Pseudospectral double excitation configuration interaction

Todd J. Martinez and Emily A. Carter
Department of Chemistry and Biochemistry, University of California, Los Angeles, 405 Hilgard Avenue, Los Angeles, California 90024-1569

(Received 10 December 1992; accepted 19 January 1993)

We present a pseudospectral formulation of the single reference, closed shell double excitation configuration interaction method using a generator state self-consistent electron pair approach. The method scales as \( O(n^3N^3) \), compared to the conventional scaling of \( O(n^2p^3 + n^3N^3) \). In no case tested does the pseudospectral energy differ by more than 0.35 mhartree from the conventional result.

INTRODUCTION

Recently, we have presented a pseudospectral formulation of the full configuration interaction (FCI) problem. The method was accurate to a mhartree after corrections, but the size of the numerical grids required made it impossible to realize any computational advantage for systems feasible with today's hardware. In this paper, we present the first application of the pseudospectral method to truncated configuration interaction, the double-excitation configuration interaction (DCI) method. Since the scaling of the conventional DCI method is not combinatorial like full CI, it is easier to realize the pseudospectral advantage.

The method we present here has several striking characteristics. First, the method will be faster than conventional DCI approaches for sufficiently large cases, as the scaling is decreased from \( O(n^4) \) to \( O(n^3) \), where \( n \) is the number of molecular orbitals. This scaling advantage will be realizable for cases of rather modest dimension, where we consider modest to mean the calculation would take < 10 min of cpu time on a CRAY YMP or equivalent using conventional methods. Second, it provides a framework within which "direct," in the sense of "direct self-consistent field (SCF)," correlated techniques can be realized. By "direct," we mean that neither the Hamiltonian matrix nor the two-electron integrals are stored on disk. This approach will become necessary as correlation treatments are applied to large systems, and the pseudospectral method is well-suited to it because of the ability to organize the required data (which will be combined to form the two-electron integrals) by blocks of gridpoints. Finally, it allows one to take advantage of locality without introducing the approximations suggested by Pulay and Sæbo. This can be done by using "cutoffs" where the requisite summations are restricted to a subset of the physical space grid. Note that the approximations of Pulay and Sæbo may be used along with "cutoffs" to gain even greater advantage. At this point, we have only coded the equations for the molecular orbital basis, and therefore we provide numerical evidence supporting only the first of these claims. A forthcoming paper will address the last two issues.

SELF-CONSISTENT ELECTRON PAIR THEORY FOR DOUBLES CI

Self-consistent electron pair (SCEP) theory provides a matrix-oriented formalism for a number of correlation methods, including configuration interaction with double excitations. The generator state formulation of Pulay, Sæbo, and Meyer is an efficient modification for the special case of a closed-shell reference state. We review the essential equations for the DCI method, but refer the reader to the original reference for more detail. Our notation conventions are that \( i,j,k,... \) denote molecular orbitals occupied (internal) in the reference state, \( \psi_0 \). Orbitals unoccupied in \( \psi_0 \) (external) are denoted as \( \alpha \beta \),... while \( p,q,r,... \) denote arbitrary orbitals, and boldface quantities denote matrices.

The generator state formalism is distinguished by a particular choice for the configuration state functions (CSF's) which removes all coupling coefficients from the equations. The CSF's chosen are pairwise nonorthogonal and are given as follows:

\[
\psi_{ij}^{\alpha\beta} = \phi_{ij}^{\alpha\beta} + \phi_{ij}^{\alpha\beta} + \phi_{ij}^{\alpha\beta} + \phi_{ij}^{\alpha\beta}, \quad i > j, \\
\psi_{ij}^{\alpha\beta} = \phi_{ij}^{\alpha\beta}, \quad (1a,b)
\]

where \( \phi_{ij}^{\alpha\beta} \) represents a determinant with two electrons excited from spin orbitals \( \alpha \) and \( \beta \) to \( \alpha \) and \( \beta \). The ansatz for the DCI wave function, assuming intermediate normalization (\( \langle \psi_0 | \psi_0 \rangle = 1 \)), is therefore

\[
\psi = \psi_0 + \sum_{i,j} \sum_{\alpha,\beta} C_{ij}^{\alpha\beta} \psi_{ij}^{\alpha\beta}. \quad (2)
\]

The reciprocal basis is defined by the criterion \( \langle \psi_1 | \psi_J \rangle = \delta_{I,J} \), where \( I \) and \( J \) denote the collective orbital indices. Explicitly,

\[
\psi_{ij}^{\alpha\beta} = \frac{1}{2} (\phi_{ij}^{\alpha\beta} + \phi_{ij}^{\alpha\beta} + 2 \phi_{ij}^{\alpha\beta} + \phi_{ij}^{\alpha\beta} + \phi_{ij}^{\alpha\beta}), \quad i > j, \\
\psi_{ij}^{\alpha\beta} = \phi_{ij}^{\alpha\beta}. \quad (3a,b)
\]

