Energies and analytic gradients for a coupled-cluster doubles model using variational Brueckner orbitals: Application to symmetry breaking in $O_4^+$

C. David Sherrill, Anna I. Krylov, Edward F. C. Byrd, and Martin Head-Gordon

Department of Chemistry, University of California, Berkeley, California 94720

(Received 20 April 1998; accepted 11 June 1998)

We describe an alternative procedure for obtaining approximate Brueckner orbitals in ab initio electronic structure theory. Whereas approximate Brueckner orbitals have traditionally been obtained by mixing the orbitals until the coefficients of singly substituted determinants in the many-electron wave function become zero, we remove singly substituted determinants at the outset and obtain orbitals which minimize the total electronic energy. Such orbitals may be described as variational Brueckner orbitals. These two procedures yield the same set of exact Brueckner orbitals in the full configuration interaction limit but differ for truncated wave functions. We consider the simplest variant of this approach in the context of coupled-cluster theory, optimizing orbitals for the coupled-cluster singles and doubles (CCD) model. An efficient new method is presented for solving the coupled equations defining the energy, doubles amplitudes, and orbital mixing parameters. Results for several small molecules indicate nearly identical performance between the traditional Brueckner CCD method and the variational Brueckner orbital CCD approach. However, variational Brueckner orbitals offer certain advantages: they simplify analytic gradients by removing the need to solve the coupled-perturbed Brueckner coupled-cluster equations for the orbital response, and their straightforward extensions for inactive orbitals suggests possible uses in size-extensive models of nondynamical electron correlation. Application to $O_4^+$ demonstrates the utility of variational Brueckner orbitals in symmetry breaking cases. © 1998 American Institute of Physics.

I. INTRODUCTION

The exact Brueckner orbitals are those orbitals for which the coefficients of singly substituted determinants in the full configuration interaction (CI) wave function become zero. Alternatively, they are the orbitals for which a single determinant has the maximum overlap with the full CI wave function. These orbitals were introduced in 1954 by Brueckner as part of a self-consistent method for the study of nuclear matter and were related to the CI method in 1958 by Nesbitt. Since the full CI wave function is too costly to compute for all but the smallest molecular systems, Brueckner orbitals are generally defined as those orbitals giving zero singles coefficients for some truncated wave function. A number of studies over the last decade indicate that ab initio electronic structure models may become more robust when based on Brueckner orbitals rather than on the standard Hartree–Fock orbitals.

In 1981, Chiles and Dykstra became the first to employ Brueckner orbitals in the context of coupled-cluster theory; using a model space of single and double substitutions, they rotated the orbitals in an iterative procedure to zero the amplitudes for single substitutions. These authors referred to this procedure as CCD ($\tilde{T}_1 = 0$); more recently, the method has been called Brueckner coupled-cluster doubles (BCCD or B-CCD) or simply Brueckner doubles (BD). The need to transform the atomic orbital (AO) integrals each orbital iteration was viewed as a significant impediment, and the method received little attention until it was revived in 1989 by Handy et al. These workers presented a detailed formulation of the method and a Møller–Plesset perturbation theory analysis indicating that BD first differs from coupled-cluster singles and doubles (CCSD) and quadratic CI with singles and doubles (QCISD) at fifth order. More complete Brueckner coupled-cluster models include connected triple excitations in complete (BCCD) or BDT or perturbative (BCCD(T) or BD(T)) fashions.

In 1991, Kobayashi et al. presented analytic gradients for BD and BD(T), enabling systematic studies of equilibrium geometries, dipole moments, vibrational frequencies, infrared intensities, and other properties of several small molecules. These studies indicate that the differences between BD and CCSD are generally small except in the case of infrared intensities. Nevertheless, in some cases Brueckner orbitals offer definite advantages. In particular, Brueckner coupled-cluster is less likely to exhibit artificial symmetry breaking than coupled-cluster based on Hartree–Fock orbitals, as shown by Stanton et al. for NO$_3$ and Barnes and Lindh for O$_3$. Additionally, Handy et al. have demonstrated for bond-breaking reactions that the point at which spin symmetry is broken (i.e., where an unrestricted solution becomes lower in energy than a restricted one) is much further out for BD than for single-reference methods based on the Hartree–Fock determinant. Spin-restricted Brueckner orbitals for high-spin...
open-shell systems have been discussed recently by Crawford et al.\textsuperscript{20}

In the full configuration interaction limit, there exists an alternative procedure for obtaining the exact Brueckner orbitals. Instead of rotating the orbitals until the singles coefficients vanish, one may delete all singly substituted determinants \textit{a priori} and find the orbitals for which the full CI (less singles) wave function gives a minimum energy; both procedures yield equivalent sets of exact Brueckner orbitals. For truncated wave functions, this strict equivalence no longer holds, so that one may consider two equally valid approximations to the exact Brueckner orbitals. In the first (discussed above), one rotates the orbitals so that there is no projection of the approximate wave function onto the space of singly substituted determinants; the orbitals obtained in this fashion might be described as projective Brueckner orbitals. In the second, one sets the coefficients of all single substitutions to zero and finds the orbitals for which the energy of the truncated wave function is a minimum; we will call these orbitals the variational Brueckner orbitals. Since these two approaches give equivalent orbitals in the full CI limit, one would expect the projective and variational Brueckner orbitals to be very similar for highly correlated wave functions.

In this paper we focus on approximate Brueckner orbitals for the CC method. The idea of variationally optimizing orbitals for the CCD wave function has been considered previously by Purvis and Bartlett\textsuperscript{9} and by Scuseria and Schaefer.\textsuperscript{21} In their classic 1982 paper on the CCSD method,\textsuperscript{9} Purvis and Bartlett noted that their implementation did not depend on the use of Hartree–Fock orbitals and commented that “applications for non-Hartree–Fock orbitals, such as optimizing orbitals so that the energy becomes stationary, are easily envisioned.” Among their various coupled-cluster and perturbation theory benchmark energies for several test cases, these authors reported a single-point energy for the H$_2$O molecule in a double-$\zeta$ (DZ) basis using CCD with optimized orbitals. In subsequent work, Purvis et al.\textsuperscript{22} suggested that variational orbitals could be useful in coupled-cluster studies of reactions where the dominant configuration changes along the reaction coordinate.

In 1987, Scuseria and Schaefer reported single-point energies for four small molecules using CCD and CCSD with optimized orbitals.\textsuperscript{21} They sketched a procedure for orbital optimization based on the fact that at convergence, the variationally optimized orbitals make no orbital response contribution to the coupled-cluster gradient;\textsuperscript{23} hence, these authors optimized orbitals by zeroing the orbital Z-vector\textsuperscript{24} using an iterative process involving successive gradient calculations followed by orbital rotations. This procedure is similar to that reported by Rendell \textit{et al.} in the context of CI singles and doubles (CISD) wave functions.\textsuperscript{23} Scuseria and Schaefer noted great difficulty in optimizing orbitals for CCSD due to the presence of $\epsilon \tilde{T}_1$, which already describes orbital relaxation; however, they reported good convergence for CCD. When these authors constructed CCSD wave functions using the CC optimized orbitals, the singles coefficients were found to be $\sim 10^{-5}$, demonstrating the similarity between variational and projective Brueckner orbitals for CCD.

