The trust-region self-consistent field method: Towards a black-box optimization in Hartree–Fock and Kohn–Sham theories

Lea Thøgersen, Jeppe Olsen, Danny Yeager, a) and Poul Jørgensen
Department of Chemistry, University of Aarhus, DK-8000 Aarhus C, Denmark

Paweł Salek
Laboratory of Theoretical Chemistry, The Royal Institute of Technology, Teknikringen 30, Stockholm SE-10044, Sweden

Trygve Helgaker
Department of Chemistry, University of Oslo, P.O. Box 1033 Blindern, N-0315 Norway

(Received 17 February 2004; accepted 5 April 2004)

The trust-region self-consistent field (TRSCF) method is presented for optimizing the total energy $E_{\text{SCF}}$ of Hartree–Fock theory and Kohn–Sham density-functional theory. In the TRSCF method, both the Fock/Kohn–Sham matrix diagonalization step to obtain a new density matrix and the step to determine the optimal density matrix in the subspace of the density matrices of the preceding diagonalization steps have been improved. The improvements follow from the recognition that local models to $E_{\text{SCF}}$ may be introduced by carrying out a Taylor expansion of the energy about the current density matrix. At the point of expansion, the local models have the same gradient as $E_{\text{SCF}}$ but only an approximate Hessian. The local models are therefore valid only in a restricted region—the trust region—and steps can only be taken with confidence within this region. By restricting the steps of the TRSCF model to be inside the trust region, a monotonic and significant reduction of the total energy is ensured in each iteration of the TRSCF method. Examples are given where the TRSCF method converges monotonically and smoothly, but where the standard DIIS method diverges. © 2004 American Institute of Physics. [DOI: 10.1063/1.1755673]

I. INTRODUCTION

The steady progress in computer technology and quantum-chemical methodology has widened the range of users of quantum-chemical software packages to include a vast number of practicing, experimental chemists. Routinely, such users perform Hartree–Fock (HF) calculations and Kohn–Sham (KS) density-functional theory (DFT) calculations for molecules of a size and complexity that, a decade ago, were beyond reach even for the most advanced research codes. This development calls for further advances in the automatization of the self-consistent field (SCF) procedure used to optimize the HF and DFT energies, so as to ensure that convergence may be reached in a routine manner even for very complex molecules.

In the original formulation, the SCF procedure consists of a sequence of Roothaan–Hall (RH) iterations.1,2 At each iteration, a Fock/KS matrix is first constructed from the current approximation to the one-electron density matrix and then diagonalized to yield an improved set of orbitals and orbital energies and thus an improved density matrix. In the subsequent iteration, this improved density matrix is then used to construct a new Fock/KS matrix, thereby establishing the iteration procedure. However, such a sequence of RH iterations converges only in simple cases. To improve upon the convergence, each RH iteration may be extended to include, in addition to the diagonalization step, also a step where the best density matrix is generated in the subspace of the density matrices of the current and preceding RH iterations. In the next RH iteration, this averaged density matrix rather than the pure density matrix obtained in the last diagonalization is used to construct the new Fock/KS matrix.

In this paper, we make improvements both to the RH diagonalization step and to the density-subspace optimization step of the SCF scheme. Our approach follows from the recognition that, in both steps, we may construct local models to the SCF energy function $E_{\text{SCF}}$ by a Taylor expansion of the energy about the current density matrix. However, since, at the point of expansion, these models have an exact gradient but only an approximate Hessian, they are valid only in a restricted region about the current approximation to the density matrix—the trust region. Therefore, when these local models are used in the course of the SCF optimization, it is essential they are used only to generate steps within their trust region. Only in this manner can it be ensured that the SCF energy is systematically and sufficiently lowered at each iteration.

In the RH diagonalization part of the SCF optimization, the improvements are obtained by introducing an energy function $E_{\text{RH}}$ that corresponds to the sum of the occupied
An unconstrained minimization of $E^{\text{RH}}$ results in the same solution (i.e., density matrix) as obtained by a diagonalization of the Fock/KS matrix. However, since, at the point of expansion, the RH energy function $E^{\text{RH}}$ has only the gradient in common with the true SCF energy $E_{\text{SCF}}$, a global minimization of $E^{\text{RH}}$ may lead to steps that are too long to be trusted. We therefore introduce a trust region where $E^{\text{RH}}$ is a good approximation to $E_{\text{SCF}}$. If a global minimization of $E^{\text{RH}}$ leads to a step outside the trust region, then the step to the minimum on the boundary of the trust region for $E^{\text{RH}}$ is taken instead. This step is found by a level-shifting technique, where the occupied molecular orbital energies effectively are shifted by some constant to increase the gap between the occupied and virtual molecular orbitals. Level shifting has previously been used to improve the convergence of the simple RH sequence of iterations. An essential feature of our implementation is to adjust the level shift in such a manner that the step is to the boundary of the trust region, recognizing that only in this manner does a lowering of $E^{\text{RH}}$ result in a lowering of $E_{\text{SCF}}$. For this reason, the resulting method is called the trust-region RH (TRRH) method.

The optimization of the density matrix in the subspace of the density matrices of the preceding RH iterations has a long history. Early on, it was recognized that a simple averaging of the density matrices of the last few RH iterations significantly improves the convergence of the RH scheme. This simple density-matrix averaging technique was later rationalized and systematized in the direct inversion in iterative subspace (DIIS) method of Pulay. In the DIIS method, an improved density matrix is obtained as a linear combination of the previous density matrices by minimizing the norm of the corresponding linear combination of gradients. The DIIS method significantly speeds up the local convergence and convergence can often be obtained to ground states of rather complex molecules with a small gap between energies of the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) and with a large number of close-lying electronic states.

Several attempts have been made to modify the DIIS algorithm so as to improve upon its global convergence behavior. Recently, Kudin, Scuseria, and Cances proposed the energy DIIS (EDIIS) method, where the DIIS gradient-norm minimization is replaced by a minimization of an approximate energy function. In EDIIS, the variational parameters, which are the linear expansion coefficients of the density matrices from the previous RH iterations, may only take on values that give densities in the convex set—that is, densities with occupation numbers between 0 and 1. As the EDIIS method is based on the minimization of an approximate energy function, it may have some advantages in the global region. However, it is worrying that a convex solution often cannot be obtained and that the observed local convergence of the EDIIS method is slower than in the standard DIIS method.

In the DIIS and EDIIS methods, an improved density matrix is obtained as a sum of the density matrices from the preceding RH diagonalization steps. Consequently, the averaged density matrix is not idempotent as required in HF and KS theories. The deviation from idempotency may be reduced using a purified density matrix as the one suggested by McWeeny. This has been done for the SCF energy minimization by several workers including Nunes and Vanderbilt and Daniels and Scuseria and for the calculation of geometrical derivatives by Ochsenfeld and co-workers. It may also be done for the EDIIS energy function. The energy function then has the same gradient as $E_{\text{SCF}}$, but also contains terms which cannot be obtained from the densities and Fock/KS matrices of the previous RH iterations. Neglecting these terms, we arrive at the density-subspace minimization (DSM) algorithm proposed in this paper. At the point of expansion, the DSM energy function $E^{\text{DSM}}$ thus has the same gradient as the true energy function $E_{\text{SCF}}$ but only an approximate Hessian. Again, a trust region may be introduced and only steps within this region are taken, ensuring that any lowering of $E^{\text{DSM}}$ also corresponds to a lowering of $E_{\text{SCF}}$. The resulting method is called the trust-region DSM (TRDSM) method.

In the next section, we first describe the standard optimization of the SCF energy function in a density-matrix formulation. The TRRH method is then discussed in Sec. II A and the TRDSM method in Sec. II B. In Sec. III, we give some numerical examples to demonstrate the performance of the resulting trust-region SCF (TRSCF) method. The last section contains some concluding remarks.

