard Young tableau T_χ^λ can be written

$$\mathcal{X}^\lambda = \varrho(\pi_{\chi 1}) \mathcal{K}_1^\lambda \pi_{\chi 1}$$

where $\varrho(\pi)$ is the signature of the permutation operator π and $\pi_{\chi 1}$ permutes the indices of T_χ^λ to obtain the standard tableau T_1^λ of Figure 1.

1	2
3	4
$\vdots$	$\vdots$
$2N_\beta - 1$	$2N_\beta$
$2N_\beta + 1$	
$\vdots$	
N	

Figure 1. The standard Young tableau T_1^λ. N_α and N_β are the numbers of boxes in the first (α) and second (β) columns, respectively.

The unnormalized Young operators $\mathcal{K}_1^\lambda$ are written as

$$\mathcal{K}_1^\lambda = \mathcal{A}_1^\lambda \mathcal{P}_1^\lambda,$$

where $\mathcal{A}_1^\lambda$ denotes an antisymmetric sum of products of permutations over indices in the columns of the tableau T_1^λ,

$$\mathcal{A}_1^\lambda = \sum_{\pi_1^\alpha, \pi_1^\beta} \varrho(\pi_1^\alpha \pi_1^\beta) \pi_1^\alpha \pi_1^\beta,$$

and $\mathcal{P}_1^\lambda$ denotes a product of symmetric sums of permutations over indices in tableau rows,

$$\mathcal{P}_1^\lambda = \prod_{\pi_1^r} (I + \pi_1^r).$$

The column operators π_1^α and π_1^β permute indices in the first (α) and second (β) columns of T_1^λ, respectively. The row operators π_1^r permute indices in the r-th row of T_1^λ, and I is the identity operator. The eigenfunctions of the spin-free Hamiltonian are approximated as

$$\Phi = \sum_{j_1,\ldots,j_N=1}^{m} C_{j_1,\ldots,j_N} \left| j_1, \cdots, j_N \right\rangle.$$

Because all Φ which belong to an irreducible representation space labelled by λ satisfy $\Phi = \mathcal{X}^\lambda \Phi$, symmetry adaptation places constraints on the variational coefficients $C_{j_1,\ldots,j_N}$. The relationships among the coefficients can be visualized by inserting the orbital indices into the Young frame to generate *Weyl tableaux* [11]. Each Weyl tableau can be associated with a particular ket and its variational coefficient; for example,

$$C_{ijk} \Longleftrightarrow \begin{array}{|c|c|}\hline \mathrm{i} & \mathrm{j} \\ \hline \mathrm{k} \\ \cline{1-1} \end{array}.$$

All coefficients whose orbital indices are related by permutations over same-column indices have equal absolute values with relative signs equal to the parity of the permutations. That is, if $(j_1, \ldots, j_N) = \pi_1^\alpha \pi_1^\beta (k_1, \ldots, k_N)$, then

$$C_{j_1,\ldots,j_N} = \varrho(\pi_1^\alpha \pi_1^\beta) C_{k_1,\ldots,k_N}.$$

When both columns of the Weyl tableau are of equal length, coefficients whose orbital indices are related by permutations exchanging Weyl tableau columns must be equal. For singlet systems, if $(j_1, \ldots, j_N) = \left(\prod_{\pi_1^r} \pi_1^r \right) (k_1, \ldots, k_N)$, then

$$C_{j_1,\ldots,j_N} = C_{k_1,\ldots,k_N}.$$

The *unique* ket coefficients are the set of all coefficients which are not mapped onto each other by permutations over same-column indices (or by column exchanges for systems with zero spin quantum number). The orbital indices corresponding to unique ket coefficients can be obtained very simply as follows: For each column of the Young frame, place all possible combinations of orbital indices in the boxes so that the indices increase going down the column and no orbital index appears in the same column twice (generating the columns of the Weyl tableau). Join the columns in all possible ways to obtain a rectangular array of Weyl tableaux. Each of these tableaux corresponds to a unique ket coefficient. The Weyl tableau generated for a doublet 3-electron system and a closed shell 4-electron system are given in Figure 2. Note that for closed shell systems pairs of kets corresponding to pairs of Weyl tableau with identical (but interchanged) columns have identical coefficients. The canonically indexed vector of unique ket coefficients (the "packed" vector) for the 3 electron system is given in Figure 3.