The fundamental quantities in the DCI problem are the doubles residuum matrices,

\[
(T_{ij})_{\alpha\beta} = \langle \psi_{ij}^{\alpha\beta} | H - E_0 + E_{\text{corr}} | \psi \rangle, \quad (4a)
\]

which are zero when \( |\psi\rangle \) is an eigenvector of \( H \) with eigenvalue \( E_0 - E_{\text{corr}} \). We follow the standard practice of designating the energy of the reference wave function as...
The pseudospectral method interprets the matrix equation (14) as a transform from spectral (function) space to physical space, multiplication by the physical space representation of the Coulomb potential operator, and subsequent back-transform to spectral space. From this vantage point\textsuperscript{32} it becomes clear that $R_I$ should be replaced by $R^{-1}$, provided the number of gridpoints equals the number of basis functions.

Friesner further improved the method by using a least-squares technique allowing the use of more gridpoints than basis functions and through the use of dealiasing functions. The dealiasing functions are included because of the recognition that the range of the physical space operator is not limited to the space spanned by the basis set. Therefore, a filtering procedure of some sort is necessary to remove the components outside of the basis set prior to the back-transformation to spectral space. This is conceptually equivalent to the phenomenon of the same name in the theory of discrete Fourier transforms. The use of a least-squares matrix $B$ in place of $R^{-1}w$ accomplishes the desired result.\textsuperscript{10} The reader is referred to the papers of Friesner and co-workers for details of the construction of this matrix.\textsuperscript{11}

**PSEUDOSPECTRAL FORMULATION OF SCEP FOR DOUBLES CI**

We introduce the pseudospectral approximation in the calculation of the two most costly terms of the residual vector, $K[C_{ij}]$ and $Q_{ij} + Q_{ji}$. Spectral calculation of these terms scales as $O(n^2 N^2 + n^3 N^3)$, where $n$ is the number of internal orbitals and $N$ is the number of external orbitals. Although the scaling behavior suggests that the first term would dominate (generally $n < N$), experience shows that this is not the case for most problems. We submit as an example the calculation on carbon dimer detailed later in this paper. With six internal orbitals and 24 external orbitals, 45\% of the time is spent calculating the first term, and 50\% is required for the second term. The remaining terms in Eqs. (4) consume only 5\% of the cpu time. Of course, this behavior is a consequence of a larger prefactor multiplying the scaling for the $Q_{ij} + Q_{ji}$ term.

The pseudospectral formula for the external exchange matrix $K[C_{ij}]$ is

$$K[C_{ij}]_{ab} = \sum_g B_{ag} \sum_d A_{bd}(g) \sum_c R_{cg} C_{ij}^{cd},$$  \hspace{1cm} (19)

where $I^n - R^I w A^n R,$

$$w_{g',g} = \delta_{g',g} w(r_g),$$

$$A_{rg}(r_g) = \delta_{g',g} \int \frac{\phi_r(r) \phi_g(r)}{|r-r_g|} dr, \hspace{1cm} (17)$$

$$R_{rg} = \phi_r(r_g). \hspace{1cm} (18)$$
number of orbitals \((n + N)\), and \(N\) is usually significantly larger than \(n\), this is roughly equivalent to \(O(n^2 N^3)\). The corresponding formula for \(Q_{ij}\) requires some rearrangement of the spectral term in Eq. (9).

\[
Q_{ij} = \sum_k \mathbf{K} \cdot \{2\mathbf{C}_{kj} - \mathbf{C}_{jk}\} - \mathbf{J}_{ik} \mathbf{C}_{kj} - \mathbf{J}_{jk} \mathbf{C}_{ik}. \tag{20}
\]

This requires two additional matrix multiplications when forming the combined quantity \(Q_{ij} + Q^T_{ji}\), but is necessary because explicit formation of the matrix \(\mathbf{K} - \frac{1}{3} \mathbf{J}_{ik}\) would obviate the pseudospectral advantage. Given this rearrangement, the final formula becomes

\[
(Q_{ij})_{ab} = \sum_g A_{ag} (g) \sum_k B_{kg} \sum_c R_{gc} (2\mathbf{C}_{kj}^a \mathbf{C}_{kj}^b)
- \sum_g B_{ag} \left[ \sum_k A_{ag} (g) \sum_c R_{gc} \mathbf{C}_{kj}^b \right]
+ \sum_g A_{ag} (g) \sum_c R_{gc} \mathbf{C}_{ik}^b. \tag{21}
\]

This scales as \(O(Mn^2 N^2)\), being equivalent by the previous arguments to \(O(n^2 N^3)\).

