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

Articles you may be interested in

Ab initio molecular orbital study of structures and energetics of Si 3 H 2 , Si 3 H 2 + , and Si 3 H 2
J. Chem. Phys. 120, 11071 (2004); 10.1063/1.1740747

Reactions of O 2 + with OH-isoprene adduct isomers: Exothermicity, product ions, and rate constants
J. Chem. Phys. 113, 8574 (2000); 10.1063/1.1318736

Ab initio study of the low-lying electronic states of Ag 3 , Ag 3 , and Ag 3 + : A coupled-cluster approach
J. Chem. Phys. 112, 9335 (2000); 10.1063/1.481553

Coupled cluster anharmonic force fields, spectroscopic constants, and vibrational energies of AIF 3 and SiF 3 +
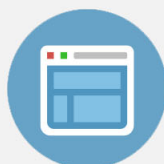J. Chem. Phys. 107, 1717 (1997); 10.1063/1.474613

The vibrational spectrum of H 2 O 2 + radical cation: An illustration of symmetry breaking
J. Chem. Phys. 106, 4877 (1997); 10.1063/1.473537

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

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 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.* [S0021-9606(98)30235-4]

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[1] and were related to the CI method in 1958 by Nesbet.[2] Since the full CI wave function is too costly to compute for all but the smallest molecular systems,[3] 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.[4–6]

In 1981, Chiles and Dykstra[7] 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 ($\hat{T}_1 = 0$); more recently, the method has been called[5,8] Brueckner coupled-cluster doubles (BCCD or B-CCD) or simply Brueckner doubles (BD).[4] 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.*[4] 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)[9] and quadratic CI with singles and doubles (QCISD)[10] at fifth order. More complete Brueckner coupled-cluster models include connected triple excitations in complete (BCCDT or BDT)[11] or perturbative [BCCD(T) or BD(T)][4,12] fashions.

In 1991, Kobayashi *et al.* presented analytic gradients for BD[13] and BD(T),[14] enabling systematic studies of equilibrium geometries, dipole moments, vibrational frequencies, infrared intensities, and other properties of several small molecules.[13–17] These studies indicate that the differences between BD and CCSD [or BD(T) and CCSD(T)] 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 artifactual symmetry breaking[18,19] than coupled-cluster based on Hartree–Fock orbitals, as shown by Stanton *et al.*[5] for NO_3 and Barnes and Lindh[6] for O_4^+. Additionally, Handy *et al.*[4] 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.*[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 *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 CCD method. The idea of variationally optimizing orbitals for the CCD wave function has been considered previously by Purvis and Bartlett[9] and by Scuseria and Schaefer.[21] In their classic 1982 paper on the CCSD method,[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_2O molecule in a double-ζ (DZ) basis using CCD with optimized orbitals. In subsequent work, Purvis *et al.*[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.[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;[23] hence, these authors optimized orbitals by zeroing the orbital Z-vector[24] using an iterative process involving successive gradient calculations followed by orbital rotations. This procedure is similar to that reported by Rendell *et al.* in the context of CI singles and doubles (CISD) wave functions.[23] Scuseria and Schaefer noted great difficulty in optimizing orbitals for CCSD due to the presence of $e^{\hat{T}_1}$, which already describes orbital relaxation; however, they reported good convergence for CCD. When these authors constructed CCSD wave functions using the CCD 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[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 CCD 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[8] achieved a breakthrough in the efficiency of projective Brueckner CCD computations by solving for the orbitals and doubles amplitudes *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[9,21] have used the simpler name optimized orbital coupled-cluster doubles (OO CCD), although this fails to convey that the orbitals are indeed approximate 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,[25] employing a more straightforward parameterization of the orbitals than used by Hampel *et al.*[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

for the optimization of inactive orbitals will be presented below. In a subsequent paper,[26] we will consider the consequences of making all nonvalence orbitals inactive (i.e., those orbitals are still variationally optimized, but excitations are not allowed out of or into them). In that work, we will demonstrate that this amounts to a size-consistent model of nondynamical correlation which differs from current multiconfigurational self-consistent-field (MCSCF) approaches only in that it employs an exponential (coupled-cluster) wave function ansatz rather than the usual linear (configuration interaction) ansatz.

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,

$$C_{\mu p} = \sum_q C^o_{\mu q} U_{qp}, \tag{1}$$

where $C_{\mu p}$ is a coefficient giving the contribution of atomic orbital (AO) μ 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 $i,j,k,\ldots$ to denote active occupied orbitals; $i',j',k',\ldots$ for inactive occupieds; $i'',j'',k'',\ldots$ for all occupieds; $a,b,c,\ldots$ for active virtuals; $a',b',c',\ldots$ for inactive virtuals; $a'',b'',c'',\ldots$ for all virtuals; and $p,q,r,\ldots$ for general orbitals. Then the matrix **U** may be written

$$\mathbf{U} = \prod_{i''a''} \mathbf{G}_{i''a''}(\theta_{i''a''}) \prod_{ii'} \mathbf{G}_{ii'}(\theta_{ii'}) \prod_{aa'} \mathbf{G}_{aa'}(\theta_{aa'}), \tag{2}$$

where $\mathbf{G}_{pq}(\theta_{pq})$ 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:

$$\mathbf{G}_{pq}(\theta_{pq})$$

$$= \begin{bmatrix} \ddots & & & & & & \\ & 1 & & & & & \\ & & \cos\theta_{pq} & & & \sin\theta_{pq} & \\ & & & 1 & & & \\ & & & & \ddots & & \\ & & & & & 1 & \\ & & -\sin\theta_{pq} & & & \cos\theta_{pq} & \\ & & & & & & 1 \\ & & & & & & & \ddots \end{bmatrix}. \tag{3}$$

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

$$\frac{\partial E}{\partial \mathbf{U}} \frac{\partial \mathbf{U}}{\partial \theta} = 0. \tag{4}$$

