On the choice of orbital bases for configuration interaction

F. Tarantelli(a) and L. S. Cederbaum

Lehrstuhl für Theoretische Chemie, Physikalisch-Chemisches Institut, Universität Heidelberg, D-6900 Heidelberg, Federal Republic of Germany

(Received 2 March 1988; accepted 16 June 1988)

We study the CI energy dependence on one-particle orbital bases obtained as SCF solutions of a general parametrized Fock operator which reduces to the Hartree–Fock mean field as a special case. In the case of single excitation CI, exact analytical expressions are derived for the curvature of the CI energy surface at the Hartree–Fock point. Numerical applications to N2 and H2O show that the convergence rate of the CI expansion can be drastically increased by an appropriate and largely system independent choice of the mean field parameters. In all configuration spaces, except single plus double excitation CI, the Hartree–Fock basis is found to be the worst orbital basis for use in CI, suggesting in general that the common-practice procedure of optimizing within a given physical model as the first step to go beyond the model itself can be a very inefficient strategy.

I. INTRODUCTION

Although full configuration interaction (FCI) calculations, mostly correlating only the outer valence electrons, have recently become feasible for a number of small molecules,1-3 the bulk of accurate molecular electronic structure studies still relies upon configuration interaction (CI) calculations carried out in an incomplete configuration space. Routinely, this space comprises all single and double excitations relative to one or a few reference configurations. Both all-electron FCI and incomplete CI types of approach suffer from the approximations inherent in the use of an underlying finite set of atomic-type basis functions in which the molecular orbitals are expanded, but, in addition to this, incomplete CI expansions are also severely affected by the choice of the molecular orbital basis used to construct the configuration functions. In particular, it is well known4,5 that use of Hartree–Fock (HF) orbitals gives rise to slow convergent CI expansions, so that the one-reference single plus double excitation CI approach (SDCI) very often recovers an inadequate (even if large) fraction of the total correlation potentially available within the given basis set. In this case, an account of the correlation effects due to selected higher order excitations is usually achieved via a multireference SDCI procedure. Over the years, techniques have been adopted to speed up the convergence of the CI expansion by using non-Hartree–Fock orbital bases,4 like natural orbitals,5 pair natural orbitals,6 or multiconfiguration SCF (MCSCF) orbitals. In particular, the development of very efficient MCSCF algorithms (notably CASSCF7-9) has made the choice of MCSCF orbital basis as reference for the CI treatment, the present most accurate approach to the calculation of correlated wave functions, as it has been generally confirmed by comparison with the available FCI data.10,11 In this case, however, since the MCSCF orbitals are optimized to describe a short CI expansion including the most important configurations, the subsequent CI calculation should be of the multireference type,3 the set of reference configurations at least comprising the MCSCF configuration space.

Due to its simplicity, the single-reference CI approach based on orbitals easily generated as self-consistent LCAO solutions of a one-particle effective operator is therefore still a very appealing procedure, besides being the only really practical one for large molecules. In this respect, the Hartree–Fock method enjoys an apparently privileged position, since it provides the lowest possible initial energy value to be improved by the CI calculation. An additional, strictly related, property of the HF wave function, stated as Brillouin’s theorem, is that it has vanishing Hamiltonian matrix elements with any singly excited configuration. However, it is obvious that the HF orbitals do not constitute the optimal unitary transformation of a given basis set for use in CI. The latter would clearly be given by an MCSCF procedure carried out in the whole configuration space of choice.

In this context, it is extremely interesting to investigate the properties of other, non-Hartree–Fock, independent particle mean field models, i.e., different Fock-like operators, to see whether they can generate one-particle orbitals affording faster CI convergence than the HF orbitals. Recently Dietz and Hess9 have reported on CI calculations based on orbitals generated by the g-Hartree model. In the present work we study a more general Fock operator, which includes both the Hartree–Fock and the g-Hartree ones as special cases. We would like to emphasize that the aim of this work is far from being that of presenting a practical method of obtaining optimal orbital bases for use in CI. We feel that too little is known yet about the behavior of non-HF mean field models. Therefore, we rather attempt, both by theoretical analysis and numerical experimentation, to establish general properties of the various Fock operators in a CI context. As a first step in this direction we have mainly focused on the simplest meaningful CI scheme, namely single-reference single excitation CI. The fact that our generalized Fock operator reduces to the HF one as a special case allows us to investigate analytically the consequences of departure from the HF basis, and in particular the loss of validity of Brillouin’s theorem in single excitation CI calculations. The theoretical analysis is supplemented by numerical applications to the molecules H2O and N2, using double-zeta plus polarization functions basis sets. The results of CI calcula-
tions on these systems using larger configuration spaces, including double and triple excitations, are also discussed, with particular emphasis on the evaluation of the various orbital bases in the wider perspective provided by our general mean field model.

II. THEORY

A. The generalized Fock operator

For an \( N \)-electron system we define the following general Fock operator:

\[
\hat{F} = \sum_{ij} \left[ h_{ij} + \sum_{k} \left( \lambda_{k,i} V_{ikj} - \bar{\lambda}_{k,i} V_{ijk} \right) \right] \hat{a}_i^\dagger \hat{a}_j ,
\]

where \( h_{ij} \) are the one-electron Hamiltonian matrix elements and \( V_{ikj}, V_{ijk} \) are the Coulomb and exchange electron repulsion integrals, respectively. \( \lambda_{k,i} \) and \( \bar{\lambda}_{k,i} \) are parameters scaling the electron repulsion potential, while \( \hat{a}_i^\dagger \) and \( \hat{a}_j \) are creation and annihilation operators, respectively, in some one-particle basis. The above Fock operator has previously been introduced and studied as a suitable unperturbed Hamiltonian in the context of perturbation theory.\(^{10,11}\) There, the idea was exploited of adjusting the \( \lambda \) parameters to incorporate correlation effects into the mean field associated with the operator. The parametrization has also been successfully applied to minimize intruder state problems which hinder the convergence of the perturbation series.\(^{10,12}\)

In the present work we wish to investigate the properties of the Fock operator in Eq. (1) in a quite different context. Instead of adjusting (nonvariationally) the \( \lambda \) parameters so as to introduce correlation effects into the mean field energy, we shall examine how the total energy obtained in a correlated treatment, e.g., single-reference CI, depends on the orbital bases which can be obtained by varying the scaling factors. Thus, in this case, the \( \lambda_{k,i} \) and \( \bar{\lambda}_{k,i} \) should be regarded as variational parameters upon which the CI energy depends and can be minimized with respect to. Obviously, in the limit of a FCI space the CI energy is independent on the choice of the scaling factors.