Scuseria and Schaefer showed\textsuperscript{21} that the CCD energies with optimized orbitals were very similar to those from CCSD using Hartree–Fock orbitals for their test cases. As the practical advantages of Brueckner orbitals were not widely appreciated at that time, there seemed little reason to pursue CCD with optimized orbitals when CCSD was available. Indeed, the CCD orbital optimization of Scuseria and Schaefer would be substantially more expensive than a CCSD computation because of the need to transform the AO integrals and evaluate the CC density matrices and amplitude Z-vector every orbital iteration. Nevertheless, over the last decade interest in Brueckner orbitals has heightened due to studies demonstrating their tendency to avoid spin and spatial symmetry breaking. In 1992, Hampel, Peterson, and Werner\textsuperscript{8} achieved a breakthrough in the efficiency of projective Brueckner CCD computations by solving for the orbitals and doubles amplitudes \textit{simultaneously}, rather than performing a complete CCD calculation each orbital iteration. Furthermore, these authors also showed how to avoid the full integral transformation every iteration. In their implementation, BD is actually slightly less expensive than standard CCSD.

Here we present a detailed treatment of the energies and analytic gradients of a method which might be fully described as variationally optimized Brueckner orbital coupled-cluster doubles. Previous studies\textsuperscript{9,21} have used the simpler name optimized orbital coupled-cluster doubles (OO CCD), although this fails to convey that the orbitals are indeed approximative Brueckner orbitals. Analogous to the shorthand “Brueckner doubles” (BD) for Brueckner CCD, we will simply employ “optimized doubles” (OD) for convenience. OD is inherently more expensive than BD for energies, since orbital gradient information is required. Nevertheless, we will show that by simultaneously solving for the orbitals, doubles amplitudes, and the amplitude response (Z-vector), we can obtain an OD gradient for roughly the same cost as a BD gradient. We accelerate convergence via the direct inversion of the iterative subspace (DIIS) extrapolation procedure of Pulay,\textsuperscript{25} employing a more straightforward parameterization of the orbitals than used by Hampel \textit{et al.}\textsuperscript{8} OD is the simplest member of a possible hierarchy of coupled-cluster methods with variationally optimized Brueckner orbitals.

We will discuss two advantages of the variational Brueckner orbitals compared to traditional projective Brueckner orbitals. First, the use of variational orbitals provides a formal simplification of the analytical energy gradients, since the wave function obeys a limited Hellmann–Feynman theorem with respect to orbital rotations. Hence, the gradients have no orbital response term, and there is no need to solve the coupled-perturbed BD equations. Determination of one-electron properties is simplified because there is no longer a distinction between “relaxed” and “unrelaxed” densities. The second advantage is that variational Brueckner orbitals are very easily extended to allow inactive orbitals. Hence, the core orbitals and (for basis sets containing them) their very high-lying virtual counterparts that would be kept doubly occupied or doubly unoccupied in the “frozen core” or “frozen virtual” approximations can be optimized for the correlated wave function. Complete details
II. THEORETICAL APPROACH

A. Orbital rotations

We desire a set of orbitals which are variationally optimal for our correlated wave function. We begin by expressing the final molecular orbitals (MOs) as a unitary transformation of an original set of orbitals,

\[
C_{\mu p} = \sum_q C_{\mu q}^a U_{qp} ,
\]

where \(C_{\mu p}\) is a coefficient giving the contribution of atomic orbital (AO) \(\mu\) to MO \(p\). Since \(U\) is a unitary matrix, it preserves the orthonormality of the MOs. The matrix \(U\) may be expressed in terms of a set of orbital rotation angles. The energy of a coupled-cluster wave function with all orbitals active is affected only by rotations which mix occupied with virtual orbitals, so that \(OV\) rotation angles are needed (\(O\) and \(V\) being the number of occupied and virtual orbitals, respectively). More generally, there may be inactive occupied and/or inactive virtual orbitals from/to which excitations are not allowed. We will use \(i,j,k,...\) to denote active occupied orbitals; \(i',j',k',...\) for inactive occupied; \(i'',j'',k'',...\) for all occupied; \(a,b,c,...\) for active virtuals; \(a',b',c',...\) for inactive virtuals; \(a'',b'',c'',...\) for all virtuals; and \(p,q,r,...\) for general orbitals. Then the matrix \(U\) may be written

\[
U = \prod_{\mu} G_{\mu \mu'} \left( \theta_{\mu \mu'} \right) G_{ii'} \left( \theta_{ii'} \right) \prod_{aa'} G_{aa'} \left( \theta_{aa'} \right) ,
\]

where \(G_{pq} \left( \theta_{pq} \right)\) is a Givens or Jacobi rotation matrix which is identical to the unit matrix except for sine and cosine functions at the \(pp, pq, qp,\) and \(qq\) positions:

\[
G_{pq} \left( \theta_{pq} \right) = \begin{pmatrix}
1 & \sin \theta_{pq} & 0 & \cdots \\
\cos \theta_{pq} & 1 & \sin \theta_{pq} & \cdots \\
0 & -\sin \theta_{pq} & 1 & \cdots \\
\cdots & \cdots & \cdots & \cdots 
\end{pmatrix}.
\]

With this parameterization, the orbital optimization leads to the following stationary condition for the energy:

\[
\frac{\partial E}{\partial U} = 0. \tag{4}
\]

Only the term \(\partial E / \partial U\) depends on the electronic structure model. The derivative of \(U\) with respect to a particular orbital rotation angle \(\theta_{pq}\) is simply

\[
\frac{\partial U}{\partial \theta_{pq}} = \begin{bmatrix}
\prod_{(rr) \neq (pq)} G_{rr} \left( \theta_{rr} \right) & \frac{\partial G_{pq}}{\partial \theta_{pq}} & \prod_{(rr) \neq (pq)} G_{rs} \left( \theta_{rs} \right)
\end{bmatrix}.
\]

B. Energy and analytic energy gradients for CCD

We now consider orbital optimization for the specific case of the coupled-cluster doubles (CCD) method.\(^{28,29}\) The CCD wave function is expressed as \(|\Psi\rangle = e^{\hat{T}_2} |\Phi_0\rangle\), where \(|\Phi_0\rangle\) is the “reference” determinant and \(\hat{T}_2\) is a double substitution operator, expressed in second quantization with spin-orbitals as

\[
\hat{T}_2 = \frac{1}{4} \sum_{ijab} a_{ij}^{ab} a_{ij}^{ab} ,
\]

where \(i,j\) (\(a,b\)) run over only active occupied (active virtual) orbitals. The reference determinant is used as a zeroth-order approximation to the electronic state of interest, and it is typically formed from the occupied orbitals of a Hartree–Fock self-consistent-field (SCF) procedure. Here, however, we will obtain molecular orbitals which are optimal not for the single determinant \(|\Phi_0\rangle\) but for the overall CCD wave function \(|\Psi\rangle\). The CCD equations for the total electronic energy \(E\) and the doubles amplitudes \(a_{ij}^{ab}\) are

\[
\langle \Phi_0 | \hat{H} | (1 + \hat{T}_2) \Phi_0 \rangle = E , \tag{8}
\]

\[
\langle \Phi_0 | \hat{H} | (1 + \hat{T}_2 + \frac{1}{2} \hat{T}_2^2) \Phi_0 \rangle = E a_{ij}^{ab} , \tag{9}
\]
Next, we expand these equations in terms of the regular one-electron integrals $h_{pq}$, the antisymmetrized two-electron integrals $\langle pq||rs\rangle$, and the matrix elements of the Fock operator,

$$f_{pq} = h_{pq} + \sum_k \langle pk''|qk''\rangle.$$  \hfill (10)