II. THEORY

For a closed-shell system with $N/2$ electron pairs, the Hartree–Fock (HF) energy excluding the nuclear–nuclear repulsion energy is given by

$$E_{\text{SCF}}(\mathbf{D}) = 2 \text{Tr} \mathbf{hD} + \text{Tr} \mathbf{DG(D)},$$

(1)

where $\mathbf{D}$ is the one-electron density matrix in the atomic-orbital (AO) basis, $\mathbf{h}$ is the one-electron Hamiltonian matrix and $\mathbf{G(D)}$ is defined as

$$G_{\mu\nu}(\mathbf{D}) = \sum_{\rho\sigma} (2g_{\mu\rho\sigma} - g_{\rho\mu\sigma}) D_{\sigma\rho},$$

(2)

where $g_{\mu\rho\sigma}$ is a two-electron integral in the AO basis. For the energy in Eq. (1) to be a valid approximation to the true HF energy, the density matrix $\mathbf{D}$ must satisfy the symmetry, trace, and idempotency conditions:

$$\mathbf{D}^T = \mathbf{D},$$

(3)

$$\text{Tr} \mathbf{DS} = \frac{N}{2},$$

(4)

$$\text{DS} = \mathbf{D}.$$  

(5)

Similar conditions apply in the Kohn–Sham (KS) theory, but the energy function of Eq. (1) must then be modified by including the exchange-correlation term and by scaling (or complete removal) of the exchange term from Eq. (2).

The traditional approach to the optimization of the HF energy is an iterative one. From the current approximation to the density matrix $\mathbf{D}_n$ in iteration $n$, a Fock matrix is built

$$\mathbf{F(D}_n) = \mathbf{h} + \mathbf{G(D}_n)$$

(6)
and, following the Roothaan–Hall (RH) procedure, the Fock matrix is diagonalized

\[ F(D_n)C_{\text{occ}} = SC_{\text{occ}} \lambda. \tag{7} \]

where \( S \) is the overlap matrix in the AO basis, to give a set of occupied molecular orbitals (MOs), from which a new approximation to the density matrix is obtained as

\[ D_{n+1} = C_{\text{occ}} C^T_{\text{occ}}. \tag{8} \]

The iteration procedure is established using \( D_{n+1} \) as the current density in Eq. (6). The final solution to the minimization problem is obtained when the \( D_n \) and \( D_{n+1} \) are the same. This self-consistent field (SCF) iteration procedure is established. In the course of the TRSCF iterations, the following matrices are set up in the local region. In practice, we define sufficiently large regions according to Eq. (9) to yield the KS matrix.

The pure RH iterations presented above often do not converge. A powerful method for handling this divergence is not to construct the Fock matrix from the density matrix, but rather from an average of all previous density matrices:

\[ \bar{D}_n = \sum_{i=1}^{n} c_i D_i. \tag{9} \]

The averaged density matrix \( \bar{D}_n \) is then used in place of the pure density matrix \( D_n \) in Eq. (6) to obtain the Fock matrix \( F(\bar{D}_n) \) as

\[ F(\bar{D}_n) = \sum_{i=1}^{n} c_i F(D_i) \tag{10} \]

and the iteration procedure is established. In the course of the TRSCF iterations, the following matrices are set up in the order indicated: \( D_1, F(D_1), D_2, F(D_2), D_3, F(D_3), \ldots \). Among these, \( D_j, F(D_j), D_k, F(D_k), D_l, F(D_l), \ldots \) are saved during the iteration procedure.

In the following, we describe improvements to the SCF diagonalization and density-subspace optimization steps. In Sec. II A, we describe how the trust-region RH (TRRH) method is used to generate new density matrices by a modification of the traditional RH method Eqs. (7) and (8). Next, in Sec. II B, we introduce the trust-region density-subspace minimization (TRDSM) method for calculating the averaged density matrix of Eq. (9). In the following, we use the indices \( i,j,k,l \) for occupied MOs and the indices \( a,b,c,d \) for the virtual MOs.

### A. The trust-region Roothaan–Hall method

As discussed in Ref. 3, the traditional RH method may be viewed as a minimization of the sum of the orbital energies of the occupied MOs

\[ E_{\text{RH}} = 2 \sum \epsilon_i = 2 \text{Tr} F(\bar{D}) D. \tag{11} \]

subject to orthonormality constraints on the occupied MOs \( \phi_i \):

\[ \langle \phi_i | \phi_j \rangle = \delta_{ij}. \tag{12} \]

Whereas \( \bar{D} \) is the current approximation to the HF/KS density matrix, usually obtained as a linear combination of the previous densities according to Eq. (9), the density matrix \( D \) to be optimized in Eq. (11) is related to the occupied MOs resulting from the diagonalization of \( F(\bar{D}) \) as

\[ D = C_{\text{occ}} C^T_{\text{occ}}. \tag{13} \]

To see this, consider the constrained minimization of \( E^\text{RH} \) in Eq. (11) expressed in terms of the Lagrangian

\[ L = 2 \text{Tr} F(\bar{D}) D - 2 \text{Tr} \lambda (C^T_{\text{occ}} C_{\text{occ}} - I_{N/2}), \tag{14} \]

where the multipliers \( \lambda_{ij} \) ensure orthonormality among the occupied MOs. Minimization of this Lagrangian leads to the standard RH equations:

\[ F(\bar{D}) C_{\text{occ}} = SC_{\text{occ}} \lambda. \tag{15} \]

However, since \( E^\text{RH} \) of Eq. (11) is only a crude model of the true energy \( E_{\text{SCF}} \) (the gradient is correct at \( \bar{D} \) assuming \( \bar{D} \) is idempotent), a global minimization of \( E^\text{RH} \) according to Eq. (15) may easily lead to steps that are too long to be trusted as they are outside the region where \( E^\text{RH} \) is a good approximation to \( E_{\text{SCF}} \). Steps outside the trust region may often not lead to a reduction of the total energy \( E_{\text{SCF}} \).

#### 1. The level-shifted Roothaan–Hall equations

To avoid too long steps, an additional constraint is imposed on the optimization of Eq. (11), namely, that the new density matrix \( D \) in Eq. (13) does not differ too much from the old matrix \( \bar{D} \). This condition is conveniently expressed in terms of the overlap between the density matrices in the metric norm

\[ (D_1, D_2) = \text{Tr} D_1 D_2 = \sqrt{a_{\text{min}}^2 \text{Tr} \bar{D}^2 D}. \tag{16} \]

where \( \text{Tr} \bar{D}\bar{D} = N/2 \) since \( \bar{D} \) is not necessarily idempotent. Note that, for \( D \) equal to (an idempotent) \( \bar{D} \), \( a \) is equal to one. For \( a \) sufficiently close to one, \( a \) will therefore be taken in the local region. In practice, we define sufficiently close to one by the parameter \( a_{\text{min}} = 0.975 \).

Introducing an undetermined multiplier \( \mu \) associated with this new constraint, we obtain the following Lagrangian:

\[ L = 2 \text{Tr} F(\bar{D}) D - 2 \mu (\text{Tr} \bar{D}^2 D - \sqrt{a_{\text{min}}^2 \text{Tr} \bar{D}^2 D}) \]

where

\[ - 2 \text{Tr} \lambda (C^T_{\text{occ}} C_{\text{occ}} - I_{N/2}). \tag{17} \]

Differentiating this Lagrangian with respect to the MO coefficients and setting the result equal to zero, we arrive at the level-shifted RH equations

\[ [F(\bar{D}) - \mu \bar{S} D] C_{\text{occ}}(\mu) = SC_{\text{occ}}(\mu) \lambda. \tag{18} \]

To interpret the level-shift term, we note that \( \bar{S} D \) projects out the component of \( C_{\text{occ}} \) that is occupied in \( \bar{D} \) (assuming idempotent \( \bar{D} \)), see Ref. 3. The level shift therefore works only on the occupied part of \( F(\bar{D}) \), shifting all the occupied orbital energies and increasing the gap between the occupied and virtual MOs, in particular the HOMO-LUMO gap.
where $\epsilon_{\mu}^{\text{HOMO}}$ and $\epsilon_{\mu}^{\text{LUMO}}$ are the HOMO and LUMO orbital energies, respectively; in Fig. 1(b), we have plotted the overlap between the old and new density matrices as given by

$$a(\mu) = \frac{\langle D(\mu) | \bar{D} \rangle_S}{\sqrt{\langle D(\mu) | D(\mu) \rangle_S \langle \bar{D} | \bar{D} \rangle_S}},$$