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

$$\frac{\partial \mathbf{U}}{\partial \theta_{pq}} = \left[\prod_{(rs)<(pq)} \mathbf{G}_{rs}(\theta_{rs}) \right] \frac{\partial \mathbf{G}_{pq}}{\partial \theta_{pq}} \left[\prod_{(rs)>(pq)} \mathbf{G}_{rs}(\theta_{rs}) \right], \tag{5}$$

where $\partial \mathbf{G}_{pq}/\partial \theta_{pq}$ is a matrix of all zeros except for elements pp, pq, qp, and qq:

$$\frac{\partial \mathbf{G}_{pq}}{\partial \theta_{pq}}$$

$$= \begin{bmatrix} \ddots & & & & & & \\ & 0 & & & & & \\ & & -\sin\theta_{pq} & & & \cos\theta_{pq} & \\ & & & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & \\ & & -\cos\theta_{pq} & & & -\sin\theta_{pq} & \\ & & & & & & 0 \\ & & & & & & & \ddots \end{bmatrix}. \tag{6}$$

An algorithm for evaluating the contribution of (5) has been given by Head-Gordon and Pople.[27]

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

$$\hat{T}_2 = \frac{1}{4} \sum_{ijab} a^{ab}_{ij} a^\dagger b^\dagger j i, \tag{7}$$

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^{ab}_{ij} are

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

$$\langle\Phi^{ab}_{ij}|\hat{H}|(1+\hat{T}_2+\tfrac{1}{2}\hat{T}^2_2)\Phi_0\rangle = E a^{ab}_{ij}. \tag{9}$$

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,

$$f_{pq}=h_{pq}+\sum_{k''}\langle pk''||qk''\rangle. \tag{10}$$

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

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

where E_0 is the reference energy,

$$E_0=\langle\Phi_0|\hat{H}|\Phi_0\rangle=\sum_{i''}f_{i''i''}-\frac{1}{2}\sum_{i''j''}\langle i''j''||i''j''\rangle. \tag{12}$$

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

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

where $\bar{\bar{u}}_{ij}^{ab}$ and $\bar{\bar{v}}_{ij}^{ab}$ are defined as[30]

$$\bar{\bar{u}}_{ij}^{ab}=\sum_c(f_{bc}a_{ij}^{ac}-f_{ac}a_{ij}^{bc})+\sum_k(f_{ki}a_{jk}^{ab}-f_{kj}a_{ik}^{ab})$$

$$+\frac{1}{2}\sum_{cd}\langle ab||cd\rangle a_{ij}^{cd}+\frac{1}{2}\sum_{kl}\langle ij||kl\rangle a_{kl}^{ab}$$

$$+\sum_{kc}[\langle ic||kb\rangle a_{jk}^{ac}-\langle ic||ka\rangle a_{jk}^{bc}+\langle jc||ka\rangle a_{ik}^{bc}$$

$$-\langle jc||kb\rangle a_{ik}^{ac}], \tag{14}$$

and

$$\bar{\bar{v}}_{ij}^{ab}=\frac{1}{4}\sum_{klcd}\langle kl||cd\rangle[a_{ij}^{cd}a_{kl}^{ab}+4a_{ik}^{ac}a_{jl}^{bd}+4a_{ik}^{bd}a_{jl}^{ac}$$

$$-2(a_{ij}^{ac}a_{kl}^{bd}+a_{ij}^{bd}a_{kl}^{ac}+a_{ik}^{ab}a_{jl}^{cd}+a_{ik}^{cd}a_{jl}^{ab})]. \tag{15}$$

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.*[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,[31–37] and gradients for projective Brueckner orbitals have been presented by Kobayashi *et al.*[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

$$E(\mathbf{I},\mathbf{f},\mathbf{a})=E_0(\mathbf{I},\mathbf{f})+\mathbf{a}^\dagger\bar{\mathbf{I}}. \tag{16}$$

$\bar{\mathbf{I}}$ 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

$$\bar{\bar{\mathbf{u}}}+\bar{\bar{\mathbf{v}}}+\bar{\mathbf{I}}=0, \tag{17}$$

or formally

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

Now consider the derivative of the energy with respect to some parameter λ,

$$\frac{\partial E}{\partial\lambda}=E^\lambda=\frac{\partial E}{\partial\mathbf{I}}\frac{\partial\mathbf{I}}{\partial\lambda}+\frac{\partial E}{\partial\mathbf{f}}\frac{\partial\mathbf{f}}{\partial\lambda}+\frac{\partial E}{\partial\mathbf{a}}\frac{\partial\mathbf{a}}{\partial\lambda}, \tag{19}$$

which may be written as inner-products:

$$E^\lambda=\mathbf{E}_I^\dagger\mathbf{I}^\lambda+\mathbf{E}_f^\dagger\mathbf{f}^\lambda+\mathbf{E}_a^\dagger\mathbf{a}^\lambda. \tag{20}$$

Here $\mathbf{E}_X$ denotes a vector whose $pqrs$th element is $\partial E/\partial X_{pqrs}$, and $\mathbf{X}^\lambda$ 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 λ. 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,[24] first applied in coupled-cluster theory by Adamowicz *et al.*[32] (similar manipulations[34] are often performed in coupled-cluster gradient theory using an operator denoted Λ). Differentiating the doubles equation (18) yields

$$\mathbf{A}^\lambda=\mathbf{A}_I\mathbf{I}^\lambda+\mathbf{A}_f\mathbf{f}^\lambda+\mathbf{A}_a\mathbf{a}^\lambda=0, \tag{21}$$

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

$$\mathbf{a}^\lambda=-\mathbf{A}_a^{-1}[\mathbf{A}_I\mathbf{I}^\lambda+\mathbf{A}_f\mathbf{f}^\lambda]. \tag{22}$$

