

A practicalist guide to Self Consistent Field

Georgia Institute of Technology

Daniel G. A. Smith

I. NOTATION

There exist many ways to label and describe quantum chemistry theories. In this section the quantities and indices used will be defined and outlined. Generalized multidimensional arrays will be referred to as "tensors" and two dimensional arrays will be referred to as "matrices".

A. Einstein summation notation

The generalized Einstein summation notation will be employed where repeated indices on the right hand side (RHS) that do not appear on left hand side (LHS) are assumed to be summed over. A matrix-matrix multiplication example:

$$\text{Matrix convention : } \mathbf{C} = \mathbf{AB} \quad (1)$$

$$\text{Explicit Sum : } C_{ij} = \sum_k A_{ik} B_{kj} \quad (2)$$

$$\text{Einstein summation : } C_{ij} = A_{ik} B_{kj} \quad (3)$$

an outer product example

$$C_{ikj} = A_{ik} B_{kj} \quad (4)$$

B. Indices

The following notation will be adhered to throughout this document. Occasionally indices will be redefined, but it should be clear in the context what the alternate meaning is.

- | | |
|---|-----------------------------|
| • General atomic orbitals | $\mu, \nu, \lambda, \sigma$ |
| • Inactive (doubly occupied) molecular orbitals | i, j, k, l |
| • Unoccupied (virtual) molecular orbitals | a, b, c, d |
| • General index molecular orbitals | p, q, r, s, m, n, o |
| • Auxiliary basis | P, Q, R, S |

C. Tensors

Common matrices and tensors utilized in this work.

| | |
|---|--|
| • Core Hamiltonian matrix | $H_{\mu\nu}$ |
| • Molecular orbital coefficient matrix | $C_{p\mu}$ |
| • Overlap matrix | $S_{\mu\nu}$ |
| • SCF one-particle density matrix | D_{pq} |
| • Atomic-orbital two-electron integral | $g_{\mu\nu\lambda\sigma} = \int \mu(\mathbf{r}_1)\nu(\mathbf{r}_1)\frac{1}{r_{12}}\lambda(\mathbf{r}_2)\sigma(\mathbf{r}_2)d^3\mathbf{r}_1d^3\mathbf{r}_2$ |
| • Molecular-orbital two-electron integral | $g_{pqrs} = C_{p\mu}C_{q\nu}g_{\mu\nu\lambda\sigma}C_{r\lambda}C_{s\sigma}$ |
| • Density-fitted tensor | $(pq A)$ |

II. SELF CONSISTENT FIELD

Here we will focus on a variant of Self Consistent Field (SCF) known as restricted Hartree-Fock (RHF) where the α and β orbitals are equal denoting a closed-shell molecule with all paired electrons. To begin, let us start with the AO Fock matrix, which is represented as^{1,2}

$$F_{\mu\nu} = H_{\mu\nu} + 2g_{\mu\nu\lambda\sigma}D_{\lambda\sigma} - g_{\mu\lambda\nu\sigma}D_{\lambda\sigma} \quad (5)$$

and the one-particle RHF density matrix D is computed from the orbitals C (assumed real).

$$D_{\lambda\sigma} = C_{i\sigma}C_{i\lambda} \quad (6)$$

It should be noted that if D is converted to the MO basis, we have a diagonal matrix of ones up to the number of occupied indices and zero thereafter

$$D_{pq} = \begin{pmatrix} 1_{ii} & 0 \\ 0 & 0 \end{pmatrix} \quad (7)$$

The total RHF energy can be expressed as a sum of the electronic and Born-Oppenheimer (BO) nuclear energies

$$E^{RHF} = E_{electronic}^{RHF} + E_{nuclear}^{BO} \quad (8)$$

$$E_{electronic}^{RHF} = (F_{\mu\nu} + H_{\mu\nu})D_{\mu\nu} \quad (9)$$

$$E_{nuclear}^{BO} = \sum_{i>j} \frac{Z_i Z_j}{r_{ij}} \quad (10)$$

where Z_i is the nuclear charge of atom i .