(22)

where $\langle D(\mu) | D(\mu) \rangle_S$ is equal to $N/2$. For sufficiently large $\mu$, the HOMO-LUMO gap Eq. (21) is linear in $\mu$. This linearity of $\Delta \epsilon_{ai}(\mu)$ for large $\mu$ arises from the dependence of the orbital energies on $\mu$ in Eq. (19), where $\mu$ is effectively subtracted from the occupied orbital energies. The MOs $\bar{C}_{\text{occ}}$ occupied in $\bar{D}$ satisfy the generalized eigenvalue equations

$$\bar{S} \bar{D} \bar{C}_{\text{occ}} = \bar{S} \bar{C}_{\text{occ}} \eta,$$

(23)

and become identical to the MOs $C_{\text{occ}}(\infty)$ obtained from Eq. (19) when $\mu$ tends to infinity. The corresponding density is denoted

$$D(\infty) = \bar{C}_{\text{occ}} \bar{C}_{\text{occ}}^T,$$

(24)

and represents a purified $\bar{D}$. In the linear regime of $\Delta \epsilon_{ai}(\mu)$, there is a continuous development of the occupied MOs from those occupied in $\bar{D}$. As $\mu$ decreases and we enter the nonlinear regime at $\mu_{\text{min}}$, the MOs in Eq. (20) no longer correspond to those in Eq. (23). Comparing plot (a) and (b) in Fig. 1, we note that the region $a(\mu) < a_{\text{min}}$ in Fig. 1(b) corresponds roughly to the region $\mu_{\text{opt}} < \mu_{\text{min}}$ in Fig. 1(a).

As we insist on a controlled, continuous development of the MOs from those occupied in $\bar{D}$, the level-shift parameter should be restricted to the linear regime $\mu_{\text{min}} < \mu < \infty$. To determine the optimal level-shift parameter $\mu_{\text{opt}}$, we therefore begin by establishing the onset of linearity $\mu_{\text{min}}$ by linear extrapolation by means of two Fock/KS matrix diagonalizations, giving the two $\Delta \epsilon_{ai}$ values marked by crosses and the linearly interpolated $\mu_{\text{min}}$ value marked with an arrow. Next, since, in the linear interval, a small $\mu$ corresponds to a large step, we investigate whether $\mu_{\text{min}}$ is acceptable by checking if $a(\mu_{\text{max}}) > a_{\text{min}}$. If this step is too long, we backtrack by increasing $\mu$ using inexact line search until an acceptable value $\mu_{\text{opt}}$ is found such that $a(\mu_{\text{opt}}) = a_{\text{min}}$, requiring a few additional Fock/KS matrix diagonalizations. In Fig. 1(b), the accepted $\mu_{\text{opt}}$ is marked with an arrow.

For a better understanding of this step, consider the Hessian of the $E_{\text{SCF}}$ energy function:

$$A_{ai,\beta} = \delta_{ij} \delta_{ab} (\epsilon_a - \epsilon_i).$$

(25)

By restricting the level-shift parameter to $\mu_{\text{min}} < \mu < \infty$ where $\epsilon_{\mu}^{\text{LUMO}}(\mu) - \epsilon_{\mu}^{\text{HOMO}}(\mu) > 0$, we ensure that the effective Hessian is positive definite and that the model energy function $E_{\text{RH}}$ is reduced. We note that the Hessian of the true energy function $E_{\text{SCF}}$ is given by the more complicated expression

$$A_{ai,\beta} = \delta_{ij} \delta_{ab} (\epsilon_a - \epsilon_i) + 4 g_{abij} - g_{abij} - g_{aijb}.$$

(26)

Often, the orbital energy difference dominates the Hessian. In such cases, we expect the above step to reduce the SCF energy $E_{\text{SCF}}$ as well as the model function $E_{\text{RH}}$. In any case, when a sufficiently large level shift is added in Eq. (19), the

FIG. 1. For the fourth iteration of the rhodium calculation described in Sec. III we have displayed as a function of the level-shift parameter $\mu$: (a) the HOMO-LUMO gap $\Delta \epsilon_{ai}$, where $\mu_{\text{min}}$ is the smallest accepted level-shift, (b) the overlap $a$ between the old and new density matrices, where $\mu_{\text{opt}}$ is the optimal level-shift, and (c) the change in the model energy $\Delta E_{\text{RH}}$ and the actual energy $\Delta E_{\text{SCF}}$.

Since the SCF energy $E_{\text{SCF}}$ is invariant with respect to an orthogonal transformation between the MOs, Eq. (18) may be transformed to the canonical basis:

$$[F(\bar{D}) - \mu SDS] C_{\text{occ}}(\mu) = SC_{\text{occ}}(\mu) \epsilon,$$

(19)

where the diagonal matrix $\epsilon$ contains the orbital energies.

### 2. Choice of the RH level-shift parameter

The density matrix generated from the restricted RH solution Eq. (19) depends on the level-shift parameter $\mu$:

$$D(\mu) = C_{\text{occ}}(\mu) C_{\text{occ}}^T(\mu).$$

(20)

To see how $\mu$ is determined, we consider the determination of $\mu$ in the fourth iteration of the rhodium-complex calculation described in Sec. III. In Fig. 1(a), we have plotted the HOMO-LUMO gap as a function of $\mu$,

$$\Delta \epsilon_{ai}(\mu) = \epsilon_{ai}^{\text{LUMO}}(\mu) - \epsilon_{ai}^{\text{HOMO}}(\mu),$$

(21)
Hessian structure of Eq. (25) becomes similar to that of the true energy function \( E_{SCF} \) in Eq. (26). The steps generated from \( E^{RH} \) with such level shifts will therefore have essentially the same direction as the ones generated from \( E_{SCF} \).

By construction, the \( E^{RH} \) energy function is lowered when \( \mu \) is chosen according to the above prescription

\[
\Delta E^{RH} = 2 \text{Tr} F(\widetilde{D})[(D(\mu) - \tilde{D})] < 0. \tag{27}
\]

Since \( E^{RH} \) is only a local model of the true energy function \( E_{SCF} \), the associated change in the true energy

\[
\Delta E^{RH}_{SCF} = E_{SCF}(D(\mu)) - E_{SCF}(\tilde{D}) \tag{28}
\]

may be either negative or positive, depending on how well \( E^{RH} \) represents \( E_{SCF} \) for the chosen step. However, for sufficiently small steps, \( \Delta E^{RH}_{SCF} < 0 \), since the model function then represents the true energy well.

Let us consider the relationship between the true lowering \( \Delta E^{RH} \) and the lowering predicted by the model function \( \Delta E^{RH} \). Introducing the (presumably small) differential density matrix

\[
\Delta = D(\mu) - \tilde{D} \tag{29}
\]

and using the identity \( \text{Tr} AG(B) = \text{Tr} BG(A) \) valid for symmetric matrices \( A \) and \( B \), we find that the change in the true energy Eq. (28) may be written in the form

\[
\Delta E^{RH}_{SCF} = 2 \text{Tr} h[D(\mu) - \tilde{D}] + \text{Tr} (\tilde{D} + \Delta) \times G(\tilde{D} + \Delta) - \text{Tr} \tilde{D} G(\tilde{D})
\]

\[
= 2 \text{Tr} h\Delta + 2 \text{Tr} \Delta G(\tilde{D}) + \text{Tr} \Delta G(\Delta). \tag{30}
\]

which shows that the changes in the true energy and in the model energy are related as

\[
\Delta E^{RH}_{SCF} = \Delta E^{RH} + \text{Tr} \Delta G(\Delta). \tag{31}
\]

If the last term (which is second order in \( \Delta \)) is negligible, the energy lowering predicted by the local model \( E^{RH} \) becomes equal to \( E^{RH}_{SCF} \). However, since the correction term is positive (strictly positive in the absence of exchange), its presence in Eq. (31) shows that, for sufficiently large steps, a lowering of the model function may not lead to a lowering of the total energy. To avoid such steps, it would be useful to provide an alternative prediction of \( \Delta E^{RH}_{SCF} \) that is less expensive than the calculation of \( \text{Tr} \Delta G(\Delta) \) itself. Section II A 3 is concerned with this problem.