Note that $k''$ runs over all occupied orbitals, active or inactive. The energy equation (8) becomes

$$E = E_0 + \frac{1}{4} \sum_{ijab} a_{ij}^{ab} \langle ij||ab\rangle,$$  \hfill (11)

where $E_0$ is the reference energy,

$$E_0 = \langle \Phi_0|H|\Phi_0\rangle = \sum_p f_{pq} - \frac{1}{2} \sum_{ij} \langle i''|j''\rangle \langle i''|j''\rangle.$$  \hfill (12)

Equation (9), which defines the doubles amplitudes, becomes

$$\langle ij||ab\rangle + \bar{u}_{ij}^{ab} + \bar{v}_{ij}^{ab} = 0,$$  \hfill (13)

where $\bar{u}_{ij}^{ab}$ and $\bar{v}_{ij}^{ab}$ are defined as

$$\bar{u}_{ij}^{ab} = \sum_c (f_{bc}a_{ij}^{cd} - f_{ac}a_{ij}^{bc}) + \sum_k (f_{ik}a_{jk}^{ab} - f_{kj}a_{ik}^{ab}) + \frac{1}{2} \sum_{cd} \langle ab||cd\rangle a_{ij}^{cd} + \frac{1}{2} \sum_{kl} \langle ij||kl\rangle a_{kl}^{ab} + \sum_{kj} \left[ (ic||kb)a_{jk}^{ac} - (ic||ka)a_{jk}^{bc} + (jc||ka)a_{jk}^{hc} - (jc||kb)a_{jk}^{ac} \right],$$  \hfill (14)

and

$$\bar{v}_{ij}^{ab} = \sum_{klcd} \langle kl||cd\rangle [a_{kl}^{cd} + 4a_{kl}^{ac}a_{kl}^{bd} + 4a_{kl}^{ac}a_{kl}^{bd} + a_{kl}^{cd}a_{kl}^{ac}].$$  \hfill (15)

Now that explicit forms for the energy and amplitude equations have been given, it is helpful to introduce a more compact notation (similar to that used by Fitzgerald et al.\cite{31}) for the development of analytic energy gradients. Gradients for coupled-cluster wave functions using Hartree–Fock-type references have been discussed by several authors,\cite{31-37} and gradients for projective Brueckner orbitals have been presented by Kobayashi et al.\cite{13} In this section we review general CCD gradient theory, and in the following two sections we present gradients with respect to orbital rotations and nuclear displacements for the present method.

Let us rewrite the energy equation (11) in vector form as

$$E(\mathbf{I}, \mathbf{f}, \mathbf{a}) = E_0(\mathbf{I}, \mathbf{f}) + \mathbf{a^T} \mathbf{I}.$$  \hfill (16)

$\mathbf{I}$ and $\mathbf{a}$ are column vectors whose $ijab$th elements are $\langle ij||ab\rangle$ and $a_{ij}^{ab}$, respectively. The vector $\mathbf{I}$ (without overbar) denotes the full set of two-electron integrals. Finally, $\mathbf{f}$ denotes a vector of the Fock matrix elements. In this notation, the doubles equation (13) becomes

$$\mathbf{u} + \mathbf{v} + \mathbf{I} = 0,$$  \hfill (17)

or formally

$$\mathbf{A}(\mathbf{I}, \mathbf{f}, \mathbf{a}) = 0.$$  \hfill (18)

Now consider the derivative of the energy with respect to some parameter $\lambda$,

$$\frac{\partial E}{\partial \lambda} = \frac{\partial E}{\partial \mathbf{I}} \frac{\partial \mathbf{I}}{\partial \lambda} + \frac{\partial E}{\partial \mathbf{f}} \frac{\partial \mathbf{f}}{\partial \lambda} + \frac{\partial E}{\partial \mathbf{a}} \frac{\partial \mathbf{a}}{\partial \lambda},$$  \hfill (19)

which may be written as inner-products:

$$E^\lambda = E^0 + E^\lambda_\mathbf{f} + E^\lambda_\mathbf{a}.$$  \hfill (20)

Here $E^\lambda_\mathbf{X}$ denotes a vector whose $pqrs$th element is $\partial E/\partial X_{pqrs}$, and $X^\lambda$ is a vector whose $pqrs$th element is $\partial X_{pqrs}/\partial \lambda$. The factor $\mathbf{a}^\lambda$ in (20) would appear to require the determination of the first-order changes of the doubles amplitudes with respect to each relevant perturbation $\lambda$. However, this can be avoided and the contribution of this term to the gradient can be computed directly using the $Z$-vector approach of Handy and Schaefer,\cite{24} first applied in coupled-cluster theory by Adamowicz et al.\cite{32} (similar manipulations\cite{34} are often performed in coupled-cluster gradient theory using an operator denoted $\Lambda$). Differentiating the doubles equation (18) yields

$$\mathbf{A}^\lambda = \mathbf{A} + \mathbf{A}^\lambda_\mathbf{f} + \mathbf{A}^\lambda_\mathbf{a} = 0,$$  \hfill (21)

which can be rearranged to solve for $\mathbf{a}^\lambda$:

$$\mathbf{a}^\lambda = -\mathbf{A}^{-1}_\mathbf{a}\mathbf{A}^\lambda_\mathbf{f}.$$  \hfill (22)

Noting that $\mathbf{a}^\lambda$ is multiplied on the left by $E^\lambda_\mathbf{f}$ in Eq. (20), we can introduce the perturbation-independent quantity

$$\mathbf{z}^\lambda = -E^\lambda_\mathbf{f}\mathbf{A}^{-1}_\mathbf{a}.$$  \hfill (23)

From Eq. (16), $E^\lambda_\mathbf{a} = \mathbf{I}$, so that $\mathbf{z}$ is alternatively given by

$$\mathbf{A}^\lambda_\mathbf{z} = -\mathbf{I}.$$  \hfill (24)

Explicitly, the $\mathbf{z}$ vector equation is

$$\mathbf{u}(\mathbf{z}) + \mathbf{v}(\mathbf{a} \times \mathbf{z}) = -\mathbf{I},$$  \hfill (25)

where $\mathbf{u}$ is the same function defined in Eq. (14), but with $\mathbf{z}$ vector elements substituted for doubles amplitudes, and where $\mathbf{v}$ is given by

$$\mathbf{v}_{ij}^{ab}(\mathbf{z} \times \mathbf{a}) = \sum_{klcd} a_{kl}^{cd} [\zeta_{ij}^{cd}(kl||ab) + \zeta_{kl}^{ij}(ij||cd) + \zeta_{kl}^{ij}(ij||cd) + \zeta_{kl}^{ij}(ij||cd)] + 2(\zeta_{ij}^{cd}(kl||ab) - \zeta_{kl}^{ij}(kl||ab) - \zeta_{kl}^{ij}(ij||cd) + 4(\zeta_{kl}^{ij}(ij||cd)).$$  \hfill (26)

The $\mathbf{z}$ vector can then be used to eliminate the need to solve for $\mathbf{a}^\lambda$ explicitly. Using Eqs. (22) and (23), the derivative of the energy with respect to $\lambda$ [Eq. (20)] becomes

$$E^\lambda = E^\lambda_\mathbf{f} + E^\lambda_\mathbf{a} + \mathbf{z}[\mathbf{A}^\lambda_\mathbf{f} + \mathbf{A}^\lambda_\mathbf{a}].$$  \hfill (27)
TABLE I. Effective one- and two-particle density matrices.