**IMPLEMENTATION AND RESULTS**

We have written a spectral generator state SCEP DCI program, defined by Eqs. (1)-(11), as well as a pseudospectral version wherein Eqs. (19) and (21) replace Eqs. (8) and (9). In the pseudospectral version, we use spectral integrals for all terms other than \(\mathbf{K}[C_{ij}]\) and \(Q_{ij} + Q^T_{ji}\). Thus the appropriate \(E_0\) is the conventional one constructed spectrally. Although we have completely ignored the lack of index symmetry in the pseudospectral integrals, we have encountered no serious convergence difficulties. The molecular orbital basis was used exclusively, and the pseudospectral quantities were generated in the atomic orbital basis using the O(n^2 N^3) scaling of the conventional two-electron integral transformation.

The primary purpose of this paper is to show that the pseudospectral DCI method is capable of chemical accuracy on very sparse grids. Therefore, we present no timing information other than to report that the pseudospectral code is \(\sim 10\%\) faster for the \(C_6\) test case and up to two times slower for the larger cases. As neither code is optimized, this information is provided only to show that the method is viable in practice. Analysis of the efficiency of the method must await optimization of the codes and tests on the larger systems for which the pseudospectral method is designed. Nevertheless, we can provide a rough estimate of 30 basis functions as the point where the pseudospectral DCI method becomes cheaper than the spectral one. An efficient code can easily perform calculations of this magnitude using \(< 10\) min of cpu time on a CRAY YMP computer.

The calculations were performed on an FPS 521-EA computer, taking advantage of the vector processor. The pseudospectral method allows for efficient vectorization over the gridpoints and thus the pseudospectral code is slightly faster than we expected from counting floating-point multiplications. This implies that the estimate of the break-even point might change when using a computer without vector processing capability.

The test cases all use 6-31G** basis sets, where the \(s\) combination of the Cartesian \(d\) functions has not been removed. The grids used are the "medium" grids of the PSGVB program. Reference wave functions \(\psi_0\) were obtained from spectral closed-shell Hartree–Fock calculations. The CSF's include all double excitations out of all internal orbitals and into all external orbitals.

We compare the spectral and pseudospectral energies for a few atomic species in Table I. The accuracy is seen to be quite satisfactory, even for \(F\), where the error is \(< 0.1\) kcal/mol. More interesting are the diatomic cases shown in Table II. Again, the accuracy is certainly sufficient for this level of theory, with the error never exceeding 0.35 Hartree (0.2 kcal/mol). However, the error varies with order of magnitude over the range of internuclear distances shown. Explanation of this requires knowledge of the behavior of the scheme used to merge atomic grids. Each atom carries its own grid, and these grids are superimposed to form the molecular grid. Then points which are too close are coalesced. This is done in order to avoid approximate linear dependencies which would make it difficult to invert \(\mathbf{R}\) accurately. Finally, some points are added to the bonding region in order to improve the quality of the quadrature in the area where the wave function is expected to have considerable amplitude. Clearly, we expect that as the internuclear distance is reduced, the quality of the quadrature will decrease unless one takes pains to replace the deleted points by ones which do not cause linear dependence. Likewise, as the internuclear distance increases, one must ensure that enough points are added to the bonding region such that the quality of the quadrature in this region remains roughly constant. Generating grids of uniform quality for arbitrary molecular geometries is a difficult task, and it is no wonder that the scheme used is not perfect. Part of the problem stems from the fact that we use excluding...
TABLE II. DCI results for various molecules using pseudospectral and spectral formulations. Basis sets are all 6-31G**, double excitations are included from all occupied orbitals to all virtual orbitals, and all molecules are in $\Sigma^+$ states.