Examining these equations, it is clear that the most computationally demanding portions are the convolution of the density matrix with the two-electron integral tensor. Therefore, it is often convenient to define the following quantities

$$J[D_{\lambda\sigma}]_{\mu\nu} = g_{\mu\nu\lambda\sigma}D_{\lambda\sigma} \quad (11)$$

$$K[D_{\lambda\sigma}]_{\mu\nu} = g_{\mu\lambda\nu\sigma}D_{\lambda\sigma} \quad (12)$$

The resulting J and K matrices are often called the Coulomb and Exchange matrices, respectively. Often quantum chemistry programs have very efficient routines to compute these equations, thus utilizing these routines is of the utmost importance.

A. Roothaan Equations

The SCF equations are often solved through the Roothaan equations², shown in matrix formalism

$$\mathbf{FC} = \mathbf{SC}\epsilon \quad (13)$$

This is a pseudo-eigenvalue equation. At every iteration we need to solve for the coefficients \mathbf{C} that diagonalize the Fock matrix. An ideal world would produce AO orbitals that are orthogonal, unfortunately, as demonstrated by non-diagonal overlap matrices, this is not the case. To overcome this problem, a orthonormalized Fock matrix is diagonalized instead, shown in matrix formalism

$$\mathbf{A} = \mathbf{S}^{-1/2} \quad (14)$$

$$\mathbf{F}' = \mathbf{A}^T \mathbf{F} \mathbf{A} \quad (15)$$

$$\mathbf{F}'\mathbf{C}' = \mathbf{C}'\epsilon \quad (16)$$

$$\mathbf{C} = \mathbf{A}\mathbf{C}' \quad (17)$$

At every iteration we construct our Fock matrix from the previous orbitals (\mathbf{C}_{n-1}) and compute new orbitals (\mathbf{C}_n) until convergence is reached.

For the rest of this section, the "physicist's water molecule" (O-H = 1.1Å, \angle HOH = 104°) will be utilized in the cc-pVDZ basis set for illustrating convergence patterns. A pure iterative diagonalization approach is shown in Fig. 1. As can be seen, the convergence for this simple molecule is quite slow and the next sections will detail convergence acceleration.

FIG. 1. A RHF computation of "physicist's water molecule" starting with a core Hamiltonian guess in the cc-pVDZ basis set. dE is the energy difference between iterations and dRMS is the root mean square of the orbital gradient.

| | | | |
|---------------|---------------------------------|-----------------|------------------|
| RHF Iteration | 1: Energy = -68.98003273414295 | dE = -6.898E+01 | dRMS = 1.165E-01 |
| RHF Iteration | 2: Energy = -69.64725442845806 | dE = -6.672E-01 | dRMS = 1.074E-01 |
| RHF Iteration | 3: Energy = -72.84030309363035 | dE = -3.193E+00 | dRMS = 1.039E-01 |
| RHF Iteration | 4: Energy = -72.89488390650019 | dE = -5.458E-02 | dRMS = 8.660E-02 |
| RHF Iteration | 5: Energy = -74.12078064688371 | dE = -1.225E+00 | dRMS = 8.646E-02 |
| RHF Iteration | 6: Energy = -74.86718194576882 | dE = -7.464E-01 | dRMS = 6.528E-02 |
| RHF Iteration | 7: Energy = -75.41490878039029 | dE = -5.477E-01 | dRMS = 5.216E-02 |
| ... | | | |
| RHF Iteration | 22: Energy = -75.98979285429830 | dE = -3.768E-06 | dRMS = 1.188E-04 |
| RHF Iteration | 23: Energy = -75.98979450314014 | dE = -1.648E-06 | dRMS = 7.860E-05 |
| RHF Iteration | 24: Energy = -75.98979522446778 | dE = -7.213E-07 | dRMS = 5.198E-05 |

III. DIRECT INVERSION OF THE ITERATIVE SUBSPACE CONVERGENCE ACCELERATION

Direct Inversion of the Iterative Subspace (DIIS) is often used for convergence acceleration³ of SCF wavefunctions. For iterative diagonalization, we discard all previous guess vectors (orbitals) and generate completely new guess vectors at each iteration. DIIS keeps previous guess vectors and builds the best guess as a linear combination of all previous guess vectors.