<table>
<thead>
<tr>
<th>( \gamma_{ab} )</th>
<th>( \gamma_{ij} )</th>
<th>( \Gamma_{abij} )</th>
<th>( \Gamma_{iajb} )</th>
<th>( \Gamma_{ijkl} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \xi_{ij} )</td>
<td>( \xi_{ji} )</td>
<td>( \xi_{ij} )</td>
<td>( \xi_{ji} )</td>
<td>( \xi_{ij} )</td>
</tr>
<tr>
<td>( \xi_{ij} )</td>
<td>( \xi_{ji} )</td>
<td>( \xi_{ij} )</td>
<td>( \xi_{ji} )</td>
<td>( \xi_{ij} )</td>
</tr>
<tr>
<td>( \xi_{ij} )</td>
<td>( \xi_{ji} )</td>
<td>( \xi_{ij} )</td>
<td>( \xi_{ji} )</td>
<td>( \xi_{ij} )</td>
</tr>
<tr>
<td>( \xi_{ij} )</td>
<td>( \xi_{ji} )</td>
<td>( \xi_{ij} )</td>
<td>( \xi_{ji} )</td>
<td>( \xi_{ij} )</td>
</tr>
<tr>
<td>( \xi_{ij} )</td>
<td>( \xi_{ji} )</td>
<td>( \xi_{ij} )</td>
<td>( \xi_{ji} )</td>
<td>( \xi_{ij} )</td>
</tr>
</tbody>
</table>

Using Eq. (16) for the energy and Eq. (17) for the doubles amplitudes, this can be expanded as

\[
E^\lambda = (E_0)^{\lambda} + (E_0)^{\lambda} + (a + z) T^\lambda + z'(\bar{u}_f + \bar{v}_f)^T^{\lambda}
\]

Introducing effective one- and two-particle density matrices with the same permutational symmetries as the Fock matrix and two-electron integrals, respectively (see Table I), we obtain

\[
E^\lambda = (E_0)^{\lambda} + (E_0)^{\lambda} + \Gamma^T^{\lambda} + \gamma^T^{\lambda}.
\]

Of course the contributions from \( E_0 \) could be added into the definitions of \( \gamma \) and \( \Gamma \), but we will keep them separate so that the indices of \( \gamma \) and \( \Gamma \) run over only active orbitals. Thus our two-particle density matrix \( \Gamma \) is entirely "nonseparable." \(^{38}\)

C. Gradients with respect to orbital rotations

Using the general expressions for the gradient of the energy given in the previous section, we now consider the case of gradients with respect to orbital rotation angles. As described in Sec. II A, the orbital rotation angles \( \theta \) parameterize a unitary transformation matrix \( U \), so that the orbitals are variationally optimized when (4) holds. Since the evaluation of \( \partial U / \partial \theta \) has already been discussed, it is sufficient here to focus on \( \partial E / \partial U \). Using the general expression (29), the derivative of the energy with respect to some particular element \( U_{pq} \) of \( U \) is:

\[
E_{U_{pq}}' = (E_0)_{I}^{U_{pq}} + (E_0)_{I}^{U_{pq}} + \Gamma^{T}_{pq} + \gamma^{T}_{pq}.
\]

As described in the Appendix, this can be evaluated to yield

\[
\frac{\partial E}{\partial U_{pq}} = 2 f_{pq} + 2 \sum_{ab} \gamma_{ab} (ap_{ab}^{\ast} || bi') + 2 \sum_{ab} \gamma_{ab} (ap_{ab}^{\ast} || bi')
\]

D. Gradients with respect to nuclear displacements

Expressions for analytic gradients of coupled-cluster wave functions using nondiagonal Fock operators have been presented previously;\(^{13,35-37}\) in the present case, we need equivalent expressions, but without the orbital response terms. In this section we provide the equation for the energy-weighted density matrix \( W \) in terms of \( \partial E / \partial U \) (see the previous section) and we elucidate the connections between gradients with respect to orbital rotations and gradients with respect to nuclear displacements. In the interest of completeness, we provide all additional equations necessary to evaluate the gradients with respect to nuclear coordinates.

For a displacement by \( \delta x \) of nuclear coordinate \( x \), the change in the (optimized) spin-orbitals may be written\(^{39}\)

\[
\phi_p \rightarrow \phi_p + \delta x \left( \sum_q \bar{U}_{qp}^x \phi_q + \phi_p^{\ast} \right) + O(\delta x^2),
\]

where \( \bar{U}_{qp}^x \) are the standard coupled-perturbed coefficients and where \( \phi_p^{\ast} \) is the molecular orbital evaluated with differentiated atomic orbitals

\[
\phi_p^{\ast} = \sum_{\mu} C_{\mu p} \left( \frac{d X_{\mu}}{dx} \right).
\]

The derivative of the electronic energy with respect to \( x \), \( E' \), is given by the general expression (29), and the only remaining steps are to determine the derivatives of the integrals \( f' \) and \( \Gamma' \). These are\(^{39}\)
It is convenient to distinguish the parts of these expressions due to mixing of molecular orbitals from those due to the changes in atomic orbitals. Following Kobayashi,\textsuperscript{13} we will denote the purely AO parts (assuming real orbitals) as

\begin{align}
\hat{h}_{pq}^\omega & = \sum_{\mu \nu} C_{\mu \nu} C_{\nu \rho} h_{\mu \sigma}^\omega, \\
\langle pq \mid rs \rangle^\omega & = \sum_{\mu \nu \sigma \rho} |C_{\mu \nu} C_{\nu \rho} C_{\sigma \rho}|^2 \langle \mu \nu \mid \rho \sigma \rangle^\omega, \\
\hat{f}_{pq}^\omega & = \sum_{\mu \nu} C_{\mu \nu} C_{\nu \rho} f_{\mu \sigma}^\omega.
\end{align}

Comparing the expressions for \( h_{pq}^\omega \) and \( \langle pq \mid rs \rangle^\omega \) to the derivatives with respect to \( U_{pq}^\omega \) in the Appendix reveals that they can be rewritten as

\begin{align}
\hat{h}_{pq}^\omega & = \hat{h}_{pq}^\omega + \sum r s U_{r s \sigma} \frac{\partial h_{pq}^\omega}{\partial U_{r s}}, \\
\langle pq \mid rs \rangle^\omega & = (\langle pq \mid rs \rangle)^\omega + \sum r s \frac{\partial \langle pq \mid rs \rangle^\omega}{\partial U_{r s}}, \\
\hat{f}_{pq}^\omega & = f_{pq}^\omega + \sum r s \frac{\partial f_{pq}^\omega}{\partial U_{r s}}.
\end{align}

Hence, the gradients with respect to nuclear displacements may be evaluated as

\begin{align}
E^\omega = E_0^\omega + \sum_{pq} \gamma_{pq} f_{pq}^\omega + \sum_{pqrs} \Gamma_{pqrs} \langle pq \mid rs \rangle^\omega \\
& + \sum_{pq} X_{pq} U_{pq}^\omega,
\end{align}

where \( E_0^\omega \) is given by

\begin{equation}
E_0^\omega = \sum_{pq} f_{pq}^\omega - \frac{1}{2} \sum_{i'j'} \left( \langle i''j'' \mid i'j' \rangle \right)^\omega.
\end{equation}

and \( X_{pq} \) is the Lagrangian in the current basis,

\begin{equation}
X_{pq} = \sum_{r} \frac{\partial E}{\partial U_{rq}} U_{rp}^\omega.
\end{equation}