Since the operator in Eq. (1) has a large number of parameters, it is convenient to start off the proposed investigation with a simplified model. Therefore, in the present work, we will restrict ourselves to the following prototype operator, obtained from Eq. (1) by setting all \( \lambda_{k,i} \) and, separately, all \( \bar{\lambda}_{k,i} \) equal to each other:

\[
\hat{F} = \sum_{ij} \left[ h_{ij} + \sum_{k} \left( \bar{\lambda} V_{ikj} - \bar{\lambda} \sum_{k} V_{ijk} \right) \right] \hat{a}_i^\dagger \hat{a}_j .
\]

(2)

Obviously, this simplified two-parameter Fock operator (\( \bar{\lambda}\bar{\lambda} \)-HF) represents a gross restriction to the flexibility afforded by the general form, Eq. (1). In particular, it forces an identical parameter for all orbitals and orbital interactions, whereas one would expect, on physical grounds, different orbitals (e.g., core orbitals and valence orbitals) to have rather different optimal parameters. However, the \( \bar{\lambda}\bar{\lambda} \)-HF operator reduces to the Hartree–Fock one for \( \bar{\lambda} = \bar{\lambda} = 1 \), and thus we hope to still be able to gain some useful insight into the problem of departing from the Hartree–Fock model as the starting point of CI calculations. Finally, we draw attention on two distinguished orthogonal cuts of the plane defined by the \( \lambda \) and \( \bar{\lambda} \) factors in Eq. (2). They are associated with two different one-parameter Fock operators, namely the scaled-Hartree–Fock (\( \lambda\lambda \)-HF),\(^{10,11}\) obtained by setting \( \bar{\lambda} = \lambda \):

\[
\hat{F} = \sum_{ij} \left[ h_{ij} + \lambda \sum_{k} V_{ikj} \right] \hat{a}_i^\dagger \hat{a}_j,
\]

(3)

and the g-Hartree (g-H),\(^{9}\) defined by \( \bar{\lambda} = 1 - \lambda = g \):

\[
\hat{F} = \sum_{ij} \left[ h_{ij} + g \sum_{k} V_{ikj} - (1-g) \sum_{k} V_{ijk} \right] \hat{a}_i^\dagger \hat{a}_j .
\]

(4)

In Eq. (3) we have used \( V_{ikj} = V_{ikj} - V_{ijk} \). Notice that the cut defined by the \( \lambda\lambda \)-HF operator also goes through the HF point at \( \lambda = 1 \). The \( \lambda\lambda \)-HF and g-H lines meet at \( \lambda = \bar{\lambda} = 0.5 \).

Having defined the \( \lambda\bar{\lambda} \)-HF operator, Eq. (2), the simplest interesting theoretical question which we wish to address is related to the loss of validity of Brillouin’s theorem occurring upon moving away from the HF orbital basis in a CI calculation. According to Brillouin’s theorem, the HF wave function is decoupled from its singly excited configurations and, therefore, its energy is invariant vs a CI carried out in the space of the single excitations. This property descends directly from, and actually can be shown to be fully equivalent to, the requirement that the HF determinant makes the Hamiltonian expectation value stationary with respect to orbital variations. Although not true in general,\(^{13}\) we will assume, as usual, that this stationary point is a minimum, and refer to the Hartree–Fock orbital basis as the HF point in the space of orbital transformations. For non-Hartree–Fock one-determinant wave functions Brillouin’s theorem no longer holds and it is, therefore, entirely possible to recover some fraction of correlation energy through the simplest class of CI calculations, namely single excitation CI (SCI). The maximum amount of correlation energy that can be obtained in this way would be achieved by allowing for all possible orbital unitary transformations, corresponding to an MCSCF calculation carried out in the full single excitation space. It is, therefore, very interesting to assess whether, and how much, correlation energy can be delivered by an SCI calculation using orbital bases generated by the very simple \( \lambda\bar{\lambda} \)-HF operator. In CI calculations using larger configuration spaces, it is well known that, for the Hartree–Fock orbital basis, double and quadruple excitations give the largest contributions to that correlation energy. Recent investigations\(^{9}\) indicate the the g-Hartree basis is more effective than the HF one in SCI calculations and when double and triple excitations are also included (SDTCC), but not in SDCI calculations. We wish also to analyze the effect of the various configuration classes in CI calculations based on the \( \lambda\bar{\lambda} \)-HF operator. This will place the previous findings in a more general perspective, and provide some new useful information. We study these points in Sec. III, by discussing the results of numerical applications to the molecules \( \text{N}_2 \) and \( \text{H}_2\text{O} \).

B. Hartree–Fock instability against single excitations

According to Brillouin’s theorem, the one-determinant energy and the corresponding SCI energy coincide at the Hartree–Fock point. (We make the assumption that the HF
energy is the lowest eigenvalue of the SCI Hamiltonian matrix built on the HF basis.) By the variational principle, it is then clear that both the HF and the HF-SCI energies must be stationary with respect to any transformation of the orbitals. Thus, in particular, they are both stationary with respect to the orbital transformations allowed by the $\lambda\lambda$-HF operator.

The one-determinant energy has a minimum at the HF point, but we lack any corresponding characterization of this stationary point on the SCI energy surface. To investigate this topic, we have used a second order perturbation expansion of the one-determinant and SCI energies around the HF point using the $\lambda\lambda$-HF orbitals. The derived exact expressions allow in principle such a characterization. The derivation is worked out in detail in the Appendix, and we discuss here the main results.