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

$$\mathbf{z}^\dagger=-\mathbf{E}_a^\dagger\mathbf{A}_a^{-1}. \tag{23}$$

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

$$\mathbf{A}_a^\dagger\mathbf{z}=-\bar{\mathbf{I}}. \tag{24}$$

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

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

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

$$\bar{\bar{\mathbf{v}}}_{ij}^{ab}(\mathbf{z}\times\mathbf{a})=\frac{1}{4}\sum_{klcd}a_{kl}^{cd}[z_{ij}^{cd}\langle kl||ab\rangle+z_{kl}^{ab}\langle ij||cd\rangle$$

$$+2(z_{ij}^{bc}\langle kl||ad\rangle-z_{ij}^{ac}\langle kl||bd\rangle-z_{ik}^{cd}\langle jl||ab\rangle$$

$$+z_{jk}^{cd}\langle il||ab\rangle+z_{kl}^{bc}\langle ij||ad\rangle-z_{kl}^{ac}\langle ij||bd\rangle$$

$$-z_{ik}^{ab}\langle jl||cd\rangle+z_{jk}^{ab}\langle il||cd\rangle)+4(z_{ik}^{ac}\langle jl||bd\rangle$$

$$-z_{jk}^{ac}\langle il||bd\rangle-z_{ik}^{bc}\langle jl||ad\rangle+z_{jk}^{bc}\langle il||ad\rangle)]. \tag{26}$$

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 λ [Eq. (20)] becomes

$$E^\lambda=\mathbf{E}_I^\dagger\mathbf{I}^\lambda+\mathbf{E}_f^\dagger\mathbf{f}^\lambda+\mathbf{z}^\dagger[\mathbf{A}_I\mathbf{I}^\lambda+\mathbf{A}_f\mathbf{f}^\lambda]. \tag{27}$$

TABLE I. Effective one- and two-particle density matrices.

$$\gamma_{ab}=\tfrac{1}{4}\Sigma_{ijc}(z_{ij}^{ac}a_{ij}^{bc}+z_{ij}^{bc}a_{ij}^{ac})$$
$$\gamma_{ij}=-\tfrac{1}{4}\Sigma_{abk}(z_{ik}^{ab}a_{jk}^{ab}+z_{jk}^{ab}a_{ik}^{ab})$$
$$\Gamma_{abcd}=\tfrac{1}{16}\Sigma_{ij}(z_{ij}^{ab}a_{ij}^{cd}+z_{ij}^{cd}a_{ij}^{ab})$$
$$\Gamma_{ijkl}=\tfrac{1}{16}\Sigma_{ab}(z_{ij}^{ab}a_{kl}^{ab}+z_{kl}^{ab}a_{ij}^{ab})$$
$$\Gamma_{iajb}=-\tfrac{1}{2}\Sigma_{kc}(z_{ik}^{bc}a_{jk}^{ac}+z_{jk}^{ac}a_{ik}^{bc})$$
$$\Gamma_{ijab}=\tfrac{1}{4}(z_{ij}^{ab}+a_{ij}^{ab})+\tfrac{1}{16}\Sigma_{klcd}z_{kl}^{cd}a_{kl}^{ab}a_{ij}^{cd}$$
$$+\tfrac{1}{4}\Sigma_{klcd}(z_{kl}^{cd}a_{ik}^{ac}a_{jl}^{bd}-z_{kl}^{cd}a_{jk}^{ac}a_{il}^{bd})$$
$$-\tfrac{1}{8}\Sigma_{klcd}(z_{kl}^{cd}a_{kl}^{ac}a_{ij}^{bd}-z_{ij}^{ab}a_{kl}^{bc}a_{kl}^{ad})$$
$$-\tfrac{1}{8}\Sigma_{klcd}(z_{kl}^{cd}a_{ik}^{cd}a_{jl}^{ab}-z_{kl}^{cd}a_{jk}^{cd}a_{il}^{ab})$$

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

$$E^{\lambda}=(\mathbf{E}_0)_I^{\dagger}\mathbf{I}^{\lambda}+(\mathbf{E}_0)_f^{\dagger}\mathbf{f}^{\lambda}+(\mathbf{a}+\mathbf{z})^{\dagger}\bar{\mathbf{T}}^{\lambda}+\mathbf{z}^{\dagger}(\bar{\bar{\mathbf{u}}}_I+\bar{\bar{\mathbf{v}}}_I)\mathbf{I}^{\lambda}$$
$$+\mathbf{z}^{\dagger}(\bar{\bar{\mathbf{u}}}_f+\bar{\bar{\mathbf{v}}}_f)\mathbf{f}^{\lambda}. \tag{28}$$

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

$$E^{\lambda}=(\mathbf{E}_0)_I^{\dagger}\mathbf{I}^{\lambda}+(\mathbf{E}_0)_f^{\dagger}\mathbf{f}^{\lambda}+\boldsymbol{\Gamma}^{\dagger}\mathbf{I}^{\lambda}+\boldsymbol{\gamma}^{\dagger}\mathbf{f}^{\lambda}. \tag{29}$$

Of course the contributions from E_0 could be added into the definitions of γ and Γ, but we will keep them separate so that the indices of γ and Γ run over only active orbitals. Thus our two-particle density matrix Γ 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 θ parameterize a unitary transformation matrix $\mathbf{U}$, so that the orbitals are variationally optimized when (4) holds. Since the evaluation of $\partial\mathbf{U}/\partial\theta$ has already been discussed, it is sufficient here to focus on $\partial E/\partial\mathbf{U}$. Using the general expression (29), the derivative of the energy with respect to some particular element U_{pq} of $\mathbf{U}$ is:

$$E^{U_{pq}}=(\mathbf{E}_0)_I^{\dagger}\mathbf{I}^{U_{pq}}+(\mathbf{E}_0)_f^{\dagger}\mathbf{f}^{U_{pq}}+\boldsymbol{\Gamma}^{\dagger}\mathbf{I}^{U_{pq}}+\boldsymbol{\gamma}^{\dagger}\mathbf{f}^{U_{pq}}. \tag{30}$$