To demonstrate the efficiency of the chosen level shift \( \mu_{opt} \) in the global region of a SCF optimization, we have for the fourth iteration of the rhodium-complex calculation plotted in Fig. 1(c), \( \Delta E^{RH}_{SCF} \) and \( \Delta E^{RH} \) as a function of \( \mu \). The energy gain \( \Delta E^{RH}_{SCF} \) is about optimal for the level shift \( \mu_{opt} \). Increasing \( \mu \) gives a smaller energy gain while decreasing \( \mu \) gives a slight increase in the energy gain and from \( \mu < 4.5 \), \( \Delta E^{RH}_{SCF} \) is actually positive. Note also that for \( \mu < \mu_{opt} \), \( \Delta E^{RH} \) and \( \Delta E^{RH}_{SCF} \) start to differ indicating that the importance of \( \text{Tr} \Delta G(\Delta) \) increases. The step representing a RH iteration where \( \mu = 0 \) is far too long to be trusted and results in a significant increase of the total energy.

### 3. Prediction of the energy close to the minimum

To develop a better prediction of \( \Delta E^{RH}_{SCF} \) than \( \Delta E^{RH} \), we note that the only part, that cannot easily be evaluated from known Fock-matrices, is the second-order contribution to Eq. (31) from that part of \( \Delta \) that does not belong to the linear space spanned by the previous density matrices \( D_i \). To see this, we decompose the current density matrix \( D(\mu) \) into two parts

\[
D(\mu) = D_i + D_\perp, \tag{32}
\]

where \( D_i \) belongs to the linear space spanned by the previous density matrices and \( D_\perp \) belongs to its orthogonal complement. We then expand \( D_i \) in the following manner:

\[
D_i(\mu) = \sum_{i=1}^{n} c_i(\mu) D_i, \tag{33}
\]

where the expansion coefficients \( c_i(\mu) \) are determined in a least-squares manner

\[
c_i(\mu) = \sum_{j=1}^{n} [M^{-1}]_{ij} \text{Tr} D_j SD_i(\mu) S_i, \quad M_{ij} = \text{Tr} D_j SD_j S_i. \tag{34}
\]

The change in the SCF energy associated with the change of density matrix from \( \tilde{D} \) to \( D(\mu) \) may be expressed as

\[
\Delta E^{RH}_{SCF}(\mu) = E_{SCF}(D_i) - E_{SCF}(\tilde{D}) + 2 \text{Tr} D_\perp F(D_i) + \text{Tr} D_\perp G(D_\perp). \tag{35}
\]

Ignoring the small term quadratic in \( D_\perp \), we may now predict the change in the SCF energy at little cost from the expression

\[
\Delta E^{RH}_{SCF}(\mu) = E_{SCF}(D_i) - E_{SCF}(\tilde{D}) + 2 \text{Tr} D_\perp F(D_i), \tag{36}
\]

using only the density matrices and Fock/KS matrices of the previous iterations. In particular in the later parts of the iteration sequence, where the space spanned by the densities of the preceding RH iterations is large, an accurate estimate of \( \Delta E^{RH}_{SCF} \) may be obtained from this formula. In the following, we shall see how we may use this prediction to determine the level shift when \( \mu_{min} = 0 \) and \( a(0) > a_{min} \).

To illustrate how \( \Delta E^{RH}_{SCF} \) is used to find the level-shift parameter, consider as an example the determination of the level-shift parameter in the ninth iteration of the rhodium-complex calculation of Sec. III. The plot of the HOMO-LUMO gap in Fig. 2(a) shows that the allowed level-shift interval is \( 0 \leq \mu < \infty \). In Fig. 2(b), we have plotted the overlap \( a(\mu) \) as a function of \( \mu \). Since \( a(0) > a_{min} \), we should, according to the discussion in Sec. II A 2, use \( \mu_{opt} = 0 \) to determine the step. In short, considerations based on the HOMO-LUMO gap and on the overlap with the averaged density matrix indicate that the next density matrix should be determined from the standard, unshifted RH equations.

However, from the nine density matrices of the previous RH iterations, we can use \( \Delta E^{RH}_{SCF}(\mu) \) to predict the change in \( E^{RH}_{SCF}(\mu) \) more accurately than with \( \Delta E^{RH}(\mu) \). Indeed, from Fig. 2(c), we see that \( \Delta E^{RH}_{SCF}(\mu) \) provides a good global representation of \( \Delta E^{RH}_{SCF}(\mu) \), with a minimum close to the minimum of \( \Delta E^{RH}(\mu) \). By contrast, the local model \( \Delta E^{RH}(\mu) \)
gives a minimum at $\mu = 0$. Clearly, $\mu = 0$ should be avoided in the calculation since it would lead to an increase in the SCF energy. Instead, the value of the level-shift parameter that corresponds to the minimum of $\Delta E^P_{\text{SCF}}$ (denoted by $\mu_{\text{opt}}$) is chosen for the calculation of the next density matrix.

This procedure may be summarized as follows. If $\mu_{\text{min}} = 0$ and $\alpha(0) > \alpha_{\text{min}}$, then we calculate the predicted energies $\Delta E^P_{\text{SCF}}(0)$ and $\Delta E^P_{\text{SCF}}(\delta)$ with $\delta > 0$. If $\Delta E^P_{\text{SCF}}(0) < \Delta E^P_{\text{SCF}}(\delta)$, then we use $D(0)$. Otherwise, we estimate the minimum $\mu_{\text{opt}}$ of $\Delta E^P_{\text{SCF}}(\mu)$ by an inexact line search and use the density matrix $D(\mu_{\text{opt}})$ at this minimum.

### B. Density-subspace minimization

#### 1. The DSM energy function

Let us assume that we have carried out $n$ RH iterations and that we have kept all previous density matrices $D_i$ and the corresponding Fock matrices $F_i$. We would now like to construct an optimal density as a linear combination of the densities from these iterations according to Eq. (9),

$$\tilde{D} = \sum_{i=1}^{n} c_i D_i.$$  \hfill (37)

Ideally, this averaged density should also fulfill the conditions Eqs. (3)–(5). The symmetry condition Eq. (3) is trivially satisfied since the averaged density Eq. (37) is a linear combination of symmetric density matrices. The trace condition Eq. (4) is also easily taken care of by imposing the restriction

$$\sum_{i=1}^{n} c_i = 1$$ \hfill (38)

on the expansion coefficients

$$\text{Tr} \tilde{D} S = \sum_{i=1}^{n} c_i \text{Tr} D_i S = \frac{N}{2}.$$ \hfill (39)

By contrast, the idempotency condition Eq. (5) cannot be imposed on the averaged density matrix. However, the idempotency may be significantly improved if, instead of working with $\tilde{D}$, we work with the purified density matrix

$$\tilde{D} = 3 D S \tilde{D} - 2 D S D S \tilde{D},$$ \hfill (40)

as proposed by Nunes and Vanderbilt.\textsuperscript{7} The electronic energy may be expressed in terms of the purified average density matrix as

$$E(\tilde{D}) = 2 \text{Tr} h D + \text{Tr} D G(\tilde{D}).$$ \hfill (41)

We note that the purified density is correct to first order in the expansion coefficients $c_i$ and that $E(\tilde{D})$ thus contains errors through second order in $c_i$. To determine the best average density matrix Eq. (37), we shall minimize Eq. (41) with respect to the expansion coefficients $c_i$ subject to the condition Eq. (38).