If we consider the situation in which we reset the reference set of orbitals to the current set (i.e., making all \( \theta_{pq} = 0 \)), it is clear from Eqs. (4)\textendash(6) that our Lagrangian \( X \) must be symmetric, just as in any other case where the orbitals are optimized for the wave function being differentiated. Maintaining the orthonormality of the orbitals to first order requires that

\begin{equation}
U_{pq}^\omega + U_{qp}^\omega + S_{pq}^\omega = 0,
\end{equation}

where \( S_{pq}^\omega \) is the normal derivative overlap matrix.\textsuperscript{39} Combined with the symmetric nature of \( X \), this means that we can completely eliminate the need to solve for the coupled-perturbed coefficients \( \hat{U}_{pq}^\omega \), either explicitly or implicitly via an orbital response \( Z \)-vector. Thus, the gradient with respect to nuclear displacements simplifies to

\begin{align}
E^\omega & = E_0^\omega + \sum_{pq} \gamma_{pq} f_{pq}^\omega + \sum_{pqrs} \Gamma_{pqrs} \langle pq \mid rs \rangle^\omega \\
& + \sum_{pq} W_{pq} S_{pq}^\omega,
\end{align}

where \( W_{pq} = -X_{pq} / 2 \). This gradient is evaluated in the usual way by back-transforming the effective one- and two-particle density matrices and \( W \) into the atomic orbital basis and contracting against the appropriate atomic orbital derivative integrals. By ensuring that the AO effective density matrices have the same permutational symmetries as the integrals, and by folding in the contributions from \( E_0^\omega \), the gradient can be cast into the usual form\textsuperscript{\textsuperscript{20}}

\begin{align}
E^\omega & = \sum_{\mu \sigma} \gamma_{\mu \sigma} \hat{f}_{\mu \sigma}^\omega + \sum_{\mu \sigma \nu \rho} \Gamma_{\mu \nu \rho \sigma} \langle \mu \nu \mid \rho \sigma \rangle^\omega \\
& + \sum_{\mu \sigma \nu \rho} \hat{W}_{\mu \nu \rho \sigma} S_{\mu \sigma}^\omega.
\end{align}

Note that \( \hat{h}^\omega \) can be used rather than \( \hat{f}^\omega \) by appropriate modification of \( \Gamma \).

E. Convergence procedure

Our variationally optimized Brueckner orbital CCD wave function is defined by a set of converged orbitals (parameterized by orbital rotation angles \( \theta \) and double substitution amplitudes \( \mathbf{a} \)). To optimize the molecular orbitals, we also require the gradient of the energy with respect to orbital rotations, which involves the amplitude response \( \mathbf{z} \). Hence, at convergence, Eqs. (4), (13), and (25) are all satisfied. The most straightforward convergence scheme would involve choosing an initial set of orbitals, converging the amplitudes \( \mathbf{a} \) and amplitude response \( \mathbf{z} \), obtaining the orbital gradient via (31)\textendash(34) and (35), taking a step in \( \theta \) space along the gradient direction, re-transforming the AO integrals, and repeating the process until the orbital gradient becomes small. However, this procedure is rather costly, as it involves the construction of the one- and two-particle density matrices and converging \( \mathbf{z} \) for every iteration required to converge the orbitals.

Computational efficiency can be dramatically increased by converging \( \mathbf{a} \), \( \mathbf{z} \), and \( \theta \) all simultaneously. We accelerate convergence by employing Pulay’s direct inversion of the iterative subspace (DIIS) procedure\textsuperscript{\textsuperscript{25}} on a compound param-

\[ h_{pq}^\omega = \sum r h_{pq}^\omega + \sum_{\mu \nu} C_{\mu \nu} C_{\nu \rho} h_{\mu \sigma}^\omega, \]

\[ \langle pq \mid rs \rangle^\omega = \sum_{\mu \nu \sigma \rho} |C_{\mu \nu} C_{\nu \rho} C_{\sigma \rho}|^2 \langle \mu \nu \mid \rho \sigma \rangle^\omega, \]

\[ f_{pq}^\omega = \sum_{\mu \nu} C_{\mu \nu} C_{\nu \rho} f_{\mu \sigma}^\omega. \]
Theoretical predictions of total energies, equilibrium geometries, and harmonic vibrational frequencies for H2O using DZP and TZ2P basis sets.

<table>
<thead>
<tr>
<th>Method</th>
<th>Energy</th>
<th>(r_e)</th>
<th>(\theta_e)</th>
<th>(\omega_1(a_1))</th>
<th>(\omega_2(a_2))</th>
<th>(\omega_3(h_2))</th>
</tr>
</thead>
<tbody>
<tr>
<td>DZP SCF</td>
<td>76.047.099</td>
<td>0.9437</td>
<td>106.63</td>
<td>4165</td>
<td>1752</td>
<td>4288</td>
</tr>
<tr>
<td>DZP CISD</td>
<td>76.258.913</td>
<td>0.9577</td>
<td>104.88</td>
<td>3970</td>
<td>1694</td>
<td>4094</td>
</tr>
<tr>
<td>DZP CCSD</td>
<td>76.267.869</td>
<td>0.9610</td>
<td>104.63</td>
<td>3915</td>
<td>1684</td>
<td>4044</td>
</tr>
<tr>
<td>DZP BD(^p)</td>
<td>76.267.659</td>
<td>0.9609</td>
<td>104.64</td>
<td>3917</td>
<td>1684</td>
<td>4047</td>
</tr>
<tr>
<td>DZP OD</td>
<td>76.267.733</td>
<td>0.9609</td>
<td>104.64</td>
<td>3917</td>
<td>1684</td>
<td>4046</td>
</tr>
<tr>
<td>TZ2P SCF</td>
<td>76.061.002</td>
<td>0.9403</td>
<td>106.31</td>
<td>4133</td>
<td>1756</td>
<td>4235</td>
</tr>
<tr>
<td>TZ2P CISD</td>
<td>76.320.581</td>
<td>0.9526</td>
<td>105.02</td>
<td>3940</td>
<td>1682</td>
<td>4044</td>
</tr>
<tr>
<td>TZ2P CCSD</td>
<td>76.331.905</td>
<td>0.9563</td>
<td>104.74</td>
<td>3879</td>
<td>1668</td>
<td>3987</td>
</tr>
<tr>
<td>TZ2P BD(^d)</td>
<td>76.331.632</td>
<td>0.9561</td>
<td>104.76</td>
<td>3884</td>
<td>1669</td>
<td>3992</td>
</tr>
<tr>
<td>TZ2P OD</td>
<td>76.331.632</td>
<td>0.9561</td>
<td>104.76</td>
<td>3882</td>
<td>1669</td>
<td>3990</td>
</tr>
<tr>
<td>Expt.(^e)</td>
<td></td>
<td>0.9578</td>
<td>104.5</td>
<td>3832</td>
<td>1649</td>
<td>3942</td>
</tr>
</tbody>
</table>

\(^a\)All electrons correlated. Bond lengths in Å, bond angles in degrees, and frequencies in cm\(^{-1}\). DZP basis set from Ref. 13, TZ2P set from Ref. 17.
\(^b\)Results from Kobayashi et al., Ref. 13.
\(^c\)Results from Kobayashi, Amos, and Handy, Ref. 17.
\(^d\)Experimental results from Hoy et al., Refs. 58 and 59.

III. RESULTS AND DISCUSSION

Results for CCD using variationally optimized Brueckner orbitals (denoted here by the shorthand OD) have been obtained for several small molecules for comparison to BD and to CCSD using Hartree–Fock orbitals. Additionally, results are reported for Hartree–Fock self-consistent-field (SCF) and configuration interaction with single and double substitutions (CISD).\(^{41–43}\) Geometries were optimized using analytic gradients for the SCF,\(^{44,45}\) CISD,\(^{46,47}\) CCSD,\(^{33,37}\) and OD methods, and using numerical gradients computed from energies for projective BD. Harmonic vibrational frequencies were determined via analytic second derivatives for SCF;\(^{38,49}\) by finite differences of analytic gradients for CISD, CCSD, and OD; and by finite differences of energies for BD.