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

$$\frac{\partial E}{\partial U_{pi'}}=2f_{p^{o}i'}+2\sum_{jk}\gamma_{jk}\langle jp^{o}||ki'\rangle$$
$$+2\sum_{ab}\gamma_{ab}\langle ap^{o}||bi'\rangle, \tag{31}$$

$$\frac{\partial E}{\partial U_{pi}}=2f_{p^{o}i}+2\sum_{jab}\Gamma_{ijab}\langle p^{o}j||ab\rangle+2\sum_{j}\gamma_{ij}f_{p^{o}j}$$
$$+2\sum_{jk}\gamma_{jk}\langle jp^{o}||ki\rangle+2\sum_{ab}\gamma_{ab}\langle ap^{o}||bi\rangle$$
$$+4\sum_{jkl}\Gamma_{ijkl}\langle p^{o}j||kl\rangle+2\sum_{jab}\Gamma_{iajb}\langle p^{o}a||jb\rangle, \tag{32}$$

$$\frac{\partial E}{\partial U_{pa}}=2\sum_{ijb}\Gamma_{ijab}\langle ij||p^{o}b\rangle+2\sum_{b}\gamma_{ab}f_{p^{o}b}$$
$$+4\sum_{bcd}\Gamma_{abcd}\langle p^{o}b||cd\rangle+2\sum_{ijb}\Gamma_{iajb}\langle ip^{o}||jb\rangle, \tag{33}$$

$$\frac{\partial E}{\partial U_{pa'}}=0, \tag{34}$$

where the effective one- and two-particle density matrices γ and Γ are defined in Table I. The superscript o's denote orbitals in the original basis [see Eq. (1)]; all other orbitals are assumed to be in the current basis. With these expressions for $\partial E/\partial\mathbf{U}$, the orbital rotation gradient can be evaluated as

$$\frac{\partial E}{\partial\theta_{pq}}=\sum_{rs}\frac{\partial E}{\partial U_{rs}}\frac{\partial U_{rs}}{\partial\theta_{pq}}. \tag{35}$$

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 $\partial E/\partial\mathbf{U}$ (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 δx of nuclear coordinate x, the change in the (optimized) spin-orbitals may be written[39]

$$\phi_p\rightarrow\phi_p+\delta x\left(\sum_q\tilde{U}_{qp}^x\phi_q+\phi_p^{\bar{x}}\right)+\mathcal{O}(\delta x^2), \tag{36}$$

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

$$\phi_p^{\bar{x}}=\sum_{\mu}C_{\mu p}\left(\frac{d\chi_{\mu}}{dx}\right). \tag{37}$$

The derivative of the electronic energy with respect to x, E^x, is given by the general expression (29), and the only remaining steps are to determine the derivatives of the integrals $\mathbf{f}^x$ and $\mathbf{I}^x$. These are[39]

$$h_{pq}^x = \sum_r h_{rq}\tilde{U}_{rp}^x + \sum_r h_{pr}\tilde{U}_{rq}^x + \sum_{\mu\nu} C_{\mu p}C_{\nu q}h_{\mu\nu}^x, \quad (38)$$

$$\langle pq||rs\rangle^x = \sum_t \left[\tilde{U}_{tp}^x \langle tq||rs\rangle + \tilde{U}_{tq}^x \langle pt||rs\rangle \right.$$
$$\left. + \tilde{U}_{tr}^x \langle pq||ts\rangle + \tilde{U}_{ts}^x \langle pq||rt\rangle \right]$$
$$+ \sum_{\mu\nu\rho\sigma} C_{\mu p}C_{\nu q}C_{\rho r}C_{\sigma s}\langle \mu\nu||\rho\sigma\rangle^x, \quad (39)$$

$$f_{pq}^x = \sum_r f_{rq}\tilde{U}_{rp}^x + \sum_r f_{pr}\tilde{U}_{rq}^x + \sum_{\mu\nu} C_{\mu p}C_{\nu q}f_{\mu\nu}^x. \quad (40)$$

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 *et al.*,[13] we will denote the purely AO parts (assuming real orbitals) as

$$h_{pq}^{\bar{x}} = \sum_{\mu\nu} C_{\mu p}C_{\nu q}h_{\mu\nu}^x, \quad (41)$$

$$\langle pq||rs\rangle^{\bar{x}} = \sum_{\mu\nu\rho\sigma} C_{\mu p}C_{\nu q}C_{\rho r}C_{\sigma s}\langle \mu\nu||\rho\sigma\rangle^x, \quad (42)$$

$$f_{pq}^{\bar{x}} = \sum_{\mu\nu} C_{\mu p}C_{\nu q}f_{\mu\nu}^x. \quad (43)$$

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

$$h_{pq}^x = h_{pq}^{\bar{x}} + \sum_{rs} \tilde{U}_{rs}^x \sum_t \frac{\partial h_{pq}}{\partial U_{ts}}U_{tr}, \quad (44)$$

$$\langle pq||rs\rangle^x = \langle pq||rs\rangle^{\bar{x}} + \sum_{tu} \tilde{U}_{tu}^x \sum_v \frac{\partial\langle pq||rs\rangle}{\partial U_{vu}}U_{vt}, \quad (45)$$

$$f_{pq}^x = f_{pq}^{\bar{x}} + \sum_{rs} \tilde{U}_{rs}^x \sum_t \frac{\partial f_{pq}}{\partial U_{ts}}U_{tr}. \quad (46)$$

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

$$E^x = E_0^{\bar{x}} + \sum_{pq} \gamma_{pq}f_{pq}^{\bar{x}} + \sum_{pqrs} \Gamma_{pqrs}\langle pq||rs\rangle^{\bar{x}}$$
$$+ \sum_{pq} X_{pq}\tilde{U}_{pq}^x, \quad (47)$$

where $E_0^{\bar{x}}$ is given by

$$E_0^{\bar{x}} = \sum_{i''} f_{i''i''}^{\bar{x}} - \frac{1}{2}\sum_{i''j''} \langle i''j''||i''j''\rangle^{\bar{x}}, \quad (48)$$