For the explanation of DIIS, we will break with all conventional terminology to stress that this is a linear algebra technique that is widely applicable to many problems besides SCF, as we will see in the next several sections. Let us assume that we have generated several state vectors (s) with a series of complementary error vectors (r) through some iterative method.

For DIIS to be applicable, we assume that both the next residual and the next state vector can be built as a linear combination of previous guess vectors with coefficients c .

$$s_{i+1} = \sum_i^n c_i s_i \quad (18)$$

$$r_{i+1} = \sum_i^n c_i r_i \quad (19)$$

$$\sum_i^n c_i = 1 \quad (20)$$

Therefore, we want to select the coefficients in a manner that minimizes the norm of the residual

$$r_{i+1} \cdot r_{i+1} = \sum_{ij}^n c_i c_j (r_i \cdot r_j) \quad (21)$$

This can be solved through the following matrix equations

$$\begin{bmatrix} B_{11} & B_{12} & \dots & B_{1n} & -1 \\ B_{21} & B_{22} & \dots & B_{2n} & -1 \\ \dots & \dots & \dots & \dots & \dots \\ B_{n1} & B_{n2} & \dots & B_{nn} & -1 \\ -1 & -1 & \dots & -1 & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \dots \\ c_n \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \dots \\ 0 \\ -1 \end{bmatrix} \quad (22)$$

where B is a matrix of inner products between error vectors

$$B_{ij} = (r_i \cdot r_j) \quad (23)$$

and λ is value of the Lagrange multiplier. It should be noted that often our error vectors are in effect tensors, e.g., $F_{\mu\nu}$, and the norm is computed as the generalized inner product between tensors.

For SCF, the state vector is always the Fock matrix and the error vector is typically the orthonormalized orbital gradient expressed in a AO matrix formalism,

$$r_{\mu\nu} = (\mathbf{A}^T (\mathbf{FDS} - \mathbf{SDF}) \mathbf{A})_{\mu\nu} \quad (24)$$

The procedure for applying DIIS to SCF at every iteration is therefore:

1. Compute the Fock (s_n) matrix using \mathbf{C}_n orbitals
2. Construct the AO gradient (r_n)
3. Obtain the DIIS Fock guess (s_{DIIS})
4. Obtain \mathbf{C}_{n+1} orbitals by diagonalizing the DIIS Fock guess (s_{DIIS})

This procedure is iterated until the desired convergence thresholds are reached.

The DIIS convergence procedure is then applied to the "physicist's water molecule" in Fig. 2. It should be noted that while DIIS can be utilized on every iteration, the results are equal to pure diagonalization until two error vectors have been constructed. Therefore, the first iteration which benefits from DIIS acceleration is iteration three. As can be seen, roughly the same energy convergence is obtained in 9 iterations compared to 24 iterations without DIIS. The cost of DIIS itself is negligible; however, the cost of building the gradient is on the order of N^3 , which is still much less than the Coulomb and Exchange matrix builds.

FIG. 2. A RHF computation of "physicist's water molecule" starting with a core Hamiltonian guess in the cc-pVDZ basis set utilizing DIIS convergence.

| | | | |
|---------------|--------------------------------|-----------------|-----------------------|
| RHF Iteration | 1: Energy = -68.98003273414295 | dE = -6.898E+01 | dRMS = 1.165E-01 |
| RHF Iteration | 2: Energy = -69.64725442845806 | dE = -6.672E-01 | dRMS = 1.074E-01 DIIS |
| RHF Iteration | 3: Energy = -75.79192914624532 | dE = -6.144E+00 | dRMS = 2.892E-02 DIIS |
| RHF Iteration | 4: Energy = -75.97218922804181 | dE = -1.802E-01 | dRMS = 7.564E-03 DIIS |
| RHF Iteration | 5: Energy = -75.98936905846086 | dE = -1.717E-02 | dRMS = 8.749E-04 DIIS |
| RHF Iteration | 6: Energy = -75.98971633493079 | dE = -3.472E-04 | dRMS = 5.356E-04 DIIS |
| RHF Iteration | 7: Energy = -75.98979323982247 | dE = -7.690E-05 | dRMS = 6.212E-05 DIIS |
| RHF Iteration | 8: Energy = -75.98979567508871 | dE = -2.435E-06 | dRMS = 1.972E-05 DIIS |
| RHF Iteration | 9: Energy = -75.98979578301157 | dE = -1.079E-07 | dRMS = 1.727E-06 DIIS |