Three basis sets of contracted Gaussian functions have been employed in the present study. The first, used for H2O, is the double-\(\zeta\) plus polarization (DZP) basis set used by Kobayashi et al.,\(^{17}\) which is comprised of the standard Huzinaga-Dunning\(^{50,51}\) double-\(\zeta\) set augmented by a set of six Cartesian \(d\)-type polarization functions for oxygen \([\alpha_d(O)=0.9]\), and by a set of three \(p\)-type polarization functions \([\alpha_p(H)=1.0]\) for hydrogen. The contraction scheme is thus \((9s5p1d4s2p1d)\) for oxygen and \((4s1p2s1p)\) for hydrogen. The second basis set, used for H2O, HF, and N\(_2\), is the triple-\(\zeta\) plus double polarization (TZ2P) basis employed by Kobayashi et al.,\(^{17}\) constructed from Dunning’s\(^{52}\) 5s4p contraction of Huzinaga’s 10s6p primitive set\(^{50}\) [3s contraction of 5s for H, exponents scaled by \((1.25)^2\)] augmented by two sets of six \(d\)-type polarization functions for first row atoms \([\alpha_d(O)=1.35, 0.45; \alpha_d(N)=1.35, 0.45; \alpha_d(F)=2.0, 0.6667]\) and two sets of \(p\)-type functions for hydrogen \([\alpha_p(H)=1.5, 0.5]\). The third and final basis set, used for O\(_2\), is the standard 6-31G* basis of Pople and co-workers,\(^{33,54}\) which was employed in previous benchmark studies of this molecule.\(^6\) Results were obtained using the Q-Chem\(^{55}\) and PSI\(^{36}\) packages of \textit{ab initio} quantum chemistry programs. Coupled-cluster predictions for O\(_2\) and all BD results were obtained using ACES II.\(^{37}\)

Table II presents total energies, equilibrium geometries, and harmonic vibrational frequencies for H2O with the DZP and TZ2P basis sets. The most remarkable feature of the table is that the present OD results are nearly identical to the BD data of Kobayashi et al.\(^{13,17}\) The energy difference is quite small, being less than a tenth of a millihartree, and the equilibrium geometries are essentially the same. The largest difference in harmonic vibrational frequencies is a mere...
force constants: the anharmonic constants and the following spectroscopic quantities derived from the Table III reports the total energies, equilibrium bond lengths, correlated SCF method, all of the coupled-cluster models are CCSD for all spectroscopic constants. In contrast to the un-

\[ \text{D} \]

predictions also lie very close

\[ \text{v} \]

to the equilibrium bond length

\[ \text{e x e} \]

and includ-

\[ \text{\& e} \]

ing double and (explicit or effective) single substitutions are generally much larger than those observed between the coupled-cluster methods considered here. Compared to the experimental results of Hoy et al.,58,59 all of the coupled-cluster methods perform rather well with the TZ2P basis.

To provide a more detailed comparison of the OD and BD potential energy surfaces near equilibrium, force constants through quartic terms were determined for HF and N₂ using the TZ2P basis by fitting a fourth-order polynomial to five energy points uniformly distributed about (and including) the equilibrium bond length (± 0.005 Å, ± 0.010 Å).

Table III reports the total energies, equilibrium bond lengths, and the following spectroscopic quantities derived from the force constants: the anharmonic constants \( \omega_e \), rotational constants \( B_e \), centrifugal distortion constants \( D_e \), and vibration-rotation coupling constants \( \alpha_e \). Once again, the BD and OD results are nearly identical, the largest difference being \( \Delta \epsilon_e = 0.0003 \) Å and \( \Delta \omega_e = 4 \) cm\(^{-1} \) for N₂. The OD predictions also lie very close (most differences < 0.2%) to CCSD for all spectroscopic constants. In contrast to the uncorrelated SCF method, all of the coupled-cluster models are in good agreement with experiment;\(^{60} \) bond lengths are slightly underestimated (by ~0.002 Å) and frequencies are overestimated by about 1% for HF and 3% for N₂.

Finally, we consider the performance of OD for a considerably more challenging case, the \( ^3 \text{B}_1 \) state of rectangular \( \text{O}_4^+ \), a molecule of importance in atmospheric ion chemistry.\(^{61} \) In a groundbreaking 1994 theoretical study, Barnes and Lindh\(^{62} \) showed that many approximate wave functions—including even complete active space self-consistent-field (CASSCF) with a 2\( p \) active space—demonstrate artifactual symmetry breaking\(^{18,19} \), whereby the wave function fails to exhibit the full point group symmetry of the molecule. If symmetry constraints are relaxed, the CASSCF wave function allows the electron hole to localize on one of the \( O_2 \) units, leading to a wave function of \( C_{2v} \) rather than \( D_{2h} \) symmetry. Symmetry breaking appears to be a problem for many radicals, including \( \text{LiO}_2, \text{N}_2^+ \), \( \text{NO}_2, \text{NO}_3 \), \( \text{NO}_5 \), allyl radical,\(^{66} \) formyloxy radical,\(^{67} \) \( \text{C}_2^+ \), \( \text{HOOH}^+ \),\(^{69} \) Evaluation of vibrational frequencies for symmetry breaking cases can easily become a theoretician’s nightmare, for at least two reasons. First, the standard finite difference procedure requires gradients at displaced geometries where the full point group symmetry can no longer be imposed; thus, it becomes technically challenging to stay on the delocalized solution when there are lower-lying localized solutions which are no longer excluded by symmetry. Second, even if the correct delocalized solution is obtained, it may exhibit unphysical force constants because of rapid mixing with the localized solutions, as documented for \( \text{NO}_2, \text{HCO}_2, \text{LiO}_2 \),\(^{19,63} \) and \( \text{O}_4^+ \).

Table IV presents energies, geometries, and frequencies for rectangular \( \text{O}_4^+ \). \( R_{\text{DD}} \) denotes the interatomic distance within each \( O_2 \) subunit, and \( R_{\text{cm}} \) denotes the distance between the two parallel \( O_2 \)'s. The UHF and ROHF frequencies for the antisymmetric \( b_{3\alpha} \) stretch in Table IV are clearly unphysical. The behavior of \( \omega_5(b_{3\alpha}) \) also appears anomalous, going from 5434 Å for UHF to 83 for ROHF. When CCSD is applied, the character of this mode changes, the frequency becoming real (154 Å \( \text{cm}^{-1} \)) for UHF CCSD and imaginary (66 Å \( \text{cm}^{-1} \)) for ROHF CCSD. For the \( b_{3\alpha} \) stretch, although CCSD lowers the unreasonably high Hartree-Fock frequencies considerably (to 1726 Å \( \text{cm}^{-1} \) and 2036 Å \( \text{cm}^{-1} \) with the UHF and ROHF references, respectively), the difference between these two results (310 Å \( \text{cm}^{-1} \)) remains unusually large. Furthermore, neither value is close to the most reliable theoretical estimate of Lindh and Barnes,\(^{62} \) who obtained 1271 Å \( \text{cm}^{-1} \) using a TZ2P basis and the CASSCF state interaction (CASSI) approach, which is a nonorthogonal CI treatment including both of the localized CASSCF solutions. Lindh and Barnes have proposed\(^{62} \) that this mode is responsible for the 1320 Å \( \text{cm}^{-1} \) band observed\(^{60} \) in matrix infrared spectra of \( \text{O}_4 \), and Jacob and Thompson have confirmed the plausibility of this assignment.\(^{70} \) We must note that, although they included no other CCSD data in their study of \( \text{O}_4 \), Barnes and Lindh\(^{62} \) reported a UHF CCSD prediction of 1220 Å \( \text{cm}^{-1} \) for the \( b_{3\alpha} \) stretch. Although this value is much more reasonable than ours (1726 Å \( \text{cm}^{-1} \)), we were unable to reproduce it despite very careful measures to ensure that we