The number of unique kets $q(m,N,S)$ is

$$q(m,N,S) = \begin{cases} \frac{1}{2}\binom{m}{N_\alpha}\left[\binom{m}{N_\beta}+1\right], & \text{for } S=0, \\ \binom{m}{N_\alpha}\binom{m}{N_\beta}, & \text{otherwise.} \end{cases}$$

For a molecule of arbitrary symmetry the number of variational coefficients required for an N-electron, m-level, spin S system in a basis of configuration state functions is given by Weyl's number [12]

$$D(m,N,S) = \frac{2S+1}{m+1}\binom{m+1}{N/2-S}\binom{m+1}{m-N/2-S}.$$

Typical values of $q(m,N,S)$ and $D(m,N,S)$ for small systems are given in Table 1. Note that $q(m,N,S)$ is larger than Weyl's number because of linear dependencies among some of the unique ket coefficients in the spin-adapted vector. When m is much greater than N, the ratio q/D approaches $\frac{1}{2}(\frac{1}{2}N+1)$ for singlet systems and $(\frac{1}{2}N+S+1)/(2S+1)$ when the spin quantum number is nonzero.

4. Method.

4.1 Large Matrix Eigenvalue Algorithms

Given the magnitude of $q(m,N,S)$ for even modest values of m and N and the structure of the Hamiltonian matrix, matrix eigenvalue methods based on matrix transformations (for example the Jacobi and QR methods) are impractical for systems with nonminimal basis sets or with more than three or four electrons. However, several methods can be profitably applied here [3].

In most of these methods it is desirable to reduce the general matrix eigenvalue problem to standard form by a congruence transformation $H' = (T^{[N]})^t H T^{[N]}$, where $T^{[N]} = \bigotimes_{i=1}^{N} T$ and T is an invertible matrix which satisfies

$$I = T^t s T.$$

The eigenvalues of H' and H are identical and the corresponding eigenvectors are related by a factor of $\left(T^{[N]}\right)^{-1}$. There are several choices for T which can accomplish the reduction to standard form [13]. We currently use the symmetric orthogonalization

$$T = T^t = s^{-1/2} = UD^{-1/2}U^t,$$

where $UDU^t = s$ is the spectral decomposition of the overlap matrix. This choice preserves the structure of the spin-adapted vectors, which can be used to advantage as will be seen.

With this transformation, the general eigenvalue problem $H\,c = E\,S\,c$ becomes

$$H'b = \left(T^{[N]}\right)^t H T^{[N]} b = Eb, \quad b = \left(T^{[N]}\right)^{-1} c,$$

where the structure of H' is

$$H' = \sum_{i=1}^{N} I^{[i-1]} \otimes (ThT) \otimes I^{[N-i]} + \sum_{i<j}^{N} P_{ij}^{-1}\left[(T\otimes T)\, g\, (T\otimes T) \otimes I^{[N-2]}\right] P_{ij}.$$

Notice that just as for the unitary group model Hamiltonian, H' is cast in terms of the molecular integrals which may be used to develop direct variational procedures.

We have adapted to our use the two-tiered inverse/Rayleigh quotient iteration scheme developed by Szyld [14]. The basic iteration produces a sequence of eigenvector estimates $x_0, x_1, x_2, \cdots$ satisfying the recurrence $(H' - \gamma_k I)x_{k+1} = \omega_{k+1} x_k$, where ω_{k+1} is chosen so that $\|x_{k+1}\| = 1$ for each k. The shift γ_k is initially set to a fixed value $\hat{\gamma}$ and so produces the usual *inverse iteration* [15]. Typically, the generated vector iterates converge linearly to an eigenvector associated with the eigenvalue closest to $\hat{\gamma}$. At some point in this process, it is possible to greatly accelerate convergence by selecting the shift $\gamma_k = r(x_k) = x_k^t H' x_k$ which effects a switch to *Rayleigh quotient iteration.* The rate of convergence becomes cubic (roughly trebling the number of significant digits per step). The feasibility of either vector iteration depends on the difficulty of solving systems $(H' - \gamma_k I)y = x$. Following Szyld, we iteratively solve these systems using the large sparse symmetric linear equation solver SYMMLQ developed by Paige and Saunders [16, 17].