For simplicity, we refer first to the case of the $\lambda$-HF operator and expand the reference determinant energy $\bar{E}_0 = \langle \Phi_0 | H | \Phi_0 \rangle$ in powers of $\theta = \lambda - 1$ around $\theta = 0$. To second order in $\theta$ this yields

$$\bar{E}^{(2)} = E_{HF} + \theta^2 C^\dagger D C,$$

where $E_{HF}$ is the Hartree–Fock energy, $D$ is a matrix defined in the space of the HF singly excited configurations, and $C$ is a vector in that space. The matrix elements of $D$ are

$$D_{a\beta} = (\epsilon_a - \epsilon_\beta) \delta_{a\beta} + V_{q(i)\beta} + V_{\beta(i)\beta},$$

where $a, \beta$ and $i, j$ denote virtual and occupied HF orbitals, respectively, and $\epsilon_a, \epsilon_\beta$ are the corresponding orbital energies. The vector $C$ in Eq. (5) is given as

$$C = D^{-1} U$$

with

$$U_{ai} = \sum_j V_{q(i)j}.$$  

It is interesting to note that an identical quadratic form of the matrix $D$, Eq. (5), was derived by Thouless for the one-determinant energy to second order in the coefficients of an arbitrary unitary transformation of the HF orbitals, of which $C$ represents the occupied/virtual mixing part. In our case $C$ is not arbitrary because the $\lambda$-HF operator determines completely (to within $\lambda$) the orbital transformation and thus fixes $C$ via Eq. (7). Since Eq. (5) is a completely general equation and $\bar{E}_0$ usually has a minimum at the HF point, we consider the matrix $D$ to be positive definite. A completely analogous equation is obtained to second order in $\theta$ for the $\lambda$-HF-SCI energy:

$$\bar{E}^{(2)}_{SCI} = E_{HF} + \theta^2 C^\dagger Q C.$$  

The only difference here with respect to Eq. (5) is that the matrix $D$ is replaced by a matrix $Q$, which is related to $D$ through

$$Q = D - D^\dagger X^{-1} D$$

and $X$ is given by

$$X = D - \Delta$$

with

$$\Delta_{a\beta} = V_{ab(i)j}.$$  

As can be immediately recognized from Eqs. (6), (11), and (12), $X$ is nothing else but the Hamiltonian matrix over the Hartree–Fock single excitations with $E_{HF}$ subtracted from the diagonal elements. It is also interesting to note that $Q$ is nonzero by virtue of the difference between the two positive definite matrices $D$ and $X$. This difference is the matrix $\Delta$ of HF double excitation matrix elements, Eq. (12). In fact, we note here that $Q$ can be rewritten as

$$Q = D^\dagger (D^{-1} - X^{-1}) D$$

or, using Eq. (11),

$$Q = - (\Delta + \Delta^\dagger X^{-1} \Delta).$$

By straightforward generalization it can be shown that similar equations obtain in the case of the $\lambda\lambda$-HF orbital variations:

$$\bar{E}^{(2)} = E_{HF} + C^\dagger (\theta, \bar{\theta}) Q C(\theta, \bar{\theta}),$$

$$C(\theta, \bar{\theta}) = \theta C_\lambda + \overline{\theta} C_\lambda,$$

where $\theta = 1 - \lambda$ and $C_\lambda$ and $C_\lambda$ define the linear terms of the orbital expansion in the $\lambda$ and $\lambda$ directions, respectively. They are obtained through the following equations:

$$C_\lambda = D^{-1} J; \quad J_{ai} = - \sum_j V_{q(i)j},$$

$$C_\lambda = D^{-1} K; \quad K_{ai} = \sum_j V_{q(i)j}. $$

It can again be shown that the key equations, Eqs. (9), (10), and (15), are actually completely general expressions. In fact, the difference $\Delta \bar{E}^{(2)}_\text{SCI}$ between the SCI energy and the Hartree–Fock energy is always given by a quadratic form of the matrix $Q$, for any arbitrary orbital transformation. The matrix $Q$ is thus a universal matrix, and specifying a particular transformation or class of transformations only fixes a particular form for the vector $C$. We see from Eq. (10) that $Q$ is given by the difference of two terms. The first term, $D$, accounts for the reference determinant energy shift while the second term arises from switching on the interaction between the reference determinant and the single excitations. As may be expected on the basis of its generality, $Q$, or equivalently $\Delta$, is generally an indefinite matrix, having both positive and negative eigenvalues. This means that, in the full space of orbital variations, the HF basis corresponds to a generalized saddle point of the SCI energy. An analysis of the eigenvalue spectrum of $Q$ shows that the eigenvectors of positive and negative curvature are approximately equal in number. But, interestingly, most of the eigenvalues are generally very small in magnitude and only some negative eigenvalues are large. This indicates that at the HF saddle point the SCI energy surface is very flat in nearly all directions, except a few ones along which it drops substantially.

The question of interest that now naturally arises is whether this conclusion applies also to the restricted case of the $\lambda$-HF and $\lambda\lambda$-HF operators, i.e., when Eq. (7) or Eqs. (17) and (18) hold, respectively. Unfortunately, even in this case, the inherent indefiniteness of the matrix $Q$ makes it very difficult to make any general, analytically provable statement about the sign of $\Delta \bar{E}^{(2)}_\text{SCI}$. In the numerical applications to be discussed in Sec. III, we have, therefore, made a special effort to analyze the SCI energy dependence on $\lambda$ and $\lambda$ around the Hartree–Fock point, in order to gain some
insight into the problem and possibly be able to reach some general conclusions. The CI calculations we have carried out with larger configuration spaces will also be of help in clarifying and generalizing our findings.

III. NUMERICAL APPLICATIONS

A. Computational details

We have studied the one-determinant and SCI energy surfaces as a function of the Coulomb ($\lambda$) and exchange ($\overline{\lambda}$) parameters with orbital bases generated by the $\lambda\overline{\lambda}$-HF operator for the molecules $\text{N}_2$ and $\text{H}_2\text{O}$. For $\text{H}_2\text{O}$ two series of calculations have been carried out, using double-zeta (DZ) and DZ plus polarization functions (DZP) Gaussian basis sets. Only the DZP basis has been used in the case of $\text{N}_2$. The polarization functions consist of one $d$ and one $p$ set of functions on the first row atoms and on hydrogen, respectively. For $\text{H}_2\text{O}$ (DZ) all-electron FCI results, and for $\text{H}_2\text{O}$ (DZP) and $\text{N}_2$ (DZP) valence FCI results are available. This will permit an accurate assessment of the amount of correlation energy obtained in our calculations. Some CI results for $\text{H}_2\text{O}$ (DZ) using orbital bases generated with the $g$-Hartree model are also available and we can thus evaluate them in the context of the more general $\lambda\overline{\lambda}$-HF operator and make comparisons with our one-parameter $\lambda$-HF model. For the $\text{N}_2$ molecule a detailed study of the SCI energy surface, involving about 800 individual calculations, has been carried out, thus permitting an accurate two-variable sixth order polynomial fit reproducing the surface. In the case of $\text{H}_2\text{O}$ a more restricted study has been performed, mainly aiming at the accurate location and characterization of the stationary points on the surface. Additional $\lambda\overline{\lambda}$-HF-SDCI and $\lambda$-HF-SDTCI calculations on the same systems have also been carried out, again with particular attention to the location of the relevant stationary points.