IV. DENSITY-FITTED COULOMB AND EXCHANGE ALGORITHMS

With density fitting⁴⁻⁷ the two-electron integrals are represented by the following

$$g_{\mu\nu\lambda\sigma} \approx (\mu\nu|\widetilde{P})[J^{-1}]_{PQ}(\widetilde{Q}|\lambda\sigma) \quad (25)$$

where the Coulomb metric $[J]_{PQ}$ and the three center overlap integral $(\widetilde{Q}|\lambda\sigma)$ are defined as

$$[J]_{PQ} = \int P(\mathbf{r}_1) \frac{1}{r_{12}} Q(\mathbf{r}_2) d^3\mathbf{r}_1 d^3\mathbf{r}_2 \quad (26)$$

$$(\widetilde{Q}|\lambda\sigma) = \int Q(\mathbf{r}_1) \frac{1}{r_{12}} \lambda(\mathbf{r}_2) \sigma(\mathbf{r}_2) d^3\mathbf{r}_1 d^3\mathbf{r}_2 \quad (27)$$

To simplify the density fitting notation, the inverse Coulomb metric is typically folded into the three center overlap tensor

$$(P|\lambda\sigma) = [J^{-\frac{1}{2}}]_{PQ}(\widetilde{Q}|\lambda\sigma) \quad (28)$$

$$g_{\mu\nu\lambda\sigma} \approx (\mu\nu|P)(P|\lambda\sigma) \quad (29)$$

The Coulomb matrix can then be computed in $\mathcal{O}(N^2 N_{aux})$ operations:

$$\chi_P = (P|\lambda\sigma) D_{\lambda\sigma} \quad (30)$$

$$J[D_{\lambda\sigma}]_{\mu\nu} = (\mu\nu|P) \chi_P \quad (31)$$

The Exchange matrix can be computed in $\mathcal{O}(N^3 N_{aux})$ operations:

$$\zeta_{P\nu\lambda} = (P|\nu\sigma) D_{\lambda\sigma} \quad (32)$$

$$K[D_{\lambda\sigma}]_{\mu\nu} = (\mu\lambda|P) \zeta_{P\nu\lambda} \quad (33)$$

however, considering the form of the density matrix, we can reduce this to $\mathcal{O}(pN^2N_{aux})$

$$D_{\lambda\sigma} = C_{p\sigma}C_{p\lambda} \quad (34)$$

$$\zeta_{P\mu p}^1 = (\mu\sigma|P)C_{p\sigma} \quad (35)$$

$$\zeta_{P\nu p}^2 = (P|\nu\lambda)C_{p\lambda} \quad (36)$$

$$K[D_{\lambda\sigma}]_{\mu\nu} = \zeta_{P\mu p}^1\zeta_{P\nu p}^2 \quad (37)$$

where p is a generalized MO index that can span any space. This technique is especially beneficial when p is an inactive or active index as these are typically small relative to the full N space. It should be noted that Eqs. 6.19-6.22 pertain to the generalized case where $C_{p\sigma} \neq C_{p\lambda}$, if both C matrices are identical, only one ζ intermediate needs to be built. The $C_{p\sigma} \neq C_{p\lambda}$ case often arises in SCF theory when either rotated, transition, or generalized density matrices are used.

Computation of the Coulomb and Exchange matrices through conventional means costs N^4 , here we see that Coulomb builds are rank reduced, but Exchange builds are of the same rank. The real benefit of density-fitted integrals for K builds comes from data locality and data storage. As N_{aux} is typically twice the size of N , we can imagine a case with 3000 AO basis functions. Conventional 4-index two-electron integrals would use 81 TB if the 8-fold symmetry of the tensor is exploited, on the other hand, the 3-index density-fitted tensor would only take up 216 GB if the 2-fold symmetry is exploited, a reduction of 375 fold.