One problem we encounter when minimizing Eq. (41) is that new Fock matrices $F_i(\tilde{D})$ need to be evaluated. To avoid this problem, we shall use an approximate form of Eq. (41). Since the purified density matrix $\tilde{D}$ is close to the original density matrix $\tilde{D}$, we can write it as

$$\tilde{D} = \tilde{D} + \Delta,$$ \hfill (42)

where $\Delta$ is the correction term. Inserting Eq. (42) into Eq. (41), we obtain

$$E = 2 \text{Tr} h D + \text{Tr} D G(\tilde{D}) + 2 \text{Tr} h \Delta$$

$$+ 2 \text{Tr} \Delta G(\tilde{D}) + \text{Tr} \Delta G(\Delta).$$ \hfill (43)

Since $\Delta$ is small, we may ignore the term quadratic in $\Delta$ and arrive at the density-subspace minimization (DSM) energy function

$$E^{\text{DSM}}(c) = 2 \text{Tr} h D + \text{Tr} D G(\tilde{D}) + 2 \text{Tr} h \Delta + 2 \text{Tr} \Delta G(\tilde{D})$$

$$= E(\tilde{D}) + 2 \text{Tr} F(\tilde{D}) (\tilde{D} - D).$$ \hfill (44)

Since $\Delta$ is first order in the expansion coefficients $c_i$, the DSM energy differs from the true energy to second and
higher orders in \( c_i \). The first contribution to the DSM energy function may for example be evaluated using the energy expression of the EDIIS algorithm,\(^5\)

\[
E(\mathbf{D}) = \sum_i c_i E_{\text{SCF}}(\mathbf{D}_i) - \frac{1}{2} \sum_{ij} c_i c_j \text{Tr}(\mathbf{F}_i - \mathbf{F}_j)(\mathbf{D}_i - \mathbf{D}_j).
\]  
(45)

Using Eq. (40), we find that the second contribution may be evaluated as

\[
2 \text{Tr} \mathbf{F}(\mathbf{D})(\mathbf{D} - \mathbf{D}) = -2 \sum_{ij} c_i c_j \text{Tr} \mathbf{F}_i \mathbf{D}_j + 6 \sum_{ijk} c_i c_j c_k \text{Tr} \mathbf{F}_i \mathbf{D}_j \mathbf{S}_k - 4 \sum_{ijkl} c_i c_j c_k c_l \text{Tr} \mathbf{F}_i \mathbf{D}_j \mathbf{S}_k \mathbf{S}_l.
\]  
(46)

All contributions to the DSM energy function are therefore easily calculated from the previous density and Fock/KS matrices.

2. The trust-region DSM minimization

We minimize the DSM energy functional by the trust-region method.\(^12\) We thus consider the second-order Taylor expansion of the DSM energy in Eq. (44) about \( c_0 \). Introducing the undetermined multipliers \( \lambda \) and \( \mu \), we restrict the step to have the length

\[
\| \Delta c \|_2 = \| v \|_2 = \| \lambda \|_2.
\]

We also introduce a trust region of radius \( h \) for \( E_{\text{DSM}}(\Delta c) \) and require that steps are always taken inside or to the boundary of this region. To determine a step to the boundary, we restrict the step to have the length \( h \) in the \( S \) metric norm of Eq. (34),

\[
\| \Delta c(\mu) \|_2^2 = \sum_{ij} \Delta c_i M_{ij} \Delta c_j = h^2.
\]  
(51)

Introducing the undetermined multipliers \( \lambda \) and \( \mu \) for the trace and step-size constraints, we arrive at the following Lagrangian for minimization on the boundary of the trust region:

\[
L(\Delta c, \mu, \lambda) = E_0 + \Delta c^T \mathbf{g} + \frac{1}{2} \Delta c^T \mathbf{H} \Delta c - \lambda \Delta c^T \mathbf{1}
\]  
(52)

where \( \mathbf{1} \) is a column vector with elements equal to 1. Differentiating this Lagrangian and setting the derivatives equal to zero, we obtain the equations

\[
\frac{\partial L}{\partial \Delta c} = \mathbf{g} + \mathbf{H} \Delta c - \mu \mathbf{M} \Delta c - \lambda \mathbf{1} = 0,
\]  
(53)

\[
\frac{\partial L}{\partial \lambda} = -\Delta c^T \mathbf{1} = 0,
\]  
(54)

\[
\frac{\partial L}{\partial \mu} = -\frac{1}{2}(\Delta c^T \mathbf{M} \Delta c - h^2) = 0.
\]  
(55)

The optimization of the Lagrangian thus corresponds to the solution of the following set of linear equations:

\[
\begin{pmatrix}
\mathbf{H} & -\mathbf{M} & -\mathbf{1} \\
-\mathbf{1}^T & 0 & 0
\end{pmatrix}
\begin{pmatrix}
\Delta c \\
\lambda
\end{pmatrix}
= -
\begin{pmatrix}
\mathbf{g} \\
0
\end{pmatrix}.
\]  
(56)

where the multiplier \( \mu \) is iteratively adjusted until the step is to the boundary of the trust region Eq. (55). The step-length restriction may be lifted by setting \( \mu = 0 \), as needed for steps inside the trust region.

To understand the behavior of the step-length function, we consider first the generalized eigenvalue problem

\[
\begin{pmatrix}
\mathbf{H} & -\mathbf{M} & -\mathbf{1} \\
-\mathbf{1}^T & 0 & 0
\end{pmatrix}
\begin{pmatrix}
\mathbf{v} \\
\lambda
\end{pmatrix}
= \omega
\begin{pmatrix}
\mathbf{0} \\
\mathbf{0}^T & \mathbf{e}
\end{pmatrix}
\begin{pmatrix}
\mathbf{v}
\end{pmatrix},
\]  
(57)

where \( \mathbf{0} \) is a column vector with zero elements, \( \mathbf{e} \) is a small positive constant, and the eigenvector is normalized such that

\[
\mathbf{v}^T \mathbf{v} + \mathbf{e}^2 = 1.
\]  
(58)

We first note that, for a finite \( \mathbf{v} \), \( \mathbf{v} \neq 0 \). Next, carrying out block multiplications in Eq. (57), we obtain

\[
\mathbf{H} \mathbf{v} - \mathbf{1} \mathbf{v} = \omega \mathbf{M} \mathbf{v},
\]  
(59)

\[
-\mathbf{1}^T \mathbf{v} = \omega \mathbf{e} \mathbf{v},
\]  
(60)

which upon elimination of \( \mathbf{v} \) from the first equation yields the relation

\[
\omega \mathbf{e} \mathbf{H} \mathbf{v} + (\mathbf{1}^T \mathbf{v}) \mathbf{1} = \omega^2 \mathbf{e} \mathbf{M} \mathbf{v}.
\]  
(61)

Since \( (\mathbf{1}^T \mathbf{v}) \mathbf{1} \) is finite, we conclude that, as \( \mathbf{e} \) tends to zero, the eigenvalue \( \omega \) tends to either plus or minus infinity \( \pm \omega \mathbf{e} \mathbf{1} \). Next, substituting these values of \( \omega \) into Eq. (60), we find that \( \mathbf{v} \) tends to the zero vector with elements proportional to \( \mathbf{e} \mathbf{1} \) and that \( \mathbf{v} \), because of the normalization Eq. (58), tends to \( \mathbf{1} \). In short, the eigenvalue problem Eq. (57) with \( \mathbf{e} = 0 \) has two eigenvalues \( \pm \omega \mathbf{e} \mathbf{1} \), whose eigenvectors have zero elements except for the last element, which is equal to \( \mathbf{1} \). Finally, invoking the Hylleraas–Undheim interchange theorem,\(^10\)^11 we conclude that the remaining \( n - 1 \) finite eigenvalues of Eq. (57) bisects the \( n \) eigenvalues of the reduced eigenvalue problem

\[
\mathbf{H} \mathbf{v} = \omega \mathbf{e} \mathbf{M} \mathbf{v}.
\]  
(62)
Let us now consider the step length $\|\Delta \mathbf{c}(\mu)\|_S$ as a function of $\mu$. In the diagonal representation of the augmented matrix in the linear equations Eq. (57), we may write these equations in the following uncoupled form:

$$
(h_i - \mu m_i)\sigma_i = -\gamma_i, \quad i = 1, 2, 3, \ldots, n+1.
$$

(63)