It was noted previously that the eigenvectors of the tensor product Hamiltonian are not restricted to any particular permutation class, so the ground state energy often does not correspond to the lowest eigenvalue of H'. We require an initial eigenvector estimate which belongs to the alternating representation and which closely approximates the true eigenvector to obtain swift convergence. Choice of the initial eigenvector estimate becomes critical as the particle number and basis set size increase. An obvious candidate is the single determinant Hartree-Fock wavefunction, spin-projected as necessary using Matsen's structure projectors. The correct spin symmetry should in principle be preserved throughout the iterative solution of the matrix eigenvalue problem because all of the operators employed are spin-free and identical integral matrices are used throughout. In practice, the use of finite precision arithmetic may allow the vector to "leak" into nonphysical subspaces. Periodic spin projection of the trial vector over the course of the iteration may then be necessary.

An important feature of these large-scale iterative matrix eigenvalue algorithms is that the Hamiltonian matrix need be referenced only via a subroutine which forms Hamiltonian matrix-vector products, allowing the tensor product structure to be used advantageously. The special tensor product structure of the Hamiltonian matrix in this context is used analogously to the way sparsity is often used for other large-scale matrix eigenvalue problems.

4.2 Formation of the Model Hamiltonian Matrix-Vector Product

In each step of the vector iterations Hamiltonian matrix-vector products are required:

$$x = H'y = H^{(1)}y + H^{(2)}y$$

where $H^{(1)}y$ and $H^{(2)}y$ are the one and two-electron parts of the product and x and y are symmetry-adapted vectors. The formation of the two-electron part

$$\sum_{i<j}^{N} P_{ij}^{-1} \left[(T \otimes T)\, g\, (T \otimes T) \otimes I^{[N-2]} \right] P_{ij} y$$

is the rate-limiting step. The expression for $H^{(2)}$ in terms of pair transpositions $p_{ij}\big(c_1 \otimes \cdots \otimes c_i \otimes \cdots \otimes c_j \otimes \cdots \otimes c_N\big) = \big(c_1 \otimes \cdots \otimes c_j \otimes \cdots \otimes c_i \otimes \cdots \otimes c_N\big)$ leads to

$$\sum_{i<j}^{N} p_{i+1,j} \big(I^{[i-1]} \otimes (T \otimes T)\, g\, (T \otimes T) \otimes I^{[N-i-1]}\big) p_{i+1,j}\, y,$$

which suggests a convenient reduction of the problem. If x and y are partitioned into subvectors x^{kl} and y^{kl} of length m^2 which cycle indices k and l while holding all other indices constant, the two-electron matrix vector product can be reduced to forming

$$x_{I''}^{kl} = (T \otimes T)\, g\, (T \otimes T) y_{I''}^{kl}$$

for all subvectors $x_{I''}^{kl}$ and $y_{I''}^{kl}$, where the fixed indices are elements of $I'' = \{i_1, \ldots, i_j, \ldots, i_N\}$, $j \neq k$, $j \neq l$. The transformation of g should be performed implicitly (by preserving the factored form of the transformation matrix) to minimize errors due to finite precision arithmetic. Multiplication by the two-factor tensor products $(T \otimes T)$ can be accomplished by noting that for matrices A, B, and C,

$$(B^t \otimes A)\, \mathbf{vec}\, C = \mathbf{vec}\, ACB,$$

where the **vec** operator produces a vector of length pq from a $p \times q$ matrix by concatenating the matrix columns [18]. Because the subvectors contain two cycling indices they can be naturally written as matrices $y_{I''}^{kl} = \mathbf{vec}\, Y_{I''}$.

While there are many subvectors, they each possess the same spin symmetries as their component kets, e.g.,

$$y_{I''}^{kl} = \varrho(\pi^\alpha \pi^\beta) y_{J''}^{ij}$$

where $\pi^\alpha \pi^\beta I = J$. Exploiting these symmetries significantly reduces the number of subvectors which must be explicitly gathered and premultiplied by the transformed two-electron integral matrix.