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

$$X_{pq} = \sum_r \frac{\partial E}{\partial U_{rq}}U_{rp}. \quad (49)$$

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)–(6) that our Lagrangian X must be sym-

metric, just as in any other case for which the orbitals are optimized for the wave function being differentiated. Maintaining the orthonormality of the orbitals to first order requires that

$$\tilde{U}_{pq}^x + \tilde{U}_{qp}^x + S_{pq}^{\bar{x}} = 0, \quad (50)$$

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

$$E^x = E_0^{\bar{x}} + \sum_{pq} \gamma_{pq}f_{pq}^{\bar{x}} + \sum_{pqrs} \Gamma_{pqrs}\langle pq||rs\rangle^{\bar{x}}$$
$$+ \sum_{pq} W_{pq}S_{pq}^{\bar{x}}, \quad (51)$$

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 $E_0^{\bar{x}}$, the gradient can be cast into the usual form[40]

$$E^x = \sum_{\mu\geq\nu} \tilde{\gamma}_{\mu\nu}f_{\mu\nu}^{\bar{x}} + \sum_{\substack{\mu\geq\nu,\rho\geq\sigma \\ \mu\nu\geq\rho\sigma}} \tilde{\Gamma}_{\mu\nu\rho\sigma}(\mu\nu|\rho\sigma)^{\bar{x}}$$
$$+ \sum_{\mu\geq\nu} \tilde{W}_{\mu\nu}S_{\mu\nu}^{\bar{x}}. \quad (52)$$

Note that $\mathbf{h}^{\bar{x}}$ can be used rather than $\mathbf{f}^{\bar{x}}$ by appropriate modification of $\tilde{\Gamma}$.

E. Convergence procedure

Our variationally optimized Brueckner orbital CCD wave function is defined by a set of converged orbitals (parameterized by orbital rotation angles θ) 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)–(34) and (35), taking a step in θ 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 θ all *simultaneously*. We accelerate convergence by employing Pulay's direct inversion of the iterative subspace (DIIS) procedure[25] on a compound param-

TABLE II. Theoretical predictions of total energies, equilibrium geometries, and harmonic vibrational frequencies for H_2O using DZP and TZ2P basis sets.[a]

Method	Energy	r_e	θ_e	$\omega_1(a_1)$	$\omega_2(a_2)$	$\omega_3(b_2)$
DZP SCF	−76.047 009	0.9437	106.63	4165	1752	4288
DZP CISD	−76.258 913	0.9577	104.88	3970	1694	4094
DZP CCSD	−76.267 869	0.9610	104.63	3915	1684	4044
DZP BD[b]	−76.267 659	0.9609	104.64	3917	1684	4047
DZP OD	−76.267 733	0.9609	104.64	3917	1684	4046
TZ2P SCF	−76.061 002	0.9403	106.31	4133	1756	4235
TZ2P CISD	−76.320 581	0.9526	105.02	3940	1682	4044
TZ2P CCSD	−76.331 905	0.9563	104.74	3879	1668	3987
TZ2P BD[c]		0.9561	104.7	3884	1669	3992
TZ2P OD	−76.331 632	0.9561	104.76	3882	1669	3990
Expt.[d]		0.9578	104.5	3832	1649	3942

[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.

eter vector comprising **a**, **z** and θ. This is similar to the algorithm of Hampel *et al.*[8] for BD, the primary difference being that the variational orbitals require the additional parameters **z**. Additionally, our orbital rotation parameters θ form a well-defined sequence and are directly extrapolable via DIIS. The same can be said of the alternative parameterization of Hampel *et al.*, although their formulation entails up to sixth-order expansions of exponentials and logarithms of operators.[8] We have found improved convergence when we scale the orbital rotation angles θ by crude approximations to the diagonal elements of the orbital Hessian.[27]

In our experience, the total number of iterations required to converge **a** and **z** is roughly the same as for a standard CCD gradient, with θ converging in earlier iterations. Obviously, once the orbitals are converged, it is no longer necessary to perform integral transformations. Although the full integral transformation each iteration can be avoided, as shown by Hampel *et al.*,[8] we have not yet implemented this feature in our program. Nevertheless, as the integral transformation is only an n^5 process, it should become relatively less important for larger molecules compared to the n^6 steps of the coupled-cluster procedure. An additional cost of our algorithm compared to a CCD gradient is that the orbital derivatives require the construction of the two-particle density matrix from **a** and **z** every iteration (an n^6 process), whereas this needs to be done only once for a CCD gradient. However, similar terms appear in the coupled-perturbed BD equations, so that, to a first approximation, the cost of a gradient for the current method is the same as for a BD gradient.

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;[48,49] 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 H_2O, is the double-ζ plus polarization (DZP) basis set used by Kobayashi *et al.*,[17] which is comprised of the standard Huzinaga-Dunning[50,51] double-ζ 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 $(9s5p1d/4s2p1d)$ for oxygen and $(4s1p/2s1p)$ for hydrogen. The second basis set, used for H_2O, HF, and N_2, is the triple-ζ 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_4^+, is the standard 6-31G* basis of Pople and co-workers,[53,54] which was employed in previous benchmark studies of this molecule.[6] Results were obtained using the Q-Chem[55] and PSI[56] packages of *ab initio* quantum chemistry programs. Coupled-cluster predictions for O_4^+ and all BD results were obtained using ACES II.[57]

Table II presents total energies, equilibrium geometries, and harmonic vibrational frequencies for H_2O 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

TABLE III. Total energies (hartrees) and spectroscopic constants for several diatomic molecules using a TZ2P basis.[a]

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

[a]Basis set from Ref. 17. All electrons correlated.
[b]Data from Huber and Herzberg, Ref. 60.

2 cm^{-1}. The differences between OD and CCSD are only slightly larger for this well-behaved case. The CISD results emphasize that differences between correlated methods including 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 N$_2$ 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 ($\omega_e x_e$), rotational constants (B_e), centrifugal distortion constants ($\tilde{D}_e$), and vibration-rotation coupling constants (α_e). Once again, the BD and OD results are nearly identical, the largest difference being $\Delta r_e = 0.0003$ Å and $\Delta \omega_e = 4$ cm^{-1} for N$_2$. 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 N$_2$.