B. Results and discussion

The energy surfaces for the $\text{N}_2$ system as a function of $\lambda$ and $\overline{\lambda}$ are shown in Fig. 1. The lower surface is the SCI energy and the upper surface, tangent to the lower one at the HF point, is the reference determinant energy. A contour plot of the SCI surface is displayed in Fig. 2. We emphasize that the corresponding surface for the water molecule, using either basis set, shows identical features and topology. This gives us substantial confidence that the following discussion exhibits general features, describing inherent properties of the $\lambda\overline{\lambda}$-HF operator.

It is clearly seen from Fig. 1 that the reference determinant energy rises very steeply upon moving away from the HF point. On the contrary, Figs. 1 and 2 show that the SCI energy lowers noticeably below $E_{\text{HF}}$ in the region immediately surrounding it and steepens up only further away. Correlation energy can thus be delivered by SCI for $\lambda$ parameter values not too different from unity. The surface exhibits two “valleys” separated by a ridge going through the HF point. The two valleys extend in a direction almost parallel to the $\lambda$ axis, and the surface as a whole is substantially flatter in this

---

**FIG. 1.** One-determinant and single excitation CI energy surfaces of $\text{N}_2$ as a function of the Coulomb and exchange scaling parameters in the $\lambda\overline{\lambda}$-HF operator. Due to technical reasons, the upper edge of the SCI surface appears flattened.

**FIG. 2.** Contour plot of the $\lambda\overline{\lambda}$-HF-SCI energy surface for $\text{N}_2$. The $\lambda$-HF cut is shown as a broken line and the $g$-$H$ cut as a dotted line.
direction than in the $\lambda$ direction, indicating a much more pronounced dependence on the Coulomb scaling than on the exchange scaling. This is not unexpected, considering that the relative magnitude of the Coulomb part of the potential is larger than the exchange one. The SCI energy rises very steeply above the HF energy for Coulomb scalings smaller than $\sim 0.8$ and larger than $\sim 1.1$ but, provided $\lambda$ is kept within these limits, $\Delta E_{\text{SCI}}$ stays negative for a wider interval of $\lambda$ values. In this respect it is interesting to note that even small negative $\lambda$ values (corresponding to a sign reversal of the exchange interaction in the mean field) permit some fraction of correlation energy to be recovered. As appears clearly in Fig. 2, a minimum of the SCI surface is found on both valleys. Interestingly, the two minima lie on opposite sides of the HF point, along a line which is very close to the scaled-HF cut, where $\lambda$ and $\lambda$ are equal. The deeper minimum corresponds to $\lambda$ and $\lambda$ values both smaller than 1. On moving away from the minima along the valleys in either direction, the two valleys become shallower and eventually each of them, on opposite sides, for $\lambda$ values larger than $\sim 2.6$ and smaller than $\sim 0.8$, respectively, merges with the ridge through a saddle point. Further away, only one valley is thus left.

The most remarkable result of our calculations is that the Hartree–Fock point is found to be a local maximum of the SCI energy. This is clearly in itself a striking piece of evidence that the optimality of the Hartree–Fock model as a stand-alone approximation becomes totally irrelevant when viewed in the wider context of correlated wave function approaches and largely justifies the quest for more suitable mean field Fock operators. Departing from the Hartree–Fock model implies the seeming disadvantage of a rapidly deteriorating first order description of the electronic structure, which in turn requires higher order corrections to become much larger if any improvement is to be achieved. Nevertheless, this is exactly what seems to happen, even for so simple a model as the $\lambda\lambda$-HF mean field.

As observed above, the minima of the SCI energy lie very close to the $\lambda$-HF cut of the surface. The $\lambda = 1 - \lambda$ cut, corresponding to the $g$-Hartree operator, results to be too far removed from HF for the two opposite effects, namely steeply increasing one-determinant energy and increasing switched-on configuration interaction, to balance out more favorably. This finding seems to deprive the $g$-Hartree model of any particularly distinguished character in the context of CI. For comparison purposes, we show in Fig. 3 a plot of the $\lambda$-HF and $g$-H sections of the SCI surface for $N_2$. It is immediately seen that the minima of the $\lambda$-HF curve are substantially deeper than the minima of the $g$-H curve and, in the case of $N_2$, even the secondary $\lambda$-HF minimum is at lower energy than the deepest $g$-H minimum. For $H_2O$, the bottom of the two valleys is somewhat flatter and, therefore, the difference between the two cuts is less pronounced. Figure 3(a) also evidences quite strikingly the large SCI energy lowering relative to the first order (one-determinant) energy shown as a broken line. At the absolute minimum of the surface for $N_2$, the correlation energy obtained is 0.678 eV, but the net CI lowering with respect to the reference determinant energy results to be 5.83 eV. This effect becomes spectacular in the case of water, where the corresponding lowering is larger than 12 eV with either basis set.