4.3 Operation Count: Subvector Algorithm

The formation of the two-electron integral matrix-vector products dominates the contribution due to the one-electron terms, hence we may fairly estimate the complexity of a single step of our strategy by restricting attention to the two-electron contribution. Each two-electron vector has length m^2 and corresponds to a subvector of the full trial vector in that a pair of orbital indices cycles from 1 to m while all other indices remain fixed. Ignoring symmetry, the number of subvectors is $\binom{N}{2} m^{N-2}$. Many of these subvectors are equal or differ only in sign; a set of nonredundant subvectors includes:

1) all subvectors that cycle two particular first-column (α) indices;
2) all subvectors that cycle one particular α and one particular second-column β index;
3) all subvectors that cycle two particular β indices.

All other subvectors are related to elements of this set by permutations that do not mix indices belonging to different columns in the Young tableau. For singlet systems, the third type of subvector is identical to the first type. So if permutational symmetry is taken into account, the number of subvectors p that need to be considered is

$$p = \begin{cases} \binom{m}{N/2}\binom{m}{N/2-2} + \binom{m}{N/2-1}^2 & \text{for singlet systems,} \\ \binom{m}{N/2+S-2}\binom{m}{N/2-S} + \binom{m}{N/2+S-1}\binom{m}{N/2-S-1} + \binom{m}{N/2+S}\binom{m}{N/2-S-2} & \text{otherwise.} \end{cases}$$

For m much greater than N, $\binom{m}{N} = \frac{m^N}{N!} - \frac{N}{(N-1)!} m^{N-1} + \mathcal{O}(m^{N-2})$ and $p = C_1 m^{N-2} - C_2 m^{N-3} + \mathcal{O}(m^{N-4})$ where

$$C_1 = \begin{cases} \frac{(N/2)(N-1)}{[(N/2)!]^2} & \text{for the singlet case,} \\ \frac{N(N-1)-(N/2+S)(N/2-S)}{(N/2+S)!(N/2-S)!} & \text{otherwise,} \end{cases}$$

and for the singlet case,

$$C_2 = \frac{(N/2)^2(N-2)^2}{[(N/2)!]^2},$$

otherwise

$$C_2 = \big[(N-1)(N-2)(N/2+S)(N/2-S) + (N/2+S)(N/2+S-1)(N/2+S-2)^2 \\ + (N/2-S)(N/2-S-1)(N/2-S-2)^2\big] \Big/ \big[(N/2+S)!(N/2-S)!\big].$$

Each subvector must be multiplied by the transformed two-electron matrix,

$$x = (T \otimes T)g(T \otimes T)y$$

where T accomplishes the basis set orthogonalization (i.e., $T = s^{-1/2}$) and $g_{[ij],[kl]} = \langle ij \mid kl \rangle$, $[ij] = i(i-1)/2 + j$. If $\mathbf{vec}\, Y = y$, and $\mathbf{vec}\, X = x$, the required operations for each subvector are:

1) Form the product $T\,Y\,T = V$.
2) Form $g\ v = w$, where $v = \mathbf{vec}\, V$ and w is a vector of length m^2.
3) Form the product $T\,W\,T = X$, where $w = \mathbf{vec}\, W$.

Without taking advantage of the structure of g or v, Step 2) requires m^4 floating point operations (flops), which we nominally consider to consist of an addition, a multiplication, and some overhead required to fetch or store a value. Steps 1) and 3) require about $4m^3$ flops. Hence, the total cost of the Hamiltonian matrix-vector multiply is $C_1 m^{N+2} + (8C_1 - C_2)m^{N+1} + \mathcal{O}(m^N)$ for m significantly larger than N. The implementation of Step 3) is complicated by the size of the two-electron integral list. The two-electron integral matrix will have permutational symmetry and often a block structure that arises from the symmetry of the molecular geometry. However, it is often not feasible to store the entire matrix in memory, even if the structure and induced sparsity of the matrix is exploited. A list of symmetry-unique integrals is put into secondary storage in some convenient order, and the integral matrix-vector multiply should access this list sequentially in as few passes as possible.

The subvectors of spin-adapted trial trial vectors have a structure that results from the constraints on the variational coefficients. Subvectors that cycle indices from the same Young tableau column can be written as skew-symmetric matrices; the rows and columns that correspond to fixed indices from the same Young tableau column are zero. Subvectors that cycle indices from different tableau columns have rows of zeros corresponding to fixed α indices and zero columns corresponding to fixed β indices.