Finally, we consider the performance of OD for a considerably more challenging case, the $^4B_{1g}$ state of rectangular 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 $2p$ 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 LiO$_2$,[19,63] N$_2^+$,[5]

NO$_2$,[5,20,64,65] NO$_3$,[5,20] allyl radical,[66] formyloxyl radical,[67] C$_3^+$,[68] and 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 NO$_2$,[64] HCO$_2$,[64] LiO$_2$,[19,63] and O$_4^+$.[6]

Table IV presents energies, geometries, and frequencies for rectangular O$_4^+$. R_{OO} denotes the interatomic distance within each O$_2$ subunit, and $R_{c.m.}$ denotes the distance between the two parallel O$_2$'s. The UHF and ROHF frequencies for the antisymmetric b_{3u} stretch in Table IV are clearly unphysical. The behavior of $\omega_5(b_{2u})$ also appears anomalous, going from $543i$ for UHF to 83 for ROHF. When CCSD is applied, the character of this mode changes, the frequency becoming real (154 cm^{-1}) for UHF CCSD and imaginary ($66i$ cm^{-1}) for ROHF CCSD. For the b_{3u} stretch, although CCSD lowers the unreasonably high Hartree-Fock frequencies considerably (to 1726 cm^{-1} and 2036 cm^{-1} with the UHF and ROHF references, respectively), the difference between these two results (310 cm^{-1}) remains unusually large. Furthermore, neither value is close to the most reliable theoretical estimate of Lindh and Barnes,[62] who obtained 1271 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 cm^{-1} band observed[61] in matrix infrared spectra of O$_4^+$, and Jacox 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 O$_4^+$, Barnes and Lindh[6] reported a UHF CCSD prediction of 1220 cm^{-1} for the b_{3u} stretch. Although this value is much more reasonable than ours (1726 cm^{-1}), we were unable to reproduce it despite very careful measures to ensure that we

J. Chem. Phys., Vol. 109, No. 11, 15 September 1998

Sherrill *et al.* 4179

TABLE IV. Theoretical predictions of total energies, equilibrium geometries, and harmonic vibrational frequencies for the $^4B_{1g}$ state of rectangular O_4^+ using the 6-31G* basis set.[a]

Method	Energy	R_{OO}	$R_{c.m.}$	$\omega_1(a_g)$	$\omega_2(a_g)$	$\omega_3(b_{1g})$	$\omega_4(a_u)$	$\omega_5(b_{2u})$	$\omega_6(b_{3u})$
UHF	−298.761 577	1.1176	2.4058	2242	255	396	197	543i	3223
ROHF	−298.739 290	1.1152	2.4073	2256	258	417	199	83	3612
UHF CCSD	−299.484 239	1.1737	2.3792	1808	269	341	178	154	1726
ROHF CCSD	−299.482 497	1.1730	2.3793	1813	269	372	179	66i	2036
BD	−299.482 212	1.1723	2.3794	1821	269	342	179	82	1194
OD	−299.482 683	1.1728	2.3793	1816	269	372	179	84	1193
UHF CCSD(T)[b]	−299.512 786	1.1846	2.3751	1713	271	372	175	97	1922
BD(T)[b]	−299.512 778	1.1846	2.3751	1713	270				1322
TZ2P CASSI[c]									1271
Expt. (Matrix IR)[d]									1320

[a]All electrons correlated. Bond lengths in Å, bond angles in degrees, and frequencies in cm^{-1}.
[b]From Barnes and Lindh, Ref. 6.
[c]From Lindh and Barnes, Ref. 62. Result of nonorthogonal configuration interaction between two symmetry-broken CASSCF wave functions using a larger TZ2P basis.
[d]From Jacox and Thompson, Ref. 70.

remained on the same UHF solution for displaced geometries. Furthermore, our frequency for this mode has been confirmed[71] 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 presence of $e^{\hat{T}_1}$ 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[74] 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 TZ2P 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 investigations[9,21] 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

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 $30\,\mathrm{cm}^{-1}$ change for the otherwise well-behaved $\omega_3(b_{1g})$ 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 evaluation. (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.[26]

ACKNOWLEDGMENTS

APPENDIX

In order to evaluate $\partial E/\partial \mathbf{U}$ 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 $\mathbf{f}^{U_{pq}}$ and $\mathbf{I}^{U_{pq}}$, respectively). Using (1), we may write a one-electron integral in the current MO basis as

$$h_{pq}=\sum_{\mu\nu tu} h_{\mu\nu}C_{\mu t}^{o}U_{tp}C_{\nu u}^{o}U_{uq}. \tag{A1}$$

The derivative is thus

$$\frac{\partial h_{pq}}{\partial U_{rs}}=\sum_{\mu\nu u} h_{\mu\nu}C_{\mu r}^{o}C_{\nu u}^{o}U_{uq}\delta_{ps}+\sum_{\mu\nu t} h_{\mu\nu}C_{\mu t}^{o}U_{tp}C_{\nu r}^{o}\delta_{qs},$$

$$=h_{r^o q}\delta_{ps}+h_{pr^o}\delta_{qs}, \tag{A2}$$

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

$$\frac{\partial\langle pq||rs\rangle}{\partial U_{xy}}=\langle x^o q||rs\rangle\delta_{py}+\langle px^o||rs\rangle\delta_{qy}+\langle pq||x^o s\rangle\delta_{ry}$$

$$+\langle pq||rx^o\rangle\delta_{sy}. \tag{A3}$$

The Fock matrix derivatives are therefore

$$\frac{\partial f_{pq}}{\partial U_{ri''}}=f_{r^o q}\delta_{pi''}+f_{pr^o}\delta_{qi''}+\langle pr^o||qi''\rangle+\langle pi''||qr^o\rangle, \tag{A4}$$

$$\frac{\partial f_{pq}}{\partial U_{ra''}}=f_{r^o q}\delta_{pa''}+f_{pr^o}\delta_{qa''}. \tag{A5}$$