V. SECOND-ORDER ORBITAL OPTIMIZATION

Diagonalization builds completely new orbitals at every iteration; an alternative is to construct new orbitals by rotating the orbitals of the previous iteration. This is often called second-order optimization and, if done exactly, the convergence will be quadratic with respect to the MO gradient⁸, e.g., $\text{grad}_{pq}^{n+1} \approx (\text{grad}_{pq}^n)^2$. As a consequence of the quadratic convergence, the second-order procedure can only be started once all elements of the gradient in the MO basis are less than one, else the error will increase.

The overall goal is to find an orbital rotation matrix, κ , that satisfies the following set of linear equations:

$$\mathbf{E}^{(2)}\kappa = -\mathbf{E}^{(1)} \quad (38)$$

where $\mathbf{E}^{(2)}$ is our orbital Hessian and $\mathbf{E}^{(1)}$ is the orbital gradient. The κ matrix is a antisym-

metric matrix of non-redundant rotations that describes the unitary transformation \mathbf{U}

$$\mathbf{U} = e^{\boldsymbol{\kappa}} \quad (39)$$

$$\mathbf{C}_{n+1} = \mathbf{C}_n \mathbf{U} \quad (40)$$

Rotations within a given space (inactive or virtual for SCF) are redundant: the only non-redundant orbital rotations for SCF are the inactive-virtual ones.

The orbital gradient can be written as

$$E_{pq}^{(1)} = 2(F_{pq} - F_{qp}) \quad (41)$$

where F is our generalized Fock matrix⁸

$$F_{in} = 2(^A F_{in} + ^I F_{in}) \quad (42)$$

$$F_{vn} = ^I F_{nw} \gamma_{vw} + Q_{vn}, \quad Q_{vn} = \Gamma_{vwxy} g_{nwx} y \quad (43)$$

$$F_{an} = 0 \quad (44)$$

and the Active ($^A F$) and Inactive ($^I F$) Fock matrices are defined as

$$^I F_{pq} = H_{pq} + 2g_{pqrs} D_{rs} - g_{prqs} D_{rs} \quad (45)$$

$$^A F_{pq} = \gamma_{tu} (g_{pqtu} - \frac{1}{2} g_{ptuq}) \quad (46)$$

In the case of RHF, the Inactive Fock matrix is identical to the AO Fock matrix given in Eq. 5, this equation can be rewritten considering the special form of the RHF density matrix with MO indices

$$^I F_{pq} = H_{pq} + 2g_{pqii} - g_{piqi} \quad (47)$$

The orbital Hessian can be built explicitly; however, this matrix is quite large and it is typically sufficient to know the product of the Hessian with a given vector. To do so, we will introduce the rotated Hamiltonian

$$\hat{H}^\kappa = h_{pq}^\kappa \gamma_{pq} + g_{pqrs}^\kappa \Gamma_{pqrs} \quad (48)$$

and the one-index transformed integrals

$$h_{pq}^\kappa = (\kappa_{po} h_{oq} + \kappa_{qo} h_{po}) \quad (49)$$

$$g_{pqrs}^\kappa = (\kappa_{po} g_{oqrs} + \kappa_{qo} g_{pors} + \kappa_{ro} g_{pqos} + \kappa_{so} g_{pqro}) \quad (50)$$