4.4 Advantages for Parallel Computation

Paging activity can be minimized during the gathering of subvector elements from the packed trial vector by organizing the trial vector as a rectangular array (square and symmetric for singlet systems) with rows (columns) corresponding to lexically ordered combinations of orbital indices over the first (second) column in the appropriate Young frame as in Figure 3. For example, for each fixed set of β (odd) indices, we obtain all necessary elements which are indexed by fixed and cycling α (even) indices. To address individual elements of the trial vector array we use Handy and Knowles' addressing array Z [6] which computes the lexicographic index of combinations of orbital indices.

At the k^{th} iteration step, the subvectors of x_k may be efficiently collected, multiplied by $(T \otimes T)g(T \otimes T)$, and scattered to different subvectors with indices dictated by spin symmetry. This procedure has several advantages for parallel computation:

1) Subvectors which differ by 3 or more noncycling indices can be constructed from distinct portions of the packed vector. This means that it is possible to simultaneously gather several subvectors from a packed vector in a shared memory multiprocessor without memory contention, or simultaneously communicate several subvectors in a distributed memory parallel computer.
2) Each of the transformed two electron integral matrix-subvector multiplications involves the same number of floating point operations. The matrix-vector multiplications can be apportioned evenly among the processors for very good load balancing.
3) The independence of the individual subvector problems requires minimal interprocessor communication.
4) The generation of the molecular integrals in parallel on different processors leaves integrals over different shells distributed among the processors. With the present tensor product method this causes no difficulty because there is no need to perform integral transformations, which would involve sorting and redistribution of the integrals among the processors.

5. Applications

The tensor product method has been implemented on an IBM 3090/600VF vector machine; parallel implementations on shared memory and distributed memory multiprocessors are being considered. The program has been preliminarily applied to several small model systems. For example, Table 2 gives total energies computed for linear H_3 with Pearson's double zeta basis set [19], with and without the p function. In Table 3, tensor product variational energies reproduce Jankowski and Paldus' [20] "H4" model hydrogen molecular dimer full CI energies to 1-2 microHartrees. The model consists of a pair of coplanar STO-3G hydrogen molecules with the atoms at the vertices of a trapezoid. Further details of the model are given by Jankowski and Paldus.

Acknowledgement.

This work was made possible by a generous grant of computer time from the Cornell Theory Center and the Cornell National Supercomputer Facility.

References.

[1] A. Szabo and N. L. Ostlund, Modern Quantum Chemistry: An Introduction to Advanced Electronic Structure Theory, Macmillan, New York (1982).

[2] S. Wilson, Electron Correlation in Molecules, Clarendon Press, Oxford (1984).

[3] E. R. Davidson, in Methods for Computational Molecular Physics, ed. G. H. F. Diercksen, S. Wilson, D. Reidel Publishing, Dordrecht, Holland (1983).

[4] C. W. Bauschlicher, J. Phys. Chem., 92, 3020 (1988).

[5] P. E. M. Siegbahn, Chem. Phys. Lett., 109, 417 (1984).

[6] P. J. Knowles, N. C. Handy, Chem. Phys. Lett., 11, 315 (1984).

[7] C. Beattie, in Numerical Treatment of Eigenvalue Problems, Int. Series Num. Math. 69, 9 (1984).

[8] C. Beattie, J. C. Schug, J. W. Viers, L. T. Watson, SIAM Conference on Numerical Linear Algebra, Raleigh, N. C. (May 1986).

[9] F. A. Matsen, A. A. Cantu, R. D. Poshusta, J. Phys. Chem., 70, 1558 (1966).

[10] D. W. Fox, SIAM Jour. Math. Anal., 3, 617 (1972).

[11] R. Pauncz, Spin Eigenfunctions, Plenum Press, New York (1979).

[12] J. Paldus, in The Unitary Group for the Evaluation of Electronic Energy Matrix Elements, ed. J. Hinze, Lecture Notes in Chemistry, 22, Springer-Verlag, Berlin (1981).

[13] B. N. Parlett, The Symmetric Eigenvalue Problem, Prentice Hall, Englewood Cliffs, New Jersey (1980).