With these definitions, it is straightforward to show[27] that the derivative of the single-determinant energy E_0 is

$$\frac{\partial E_0}{\partial U_{pi''}}=2f_{p^o i''}. \tag{A6}$$

[1] K. A. Brueckner, Phys. Rev. **96**, 508 (1954).
[2] R. K. Nesbet, Phys. Rev. **109**, 1632 (1958).
[3] C. D. Sherrill and H. F. Schaefer, Adv. Quantum Chem. (in press).
[4] N. C. Handy, J. A. Pople, M. Head-Gordon, K. Raghavachari, and G. W. Trucks, Chem. Phys. Lett. **164**, 185 (1989).
[5] J. F. Stanton, J. Gauss, and R. J. Bartlett, J. Chem. Phys. **97**, 5554 (1992).
[6] L. A. Barnes and R. Lindh, Chem. Phys. Lett. **223**, 207 (1994).
[7] R. A. Chiles and C. E. Dykstra, J. Chem. Phys. **74**, 4544 (1981).
[8] C. Hampel, K. A. Peterson, and H.-J. Werner, Chem. Phys. Lett. **190**, 1 (1992).
[9] G. D. Purvis and R. J. Bartlett, J. Chem. Phys. **76**, 1910 (1982).
[10] J. A. Pople, M. Head-Gordon, and K. Raghavachari, J. Chem. Phys. **87**, 5968 (1987).
[11] K. K. Raghavachari, J. A. Pople, E. S. Replogle, M. Head-Gordon, and N. C. Handy, Chem. Phys. Lett. **167**, 115 (1990).
[12] G. E. Scuseria, Chem. Phys. Lett. **226**, 251 (1994).
[13] R. Kobayashi, N. C. Handy, R. D. Amos, G. W. Trucks, M. J. Frisch, and J. A. Pople, J. Chem. Phys. **95**, 6723 (1991).
[14] R. Kobayashi, R. D. Amos, and N. C. Handy, Chem. Phys. Lett. **184**, 195 (1991).
[15] T. J. Lee, R. Kobayashi, N. C. Handy, and R. D. Amos, J. Chem. Phys. **96**, 8931 (1992).
[16] R. Kobayashi, H. Koch, P. Jørgensen, and T. J. Lee, Chem. Phys. Lett. **211**, 94 (1993).
[17] R. Kobayashi, R. D. Amos, and N. C. Handy, J. Chem. Phys. **100**, 1375 (1994).
[18] E. R. Davidson and W. T. Borden, J. Phys. Chem. **87**, 4783 (1983).
[19] W. D. Allen, D. A. Horner, R. L. DeKock, R. B. Remington, and H. F. Schaefer, Chem. Phys. **133**, 11 (1989).
[20] T. D. Crawford, T. J. Lee, N. C. Handy, and H. F. Schaefer, J. Chem. Phys. **107**, 9980 (1997).
[21] G. E. Scuseria and H. F. Schaefer, Chem. Phys. Lett. **142**, 354 (1987).
[22] G. D. Purvis, R. Shepard, F. B. Brown, and R. J. Bartlett, Int. J. Quantum Chem. **23**, 835 (1983).
[23] A. P. L. Rendell, G. B. Backsay, N. S. Hush, and N. C. Handy, J. Chem. Phys. **87**, 5976 (1987).

[24] N. C. Handy and H. F. Schaefer, J. Chem. Phys. **81**, 5031 (1984).

[25] P. Pulay, Chem. Phys. Lett. **73**, 393 (1980).

[26] A. I. Krylov, C. D. Sherrill, E. F. C. Byrd, and M. Head-Gordon (manuscript in preparation).

[27] M. Head-Gordon and J. A. Pople, J. Phys. Chem. **92**, 3063 (1988).

[28] J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, Int. J. Quantum Chem. **14**, 545 (1978).

[29] R. J. Bartlett and G. D. Purvis, Int. J. Quantum Chem. **14**, 561 (1978).

[30] Our definitions of $\bar{\bar{u}}_{ij}^{ab}$ and $\bar{\bar{v}}_{ij}^{ab}$ 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 $\bar{\bar{u}}_{ij}^{ab}$.

[31] G. Fitzgerald, R. J. Harrison, and R. J. Bartlett, J. Chem. Phys. **85**, 5143 (1986).

[32] L. Adamowicz, W. D. Laidig, and R. J. Bartlett, Int. J. Quantum Chem., Symp. **18**, 244 (1984).

[33] A. C. Scheiner, G. E. Scuseria, J. E. Rice, T. J. Lee, and H. F. Schaefer, J. Chem. Phys. **87**, 5361 (1987).

[34] E. A. Salter, G. W. Trucks, and R. J. Bartlett, J. Chem. Phys. **90**, 1752 (1989).

[35] J. Gauss, W. J. Lauderdale, J. F. Stanton, J. D. Watts, and R. J. Bartlett, Chem. Phys. Lett. **182**, 207 (1991).

[36] J. Gauss, J. F. Stanton, and R. J. Bartlett, J. Chem. Phys. **95**, 2639 (1991).

[37] J. Gauss, J. F. Stanton, and R. J. Bartlett, J. Chem. Phys. **94**, 2623 (1991).

[38] M. J. Frisch, M. Head-Gordon, and J. A. Pople, Chem. Phys. Lett. **166**, 275 (1990).

[39] Y. Yamaguchi, Y. Osamura, J. D. Goddard, and H. F. Schaefer, *A New Dimension to Quantum Chemistry: Analytic Derivative Methods in Ab initio Molecular Electronic Structure Theory* (Oxford, New York, 1994).

[40] J. E. Rice and R. D. Amos, Chem. Phys. Lett. **122**, 585 (1985).