We summarize our results in Table I, reporting the location and correlation energy of the minima on the overall surface and on the two one-parameter cuts for the three cases investigated. It is again immediately seen that no significant differences exist in the amount of correlation energy obtainable employing the $\lambda\lambda$-HF operator and the simpler $\lambda$-HF operator ($\sim 0.002$ eV in all cases), whereas the $g$-H results deviate from the $\lambda\lambda$-HF ones by $0.1-0.2$ eV. One eye-catching point emerging from the results in Table I is the large correlation energy difference between $\text{H}_2\text{O}$ and $N_2$. The 1.933 eV correlation recovered for $\text{H}_2\text{O}$ (DZ) correspond to 48% of the all-electron FCI result, which is a surprisingly high fraction to be obtained through a simple single-reference SCI calculation. Increasing the basis set size should reduce the fraction of correlation reachable by a limited space CI procedure. It is, however, somewhat unexpected that the correlation energy we have obtained for $\text{H}_2\text{O}$ (DZP) is slightly smaller than the DZ value already in absolute value. The valence FCI correlation for water with the DZP basis set is 5.88 eV. We have computed the difference between the all-electron and valence SDCI correlation to be 0.40 eV. We can thus estimate the all-electron FCI correlation to be not much larger than 6.3 eV (the correlation energy given by the Davidson correction to the all-electron SDCI results to be 6.17 eV). This indicates that the 1.888 eV corre-

![Figure 3](http://example.com/fig3.png)

**FIG. 3.** (a) Full line: scaled-Hartree–Fock curve of the SCI energy for $N_2$. Broken line: one-determinant energy; (b) $g$-Hartree curve of the SCI energy for $N_2$.
TABLE I. Parameter values at the minima of the CI energy surface and correlation energy obtained in single excitation CI calculations using $\lambda\lambda$-HF, $\lambda$-HF, and $g$-H orbitals. All energies in eV.

<table>
<thead>
<tr>
<th>Operator</th>
<th>$\lambda$</th>
<th>$\bar{\lambda}$</th>
<th>Corr. energy</th>
<th>$\lambda$</th>
<th>$\bar{\lambda}$</th>
<th>Corr. energy</th>
<th>$\lambda$</th>
<th>$\bar{\lambda}$</th>
<th>Corr. energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda\lambda$-HF</td>
<td>0.824</td>
<td>0.718</td>
<td>1.933 (12.39*)</td>
<td>0.830</td>
<td>0.770</td>
<td>1.888 (12.21*)</td>
<td>0.878</td>
<td>0.820</td>
<td>0.678 (5.83*)</td>
</tr>
<tr>
<td>$\lambda$-HF</td>
<td>1.203</td>
<td>1.467</td>
<td>1.597</td>
<td>1.202</td>
<td>1.465</td>
<td>1.547</td>
<td>1.104</td>
<td>1.119</td>
<td>0.541</td>
</tr>
<tr>
<td>$g$-H</td>
<td>0.840</td>
<td>0.840</td>
<td>1.931</td>
<td>0.839</td>
<td>0.839</td>
<td>1.885</td>
<td>0.886</td>
<td>0.886</td>
<td>0.676</td>
</tr>
<tr>
<td></td>
<td>1.167</td>
<td>1.167</td>
<td>1.565</td>
<td>1.167</td>
<td>1.167</td>
<td>1.514</td>
<td>1.103</td>
<td>1.103</td>
<td>0.541</td>
</tr>
<tr>
<td></td>
<td>0.763</td>
<td>0.237</td>
<td>1.868</td>
<td>0.762</td>
<td>0.238</td>
<td>1.811</td>
<td>0.808</td>
<td>0.192</td>
<td>0.496</td>
</tr>
<tr>
<td></td>
<td>1.010</td>
<td>-0.010</td>
<td>1.015</td>
<td>1.005</td>
<td>-0.005</td>
<td>0.926</td>
<td>0.964</td>
<td>0.036</td>
<td>0.240</td>
</tr>
</tbody>
</table>

*CI energy lowering relative to the reference configuration energy.

The situation changes dramatically in the case of $N_2$. The outer valence ($2p$) FCI correlation is here 5.49 eV$^3$ (at the bond distance we have used of 2.15 a.u.). Adding to this figure the computed difference of 4.14 eV between the all-electron and $2p$-only SDCI correlation yields a lower bound to the true FCI correlation energy of 9.6 eV. Again, this is consistent with the SDCI + Davidson correction estimate of 9.53 eV. Thus, in this case, the 0.678 eV correlation we have obtained at the minimum of the $\lambda\lambda$-HF surface is possibly less than 7% of the total. It is not easy to give a specific explanation of these gross differences between the $H_2O$ and $N_2$ cases. On one side, it seems clear that the inherent contribution of single excitations to electron correlation becomes increasingly less relevant as the complexity of the system under study increases, that is to say, the full MCSCF correlation energy in the single excitation space is presumably a much smaller fraction for $N_2$ than for $H_2O$. On the other hand, it cannot be excluded that the severely restricted class of orbital transformations allowed for by the $\lambda\lambda$-HF operator may have rather diverse effectiveness depending on the characteristics of the electronic structure. In this latter respect, a gradual relaxation of the constraints on the mean field scaling parameters (e.g., allowing for different treatment of the $\sigma$ and $\pi$ electrons in $N_2$) would give rewarding information. As we have stressed, however, this is beyond the scope of the present work, whose main concern was to establish some fundamental theoretical properties of prototype mean field models.

We would like finally to discuss the results of some applications of the $\lambda\lambda$-HF and $\lambda$-HF operators in CI calculations with larger configuration spaces. We have used $\lambda\lambda$-HF orbitals in SDCI calculations on $H_2O$ (DZ) and $\lambda$-HF orbitals in SDCI calculations on $N_2$ (DZP) and SDTCI on $H_2O$ (DZ). The results for the water molecule are summarized in Table II. The SDCI results indicate that the amount of extra correlation that can be obtained with respect to the HF basis when double excitations are included is very small (some fraction of eV). This is expected to be a rather general feature, which can be qualitatively explained by inspecting the exponential ansatz of orbital transformations (Thouless theorem$^{13}$). Using a suitably chosen non-Hartree–Fock orbital basis and including single excitations contributions in the wave function permits to take some restricted account, principally, of double excitation terms of the corresponding HF-based CI expansion. These can give large correlation contributions. On the other hand, a SDCI expansion on a non-HF basis corresponds essentially to additionally include less important triple excitation terms of the HF-CI wave function. One interesting finding from our SDCI calculations on $H_2O$ is, however, that the optimal SDCI parameters for the $\lambda\lambda$-HF operator are both slightly larger than 1 and,

TABLE II. Optimal parameters and percentage of correlation energy obtained in various CI calculations on $H_2O$ (DZ), using HF, $\lambda\lambda$-HF, $\lambda$-HF, and $g$-H orbitals.