[14] D. B. Szyld, in Innovative Numerical Methods in Engineering, ed. R. P. Shaw, J. Periaux, A. Chaudouet, J. Wu, C. Marino, C. Brebbia, Springer Verlag, Berlin (1986).

[15] G. H. Golub, C. F. van Loan, Matrix Computations, John Hopkins University Press, Baltimore (1983).

[16] J. G. Lewis, Technical Report STAN-CS-77-595, Computer Science Department, Stanford University, Stanford, CA (March 1977).

[17] C. C. Paige, M. A. Saunders, SIAM J. Numer. Anal., 12, 617 (1975).

[18] A. Graham, Kronecker Products and Matrix Calculus with Applications, Halsted Press, New York (1981).

[19] P. K. Pearson, Ph.D. Dissertation, University of California, Berkeley, CA, (1974).

[20] K. Jankowski, J. Paldus, Int. J. Quant. Chem.,18, 1243 (1980).

[21] J. S. Binkley, D. J. De Frees, M. J. Frisch, K. Raghavachari, R. A. Whiteside, H. B. Schlegel, E. M. Fluder, J. A. Pople, *Gaussian 82 (release H version)*, Carnegie Mellon University, Pittsburgh, PA, (1982).

TABLE 1. Full space and subspace dimensions for typical systems with N electrons, m basis functions, and $S = 0$.

m	N	Full Space Dimension m^N	Number of Determinants $\binom{2m}{N}$	Number of Unique Kets $q(m, N, S)$	Weyl's Number $D(m, N, S)$
10	4	10000	4845	1035	825
20	4	160000	91390	18145	13300
30	4	810000	487635	94830	67425
40	4	2560000	1581580	304590	213200
50	4	6250000	3921225	750925	520625
20	8	2.56×10^{10}	76904685	11739435	5799465
30	8	6.56×10^{11}	2.56×10^{9}	3.76×10^{8}	1.72×10^{8}
40	8	6.55×10^{12}	2.89×10^{10}	4.18×10^{9}	1.85×10^{9}
10	10	1.00×10^{10}	184756	31878	19404
50	10	9.77×10^{16}	1.73×10^{13}	2.24×10^{12}	8.3×10^{11}
75	10	5.63×10^{18}	1.17×10^{15}	1.48×10^{14}	5.3×10^{13}

TABLE 2. Total energies (H) for linear H_3, with $R_{H_1-H_2} = R_{H_2-H_3} = 1.7924$ bohrs, using Pearson's [2s] and [2s1p] basis sets.

Method	(4s)/[2s]	(4s1p)/[2s1p]
UHF	-1.597903	-1.601412
MP2	-1.615693	-1.631693
MP3	-1.622184	-1.638914
SD-CI	-1.627822	-1.643855
SD-CI+SCC	-1.628111	-1.644279
This work	-1.627969	-1.644761

UHF: Unrestricted Hartree-Fock energy.

MP2, MP3: Second and third order Möller-Plesset energies.

SD-CI: Configuration interaction including single and double excitations only.

SCC: Davidson's size-consistency correction.

UHF, MP2, MP3, SD-CI energies computed using Gaussian 82 [21].

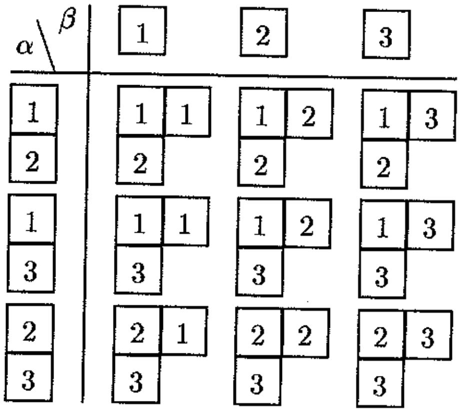
TABLE 3. CI correlation energies for the H_4 model (in mH, all signs reversed).

a	SD–CI	DQ–CI	F–CI	This work
0.500	—	53.511	53.690	—
0.200	55.8710	57.168	57.260	57.2591
0.100	63.5454	65.227	65.321	65.3203
0.050	73.7890	76.401	76.429	76.4273
0.020	88.8065	92.124	92.148	92.1462
0.015	93.2459	96.686	96.711	96.7099
0.005	105.5560	109.188	109.196	109.1947
0.000	113.9716	—	—	117.6195

SD-CI: Configuration interaction including single and double excitations only, computed using Gaussian 82 [21].