[41] I. Shavitt, Int. J. Quantum Chem., Symp. **12**, 5 (1978).

[42] B. R. Brooks and H. F. Schaefer, J. Chem. Phys. **70**, 5092 (1979).

[43] P. Saxe, D. J. Fox, H. F. Schaefer, and N. C. Handy, J. Chem. Phys. **77**, 5584 (1982).

[44] P. Pulay, Mol. Phys. **17**, 197 (1969).

[45] J. D. Goddard, N. C. Handy, and H. F. Schaefer, J. Chem. Phys. **71**, 1525 (1979).

[46] B. R. Brooks, W. D. Laidig, P. Saxe, J. D. Goddard, Y. Yamaguchi, and H. F. Schaefer, J. Chem. Phys. **72**, 4652 (1980).

[47] J. E. Rice, R. D. Amos, N. C. Handy, T. J. Lee, and H. F. Schaefer, J. Chem. Phys. **85**, 963 (1986).

[48] J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, Int. J. Quantum Chem., Symp. **13**, 225 (1979).

[49] P. Saxe, Y. Yamaguchi, and H. F. Schaefer, J. Chem. Phys. **77**, 5647 (1982).

[50] S. Huzinaga, J. Chem. Phys. **42**, 1293 (1965).

[51] T. H. Dunning, J. Chem. Phys. **53**, 2823 (1970).

[52] T. H. Dunning, J. Chem. Phys. **55**, 716 (1971).

[53] W. J. Hehre, R. Ditchfield, and J. A. Pople, J. Chem. Phys. **56**, 2257 (1972).

[54] P. C. Hariharan and J. A. Pople, Theor. Chim. Acta **28**, 213 (1973).

[55] B. G. Johnson, P. M. W. Gill, M. Head-Gordon, C. A. White, D. R. Maurice, T. R. Adams, J. Kong, M. Challacombe, E. Schwegler, M. Oumi, C. Ochsenfeld, N. Ishikawa, J. Florian, R. D. Adamson, J. P. Dombroski, R. L. Graham, and A. Warshel, Q-Chem 1.1, Q-Chem, Inc., Export, PA, 1997.

[56] C. L. Janssen, E. T. Seidl, G. E. Scuseria, T. P. Hamilton, Y. Yamaguchi, R. B. Remington, Y. Xie, G. Vacek, C. D. Sherrill, T. D. Crawford, J. T. Fermann, W. D. Allen, B. R. Brooks, G. B. Fitzgerald, D. J. Fox, J. F. Gaw, N. C. Handy, W. D. Laidig, T. J. Lee, R. M. Pitzer, J. E. Rice, P. Saxe, A. C. Scheiner, and H. F. Schaefer, PSI 2.0.8, PSITECH, Inc., Watkinsville, GA 30677, U.S.A., 1995. This program is generally available for a handling fee of $100.

[57] ACES II program is a product of the Quantum Theory Project, University of Florida, J. F. Stanton, J. Gauss, J. D. Watts, M. Nooijen, N. Oliphant, S. A. Perera, P. G. Szalay, W. J. Lauderdale, S. R. Gwaltney, S. Beck, A. Balková, D. E. Bernholdt, K.-K. Baeck, H. Sekino, P. Rozyczko, C. Huber, and R. J. Bartlett. Integral packages included are VMOL (J. Almlöf and P. R. Taylor), VPROPS (P. R. Taylor), and a modified version of the ABACUS integral derivative package (T. U. Helgaker, H. J. Aa. Jensen, J. Olsen, P. Jørgensen, and P. R. Taylor).

[58] A. R. Hoy, I. M. Mills, and G. Strey, Mol. Phys. **24**, 1265 (1972).

[59] A. R. Hoy and P. R. Bunker, J. Mol. Spectrosc. **74**, 1 (1979).

[60] K. P. Huber and G. Herzberg, *Constants of Diatomic Molecules* (Van Nostrand Reinhold, New York, 1979).

[61] W. E. Thompson and M. E. Jacox, J. Chem. Phys. **91**, 3826 (1989).

[62] R. Lindh and L. A. Barnes, J. Chem. Phys. **100**, 224 (1994).

[63] T. D. Crawford, J. F. Stanton, W. D. Allen, and H. F. Schaefer, J. Chem. Phys. **107**, 10626 (1997).

[64] N. A. Burton, Y. Yamaguchi, I. L. Alberts, and H. F. Schaefer, J. Chem. Phys. **95**, 7466 (1991).

[65] T. D. Crawford, J. F. Stanton, P. Szalay, and H. F. Schaefer, J. Chem. Phys. **107**, 2525 (1997).

[66] D. B. Cook, J. Chem. Soc., Faraday Trans. 2 **82**, 187 (1986).

[67] A. D. McLean, B. H. Lengsfield, J. Pacansky, and Y. Ellinger, J. Chem. Phys. **83**, 3567 (1985).

[68] R. S. Grev, I. L. Alberts, and H. F. Schaefer, J. Phys. Chem. **94**, 3379 (1990).

[69] Y. Xie, W. D. Allen, Y. Yamaguchi, and H. F. Schaefer, J. Chem. Phys. **104**, 7615 (1996).

[70] M. E. Jacox and W. E. Thompson, J. Chem. Phys. **100**, 750 (1994).

[71] T. D. Crawford and J. F. Stanton, private communication.

[72] J. F. Stanton and J. Gauss, Analytic evaluation of second derivatives of the energy: Computational strategies for the CCSD and CCSD(T) approximations, in *Recent Advances in Coupled-Cluster Methods*, edited by R. J. Bartlett, volume 3 of *Recent Advances in Computational Chemistry* (World Scientific, Singapore, 1997), pp. 49–79.

[73] E. A. Salter, H. Sekino, and R. J. Bartlett, J. Chem. Phys. **87**, 502 (1987).

[74] Y. Yamaguchi, I. L. Alberts, J. D. Goddard, and H. F. Schaefer, Chem. Phys. **147**, 309 (1990).

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

[76] 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.