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>Pseudospectral</th>
<th>Spectral</th>
<th>$E_0$</th>
<th>Error(kH)</th>
<th>Grid size</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_2$</td>
<td>0.50</td>
<td>0.032 284</td>
<td>0.032 212</td>
<td>-1.061 482</td>
<td>-72</td>
</tr>
<tr>
<td></td>
<td>0.60</td>
<td>0.032 822</td>
<td>0.032 789</td>
<td>-1.113 933</td>
<td>-33</td>
</tr>
<tr>
<td></td>
<td>0.70</td>
<td>0.033 046</td>
<td>0.033 664</td>
<td>-1.131 292</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>0.80</td>
<td>0.034 144</td>
<td>0.034 162</td>
<td>-1.128 434</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>0.035 286</td>
<td>0.035 294</td>
<td>-1.116 516</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>0.036 914</td>
<td>0.036 926</td>
<td>-1.099 475</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>1.10</td>
<td>0.039 153</td>
<td>0.039 176</td>
<td>-1.079 944</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>1.20</td>
<td>0.042 097</td>
<td>0.042 177</td>
<td>-1.059 423</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>2.00</td>
<td>0.092 279</td>
<td>0.092 736</td>
<td>-0.996 740</td>
<td>157</td>
</tr>
<tr>
<td>$HF$</td>
<td>0.50</td>
<td>0.161 471</td>
<td>0.161 692</td>
<td>-0.993 079</td>
<td>221</td>
</tr>
<tr>
<td></td>
<td>0.60</td>
<td>0.167 269</td>
<td>0.167 586</td>
<td>-0.993 978</td>
<td>317</td>
</tr>
<tr>
<td></td>
<td>0.70</td>
<td>0.172 529</td>
<td>0.172 869</td>
<td>-0.992 098</td>
<td>340</td>
</tr>
<tr>
<td></td>
<td>0.80</td>
<td>0.177 325</td>
<td>0.177 632</td>
<td>-0.999 570</td>
<td>307</td>
</tr>
<tr>
<td></td>
<td>0.91</td>
<td>0.182 030</td>
<td>0.182 292</td>
<td>-0.100 117</td>
<td>262</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>0.185 699</td>
<td>0.185 886</td>
<td>-0.100 001</td>
<td>187</td>
</tr>
<tr>
<td></td>
<td>1.10</td>
<td>0.189 539</td>
<td>0.189 645</td>
<td>-0.997 492</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td>1.20</td>
<td>0.193 407</td>
<td>0.193 422</td>
<td>-0.995 123</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>1.40</td>
<td>0.201 464</td>
<td>0.201 489</td>
<td>-0.989 528</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>2.00</td>
<td>0.231 427</td>
<td>0.231 329</td>
<td>-0.974 453</td>
<td>-98</td>
</tr>
<tr>
<td>$C_2$</td>
<td>1.10</td>
<td>0.278 112</td>
<td>0.278 140</td>
<td>-0.753 397</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>1.20</td>
<td>0.278 733</td>
<td>0.278 748</td>
<td>-0.753 661</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>1.24</td>
<td>0.278 805</td>
<td>0.278 809</td>
<td>-0.753 791</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>1.26</td>
<td>0.278 766</td>
<td>0.278 771</td>
<td>-0.753 683</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>1.28</td>
<td>0.278 698</td>
<td>0.278 707</td>
<td>-0.753 268</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>1.29</td>
<td>0.278 665</td>
<td>0.278 682</td>
<td>-0.753 186</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>1.30</td>
<td>0.278 582</td>
<td>0.278 601</td>
<td>-0.753 878</td>
<td>19</td>
</tr>
</tbody>
</table>

Internuclear distance in angstroms.

DCI correlation energy in atomic units. DCI total energy = $E_0 - E_{corr}$.

Reference energy in atomic units.

$\mu$H = 10$^{-6}$ hartree. Error = $E_{corr}$ (spectral) - $E_{corr}$ (pseudospectral).

Number of points in grid used.

We have presented a pseudospectral formulation of the generator state SCEP method applicable to DCI wave functions. The scaling is improved by a factor of $n$, the number of internal molecular orbitals. The method has been shown to be sufficiently accurate for a number of test cases, with errors not exceeding 0.35 mhartree for any case tested. Our pseudospectral code is slightly faster than the spectral version for $C_2$ with 30 basis functions, and thus we expect the method to be useful for DCI calculations on systems with 30 or more basis functions (molecular orbitals). As DCI calculations of this size are not difficult, we expect the pseudospectral DCI method to open the doors to correlated calculations on molecules of unprecedented size in the near future. In particular, we are excited about the ease with which the generator-state formalism for DCI may be cast in terms of atomic orbitals. This will allow the use of "cutoffs," and it will also make the implementation of "atomic corrections" trivial. We are currently formu-
lating the code in terms of atomic orbitals in order to implement these improvements and further reduce the grid size for a given accuracy.

ACKNOWLEDGMENTS

This work was supported in part by the Air Force Office of Scientific Research. E.A.C. acknowledges generous support from the National Science Foundation and the Camille and Henry Dreyfus Foundation through Presidential Young Investigator and Teacher–Scholar Awards, respectively. One of us (T. J. M.) is grateful for partial support through a NSF graduate assistantship. We also thank Richard Friesner for supplying us with a pseudospectral Hartree–Fock code.

9Note that in the discrete variable representation (DVR) method of Light and co-workers [see, for example, J. C. Light and Z. Bacic, J. Chem. Phys. 85, 4594 (1986)], this is ensured by choosing gridpoints such that \( R \) is an orthogonal matrix. Hence the close connection between the DVR and pseudospectral methods.
10Friesner and co-workers designate the matrix \( B \) as \( Q \). We have followed our previous paper in altering the pseudospectral notation where there is conflict between the notation used in the spectral CI methods and that used to denote the fundamental pseudospectral quantities.