Here, the $h_i$ and $m_i$ are the diagonal elements of the Hessian and metric matrices, respectively, of the generalized eigenvalue problem Eq. (57), whereas the $\sigma_i$ and $\gamma_i$, respectively, are the corresponding elements of the solution and gradient vectors of Eq. (56). Since the last element of the gradient vector in Eq. (56) is zero, the gradient vector has no contributions from the eigenvectors with infinite eigenvalues

$$
\gamma_1 = \gamma_{n+1} = 0, \quad \omega_1 = -\omega_{n+1} = -\infty
$$

assuming that the eigenvalues are sorted in increasing order $\omega_1 < \omega_2 < \cdots < \omega_{n+1}$. In the diagonal representation, therefore, we may write the step norm in the form

$$
\|\Delta \mathbf{c}(\mu)\|_S = \sqrt{\sum_{i=2}^{n} \frac{m_i \gamma_i^2}{(h_i - \mu m_i)^2}}.
$$

(65)

From this expression, we note that the step function consists of $n$ branches separated by $n - 1$ asymptotes at the finite eigenvalues $\omega_i$. Moreover, it increases monotonically from zero to infinity as $\mu$ increases from minus infinity and approaches the lowest finite eigenvalue $\omega_2$. Therefore, there is always one and only one $-\infty < \mu < \omega_2$ that gives rise to a step of length $h$. As shown by Fletcher, this value of $\mu$ corresponds to the global minimum on the boundary of the trust region.

In practice, we cannot easily determine the eigenvalues $\omega_i$ of the augmented eigenvalue problem Eq. (57). Instead, we determine the eigenvalues $\nu_i$ of the reduced problem Eq. (62) and restrict our search of $\mu$ to the smaller monotonic interval $-\infty < \mu < \nu_1$. Since $\nu_1 < \omega_2$, it is possible that no solution exists in this reduced interval. Mostly, however, this restriction is mild since the two eigenvalues are usually close. If no solution is found, we choose instead the slightly shorter step obtained with $\mu = \nu_1$.

To illustrate how the level-shift parameter $\mu$ in Eq. (56) is determined, we consider the first [Fig. 3(a)] and third [Fig. 3(b)] DSM step in the eighth iteration of the rhodium-complex calculation in Sec. III. We have plotted the step-length function $\|\Delta \mathbf{c}(\mu)\|_S$ as a function of $\mu$. The plots consist of a series of branches between asymptotes where $\mu$ makes the matrix on the left-hand side of Eq. (56) singular. The lowest eigenvalue $\nu_1$ is marked with a vertical dashed line in Figs. 3(a) and 3(b). For minimization, the level-shift parameter is chosen in the interval $-\infty < \mu < \min(\nu_1, 0)$, where $\nu_1$ is the lowest eigenvalue of Eq. (62). The proper value is found where the step-length function crosses the line representing the trust radius $h$, as marked with a cross in Fig. 3(a). If the step that minimizes $E_{23}$ is inside the trust region, $\mu = 0$ is chosen as marked with a cross in Fig. 3(b). The trust region is updated during the iterative procedure.

FIG. 3. The step-length function $\|\Delta \mathbf{c}(\mu)\|_S$ is plotted as a function of $\mu$ for the first (a) and third (b) DSM step in the eighth iteration of the rhodium calculation described in Sec. III. The trust radius $h$ is represented by a horizontal line. The proper $\mu$ value is marked with a cross.

3. Global optimization of the DSM function

The optimization of the $E_{\text{DSM}}$ energy is carried out in the usual manner, requiring several trust-region steps, each of which involves the construction of the gradient $\mathbf{g}$ and the Hessian $\mathbf{H}$, and the solution of the modified level-shifted Newton equations Eq. (56). After $p$ iterations, the density is calculated from the coefficients

$$
\mathbf{c}_p = \mathbf{c}^{(0)} + \sum_{i=1}^{p} \Delta \mathbf{c}_i.
$$

(66)

However, since $E_{\text{DSM}}$ itself is a rather crude model of the true energy function $E_{\text{SCF}}$, it resembles $E_{\text{SCF}}$ only in a small region about the initial point $\mathbf{c}^{(0)}$. The DSM iterations are therefore terminated when the total step length $\|\mathbf{c}_p - \mathbf{c}^{(0)}\|$ exceeds some preset value $k$. If a minimum of $E_{\text{DSM}}$ is found inside the trust region $\|\mathbf{c}_p - \mathbf{c}^{(0)}\| < k$, then the step to the minimum is taken and the iterations are terminated. This is often the case.

Occasionally, the iterations start where the lowest eigenvalue of the Hessian in Eq. (62) is negative. In the course of the iterations, the Hessian can become positive definite and a minimum is reached. In a few cases, however, a negative Hessian eigenvalue may persist, changing little from iteration to iteration. In our experience, a step along the eigenvector corresponding to the negative eigenvalue cannot be trusted. This direction is therefore projected out from the step and the DSM function is minimized in the orthogonal subspace.

As an illustration, consider the first DSM step of the tenth SCF iteration of the rhodium-complex calculation in Sec. III. In Fig. 4, we have, for comparison, plotted the step-length functions with the negative component kept and projected out. The level shifts resulting from the two situations
are marked with crosses in Fig. 4. The level shift used in the
DSM optimization is, in this particular case, $\mu = 0$.

When the trust-region minimization is terminated, a new
RH iteration is initiated by constructing a new density and
associated Fock matrix

$$\tilde{D} = \sum_{i=1}^{n} c_i D_i, \quad \tilde{F} = \sum_{i=1}^{n} c_i F(D_i),$$

(67)

where we have used the fact that the Fock matrix is linear in
the density. By construction $E_{\text{DSM}}^{(c)}$ is lowered at each it-
eration of the trust-region minimization. The total energy
lowering at the $p$th iteration is given by

$$\Delta E_{\text{DSM}} = E_{\text{DSM}}^{(c_p)} - E_{\text{DSM}}^{(c^{(0)})}.$$  

(68)

Since $E_{\text{DSM}}$ is a local model to the true energy $E_{\text{SCF}}$, the
lowering of $E_{\text{DSM}}$ will also lead to a lowering of $E_{\text{SCF}}$
provided the total step is sufficiently short to be in the local
region.

4. Relationship to the DIIS method

The optimal density has previously been determined us-
ing the DIIS scheme of Pulay. In the DIIS method, the
improved density matrix is obtained as a linear combination of
the previous density matrices where the expansion coeffi-
cients are determined by minimizing the norm of the error
vector, using the gradients of the previous iterations as error
vectors. To highlight the difference between TRDSM and
DIIS, we give below an alternative derivation of the DIIS
algorithm.

In an SCF calculation, the electronic gradient with the
averaged density matrix $\tilde{D}$ in Eq. (37) may be expressed in
the form,

$$g(\tilde{D}) = 4(\tilde{D}SF(\tilde{D}) - F(\tilde{D})SD\tilde{D}).$$

(69)

To determine the best linear combination of densities $D_i$, we
minimize the norm of the squared gradient

$$\|g(\tilde{D})\|^2 = 16 \text{Tr}[\tilde{D}SF(\tilde{D}) - F(\til{D})SD\tilde{D}].$$

(70)

Inserting the expansion Eq. (37), we obtain a quartic poly-
nomial in $c_j$,

FIG. 4. The step-length function $|\Delta c(\mu)|_2$ is plotted as a function of $\mu$ with
the direction corresponding to the negative Hessian eigenvalue kept (—) and
projected out (— —), respectively. The $\mu$ values resulting from the two
situations are marked with crosses.

FIG. 5. The convergence of calculations on the rhodium complex using
AhlrichsVDZ basis (Ref. 16) combined with STO-3G for Rh. The error in
the total energy is given for the TRSCF, the standard DIIS, and the QRHF
method as a function of the iteration number. Furthermore results are given
where DIIS is applied after nine TRSCF iterations.

$$\|g(\tilde{D})\|^2 = 16 \text{Tr}\left[\sum_{i} c_i g(D_i) + \sum_{i,j} c_i c_j (D_iSF(D_j - D_i))
- F(D_j - D_i)SD_{i,j}\right]^2.$$

(71)

To simplify this expression, we neglect all cubic and quartic
terms

$$\|g(\til{D})\|^2_{\text{app}} = \sum_{i,j} c_i c_j g(D_i)g(D_j).$$

(72)

Optimization of Eq. (72) subject to the constraint Eq. (38)
gives the DIIS expression of the expansion coefficients in
Eq. (37).