<table>
<thead>
<tr>
<th>Operator</th>
<th>$\lambda$</th>
<th>$\bar{\lambda}$</th>
<th>% corr. energy</th>
<th>$\lambda$</th>
<th>$\bar{\lambda}$</th>
<th>% corr. energy</th>
<th>$\lambda$</th>
<th>$\bar{\lambda}$</th>
<th>% corr. energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>HF</td>
<td>1.000</td>
<td>1.000</td>
<td>0.0</td>
<td>1.000</td>
<td>1.000</td>
<td>94.7</td>
<td>1.000</td>
<td>1.000</td>
<td>95.5</td>
</tr>
<tr>
<td>$\lambda\lambda$-HF</td>
<td>0.824</td>
<td>0.718</td>
<td>48.0</td>
<td>1.007</td>
<td>1.105</td>
<td>94.8</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$\lambda$-HF</td>
<td>0.840</td>
<td>0.840</td>
<td>47.9</td>
<td>0.993</td>
<td>0.993</td>
<td>49.8</td>
<td>0.850</td>
<td>0.850</td>
<td>98.7</td>
</tr>
<tr>
<td>$g$-H$^b$</td>
<td>0.763</td>
<td>0.237</td>
<td>46.4</td>
<td>0.88</td>
<td>0.12</td>
<td>93.2</td>
<td>0.770</td>
<td>0.230</td>
<td>98.6</td>
</tr>
</tbody>
</table>

*a Here, the search in the $\lambda\lambda$-HF plane has not been carried out. We expect a slight improvement over $\lambda$-HF.

*b Reference 9.
again, close to the scaled-HF cut of the surface. Indeed, both for H$_2$O (DZ) and N$_2$ (DZP), use of the $\lambda$-HF operator itself leads to a better SDCI energy than obtained with the HF one. It should be noted, in this respect, that the $g$-Hartree orbital basis, on the contrary, fails to improve over Hartree–Fock.\(^9\)

On the grounds of the above interpretation of the poor SDCI results, it should be expected that our mean field models can instead be very effective in CI expansions including also triple excitations (SDTCI), which may account for important quadruple excitation contributions to the corresponding HF-CI. Our SDCI calculations using the $\lambda$-HF operator on H$_2$O largely confirm this. While the HF-SDTCI calculation recovers 95.5% of the correlation energy (improving by only 0.8% over the SDCI), the $\lambda$-HF-SDTCI result approaches 99%. This means that more than 71% of the residual correlation left out by the HF-SDTCI treatment is accounted for by our calculation, i.e., practically the full quadruple excitation contribution to the HF-CI wave function. In this case (see Table II), employing the $g$-H orbitals produces results which are almost as good, reflecting the already noted fact that, as soon as one departs from the HF point, the $\lambda\lambda$-CI energy surfaces for H$_2$O (DZ) are rather flat along the bottom of the valleys. A more pronounced difference in favor of the $\lambda$-HF orbitals is probably to be expected in the case of nitrogen.

It is extremely interesting to analyze the SDTCI energy dependence on the parameter $\lambda$, shown in Fig. 4. The curve shows a very remarkable similarity to the SCI curve of Fig. 3(a), which is strongly suggestive of an overall similarity of the whole $\lambda\lambda$ surface. Due to the interaction with the double excitations, the HF point is not stationary on the SDTCI surface, but, in fact, results to be extremely close to the actual maximum of the $\lambda$-HF curve (located at $\lambda = 0.994$) and, presumably, very close to a local maximum of the whole surface. Thus, we arrive at the conclusion that, even in SDTCI calculations, the Hartree–Fock orbital basis is practically the worst basis one could select. We can make this argument stronger. Double excitations are the only class of excitations which couples directly to the HF wave function. As mentioned above, and can be easily proved, this is responsible for the fact that the CI energy is not stationary at the HF point when double excitations are included. We have, therefore, also performed CI calculations with $\lambda\lambda$-HF orbital bases around the HF point in a configuration space comprising only single and triple excitations. The results unambiguously prove that in this space the CI energy has again a true local maximum at $\lambda = \lambda = 1$. Finally, we note that the close similarity between the SCI and SDTCI surfaces also implies that the optimal SCI $\lambda$ and $\lambda$ parameters, which can be quite inexpensively determined, provide an extremely good starting guess of the optimal parameters for SDTCI calculations which can deliver near-FCI energies.

IV. SUMMARY AND CONCLUDING REMARKS

The configuration interaction wave function expansion is known to be a slowly convergent expansion, whose rate of convergence depends substantially on the molecular orbital basis used to express the configuration functions. In this respect, the most commonly used orbital basis, namely the Hartree–Fock basis, is clearly not the optimal choice. Guided by this simple idea, we have investigated the capabilities of a generalized mean field model to generate more suitable orbitals for use in single reference CI calculations. The proto-type Fock operator we have used ($\lambda\lambda$-HF) is characterized by two independent parameters scaling the Coulomb and exchange potentials and, thus, includes the Hartree–Fock and $g$-Hartree operators as special cases. We have, therefore, studied theoretically the CI energy surface as a function of the $\lambda$ and $\lambda$ scaling parameters, devoting particular attention to the consequences of departure from the HF point on the surface. This problem is of particularly simple formulation in the case of single excitation CI, where use can be made of Brillouin’s theorem, and we have derived general exact expressions for the curvature of the SCI surface at the HF stationary point.

Using $\lambda\lambda$-HF orbital bases, we have carried out CI calculations for the molecules H$_2$O and N$_2$ in configuration spaces comprising single, single and double, single, double, and triple, and single and triple excitations. The results show that the features and topology of the computed surfaces are essentially identical for both molecular systems and seem, therefore, only dependent on the properties of the $\lambda\lambda$-HF operator. Single reference CI calculations based on orbitals generated by our mean field model can recover very large fractions of correlation energy, approaching 99% in the case of SDTCI on H$_2$O. The one-parameter scaled–Hartree–Fock operator, which is also a special case of the $\lambda\lambda$-HF model, is already more effective than Hartree–Fock and $g$-Hartree in all cases. Viewed in the broader context of our general model, the $g$-Hartree operator describes only a curve of the $\lambda\lambda$ surface and does not possess any distinguished quality in providing orbital bases for CI calculations.