Applying this Hamiltonian to our SCF wavefunction results in a rotated Fock matrix (F^κ) where all one- and two-electron integrals have been replaced by their respective one-index transformed counterparts. This rotated Fock matrix is the Hessian vector product ($\mathbf{E}^{(2)}\boldsymbol{\kappa}$) and can be substituted for the LHS of Eq. 38:

$$2(F_{pq}^\kappa - F_{qp}^\kappa) = -2(F_{pq} - F_{qp}) \quad (51)$$

Considering the special structure of the generalized Fock matrix for RHF calculations (only inactive-virtual rotations are non-redundant and F_{ai} is zero), we can write our set of linear equations for the last time as

$$-4^I F_{ia}^\kappa = 4^I F_{ia} \quad (52)$$

which must be solved iteratively. As shown below, this is the most computationally efficient form of these equations. For demonstration purposes, the $^I F_{ia}^\kappa$ tensor need not be contracted with $\boldsymbol{\kappa}$ to yield iterative form shown in Eq. 52. Instead, the full 4-index orbital Hessian can be constructed and Eq. 38 can then be solved exactly. However, when solved exactly, the cost of inverting the full orbital Hessian is $o^3 v^3$, unless o is very small, we have effectively changed the cost of SCF from N^4 to N^6 , or roughly the cost of CCSD. Solving either Eq. 38 or Eq. 52 will be denoted second-order SCF (SOSCF).

FIG. 3. A RHF computation of "physicist's water molecule" starting with a core Hamiltonian guess in the cc-pVDZ basis set utilizing DIIS and SOSCF convergence. For the SOSCF step, the Eq. 38 has been solved exactly by inverting the full orbital Hessian. The right most abbreviations indicate the type of step taken in each iteration.

| | | | | |
|---------------|--------------------------------|-----------------|------------------|-------|
| RHF Iteration | 1: Energy = -68.98003273414295 | dE = -6.898E+01 | dRMS = 1.165E-01 | |
| RHF Iteration | 2: Energy = -69.64725442845806 | dE = -6.672E-01 | dRMS = 1.074E-01 | DIIS |
| RHF Iteration | 3: Energy = -75.79192914624532 | dE = -6.144E+00 | dRMS = 2.892E-02 | DIIS |
| RHF Iteration | 4: Energy = -75.97218922804181 | dE = -1.802E-01 | dRMS = 7.564E-03 | DIIS |
| RHF Iteration | 5: Energy = -75.98973929215161 | dE = -1.755E-02 | dRMS = 3.049E-04 | SOSCF |
| RHF Iteration | 6: Energy = -75.98979578473095 | dE = -5.649E-05 | dRMS = 1.231E-06 | SOSCF |
| RHF Iteration | 7: Energy = -75.98979578551825 | dE = -7.873E-10 | dRMS = 1.901E-11 | SOSCF |
| RHF Iteration | 8: Energy = -75.98979578551835 | dE = -9.948E-14 | dRMS = 2.395E-15 | SOSCF |

Returning to our "physicist's water molecule", the convergence pattern for SOSCF is demonstrated in Fig. 3. We can observe that three regular DIIS steps were required before all elements of the gradient were less than one. The RHF wavefunction is then converged to within machine precision in terms of energy and density within 4 SOSCF iterations. SOSCF is quite beneficial in terms of the number of iterations, but is not currently competitive in terms of cost.

Let us return our attention to Eq. 52 and to the implementation of an efficient rotated Inactive Fock (${}^IF^\kappa$) build

$${}^IF_{mn}^\kappa = h_{mn}^\kappa + 2g_{mnii}^\kappa - g_{miin}^\kappa \quad (53)$$

Inserting our equations for one-index transformed integrals, we obtain

$$\begin{aligned} {}^IF_{mn}^\kappa = & (\kappa_{mp}h_{pn} + \kappa_{np}h_{mp}) + \\ & 2(\kappa_{mo}g_{onii} + \kappa_{no}g_{moui} + \kappa_{io}g_{mnio} + \kappa_{io}g_{mnio}) - \\ & (\kappa_{mo}g_{oiin} + \kappa_{io}g_{moin} + \kappa_{io}g_{mion} + \kappa_{no}g_{miio}) \end{aligned} \quad (54)$$