F-CI, DQ-CI: Full and truncated CI including double and quadruple excitations taken from Jankowski and Paldus [20].

Figure 2. Weyl tableau construction of the trial vector: $N = 3$, $m = 3$, $S = 1/2$.



Weyl tableau construction of the trial vector: $N = 4$, $m = 4$, $S = 0$.

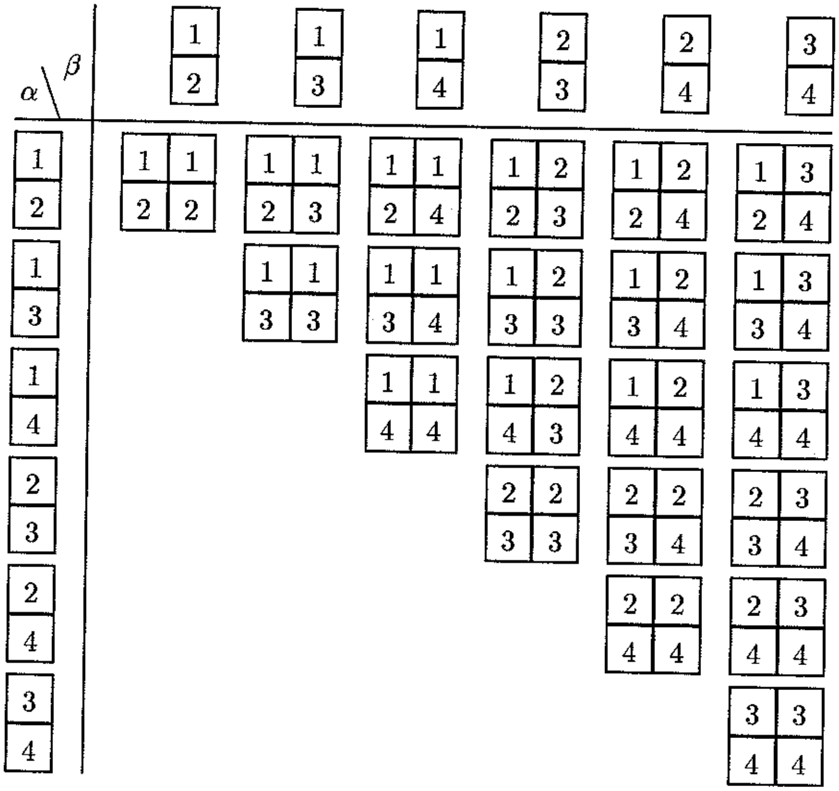


Figure 3. The full and packed spin-adapted trial vectors C and $\tilde{C}$ for $N = 3$, $m = 3$, $S = 1/2$.

Index i	Ket	C_i	$\tilde{C}_i$
1	$\|111\rangle$	0	c_{112}
2	$\|112\rangle$	c_{112}	c_{113}
3	$\|113\rangle$	c_{113}	c_{213}
4	$\|121\rangle$	0	c_{122}
5	$\|122\rangle$	c_{122}	c_{123}
6	$\|123\rangle$	c_{123}	c_{223}
7	$\|131\rangle$	0	c_{132}
8	$\|132\rangle$	c_{132}	c_{133}
9	$\|133\rangle$	c_{133}	c_{233}
10	$\|211\rangle$	$-c_{112}$	
11	$\|212\rangle$	0	
12	$\|213\rangle$	c_{213}	
13	$\|221\rangle$	$-c_{122}$	
14	$\|222\rangle$	0	
15	$\|223\rangle$	c_{223}	
16	$\|231\rangle$	$-c_{132}$	
17	$\|232\rangle$	0	
18	$\|233\rangle$	c_{233}	
19	$\|311\rangle$	$-c_{113}$	
20	$\|312\rangle$	$-c_{213}$	
21	$\|313\rangle$	0	
22	$\|321\rangle$	$-c_{123}$	
23	$\|322\rangle$	$-c_{223}$	
24	$\|323\rangle$	0	
25	$\|331\rangle$	$-c_{133}$	
26	$\|332\rangle$	$-c_{233}$	
27	$\|333\rangle$	0	