III. APPLICATIONS

In this section, we examine the convergence characteris-
tics of the TRSCF algorithm. First, we consider a rhodium-
complex optimization as an example of a difficult case; next,
as a simpler case, we consider a calculation on H$_2$O with the
OH bond lengths stretched to double length. For comparison,
we also give the convergence characteristics of the DIIS
algorithm and the quadratically convergent restricted step
Hartree–Fock (QRHF) method. All calculations are car-
ried out using a local version of the DALTON program
package.
In the RH diagonalization step, a global optimization is carried out for \( E^{RH} \), a level-shifted Fock eigenvalue equation is solved where the level-shift parameter effectively introduces a shift in the Hessian of \( E^{RH} \) [Eq. (25)]. The similarity of the Hessians of \( E_{SCF} \) and \( E^{RH} \) makes the directions of the steps taken by the QRHF and RH methods very similar for sufficiently large level shifts, the essential difference being the global character of the RH steps and the local character of the QRHF steps. It is this local character of the QRHF steps that prevents the QRHF method from being efficient for systems as difficult as the rhodium complex.

Let us now consider the individual TRSCF iterations as listed in Table II. The optimization begins with orbitals that diagonalize the one-electron Hamiltonian, giving a start energy of \(-5.466.530.208.964.75 \, E_n\). In Table II, the SCF energy lowering \( \Delta E_{SCF} \) is divided into two contributions, one from the RH step and one from the DSM step. Recalling from Eq. (24) that \( D(\infty)_n \) is the purified \( D_n \),

\[
\Delta E_{SCF_{n+1}}^{DSM} = E_{SCF}[D(\infty)_n] - E_{SCF}[D_n] \tag{73}
\]

becomes a realistic measure of the energy change in the DSM part of the iteration. Similarly,

\[
\Delta E_{SCF_{n+1}}^{RH} = E_{SCF}[D_{n+1}] - E_{SCF}[D(\infty)_n] \tag{74}
\]

becomes a realistic measure of the change in the RH part. Clearly, the sum of Eqs. (73) and (74) is equal to the total change \( \Delta E_{SCF} \). These exact energy changes should be compared with the energy changes in the local models \( \Delta E^{RH} \) and \( \Delta E^{DSM} \) given in Eqs. (27) and (68), respectively, also listed in the table. Note that, to obtain \( E_{SCF}[D(\infty)_n] \), we must carry out an additional energy calculation, which is here done only for the purpose of this analysis.

For the DSM method, we have also indicated in Table II how the trust-region optimization was terminated (exit\( \text{DSM} \)): (M) indicates that a minimum was determined in the full space; (PM) indicates that a minimum was obtained in the reduced space with the direction corresponding to the negative Hessian eigenvalue projected out; and (L) indicates that the iterations were terminated because the maximum step length \( k \) was reached. For the RH steps, we have also listed the level-shift parameter \( \mu_{opt} \) and the corresponding overlap \( a(\mu_{opt}) \) of Eq. (22).

The TRSCF iterations converge linearly, with a reduction in the error of about a factor 2–4 at each iteration. Moreover, the energy lowerings of the local models \( \Delta E^{RH} \) and \( \Delta E^{DSM} \) are in good agreement with the actual SCF energy changes, in the local as well as in the global part of the optimization. Both the predicted and the actual energy changes are negative in all iterations. In the global region, \( \Delta E_{SCF}^{RH} \) is usually significantly larger than \( \Delta E_{SCF}^{DSM} \), whereas, in the local region, they have similar sizes.

Except for three iterations in the global part of the SCF optimization, the DSM trust-region method finds a minimum within the step-length limit \( k \). In the intermediate region, we encounter components of the step vector that cannot be trusted and have been projected out as described in Sec. II B 3. The DSM iterations then reach a minimum in the orthogonal subspace.
In the beginning of the SCF optimization, large level shifts are applied in the RH diagonalization to ensure a continuous development of the MOs. Thus, in the first few iterations, the overlap constant \( \alpha (\mu_{\text{opt}}) \) is significantly larger than the minimum accepted overlap of 0.975. However, the level-shift parameter decreases during the subsequent SCF iterations, until in the local region, no level shift is required and conventional RH iterations are carried out. To summarize, the TRSCF method gives a monotonic and significant energy lowering both in the RH and in the DSM part of the optimization.

### B. The water calculation

To demonstrate the performance of the TRSCF method in a simpler case, we consider optimizations of \( \text{H}_2\text{O} \) with the OH bonds stretched to twice the equilibrium value (195.10 pm). In Figs. 6(a) and 6(b), we have plotted the errors in the energy during TRSCF, DIIS, and QRHF optimizations in the cc-pVDZ basis. In Fig. 6(a), the initial guess of the orbitals are the Hückel orbitals as implemented in the DALTON program. With these initial orbitals, the TRSCF and DIIS methods converge in a very similar manner to within a threshold of \( 10^{-10} \) in ten iterations. In this case, therefore, gradient information is sufficient for convergence. Although the QRHF method outperforms the TRSCF and DIIS methods in terms of iterations, this is of no practical value since, in each QRHF step, about the same number of new Fock matrices are needed to solve the Newton equations as is required to find the optimized Hartree–Fock wave function with the TRSCF and DIIS methods.

#### TABLE II. Convergence details for the TRSCF calculation on the rhodium complex using AhlrichsVDZ basis combined with STO-3G on Rh. Energies given in atomic units.