The most remarkable theoretical result that our calculations have provided is that the Hartree–Fock point is, in fact, a true local maximum of the CI surface in all cases when double excitations are not included in the configuration space, and extremely close to a maximum in SDTCI. These results are not yet analytically generalized but constitute very strong evidence for conjecturing that they reflect gen--
eral inherent properties of the $\lambda\tilde{A}$-HF operator.

Our findings lend themselves to be interpreted in a more general perspective, allowing us to make some provocative considerations. The Hartree–Fock scheme provides an optimal independent-particle model, i.e., the best first order electronic wave function. This model is, particularly for small molecules such as the ones here studied, often a satisfactory one, correlation effects constituting a minor perturbation. The optimality property obviously refers to Hartree–Fock as a stand-alone physical model, but it is often implicitly assumed that it also necessarily renders this model a good (if clearly not the best) starting point for theoretical treatments which incorporate electron correlation effects. Our study proves that this assumption can be utterly misleading and that, indeed, the Hartree–Fock model can be the worst choice, among a suitable class of mean field models, to generate the first order wave function to be improved through a CI procedure. It is tempting to cast this into the more general conjecture that optimization within a given physical or computational model, which amounts to minimizing the perturbations to that model, can actually constitute the less efficient procedure to start recovering those perturbation effects, i.e., to overcome the deficiencies of the model itself. This speculation, to which the present work provides preliminary supporting evidence, might have some impact on the whole approach to the electron correlation problem and apply to a wide range of methods, going from Møller–Plesset perturbation theory to MCSCF optimization followed by multireference CI. It seems clear that further theoretical study in the directions outlined by the present work could be highly rewarding.

ACKNOWLEDGMENTS

The authors wish to thank H.-D. Meyer and A. Sgamelotti for very helpful discussions and R. Sawatzki for his continuous collaboration. F. T. thanks the A. von Humboldt Stiftung for his scholarship.

APPENDIX

We give here an explicit derivation of Eqs. (5) and (9) for the second derivative of the one-determinant energy and of the SCI energy with respect to the parameter $\lambda$ of the scaled-HF operator, computed at $\lambda = 1$. We change our independent variable $\lambda$ to $\theta = \lambda - 1$, in order to expand all the quantities of interest to second order in powers of $\theta$ around $\theta = 0$. In the following we will denote $\theta$-dependent quantities of a tilde, which will be omitted for the corresponding Hartree–Fock ($\theta = 0$) quantities. By a standard partitioning procedure$^{16}$ of the SCI Hamiltonian eigenvalue equation we obtain the following formal expression for the SCI energy at any value of $\theta$:

$$\tilde{E}_{\text{SCI}} = \tilde{E}_0 + \tilde{T} (\tilde{E}_{\text{SCI}} 1 - \tilde{H}_0) \tilde{T}^{-1}. \quad (A1)$$

This equation $\tilde{E}_0$ is the energy of the reference determinant, $\tilde{T}$ is the vector of Hamiltonian matrix elements relative to the reference determinant and its single excitations, and $\tilde{H}_0$ is the Hamiltonian matrix over the single excitations alone. Upon expansion in powers of $\theta$ around $\theta = 0$, $\tilde{E}_0$ will have a zero order term given by the Hartree–Fock energy, $E_{\text{HF}}$, and a vanishing linear term, due to its stationarity. To second order in $\theta$ we thus write

$$\tilde{E}_0 = E_{\text{HF}} + \theta^2 E_{\text{HF}}^\prime, \quad (A2)$$

where $E_{\text{HF}}^\prime$ is a second order correction. Owing to Brillouin's theorem, the vector $\tilde{T}$ has no zero order term and, therefore, to obtain a second order expansion of Eq. (A1), we only need the first order expression for $\tilde{T}$:

$$\tilde{T} = \theta \tilde{T}^\prime \quad (A3)$$

and the zeroth order (i.e., Hartree–Fock) expression for the inverse matrix:

$$(\tilde{E}_{\text{SCI}} 1 - \tilde{H}_0)^{-1} = (E_{\text{HF}} 1 - H_0)^{-1}. \quad (A4)$$

We can now subtract $E_{\text{HF}}$ from the diagonal of the SCI matrix, thus obtaining directly the second order expression for the difference between the Hartree–Fock energy and the SCI energy:

$$\Delta \tilde{E}^{(2)}_{\text{SCI}} = \theta^2 (E_{\text{HF}}^\prime - \tilde{T}^\prime \tilde{X}^{-1} \tilde{T}^\prime), \quad (A5)$$

where

$$X = H_0 - E_{\text{HF}} 1. \quad (A6)$$

As anticipated, no first order term appears in Eq. (A5).

We would like now to derive explicit expressions for the quantities $\tilde{T}^\prime$ and $E_{\text{HF}}^\prime$ appearing in Eq. (A5), and to do this we must first analyze the perturbation expansion of the scaled-HF equations:

$$\tilde{F} \phi = \tilde{E} \phi. \quad (A7)$$

For ease of notation, we will henceforth denote occupied orbital indices by $ij, k, \ldots$ and virtual orbital indices by $a, b, c, \ldots$. When the orbital occupation number is irrelevant we will instead use the indices $r, s, t, \ldots$. We further assume throughout to be dealing with real orbitals. We start with re-writing the $\lambda$-HF operator as a function of $\theta$ in first quantized notation:

$$\tilde{F} = \hat{h} + (\theta + 1) \sum_i \tilde{V}_i, \quad (A8)$$

where $\tilde{V}_i$ is the usual difference between the Coulomb and exchange operators. We further write the second order expansion in $\theta$ for the $\lambda$-HF orbitals and orbital energies:

$$\tilde{\phi}_i = \phi_i + \theta \phi_i^\prime + \theta^2 \phi_i^\prime', \quad (A9)$$

$$\tilde{e}_i = e_i + \theta e_i^\prime + \theta^2 e_i^\prime', \quad (A10)$$

where $\phi$ and $e$ are the HF orbitals and orbital energies, and the primed and doubly primed quantities are again first and second order corrections, respectively. By substituting Eq. (A9) into Eq. (A8) we obtain, to second order,