Simplifying and collecting these terms yields

$${}^IF_{mn}^\kappa = ({}^IF_{mp}\kappa_{np} + {}^IF_{pn}\kappa_{mp}) + \kappa_{ip}(4g_{mnip} - g_{mpin} - g_{npim}) \quad (55)$$

which can be computed using conventional AO-based J and K routines:

$$\vartheta_{\lambda\sigma} = C_{i\lambda}\kappa_{ip}C_{p\sigma} \quad (56)$$

$$\begin{aligned} {}^IF_{mn}^\kappa = & ({}^IF_{mp}\kappa_{np} + {}^IF_{pn}\kappa_{mp}) \\ & + C_{m\mu}(4J[\vartheta_{\lambda\sigma}]_{\mu\nu} - K[\vartheta_{\lambda\sigma}]_{\mu\nu} - K[\vartheta_{\lambda\sigma}]_{\nu\mu})C_{n\nu} \end{aligned} \quad (57)$$

The cost of solving Eq. 52 is now N^4 when the left hand side is computed from Eq. 57 and each iteration of Eq. 52 (often called a microiteration) is equivalent to the cost of a normal RHF iteration. The term macroiteration refers to a step in which the overall SCF energy is computed, this step collects all Fock builds, microiterations, and orbital rotations.

FIG. 4. A RHF computation of "physicist's water molecule" starting with a core Hamiltonian guess in the cc-pVDZ basis set utilizing DIIS and SOSCF convergence. For the SOSCF step, Eq. 52 has been solved iteratively through a conjugate gradient method limited to four microiterations. The rightmost abbreviations indicate the type of step taken on each iteration.

| | | | |
|---------------|--------------------------------|-----------------|------------------------|
| RHF Iteration | 1: Energy = -68.98003273414295 | dE = -6.898E+01 | dRMS = 1.165E-01 |
| RHF Iteration | 2: Energy = -69.64725442845806 | dE = -6.672E-01 | dRMS = 1.074E-01 DIIS |
| RHF Iteration | 3: Energy = -75.79192914624532 | dE = -6.144E+00 | dRMS = 2.892E-02 DIIS |
| RHF Iteration | 4: Energy = -75.97218922804181 | dE = -1.802E-01 | dRMS = 7.564E-03 DIIS |
| RHF Iteration | 5: Energy = -75.98970327666461 | dE = -1.751E-02 | dRMS = 5.908E-04 SOSCF |
| RHF Iteration | 6: Energy = -75.98979576713703 | dE = -9.249E-05 | dRMS = 8.777E-06 SOSCF |
| RHF Iteration | 7: Energy = -75.98979578513478 | dE = -1.799E-08 | dRMS = 8.237E-07 SOSCF |
| RHF Iteration | 8: Energy = -75.98979578551806 | dE = -3.832E-10 | dRMS = 1.648E-08 SOSCF |

Our "physicist's water molecule" was again optimized using iterative SOSCF in Fig. 4. We observe that while the convergence is significantly faster than for the DIIS acceleration alone, the

overall convergence is much slower than when Eq. 52 was solved exactly. This is due to the fact that the number of microiterations used is insufficient to solve Eq. 52 exactly. More microiterations can be utilized, but a balance between macro- and microiterations must be computationally efficient. However, it should be noted that even though each microiteration of SOSCF is equivalent to the cost of a normal RHF step, the overall computational effort expended for SOSCF is generally greater than the DIIS methods for a given convergence criterion. SOSCF should therefore only be utilized for difficult to converge SCF cases.

REFERENCES

- ¹A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry*, Introduction to Advanced Electronic Structure Theory. Courier Corporation, 1996.
- ²I. N. Levine, *Quantum Chemistry*. Prentice-Hall, New Jersey, 5th edition, 2000.
- ³P. Pulay, Chem. Phys. Lett. **73**, 393 (1980).
- ⁴F. Weigend, Phys. Chem. Chem. Phys. **4**, 4285 (2002).
- ⁵O. Vahtras, J. Almlöf, and M. W. Feyereisen, Chem. Phys. Lett. **213**, 514 (1993).
- ⁶B. I. Dunlap, J. W. D. Connolly, and J. R. Sabin, J. Chem. Phys. **71**, 3396 (1979).
- ⁷J. L. Whitten, J. Chem. Phys. **58**, 4496 (1973).
- ⁸T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular electronic-structure theory*. John Wiley & Sons Inc, 2000.