<table>
<thead>
<tr>
<th>Iter.</th>
<th>( \Delta E_{\text{SCF}} )</th>
<th>( \Delta E_{\text{DM}} )</th>
<th>( \Delta E_{\text{RH}} )</th>
<th>( \rho_{\text{RH}} )</th>
<th>( \alpha_{\mu_{\text{opt}}} )</th>
<th>ExiDSM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-18.94647615033</td>
<td>-0.0000000009</td>
<td>-0.0000000000</td>
<td>-19.21320649447</td>
<td>17.47</td>
<td>0.99382</td>
</tr>
<tr>
<td>2</td>
<td>-45.4588825211</td>
<td>-8.95768890498</td>
<td>-7.1039975657</td>
<td>-36.50089934714</td>
<td>14.44</td>
<td>0.98630</td>
</tr>
<tr>
<td>3</td>
<td>-59.8103738071</td>
<td>-12.9365160037</td>
<td>-8.85502694483</td>
<td>-46.8738570361</td>
<td>11.68</td>
<td>0.97940</td>
</tr>
<tr>
<td>4</td>
<td>-63.3446220663</td>
<td>-24.25263285599</td>
<td>-21.6371685856</td>
<td>-39.0922935064</td>
<td>7.28</td>
<td>0.97288</td>
</tr>
<tr>
<td>5</td>
<td>-30.2287541345</td>
<td>-12.8783382045</td>
<td>-12.23686358427</td>
<td>-17.4109279300</td>
<td>2.63</td>
<td>0.97384</td>
</tr>
<tr>
<td>6</td>
<td>-11.5661105704</td>
<td>-5.6494046510</td>
<td>-4.74940263979</td>
<td>-5.9115641194</td>
<td>0.90</td>
<td>0.97555</td>
</tr>
<tr>
<td>7</td>
<td>-4.61334096599</td>
<td>-1.0022093646</td>
<td>-1.51155032176</td>
<td>-1.7683303184</td>
<td>0.24</td>
<td>0.97792</td>
</tr>
<tr>
<td>8</td>
<td>-2.1620415323</td>
<td>-0.44637212140</td>
<td>-0.4484960108</td>
<td>-1.71633203184</td>
<td>0.07</td>
<td>0.97876</td>
</tr>
<tr>
<td>9</td>
<td>-0.60805181167</td>
<td>-0.29078332276</td>
<td>-0.21298647367</td>
<td>-0.3172684889</td>
<td>1.30</td>
<td>0.99082</td>
</tr>
<tr>
<td>10</td>
<td>-0.1667264229</td>
<td>-0.00294157325</td>
<td>-0.00194422453</td>
<td>-0.1659813594</td>
<td>0.70</td>
<td>0.99934</td>
</tr>
<tr>
<td>11</td>
<td>-0.05893026674</td>
<td>-0.00782290321</td>
<td>-0.00662821837</td>
<td>-0.0511071237</td>
<td>0.00</td>
<td>0.99955</td>
</tr>
<tr>
<td>12</td>
<td>-0.01821537974</td>
<td>-0.00935849099</td>
<td>-0.00823957093</td>
<td>-0.0085688885</td>
<td>0.00</td>
<td>0.99989</td>
</tr>
<tr>
<td>13</td>
<td>-0.00829012952</td>
<td>-0.00417698583</td>
<td>-0.0032848541</td>
<td>-0.0041317118</td>
<td>0.00</td>
<td>0.99995</td>
</tr>
<tr>
<td>14</td>
<td>-0.00336722651</td>
<td>-0.0024626574</td>
<td>-0.00222734467</td>
<td>-0.0009146077</td>
<td>0.00</td>
<td>0.99998</td>
</tr>
<tr>
<td>15</td>
<td>-0.00144195016</td>
<td>-0.00106346997</td>
<td>-0.0009168267</td>
<td>-0.00037834519</td>
<td>0.00</td>
<td>1.00000</td>
</tr>
<tr>
<td>16</td>
<td>-0.00049317801</td>
<td>-0.00040627140</td>
<td>-0.00039284830</td>
<td>-0.00008090661</td>
<td>0.00</td>
<td>1.00000</td>
</tr>
<tr>
<td>17</td>
<td>-0.00005633666</td>
<td>-0.00003203696</td>
<td>-0.0000263768</td>
<td>-0.0000243097</td>
<td>0.00</td>
<td>1.00000</td>
</tr>
<tr>
<td>18</td>
<td>-0.0001495199</td>
<td>-0.00000969523</td>
<td>-0.00000917530</td>
<td>-0.0000050495</td>
<td>0.00</td>
<td>1.00000</td>
</tr>
<tr>
<td>19</td>
<td>-0.00000549749</td>
<td>-0.00000312992</td>
<td>-0.00000279715</td>
<td>-0.00000236757</td>
<td>0.00</td>
<td>1.00000</td>
</tr>
<tr>
<td>20</td>
<td>-0.00000196603</td>
<td>-0.00000126150</td>
<td>-0.00000125156</td>
<td>-0.00000070454</td>
<td>0.00</td>
<td>1.00000</td>
</tr>
<tr>
<td>21</td>
<td>-0.00000038264</td>
<td>-0.00000022641</td>
<td>-0.00000015423</td>
<td>-0.00000015423</td>
<td>0.00</td>
<td>1.00000</td>
</tr>
<tr>
<td>22</td>
<td>-0.00000008729</td>
<td>-0.00000004496</td>
<td>-0.00000004496</td>
<td>-0.00000004496</td>
<td>0.00</td>
<td>1.00000</td>
</tr>
<tr>
<td>23</td>
<td>-0.00000002788</td>
<td>-0.00000001171</td>
<td>-0.00000001617</td>
<td>-0.00000001617</td>
<td>0.00</td>
<td>1.00000</td>
</tr>
<tr>
<td>24</td>
<td>-0.00000001286</td>
<td>-0.00000000813</td>
<td>-0.00000000813</td>
<td>-0.00000000813</td>
<td>0.00</td>
<td>1.00000</td>
</tr>
<tr>
<td>25</td>
<td>-0.00000000924</td>
<td>-0.00000000131</td>
<td>-0.00000000131</td>
<td>-0.00000000131</td>
<td>0.00</td>
<td>1.00000</td>
</tr>
<tr>
<td>26</td>
<td>-0.00000000119</td>
<td>-0.00000000073</td>
<td>-0.00000000073</td>
<td>-0.00000000073</td>
<td>0.00</td>
<td>1.00000</td>
</tr>
<tr>
<td>27</td>
<td>-0.00000000035</td>
<td>-0.00000000019</td>
<td>-0.00000000019</td>
<td>-0.00000000019</td>
<td>0.00</td>
<td>1.00000</td>
</tr>
</tbody>
</table>

![FIG. 6. The convergence of calculations on water with stretched bonds using the cc-pVDZ basis and (a) a Hückel start guess and (b) a one-electron Hamiltonian start guess. The error in the total energy is given for the TRSCF, the standard DIIS and the QRHF method as a function of the iteration number.](image-url)
In Fig. 6(b), we have plotted the error of the energy in H₂O optimizations starting with the orbitals that diagonalize the one-electron Hamiltonian. In this case, convergence to 10⁻¹⁰ is reached in 13 iterations with the TRSCF method and in 18 iterations with the DIIS method. The main reason for the better performance of the TRSCF algorithm is that, in the global region, it gives a significant energy lowering in each step, whereas the DIIS algorithm shows a much less systematic behavior.

IV. CONCLUSION

A conventional SCF optimization consists of a sequence of iterations, each of which begins with a Roothaan–Hall (RH) diagonalization step, where a Fock/KS matrix is diagonalized to obtain an improved density matrix, followed by an averaging step, where the optimal density matrix is determined in the subspace of the density matrices of the previous RH diagonalization steps. In this paper, we have introduced a trust-region SCF (TRSCF) algorithm, where improvements have been made to both the diagonalization and the averaging steps. In both steps, local energy model functions are constructed which have the same gradient as the true energy function $E_{SCF}$ but approximate Hessians. Recognizing the locality of these energy functions, trust regions are introduced as regions where they represent a good approximation to $E_{SCF}$ and only steps inside these trust regions are allowed.

For the density-subspace minimization step, an energy function is constructed and minimized with respect to the coefficients of the linear combination of the previous density matrices. Its functional form is based on a purified averaged density matrix that is idempotent to first order. The advantages of this model compared to EDIIS is the built-in density purification, which helps to avoid problems arising from non-idempotency. In addition, information about the Hessian is extracted and used, leading to a monotonic and stable convergence.

The RH diagonalization step corresponds to a minimization of an energy function $E^{RH}$ that represents the sum of the orbital energies of the occupied MOs. Since this very simple energy function is a local model function for $E_{SCF}$, large steps cannot be trusted. To generate steps to the boundary of the trust region, level-shifted RH equations are solved where the level shifts are determined in a systematic and general manner, leading to a decrease in the model energy at each iteration. If sufficiently small steps are taken, a similar decrease is obtained in the SCF energy.

In the TRSCF algorithm a few diagonalizations are required in each SCF iteration to obtain solutions for the level-shifted RH equations in order to determine the optimal density matrix. The number of diagonalizations may be reduced in the local SCF region solving RH equations with zero level shift with little consequence for the convergence. In the local SCF region one may also safely use the DIIS algorithm if desired.

The advantages of the TRSCF algorithm are demonstrated by calculations on a rhodium complex and on a water molecule with stretched bonds. In the rhodium-complex optimization, the TRSCF algorithm converges monotonically and fast, with a significant decrease in the energy in both the RH part and DSM part at each iteration. By contrast, convergence is not obtained with the DIIS method for this complex. For the simpler water molecule, the TRSCF and DIIS methods behave in a more similar manner, the TRSCF method converging slightly faster than the DIIS method when the initial orbitals are obtained by diagonalizing the one-electron Hamiltonian. With the Hückel guess, the water convergence is essentially obtained in the same number of steps for the TRSCF and DIIS methods. In short, it appears that the TRSCF algorithm, and its use of local energy model functions to obtain significant reductions in $E_{SCF}$ in each iteration, constitutes a significant step towards a black-box optimization of SCF wave functions.

ACKNOWLEDGMENTS

This work has been supported by the Danish Natural Research Council (Grant No. 21-02-0467) and the Carlsberg-fondet. We also acknowledge support from the Danish Center for Scientific Computing (DCSC). D.Y. acknowledges support from the Robert A. Welch Foundation, Grant No. A-770.