$$\tilde{F} = \tilde{F} + \theta \tilde{F}^\prime + \theta^2 \tilde{F}^\prime', \quad (A11)$$

where $\tilde{F}$ is the HF operator and

$$\tilde{F}^\prime = \frac{\delta \tilde{F}}{\delta \theta}_{\theta = 0} = \sum_i (\tilde{V}_i + \tilde{V}_i' + \tilde{V}_i''), \quad (A12)$$

$$\tilde{F}^\prime' = \frac{1}{2} \left[ \frac{\delta^2 \tilde{F}}{\delta \theta^2} \right]_{\theta = 0} = \sum_i (\tilde{V}_i + \tilde{V}_i' + \tilde{V}_i'' + \tilde{V}_i''' + \tilde{V}_i'''). \quad (A13)$$

In Eqs. (A12) and (A13) the primed and doubly primed
indices indicate that the corresponding orbital expansion terms of Eq. (A9) appear in the Coulomb and exchange operators. We now introduce Eqs. (A9)–(A11) into Eq. (A7) to obtain the desired expansion of the $\lambda$-HF equations. The zeroth order equations coincide of course with the HF equations. From the first order equations
\[
(\hat{F} - e_r)\phi_r' = (e_r' - \hat{F}^r)\phi_r,
\]
we get an expression for the first order term of the orbital energies:
\[
e_r' = \langle \phi_r | \hat{F}^r | \phi_r \rangle = \sum_i [V_{ri}n_{i1} + 2V_{ri}n_{i1}] + \frac{1}{2} \sum_{ij} V_{ij}
\]
We now take the first order terms $\phi_r'$ to be orthogonal to the corresponding HF orbitals and expand them in the HF basis:
\[
\phi_r' = \sum_{s \neq r} \phi_s Y_{sr}.
\]
From Eq. (A14), for $s \neq r$, we then obtain:
\[
Y_{sr} = \frac{\langle \phi_r | \hat{F}^r | \phi_s \rangle}{e_r - e_s}.
\]
Since $\hat{F}^r$ is an Hermitian operator, we see, as expected, that $Y$ is an anti-Hermitian (antisymmetric) matrix, i.e., $Y^\dagger = -Y$. Making now use of Eqs. (A12) and (A16) in Eq. (A17) leads to a linear equation system for the mixing coefficients $Y_{sr}$:
\[
(e_r - e_s) Y_{sr} + \sum_{a} [V_{ra}n_{a1} + V_{ra}n_{a1}] Y_{ar} = \sum_{i} V_{si}
\]
where all the quantities are defined in the HF basis. Note that the sum at the second term of the left-hand side of Eq. (A18) is restricted to occupied/virtual mixing by virtue of the antisymmetry of $Y$. From the form of Eq. (A18) it is immediately seen that the occupied/virtual mixing coefficients decouple. We can regard the part of the lower triangle of $Y$ containing them as a vector $C$ in the space of the single excitations and rewrite Eq. (A18) in matrix form as follows:
\[
DC = U,
\]
where the matrix elements of $D$ and $U$ are given by Eqs. (6) and (8) of the text, respectively, i.e.,
\[
D_{a,b} = (e_a - e_b) \delta_{ab} \delta_{\gamma} + V_{a\gamma}n_{\gamma1} + V_{a\gamma}n_{\gamma1},
\]
\[
U_{a} = \sum_{j} V_{aj}\gamma^j.
\]
The occupied/occupied and virtual/virtual mixing coefficients are then obtained from $C$ and $\hat{T}_0$ 9. (A18). The second order part of the $\lambda$-HF equations reads
\[
\hat{F} \phi_r'' + \hat{F} \phi_r' + \hat{F}^r \phi_r = e_r \phi_r'' + e_r' \phi_r' + e_r'' \phi_r,
\]
from which it is useful to extract an expression for the second order term of the orbital energies:
\[
e_r'' = \langle \phi_r | \hat{F}^r | \phi_r \rangle + \langle \phi_r | \hat{F} | \phi_r \rangle = \sum_i [V_{ri}n_{i1} + 2V_{ri}n_{i1}] + \frac{1}{2} \sum_{ij} V_{ij}
\]
We now return to the reference determinant energy $\tilde{E}_0$ which, as a function of $\theta$, can be written
\[
\tilde{E}_0 = \sum_i (\theta + \frac{1}{2}) \sum_{ij} V_{ij}
\]
By inserting here Eqs. (A9), (A10), and (A22) one readily obtains the desired expression for the second order term $\tilde{E}_0$: 10
\[
\tilde{E}_0'' = \sum_{ij} V_{ij}\gamma^j.
\]
Making use of Eq. (A16) and of the antisymmetry of $Y$, this reduces to
\[
E_0'' = \sum_{ij} V_{ij}\gamma^j U U^\dagger
\]
or
\[
E_0'' = C^\dagger D C
\]
which, together with Eq. (A2), yields Eq. (5) of the text.
We turn now to the vector $\hat{T}_0$ of the Hamiltonian matrix elements between the reference configuration and its single excitations, which is defined by
\[
\hat{T}_0 = \langle \phi_0 | \hat{H}_a | \phi_0 \rangle = h_{aa} + \sum_{ij} V_{ij}\gamma^j.
\]
where $\Phi_0$ being the reference determinant. By introducing here Eqs. (A9), (A15), and (A16), it can be verified that the zeroth order term vanishes and the following simple expression holds for the first order term:
\[
\hat{T}_0 = DC = U.
\]
We can now replace Eqs. (A25) and (A28) into Eq. (A5) to obtain the final desired expression for the second order SCI energy shift:
\[
\Delta E_{SCI}^{(2)} = \theta^2 (U U^\dagger - U X^{-1} U).
\]
This is more conveniently rewritten as a quadratic form:
\[
\Delta E_{SCI}^{(2)} = \theta^2 (U^\dagger (D^{-1} - X^{-1}) U
\]
which, written in terms of the vector $C$, yields immediately Eqs. (9) and (10) of the text.

14. L. S. Cederbaum and F. Tarantelli (to be published).