\[ \begin{array}{cccccc}
\text{Method} & \text{Energy} & r_e & \omega_e & \omega_{e,x} & B_e & D_e \\
& (\text{hartrees}) & (\AA) & (\text{cm}^{-1}) & (\text{cm}^{-1}) & (\text{cm}^{-1}) & (10^{-6} \text{ cm}^{-1}) \\
\hline
\text{HF} & \text{SCF} & -100.063 & 742 & 0.8977 & 4469 & 86.1 & 21.860 & 2092 & 0.756 \\
& \text{CCSD} & -100.342 & 472 & 0.9148 & 4184 & 92.2 & 21.048 & 2130 & 0.803 \\
& \text{BD} & -100.342 & 059 & 0.9145 & 4191 & 92.3 & 21.063 & 2129 & 0.801 \\
& \text{OD} & -100.342 & 171 & 0.9146 & 4189 & 92.3 & 21.059 & 2129 & 0.802 \\
& \text{Expt.} & & 0.9168 & 4138 & 89.9 & 20.956 & 2151 & 0.798 \\
\hline
\text{N}_2 & \text{SCF} & -108.986 & 905 & 1.0670 & 2730 & 11.0 & 2.115 & 5.08 & 0.014 \\
& \text{CCSD} & -109.388 & 073 & 1.0960 & 2426 & 13.3 & 2.004 & 5.47 & 0.016 \\
& \text{BD} & -109.386 & 999 & 1.0951 & 2437 & 13.1 & 2.008 & 5.45 & 0.016 \\
& \text{OD} & -109.387 & 279 & 1.0954 & 2433 & 13.2 & 2.007 & 5.46 & 0.017 \\
& \text{Expt.} & & 1.0977 & 2359 & 14.3 & 1.998 & 5.76 & 0.017 \\
\hline
\end{array} \]

*Basis set from Ref. 17. All electrons correlated.

*Data from Huber and Herzberg, Ref. 60.
remained on the same UHF solution for displaced geometries. Furthermore, our frequency for this mode has been confirmed using analytic CCSD second derivative methods.72

These anomalously high CCSD vibrational frequency predictions for the \(b_{3u}\) stretch seem surprising given the well-known insensitivity of CCSD to the molecular orbitals; the the presence of \(\epsilon^{T}\) accounts for most of the effects of orbital relaxation.73 One possible explanation lies in the recent work of Crawford et al.,63 who demonstrated that unusually high vibrational frequencies in symmetry breaking cases can result when eigenvalues of the molecular orbital Hessian approach zero; this phenomenon affects both the Hartree–Fock method and correlated methods based on a Hartree–Fock reference. Force constants from many-body perturbation theory were affected over the largest range of geometries about the MO Hessian singularity, while CCSD(T) was affected over a smaller range, and CCSD was affected only over a very narrow region. We performed a Hartree–Fock stability analysis at the UHF CCSD and ROHF CCSD optimum geometries and obtained MO Hessian eigenvalues of \(-0.04\) and \(-0.01\), respectively, for \(b_{3u}\) rotations mixing the localized \(C_{2v}\) and delocalized \(D_{2h}\) solutions.75 It is not obvious that these eigenvalues are sufficiently close to zero to have a large, direct effect on the \(b_{3u}\) antisymmetric stretching frequency. Moreover, UHF stability analysis shows an eigenvalue of much smaller magnitude \((-0.004)\) of \(b_{1g}\) symmetry, but the \(b_{1g}\) vibration seems less problematic than the \(b_{3u}\) mode. Hence, it is unclear whether the overestimation of the \(b_{3u}\) stretching frequency by CCSD is driven by near-singularities in the MO Hessian at the CCSD equilibrium geometry.76

In contrast to CCSD, there was no difficulty in obtaining the BD and OD frequencies, since the same set of Brueckner orbitals will be obtained regardless of which UHF solution is used as an initial guess. The BD and OD methods predict substantially lower values for the antisymmetric \(b_{3u}\) stretch (1194 cm\(^{-1}\) and 1193 cm\(^{-1}\), respectively). These are in much better agreement with the TZP CASSI result and with experiment; a closer match is precluded by the use of the small 6-31G\(^*\) basis set. Note that, even for this demanding case, the BD and OD results are remarkably similar except for the \(b_{1g}\) mode, for which they differ by 30 cm\(^{-1}\). This single difference seems somewhat surprising, given that the other frequencies differ by 5 cm\(^{-1}\) or less, and that the \(b_{1g}\) mode does not seem theoretically challenging like the \(b_{2u}\) or \(b_{3u}\) modes. However, the OD value for \(\omega_3(b_{1g})\) exactly matches the ROHF CCSD prediction, and indeed the OD and ROHF CCSD results are all nearly identical except for the troublesome \(b_{2u}\) and \(b_{3u}\) vibrations. Furthermore, the same can be said for the comparison between BD and UHF CCSD, which differ in their prediction of \(\omega_3(b_{1g})\) by 1 cm\(^{-1}\). Finally, as reported by Barnes and Lindh,6 although the CCSD(T) method is unable to overcome the effects of symmetry breaking on the antisymmetric \(b_{3u}\) stretch (yielding 1922 cm\(^{-1}\)), the use of Brueckner orbitals [BD(T)] gives a much more reasonable 1322 cm\(^{-1}\). The present BD and OD predictions are in reasonable agreement with the latter result.

**IV. CONCLUSIONS**

We have presented the variational optimization of molecular orbitals for a correlated wave function lacking single excitations as an alternative method for obtaining approximate Brueckner orbitals. Equations for the energy and analytic energy derivative, and a description of the computational procedure for solving them, have been presented for the coupled-cluster doubles (CCD) model using these variational Brueckner orbitals. The present paper has expanded upon previous investigations9,21 of orbital optimization for CCD in the following ways: (1) We have noted a theoretical connection between orbitals optimized for CCD and BD orbitals; namely, they become equivalent when the correlation treatment is improved to the full CI limit. (2) We have presented detailed, explicit procedures for evaluating the OD energy; moreover, the present approach is dramatically more efficient than the test algorithm sketched by Scuseria and Schaefer.21 (3) We have presented the details necessary to allow for the optimization of inactive orbitals. (4) Although they are straightforward to obtain, we have included explicit

### Table IV. Theoretical predictions of total energies, equilibrium geometries, and harmonic vibrational frequencies for the \(^{18}\)O\(^{2}\) state of rectangular O\(^{2}\) using the 6-31G\(^*\) basis set.

<table>
<thead>
<tr>
<th>Method</th>
<th>Energy</th>
<th>(R_{\infty})</th>
<th>(R_{cm})</th>
<th>(\omega_1(a_u))</th>
<th>(\omega_2(a_u))</th>
<th>(\omega_3(b_{1g}))</th>
<th>(\omega_4(b_{3u}))</th>
<th>(\omega_5(b_{2u}))</th>
<th>(\omega_6(b_{1g}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>UHF</td>
<td>-298.76</td>
<td>577</td>
<td>1.1176</td>
<td>2.4058</td>
<td>2242</td>
<td>255</td>
<td>396</td>
<td>197</td>
<td>543i</td>
</tr>
<tr>
<td>ROHF</td>
<td>-298.739</td>
<td>290</td>
<td>1.1152</td>
<td>2.4073</td>
<td>2256</td>
<td>258</td>
<td>417</td>
<td>199</td>
<td>83</td>
</tr>
<tr>
<td>UHF CCSD</td>
<td>-298.484</td>
<td>239</td>
<td>1.1737</td>
<td>2.3792</td>
<td>1808</td>
<td>269</td>
<td>341</td>
<td>178</td>
<td>154</td>
</tr>
<tr>
<td>ROHF CCSD</td>
<td>-298.482</td>
<td>497</td>
<td>1.1730</td>
<td>2.3793</td>
<td>1813</td>
<td>269</td>
<td>372</td>
<td>179</td>
<td>66i</td>
</tr>
<tr>
<td>BD</td>
<td>-298.482</td>
<td>212</td>
<td>1.1723</td>
<td>2.3794</td>
<td>1821</td>
<td>269</td>
<td>342</td>
<td>179</td>
<td>82</td>
</tr>
<tr>
<td>OD</td>
<td>-298.482</td>
<td>683</td>
<td>1.1728</td>
<td>2.3793</td>
<td>1816</td>
<td>269</td>
<td>372</td>
<td>179</td>
<td>84</td>
</tr>
<tr>
<td>UHF CCSD(T)</td>
<td>-299.512</td>
<td>786</td>
<td>1.1846</td>
<td>2.3751</td>
<td>1713</td>
<td>271</td>
<td>372</td>
<td>175</td>
<td>97</td>
</tr>
<tr>
<td>BD(T)</td>
<td>-299.512</td>
<td>778</td>
<td>1.1846</td>
<td>2.3751</td>
<td>1713</td>
<td>270</td>
<td>372</td>
<td>175</td>
<td>97</td>
</tr>
<tr>
<td>TZP CASSI</td>
<td>-298.739</td>
<td>286</td>
<td>1.1152</td>
<td>2.4073</td>
<td>2256</td>
<td>258</td>
<td>417</td>
<td>199</td>
<td>83</td>
</tr>
</tbody>
</table>

*All electrons correlated. Bond lengths in Å, bond angles in degrees, and frequencies in cm\(^{-1}\).*

*From Barnes and Lindh, Ref. 6.

*From Lindh and Barnes, Ref. 62. Result of nonorthogonal configuration interaction between two symmetry-broken CASSCF wave functions using a larger TZP basis.

*From Jacox and Thompson, Ref. 70.
expressions for analytic energy gradients with respect to nuclear coordinates for the OD wave function. Furthermore, we have presented the first results for OD optimized geometries, harmonic vibrational frequencies, and other spectroscopic constants.

For well-behaved molecules, OD predictions of molecular properties appear to be virtually identical to those from the standard projective BD procedure. Furthermore, the CCSD results are also quite similar. Symmetry breaking effects in $O_4^+$ magnify these differences and plague attempts to obtain reliable vibrational frequencies using CCSD. In contrast, the BD and OD frequencies are straightforward to obtain, owing to their independence from the choice of starting orbitals; moreover, they yield uniformly reasonable frequencies. The largest difference between the two Brueckner approaches (BD and OD) is a 30 cm$^{-1}$ change for the otherwise well-behaved $\alpha(b_1\gamma)$ frequency, where the OD result is the same as the ROHF CCSD result and the BD result is virtually the same as UHF CCSD.

Although they yield very similar numerical results, the variational Brueckner orbitals offer the following advantages compared to standard (projective) Brueckner orbitals: (1) They provide a formal simplification of analytic energy gradients because, like the exact wave function, the OD wave function satisfies the Hellmann-Feynman theorem with respect to orbital rotations. (2) Because there are no orbital response terms, there is no need to distinguish between relaxed and unrelaxed densities for one-electron property derivatives. (3) The variational Brueckner orbitals are easily extended to allow for inactive orbitals, which are variationally optimized but not allowed to vary their occupancy in the correlation procedure. Hence, the present work immediately allows for models of nondynamical correlation which employ a CCD expansion in the active space. We will describe such approaches in a subsequent publication.

ACKNOWLEDGMENTS

This work was sponsored by the U.S. National Science Foundation through Grant No. CHE-9357129 and through a postdoctoral fellowship to C.D.S. (Grant No. CHE-9626094). A.I.K. acknowledges support from a M. Kennedy-Li Prize. M.H.-G. acknowledges support by a fellowship from the David and Lucile Packard Foundation. We thank Dr. Steve Gwaltney for assistance with the CCSD predictions for $O_4^+$. C.D.S. wishes to thank Dr. Yukio Yamaguchi for assistance in interfacing the OD gradients to the derivative integrals in PSI, and Dr. Daniel Crawford for helpful discussions and for providing analytic UHF CCSD frequencies for $O_4^+$.

APPENDIX

In order to evaluate $\partial E / \partial U$ explicitly, we must determine the derivatives of the Fock matrix elements and two-electron integrals with respect to a general element $U_{pq}$ (i.e., the terms $U_{pq}^f$ and $U_{pq}^g$, respectively). Using (1), we may write a one-electron integral in the current MO basis as

$$h_{pq} = \sum_{\mu \nu \alpha \beta} h_{\mu \nu} \alpha \beta U_{\mu \alpha} \beta \nu C_{\nu \gamma} U_{\gamma \alpha} \beta \nu .$$

(A1)

The derivative is thus

$$\frac{\partial h_{pq}}{\partial U_{rs}} = \sum_{\mu \nu \alpha \beta} h_{\mu \nu} \alpha \beta U_{\mu \alpha} \beta \nu \delta_{pqrs} + \sum_{\mu \nu} h_{\mu \nu} \alpha \beta U_{\mu \alpha} \beta \nu \delta_{qrs} ,$$

(A2)

where we employ a superscript $\alpha$ to denote an orbital in the original (unrotated) basis. Similarly, the partial derivatives of the two-electron integrals are

$$\frac{\partial \langle pq || rs \rangle}{\partial U_{xy}} = \langle x^\alpha q || rs \rangle \delta_{pqxy} + \langle p x^\alpha || rs \rangle \delta_{qxy} + \langle pq || x^\alpha s \rangle \delta_{xy} .$$

(A3)

The Fock matrix derivatives are therefore

$$\frac{\partial f_{pq}}{\partial U_{mn}} = f_{pq} \delta_{mn} + f_{pm} \delta_{qmn} + \langle pm || nq \rangle + \langle qn || mp \rangle ,$$

(A4)

$$\frac{\partial f_{pq}}{\partial U_{pq}} = f_{pq} \delta_{pq} + f_{pm} \delta_{qmp} .$$

(A5)

With these definitions, it is straightforward to show that the derivative of the single-determinant energy $E_0$ is

$$\frac{\partial E_0}{\partial U_{pq}} = 2 f_{pq} .$$

(A6)

Our definitions of $\delta_{ij}^{ab}$ and $\tilde{\delta}_{ij}^{ab}$ correspond to those of Ref. 28, except that we have dropped all terms involving singles amplitudes (which are set equal to zero in our method) and we have absorbed the Fock matrix elements into $\tilde{u}$.


T. D. Crawford and J. F. Stanton, private communication.


The UHF stability analysis was performed using ACES II, while the ROHF stability analysis was obtained using PSI and allowed only spin-conserving rotations.

A preliminary investigation indicates that the analytic ROHF second derivatives are not rapidly varying as a function of nuclear configuration, as would be expected for cases where anomalous vibrational frequencies are directly driven by Hartree-Fock instabilities.